춘계학술대회 프로시딩

한국지능정보시스템학회

2024년 춘계 학술대회

Human-Centric AI: Nurturing Innovtion, Sustainability, and Inclusivity

날짜 : 2024년 5월 10일(금) ~ 11일(토)

장소 : BPEX(부산항국제전시컨벤션센터)

주관 : 한국지능정보시스템학회

후 원

(주)우리아이티 Leading company in IT convergence
LG CNS
K INNOVATION
KISTI
KIWOONG INFORM & COMM
easyM Soft Easy Management Software Group
kt
대신정보통신(주)
MarkAny
woojoo telecom
투이컨설팅
Flexsys
SK broadband
쌍용정보통신주식회사 SsangYong Information & Communications Corp.
kakao
DAEBO
올포랜드
MT엠티데이타
영림원소프트랩
elim 세림TSG
WHITESCAN
ReDWit
(주)선도소프트

우리아이티를 통해 이루는 여러분의 Work Load

More Early | More Safely | More Fully

Partner

vmware SUSE AXGATE Security and Beyond CISCO

DELL Technologies NUTANIX N-N-N Cloud

WiT (주)우리아이티
Leading company in IT convergence
부산시 동구 조방로 14 범일동, 동일타워 415호
TEL 051-637-2386 | FAX 050-5964-5555
www.wooriit.kr

LG CNS

CloudXper
클라우드

DAP
AI빅데이터

Factova
스마트팩토리

Monachain
블록체인

INFioT
사물인터넷

SecuXper
보안서비스

Cityhub
스마트시티

별이 되겠습니다
길이 되겠습니다

클라우드, AI빅테이터, 스마트팩토리, 블록체인은 물론,
사물인터넷, 보안, 스마트시티, 스마트물류, 금융DX 까지 -
밤하늘에 빛나는 북두칠성이 길잡이가 되어주듯,
LG CNS의 7개 전략적 핵심 브랜드가 다가올 미래에
DX 혁신의 길잡이가 되고 있습니다

DIGITAL GROWTH PARTNER LG CNS

INNOVATION

GROW YOUR BUSINESS WITH US!

수출을 희망하시는 중소기업을 찾고있습니다.
글로벌 시장의 선두 주자, JK INNOVATION과 함께
성장의 기회를 잡으세요.

1. 글로벌 파트너사 확보
2. 수출 희망 중소기업의 파트너
3. 글로벌 시장연결의 가교

토목건설장비 부대품

- 토목건설장비 및 부대품 수출
- 전세계 20개국 바이어 확보
- 풀림방지볼트 브라질 신규 수출
- 특수목적용 장비 수출 최적화

의류 부자재

- 재귀반사제품 및 씸씰링 테이프 수출 혁신적인 원단이나 필름 소재
- 수출상담환영

30개국 수출경험

- 유럽, 미국, 동남아시아, 중동, 브라질 등 30여 개국 수출 경험
- 현재 20여 개국 수출중
- 수출상담 대 환영

055-321-4055 www.jkinnovation.com info@jkinnovation.com

각종 자금 사고에 노출된 아파트관리비 집행!

아파트 관리사무소는 각종 규제와 업무 처리에 어려움을 겪고 있고,
입주민들은 혹시나 사고가 나지 않을까 하는 불신감을 갖고 있습니다.

아파트뱅크는 전자금융 시스템과 전자장부 시스템을 통합하여
투명한 자금집행이 가능한 클린 자금관리서비스입니다.

이체 관련 이상거래가 있는지, 모니터링하여 보고하는 체계를 갖추고 있습니다.

아파트관리비를 위한 디지털혁신 서비스!

아파트뱅크

전국의 1,000만 세대가 관리비 사고의 걱정 없이, 열악한 아파트 관리 부담을 크게 줄인
아파트뱅크서비스를 국내 최초로 아파트에 적용할 수 있게 되었습니다.

KB국민은행의 지원으로 원하는 모든 아파트는 무료로 이용할 수 있습니다.

주요 기능

아파트ERP 연계를 통한
자금조회 및 이체 / 지출처리 / 이상거래 자동감지 및 보고 / 전자증빙 자동 보관

특장점

디지털 소외 영역에 디지털 혁신 체감
은행 지점 방문 불필요 / 아파트관리 업무 대폭 간소화 / 무상 이용 가능 (은행 지원)

본 서비스는 기웅정보통신㈜이 제공하고 있습니다.

기웅정보통신은 10년 이상 축적해온 차별적인 데이터 처리 기술과 지능형 업무 처리 역량을 활용하여
영리·비영리 사업자를 위한 BPA[1]자동화 플랫폼과 서비스를 제공하는 전문기업입니다.

주1) BPA (Business Process Automation): 태스크 중심의 단순 정형업무 자동화하는 RPA(Robotics Process Automation)가 진화된 개념으로, 비정형 업무와 Task를 묶는 업무 절차 전체를 자동화하여 디지털 혁신을 지원하는 기술 (Gartner)

KIWOONG INFORM & COMM

문의전화 1566-0174

요양시장 유일한 요양 ERP

「인사-급여-평가관리-세무·재무-노무-리포트 관리」를
요양시장에서 유일하게 One-Stop으로 제공

자체 개발한 요양 전용 솔루션 보유

기관의 인증서 없이 「공단·W4C·희망이음」 등
보건복지부 시스템과 직접 연계 가능

고객의 목소리로 만들어진 프로그램

수년간 이용자의 필요한 기능과 요구사항을
반영하여 성장하였으며, 서비스를 넘어 플랫폼으로 진화 中

문의전화

1644-0340

이지케어를 검색하세요!

이지케어

ezcare.easyms.co.kr

easyM Soft
Easy Management Software Group

큰大믿을信

대신정보통신주식회사 Daishin Information & Communications Co., Ltd.

서울특별시 금천구 가산디지털2로 169-16, 6층 (가산동, 하우스디가산퍼스타) Tel_02-2107-5000 Fax_02-2107-0515 www.dsic.co.kr

고객의 더 나은 삶을 위하여
미래를 만들어갑니다

미래의 교통수단 초고속 철도가
더 안전하고 빠르게 달릴 수 있게

전국을 연결하는 고속도로가
더 좋은 주행 환경을 갖출 수 있게

대한민국을 이끌어갈 청소년들이
더 큰 꿈을 키울 수 있게

어떤 재난 상황에도 모두가
더 안전할 수 있게

우주텔레콤의 도전은 계속됩니다.

통합공공망(LTE-X)

철도통합무선망(LTE-R) 및
해상무선 통신망(LTE-M),
재난안전통신망 (PS-LTE) 등
통합공공 무선망 구축

- 유/무선(LTE/WIFI) 통신망
- 정보시스템 개발

정보/통신 인프라

네트워크 설계 및 기반
인프라 구축으로 운용시스템
최적화

- 유/무선(LTE/WIFI) 통신망
- IOT 설비/LoRa 통신망

자율협력주행

최적의 자율주행 환경을
위한 통신환경 설계 및 구축

- C-ITS
- Big Data & AI

학교무선통신망

전국 초·중학교 대상
무선인프라 확충 및
운영체계 구축

- 유/무선(LTE/WIFI) 통신망
- 통신인프라 고도화

재난안전 시스템

재난안전통신망 전환을 위한
치안업무용 무선시스템 구축 및
긴급상황센터, 13개 검역소
상황실 시스템 구축

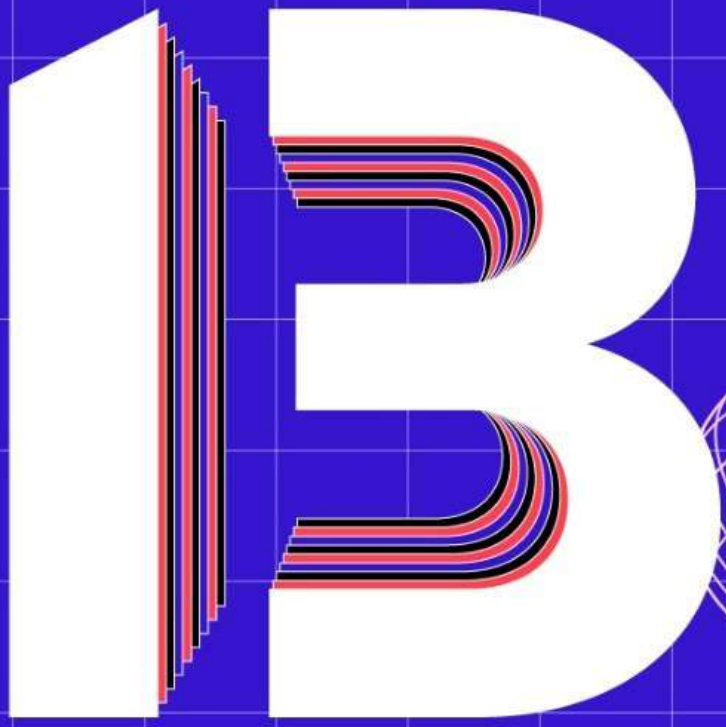
- 통합전산센터 환경 구축
- 데이터 센터 설계

SK broadband

NCSI 국가고객만족도
초고속인터넷 / IPTV
13년 연속 1위

BRAVO!

SK Broadband



IPTV, 초고속인터넷 2개부문

국가고객만족도 13년 연속 1위

감사합니다. 한결같은 사랑에 더 큰 감동으로 보답하겠습니다.

SK broadband

디지털트랜스포메이션 너머의 미래
쌍용정보통신과 함께 설계하세요!

No.1 Cloud Native IT Service Company

DX Innovator, 클라우드 · AI 전문가 그룹

쌍용정보통신주식회사
SsangYong Information & Communications Corp.
member of ITCenGroup

아이티센그룹
ITCenGROUP

모두가 행복한 ICT 세상!

풍요롭고 행복한 세상을 만드는 ICT기업

대보정보통신

IT컨설팅에서 시스템 통합 (SI) 및 유지관리 (SM)까지 고객이 필요로 하는 최적화된 솔루션과 서비스로 고객의 성공비즈니스를 만들어 갑니다.

System Integration

최적화된 시스템을 구축합니다.

Industry Solution (전자정부, 공공, 공항, 국방, 교육 등) /
Ubiquitous Solution (U-City, ITS 등) / Network Intergration

System Management

IT시스템의 효율을 극대화합니다.

IT Outsourcing (시스템, 데이터센터, 보안서비스) / ITS 운영관리 /
장대터널 운영관리

Solutions

한발 앞선 솔루션을 제공합니다.

ITS기반 솔루션 / Smart Highway / 데이터 분석 / 검색엔진 / 보안 /
Mobile App / U-BIZ / Hi-pass

대보정보통신(주) | **본사** Add. 서울시 강남구 광평로 280 로즈데일빌딩 6층 I Tel. 02.3470.7700 I Fax. 02.3470.7719 I www.dbcs.co.kr
대보그룹 대보건설(주) 대보실업(주) 대보유통(주) 서원밸리컨트리클럽

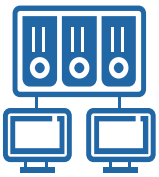
MT엠티데이타

고객과 함께 행복한 IT미래를 만들어 가겠습니다.

SI (System Integration)

정보시스템 구축·운영

Cloud·VDI 구축



IT Consulting

BigData

AI Solution·Service

IT Service

IT Outsourcing

IDC·전산실 구축

본사 / 경기도 성남시 분당구 판교로255번길9-22,514호(삼평동,우림W-CITY)

대구지사 / 대구광역시 달서구 성서로 329, 703~704호(갈산동, 동원비즈플랫폼)

울산지사 / 울산광역시 북구 명촌7길 15, 2층(명촌동)

부산지사 / 부산광역시 금정구 식물원로9번길 29(장전동)

www.mtdata.co.kr TEL : 031-628-8180 FAX : 031-628-8188

[전체 프로그램]

시간	순서					
08:30-09:00	등록 및 접수			이사회		
	Session A 컨퍼런스홀 A	Session B 컨퍼런스홀 F	Session C 회의실 1	Session D 회의실 2	Session E 회의실 3	Session F 회의실 4
09:00-10:20	**A1** **학술세션**	**B1** **학술세션**	**C1** **학술세션**	**D1** **학술세션**	**E1** **학술세션**	**F1**
	AI와 인간의 상호작용 좌장: 안현철 (국민대)	디지털 데이터를 활용한 소비자 인사이트 분석 좌장: 이형용 (한성대)	언어 모델의 최적화와 응용 좌장: 강주영 (아주대)	AI기반 추천시스템 좌장: 고준 (전남대)	데이터 주도 에너지 및 환경 정책 연구 좌장: 방영석 (연세대)	이사회
10:20-10:30	Coffee Break					
10:30-11:50	총회 및 개회식 (컨퍼런스홀 A)					
11:50-13:10	점심식사(3층)					
13:10-14:30	A2 [튜토리얼]	B2 [산업체세션]	C2 [특별세션]	D2 [특별세션]	E2 [특별세션]	F2 [특별세션]
	대학원생을 위한 연구방법론 좌장: 김태경 (경희대)	취창업을 위한 산업체 동향 좌장: 한동욱 (전주대)	AI 신뢰성의 과거 및 현재 이슈와 미래 전망 좌장: 허용석 (영남대)	안전한 의료데이터 활용 좌장: 한현욱 (차의과대학)	데이터로 보는 글로벌 R&D 인사이트 좌장: 전승표 (KISTI)	Comparative Genomics 2.0 and AI (I) 좌장: 유혁 (고려대)
14:30-14:50						
14:50-16:10	A3 [튜토리얼]	B3 [학술세션]	C3 [학술세션]	D3 [학술세션]	E3 [특별세션]	F3 [특별세션]
	신진연구자들을 위한 텍스트 마이닝의 활용 좌장: 김은미 (부산대)	블록체인 기술의 응용 좌장: 김호원 (부산대)	디지털 금융 혁신: AI와 딥러닝 좌장: 고지훈 (부산대)	인공지능과 4차산업혁명 공진화 좌장: 권오병 (경희대)	과학기술사업화 AI플랫폼 개발과 활용전략 좌장: 최윤정 (KISTI)	Comparative Genomics 2.0 and AI (II) 좌장: 정대균 (경희대)
16:10-16:25						
16:25-17:45	A4 [학술세션]	B4 [학술세션]	C4 [학술세션]	D4 [학술세션]	E4 [학술세션]	F4 [학술세션]
	디지털 미디어와 소비자 행동 좌장: 장문경 (가천대)	비지도 학습을 통한 이상 탐지 기술의 혁신 좌장: 이철호 (KAIST)	인공지능을 활용한 지식 기반 질의 응답과 사용자 경험 최적화 좌장: 양성병 (경희대)	산업별 데이터 기반 수요 예측 모델 좌장: 류민호 (동아대)	AI를 활용한 의료 및 건강 데이터의 혁신적 응용 좌장: 김상진 (동아대)	AI 기반 첨단 이미지 및 패턴 분석 기술 좌장: 김은미 (부산대)
17:50-18:10	폐회식(컨퍼런스홀 A)					

[세부 프로그램]

Session A (컨퍼런스홀A)

A1. [학술세션] AI와 인간의 상호작용 (09:00~10:20) ---------- 좌장: 안현철(국민대)

A1.1 A Study on the Influence of ChatGPT Characteristics on Acceptance Intention in the Financial Sector: Exploring the Nexus between ChatGPT and Finance with a Focus on Moderation Effects ····· 1

박수진(경기대)

A1.2 The Impact of AI Utilization and Digital Literacy on User Experience in Digitalized Tourism Platforms ····· 2

강명준, 김민서, 구철모(경희대)

A1.3 Hello, I'm Here: The impact of Social Chatbot Conversation Design oin Mental Health ····· 6

민동규, 이철호 (KAIST), 정성훈 (GWU), Xitong Li (HEC Paris)

A1.4 기술수용모델(TAM)을 통한 디지털 휴먼 이용의도에 관한 연구 -준사회적 상호작용을 중심으로- ····· 9

고상훈, 정경호, 안현철(국민대)

A1.5 중소기업의 생성형 인공지능 챗봇 서비스 전환의도에 대한 연구 : PPM(Push-Pull-Mooring) 프레임워크 기반- ····· 11

조종현, 성행남, 홍태호(부산대)

A2. [튜토리얼] 대학원생을 위한 연구방법론 (13:10~14:30) ---------- 좌장: 김태경(경희대)

A2.1 행동 연구의 최신 동향과 방법론 ·····

양성병(경희대)

A2.2 기술적 연구의 최신 동향과 방법론 ·····

김태경(경희대)

A3. [튜토리얼] 신진연구자를 위한 텍스트마이닝의 활용 (14:50~16:10) ---------- 좌장: 김은미(부산대)

A3.1 텍스트마이닝을 활용한 온라인 리뷰의 감성분석 ·····

김은미(부산대)

A3.2 텍스트 마이닝 최신 기법의 동향과 이해 ·····

전성민(가천대)

A3.3 텍스트 마이닝을 이용한 응용 연구 ·····

안재영(연세대)

A4. [학술세션] 디지털 미디어와 소비자 행동 (16:25~17:45)---------- 좌장: 장문경(가천대)

A4.1 SNS 어포던스가 선의적 선망과 악의적 선망에 미치는 영향 ····· 12
이혜원, 이새롬(경북대), 곽동헌(Kent State Univ.)

A4.2 OTT 플랫폼 고착화와 구독 지속 의도에 관한 연구 ····· 15
이서현, 이새롬(경북대)

A4.3 숏폼 영상의 특성이 중독에 미치는 영향 및 결과 : 상대과정 이론 관점 ····· 17
한민정, 이새롬(경북대)

A4.4 Keep an Eye on Me: The Role of Online Communities as a Commitment for Healthier Behavior ····· 20
김가윤, 민동규, 이철호(KAIST)

A4.5 리뷰 작성자의 문화적 배경요인이 리뷰 유용성에 미치는 영향 ····· 21
김태이, 올가 체르냐예바, 야오즈옌, 홍태호(부산대)

Session B (컨퍼런스홀F)

B1. [학술세션] 디지털 데이터를 활용한 소비자 인사이트 분석 (09:00~10:20)---- 좌장: 이형용(한성대)

B1.1 Exploring the impact of informational factors in online reviews on review helpfulness ····· 22
야오즈옌, 홍태호, 김은미(부산대), 김종원(동의대)

B1.2 Enhancing Fake Review Detection through Verification Status and Explainable AI ····· 23
Chernyaeva Olga , Hong Taeho (부산대), Kim, Jong-weon (동의대)

B1.3 SNS 분석 기반 팝업스토어 고객 경험 분석 및 몰입개념을 통한 성공요인 분석 ····· 25
황지선, 황은혜, 강주영(아주대)

B1.4 A Study on Factors Affecting Academic Achievement of College Students : Focusing on Fuzzy set Qualitative Comparative Analysis- ····· 26
권승인, 박영식, 이형용(한성대), 박정민(한국발명진흥회)

B2. [산업체세션] 취 • 창업을 위한 산업체 동향 (13:10~14:30) ---------- 좌장: 한동욱(전주대)

B2.1 No-code/Low-code 개발 툴 Flextudio로 수준 높은 앱 개발하기 ····· 27
정한중(영림원소프트랩)

B2.2 xEV 전기차 부품 '모터풀리' 외관검사 및 불량검출 인공지능 시스템 개발 ····· 43
조영우,장마가(엠티데이타)

B2.3 국내 법률 대상 다국어 번역을 위한 인공지능 학습 모델 연구 ····· 44
이정기(타임게이트), 정회창, 전민기(엠티데이타)

B2.4 블록체인 기반 연합학습 리소스 공유 시스템 ····· 45
김요한(스마트엠투엠)

B2.5 (제조)기업 사례를 통한 AI, 빅데이터 적용방안 ·····
김도현(일주지앤에스)

B3. [학술세션] 블록체인 기술의 응용 (14:50~16:10) ---------- 좌장: 김호원(부산대)

B3.1 블록체인 기반 ESG 정보 공시 플랫폼 구축 방안 연구 ····· 46
최하눌(계명대)

B3.2 블록체인 기반 배터리 여권 기술 동향 분석 ····· 47
이상명, 김재현, 정한호, 조재한, 김호원(부산대)

B3.3 데이터 양자화를 통한 FPGA기반 이진 신경망 경량 설계 ····· 50
윤동욱, 강은세, 김호원(부산대)

B3.4 블록체인 기반 LLM 생성 컨텐츠 필터링 및 모니터링 시스템 ····· 53
김재현, 이상명, 조재한, 정한호, 김호원(부산대)

B3.5 전문가 혼합 신경망 적용을 통한 객체 탐지 트랜스포머 성능 개선 방안 연구 ····· 56
박지성, 김용수, 강효은(스마트엠투엠), 김호원(부산대)

B4. [학술세션] 비지도 학습을 통한 이상 탐지 기술의 혁신 (16:25~17:45) ------ 좌장: 이철호(KAIST)

B4.1 멀티 모달 데이터를 위한 Flow-based Generative Models 기반 비지도 학습 다변량 시계열 이상 탐지 알고리즘 ····· 59
하은겸, 김우주(연세대)

B4.2 트랜스포머 기반 오토인코더를 활용한 전류 센서 데이터의 비지도 학습 이상탐지 ····· 60
박종호, 이제혁(국민대)

B4.3 EfficientAD의 손실함수를 활용한 비지도 이상 탐지 성능 개선 ····· 63
김명섭, 오홍규, 이승재, 고병진, 윤종완, 박태준(한양대)

B4.4 특징 어댑터를 사용한 비지도 이상 탐지를 위한 Reverse Distillation 모델의 편향 완화 … 64
오홍규, 김명섭, 이승재, 고병진, 윤종완, 박태준(한양대)

Session C (회의실 1)

C1. [학술세션] 언어모델의 최적화와 응용 (09:00~10:20) ---------- 좌장: 강주영(아주대)

C1.1 한국어 특화 대규모 언어 모델의 토크나이저 심층분석 ····· 65
최창우, 김용수, 강효은(스마트엠투엠), 김호원(부산대)

C1.2 Text-to-Gloss를 위한 사전학습 언어모델 단어사전 제어 연구 ····· 68
이명진, 김우영, 김태용, 손의성, 김우주(연세대)

C1.3 LLM과 RAG에 기반한 대화형 매칭 에이전트 프롬프트 엔지니어링 과정과 최적화 이슈 ····· 69
이경전, 윤이지, 이수영, 정백, 심민준, 정규윤, 옥근우(경희대), 안은정(서울여대),김담, 박동주(하렉스인포텍)

C1.4 텍스트 마이닝을 활용한 반려동물 관련 연구의 동향 연구 ····· 71
이동현, 강주영(아주대)

C2. [특별세션] AI 신뢰성의 과거 및 현재 이슈와 미래 전망(AI 신뢰성 연구회 창립 총회 및 창립 세미나) (13:10~14:30) ---------- 좌장: 허용석(영남대)

C2.1 AI 신뢰성 연구회 창립 총회 ·····
허용석(영남대)

C2.2 AI 신뢰성의 과거 및 현재 이슈와 미래 전망(정책 분야 전문가 초청 강연 및 패널 토론) ·····
조원영(소프트웨어 정책연구소)

C2.3 AI 신뢰성의 과거 및 현재 이슈와 미래 전망(산업 분야 전문가 초청 강연 및 패널 토론) ·····
정창모(현대모비스)

C3. [학술세션] 디지털 금융 혁신: AI와 딥러닝 (14:50~16:10) ---------- 좌장: 고지훈(부산대)

C3.1 Anchoring bias and the cross-section of stock return patterns ····· 72
김동훈(KAIST), 고지훈(부산대)

C3.2 ChatGPT의 지시문 조정을 통한 한국 주식시장 투자 전략 구축 ····· 74
이재욱, 도영훈, 정다훈, 김태현, 이현빈, 이민혁(부산대)

C3.3 Designing Mobile Message to Reduce Loan Repayment Default in Microfinance: Evidence from Field Experiment ····· 76
양우진, Nigusse Gebre, 이철호(KAIST)

C3.4 동백전 빅데이터를 활용한 인공지능 기반 신흥 상권 도출에 관한 연구 ····· 77
송위창, 김상진(동아대)

C3.5 김치 프리미엄 현상과 가상화폐 2차 시장 간의 매개효과 ····· 80
조승연,임규건(한양대)

C4. [학술세션] 인공지능을 활용한 지식 기반 질의응답과 사용자 경험 최적화 (16:25~17:45) ---------- 좌장: 양성병(경희대)

C4.1 라이브커머스 카테고리별 소비자 행태 분석: 실시간 채팅을 활용한 탐색적 연구 …… 82
강은경, 김진수, 장하렴, 양성병(경희대)

C4.2 RGQA : Reasoning Guideline을 활용한 LLM 기반 KGQA …… 83
남정재, 이태훈, 왕수연, 김우주(연세대)

C4.3 Knowledge-Augmented Prompting Noise 감소 전략 : Large Language Models을 활용한 KGQA의 성능 향상을 중심으로 …… 84
정세영, 김대식, 김우영, 김우주(연세대)

C4.4 인공지능 법령 질의응답시스템을 위한 규범 기반 사례 데이터 생성과 활용에 관한 연구 …. 85
박예린(경희대), 박가림(이화여대), 이유봉(한국법제연구원), 손권상(인하대)

Session D (회의실 2)

D1. [학술세션] AI기반 추천시스템 (09:00~10:20) ---------- 좌장: 고준(전남대)

D1.1 Distributional Reinforcement Learning for Recommender System Addressing Randomness in Returns ····· 87
류명진, 강윤철(이화여대)

D1.2 The effect of applying Aristotle's rhetoric on persuasive cues in AI's travel recommendation on user trust ····· 90
정한나, 구철모(경희대), 신승훈(HongKong Polytechnic Univ.)

D1.3 딥러닝 추천시스템: Wide & Deep 모델의 적용 ····· 95
박준석, 조현진, 최지훈, 이지윤(국민대)

D1.4 Advancing Tourist Experience Through Personalized Menu Recommendations with Auto-Tagging ····· 96
김민혜, 박병학, 이호진(레드테이블), 이민정(세종사이버대), 이주민(경희사이버대)

D1.5 Firm's cloud adoption decision according to environmental policy ····· 97
서정권, 이철호, 이나래(KAIST)

D2. [특별세션] 안전한 의료데이터 활용 (13:10~14:30) ---------- 좌장: 한현욱(차의과대학)

D2.1 의료데이터 안전활용 정책 ····· 99
심동욱(한국인터넷진흥원)

D2.2 의료기기 네트워크 환경의 사이버보안 위협분석 기술 ····· 124
정일안(휴네시온)

D2.3 자연어처리 기반 우울증 중증도 진단 분류 모델 개발 ····· 135
김명관(차의과대학)

D2.4 In-context learning for bioelectrical signals ····· 146
최호식(서울시립대)

D2.5 안전한 의료데이터 활용을 위한 PET: 동형암호 활용사례 ····· 168
신성철(크립토랩)

D3. [학술세션] 인공지능과 4차산업혁명 공진화 (14:50~16:10) ---------- 좌장: 권오병(경희대)

D3.1 식품 위해 사전과 모델 성능 고도화 방안을 통한 수입식품 위해도 예측모형 성능 제고 ····· 178
이다현, 이은서, 권상지, 신윤종, 김정연, 권오병(경희대)

D3.2 Engaging Audiences with Varicose Veins Information: Insights from YouTube Content Analysis ····· 180
이홍주, 김연정, 이준우(가톨릭대)

D3.3 생성형AI에 의한 이미지 생성이 미술관 방문의도에 미치는 영향연구: 작가와 미술관 조절효과를 중심으로 ····· 182
최륜희, 이은서, 권오병(경희대)

D3.4 가짜뉴스가 주식시장에 미치는 영향: 과거 사례의 조절효과 ····· 186
박준휘, 김종우(한양대)

D3.5 메타데이터 기반 3D 애셋 검색 기능 강화를 위한 사용자 친화적 추천 시스템 설계 방안 연구 ····· 188
권상지, 허지혜, 이은서, 최상호, 최륜희, 권오병(경희대)

D4. [학술세션] 산업별 데이터 기반 수요 예측 모델 (16:25~17:45) ---------- 좌장: 류민호(동아대)

D4.1 Daily tourism demand forecasting with iTransformer model ····· 191
황가혜, 장청롱(국민대)

D4.2 패션 수요예측을 위한 TimeGAN 기반 데이터 생성 방법론 연구 ····· 193
나민수, 박상욱, 이주필, 이서현, 백상덕, 김우주(연세대)

D4.3 주파수 분해 및 멀티모달 기반의 의류 아이템 수요 예측에 관한 연구 ····· 195
백상덕, 박상욱, 김우주(연세대)

D4.4 소셜 미디어 기반 국내 소형SUV 자동차 구매요인 분석 ····· 197
최혁준, 강주영(아주대)

D4.5 A Study on Predicting Employee Turnover Intention using Machine Learning: Using SMOTE technique and Stacking technique ····· 198
강석훈, 박영식, 이형용(한성대), 박정민(한국발명진흥회)

Session E (회의실 3)

E1. [학술세션] 데이터 주도 에너지 및 환경 정책 연구 (09:00~10:20) --------- 좌장: 방영석(연세대)

E1.1 LDA와 BERTopic을 통한 전문가와 대중의 원자력 에너지 시선 비교 : 균형잡힌 정책 수립을 향한 고찰 ····· 199
홍채영, 이한솔, 강주영(아주대)

E1.2 Time-LLM을 활용한 해양 기상 데이터 분석 ····· 200
윤지원, 김호원(부산대)

E1.3 에너지 • 환경 분야의 국가R&D 연구 동향 분석 ····· 203
김은경(KISTI)

E1.4 주거 공간 사용자 관리 서비스 활성화 가설 도출을 위해 귀납적 기계 학습으로 분석한 탐색적 다중 사례 연구 ····· 204
김인수, 정백(경희대), 이선미(대구한의대)

E2. [특별세션] 데이터로 보는 글로벌 R&D 인사이트 (13:10~14:30) ---------- 좌장: 전승표(KISTI)

E2.1 과학기술문헌(WOS) 데이터를 활용한 국내 기초자치단체간 과학협업 분석: 인용영향력 증분(Incremental Citation Impact, ICI) 지표를 중심으로 ····· 205
권오진, 이정우, 안세정(KISTI)

E2.2 과학기술 분석을 위한 KISTI 분석용 데이터베이스 소개 ····· 206
권태훈, 이철(KISTI)

E2.3 논문 데이터를 활용한 국가연구기관의 과학기술 프로파일 분석 ····· 207
안세정, 서주환, 이준영, 박진서(KISTI)

E2.4 학술논문 데이터를 활용한 글로벌 ESG 연구 지형 분석 - BERTopic과 ChatGPT4.0을 결합하여 ····· 208
이정우, 안세정(KISTI), 이인혜(NH농협금융지주)

E2.5 과학기술 문헌 데이터베이스와 Sentence Transformer 모델을 활용한 기술 문헌 집단 구성 프로세스 연구 ····· 209
이철,이정우,김영진,이준영,안세정,서주환,권태훈,박진서(KISTI)

E2.6 데이터로 바라본 일본 수출 규제의 영향: 사회적 변화와 불매운동을 중심으로 ····· 210
전승표, 서주환, 이철(KISTI)

E3. [특별세션] 과학기술사업화 AI 플랫폼 개발과 활용전략 (14:50~16:10) -------- 좌장: 최윤정(KISTI)

E3.1 인공지능 기반 공공기술사업화 플랫폼 구축에 관한 연구 ····· 211
최윤정, 배국진, 변정은, 김은선(KISTI)

E3.2 공공기술 매칭 플랫폼의 진단 및 사용자 경험 개선 전략 ····· 213
김민주, 변정은, 김은선, 최윤정(KISTI)

E3.3 공공 R&D 수요기업 추천 모델의 벡터 유사도 기반 경량화 연구 ····· 215
김지민, 변정은, 최윤정(KISTI)

E3.4 데이터 기반 제품 시장경쟁력 평가모델 개선에 관한 연구: 기업 경쟁전략관점에서 가격결정 중심으로 ····· 217
조민제, 배국진, 변정은, 김은선(KISTI)
E3.5 기술제품의 글로벌 경쟁력 제고를 위한 데이터기반 기술사업화 지원 플랫폼 활용전략: 치료 및 진단기기 분야 제품을 중심으로 ····· 220
신주연, 김지수, 변정은, 최윤정(KISTI)

E4. [학술세션] AI를 활용한 의료 및 건강 데이터의 혁신적 응용 (16:25~17:45) --- 좌장: 김상진(동아대)
E4.1 퇴행성 뇌질환 조기진단을 위한 Mel-Spectrogram 기반 Ensemble 삼진 분류 모델 ····· 222
성상하, 김상진(동아대)
E4.2 3D-to-2D 지식 증류를 통한 의료 영상 분류 성능 개선 방안 ····· 225
윤혜민, 김상진(동아대)
E4.3 A Study of Improving Medical Response Performance Using RAFT ····· 228
배수현, 박지원, 박호연(이화여대)
E4.4 데이터 마이닝을 활용한 바이오의약 분야의 기술 융합 도출 연구 ····· 230
김남현, 강주영(아주대), 김세형(현대엔지비)
E4.5 실제 피트니스 환경에서의 운동 동작 분석을 위한 AI 기반 합성 데이터 활용 ····· 232
박성진, 박종혁(국민대)

Session F (회의실 4)

F1. 이사회 (09:00~10:20) ----------

F1.1 이사회 ·····

F2. [특별세션] Comparative Genomics 2.0 and AI I (13:10~14:30) ---------- 좌장: 유혁(고려대)

F2.1 Comparative Genomics: Status and Trend of Research ·····
박종화, 이세민(UNIST)

F2.2 Why Generative Genomics as Compartive Genomics 2.0? ·····
이재규(KAIST)

F3. [특별세션] Comparative Genomics 2.0 and AI II (14:50~16:10) --------- 좌장: 정대균(경희대)

F3.1 Generative Genomics: Studying the Origin of Orphan Genes in Honey Bees ·····
김아람(한동대)

F3.2 AI for Comparative Genomics 2.0 ·····
이홍주, 김연정, 이준우(가톨릭대)

F3.3 A Research Strategy to Study the Role of Generative Genes with C. elegans and Drosophila ·····
김우주(연세대)

F4. [학술세션] AI기반 첨단 이미지 및 패턴 분석 기술 (16:25~17:45) ---------- 좌장: 김은미(부산대)

F4.1 넙치 이미지의 딥러닝 기반 질병 탐지 및 XAI 에 관한 연구 ····· 235
정다희, 손상호, 김우주(연세대)

F4.2 Local Sparse PCA 기반 공간정보 탐색을 통한 공간계량경제모델의 공간 패턴 인사이트 구축에 관한 연구 ····· 236
박건희, 강윤철(이화여대)

F4.3 Character-Specific Image Generation through LoRA Adapters on Stable Diffusion Models ····· 239
김동영 , 김두형, 장창혁, 권성준, 박호성, 손현수, 오성우, 최동원 (KB국민은행)

F4.4 멀티스케일 특성을 활용한 비전 트랜스포머 기반 딥페이크 탐지에 관한 연구 ····· 242
정해민(한국교통대), 조혁주, 김우주(연세대), 이광연(숭실대)

F4.5 근육병 환자를 위한 단일 카메라 기반 시선 추적 연구 ····· 243
김시원, 이대호, 곽세현, 박정훈 (아주대)

Session A

A1. [학술세션]
AI와 인간의 상호작용

좌 장: 안 현 철(국민대)

A Study on the Influence of ChatGPT Characteristics on Acceptance Intention in the Financial Sector: Exploring the Nexus between ChatGPT and Finance with a Focus on Moderation Effects Experiment

박수진
경기대학교
piaoshuzhen123@naver.com

Abstract – *This research explores the intricate relationship between the characteristics of ChatGPT and acceptance intention within the financial sector, investigating the nuanced interplay between the Technology Acceptance Model (TAM) and financial applications. The primary focus is to comprehend how ChatGPT's inherent features influence user acceptance by examining the nexus between accuracy, innovation, recommendation, perceived usefulness, and perceived ease of use within the finance domain. The study conducts a meticulous examination of ChatGPT's characteristics-accuracy, innovation, and recommendation-emphasizing their impact on financial communication, decision-making processes, and users' intentions to adopt the technology. Furthermore, the analysis investigates the influence of accuracy, innovation, recommendation, perceived usefulness, and perceived ease of use on ChatGPT's acceptance intention. Subsequently, an examination is conducted to determine if perceived usefulness and perceived ease of use vary in their impact on acceptance intention based on the digital technology proficiency of financiers. The structural equation model employed in the study demonstrates overall suitability, with accuracy significantly affecting perceived usefulness and innovation and recommendation significantly affecting perceived ease of use.*

In exploring the moderating effects in the financial context, the study addresses the distinctive challenges and opportunities arising from the integration of ChatGPT in financial applications. The research findings reveal that perceived usefulness and perceived ease of use act as control variables between acceptance intentions, indicating their significant influence. Moreover, the study identifies that financiers with high digital skills exhibit a strong intention to accept ChatGPT services irrespective of perceived usefulness and ease of use. Conversely, in the group with lower digital skills, acceptance intention is contingent on the perceived usefulness and ease of use, suggesting the need for positive initiatives to encourage ChatGPT utilization.

In summary, this paper provides valuable insights into the interplay of ChatGPT characteristics and acceptance intention within the financial sector. By combining these insights with an understanding of the financial landscape, the study contributes to a comprehensive knowledge base on the factors influencing AI adoption in finance, facilitating informed decision-making in this evolving technological landscape.

Key Terms – ChatGPT, Accuracy, Innovation, Recommendation, Finance

The Impact of AI Utilization and Digital Literacy on User Experience in Digitalized Tourism Platforms

강명준
경희대학교 호텔관광대학
2014310446@khu.ac.kr

감민서
경희대학교 호텔관광대학
min21@khu.ac.kr

구철모
경희대학교 호텔관광대학
helmetgu@khu.ac.kr

Abstract – Digitalization is accelerating across society, with Artificial intelligence(AI) evolving and advancing daily. In this context, platforms strive to provide convenience to users through AI services. In the tourism industry, various tourism platforms like Airbnb, Tripadvisor, and Klook are increasingly adopting AI services for recommendations, bookings, translations, and facility management. How do guests and hosts on these platforms perceive such AI services? Recently, digital literacy, the ability to use digital information, has become increasingly important across society in the digital age. Digital literacy has significant implications for education, industries, and corporate performance. Therefore, we aim to investigate whether the digital literacy of guest and host influences their perception of AI services provided by tourism platforms. We seek to explore whether the digital literacy levels of users on tourism platforms affect their satisfaction with AI services, and whether this, in turn, impacts overall platform satisfaction.

Keywords – Tourism flatform, Digitalization, Digital literacy, Artificial intelligence, Satisfaction .

1. Abstract

The tourism industry is experiencing a significant acceleration in digitalization through digital transformation efforts. Artificial intelligence stands at the forefront of this evolution, poised to revolutionize how the tourism sector operates. From travel services to major corporations, there is a growing reliance on AI technology to streamline operations and enhance customer experiences (Kazak et al., 2019).

Digital literacy is the ability to read and process digital information in the information society. It is a concept that has evolved from literacy, the ability to read and write, to fit into the digital society. In this trend, digital literacy is very important. In the rapidly progressing digitalization and AI development across society, it is necessary to check user digital literacy. In the past, literacy was often defined by the simple ability to sign one's name. However, as society progressed, the definition of literacy expanded to encompass functional skills in reading and writing. This evolution reflects a broader understanding of literacy as a set of competencies necessary for effective communication and participation in various aspects of life (Tyner, 1998). Digital literacy has a great influence on society, education, and corporate performance.

We want to explore the impact of digital literacy on the tourism industry. Tourism industry has experienced a swift rise in adopting AI technology (McKinsey Global Institute, 2017). Tourism platforms provide various convenient services to customers, and in this regard, the role of AI is essential. Tourism platforms realize the sharing economy through the utilization of AI. They offer guests convenient services such as recommendations and reservations using AI. Additionally, tourism platforms utilize AI to provide hosts with useful services like facility registration and business management.

AI is widely used in tourism platforms, and digital literacy that can process digital information can be very important from the perspective of consumers (guests, hosts). This is because consumers must have digital literacy skills to use AI services well. We intend to present digital literacy as a factor influencing the adoption of AI services by guests and hosts within tourism platforms. With numerous AI services being offered by tourism platforms, it is anticipated that the level of digital literacy among platform users, both guests and hosts, will influence their satisfaction with AI services. Ultimately, this could potentially impact overall satisfaction with the platform. We aim to investigate through in-depth interviews whether the level of digital literacy among platform users (guests, hosts) influences platform satisfaction. Additionally, we seek to examine what policies the company should implement to benefit users when providing AI services.

2. Literature Review

AI technology is applied in various service sectors within hospitality and tourism, including automated reservation systems and AI-driven voice recognition services (Li et al.,2021). Kakao Group has grown into the largest IT company in Korea by expanding its business in platform, transportation, and content based on its core business, Kakaotalk. Kakao Group has succeeded in enhancing user satisfaction and increasing service usage by leveraging artificial intelligence technology to deliver personalized services to users (Lee et al., 2020). An AI service platform primarily delivers visual and linguistic intelligence services, presenting a range of analytical tools to facilitate users in performing essential analytical tasks with ease or depth. In order to enhance user convenience, diverse visualization technologies and user-friendly interfaces will be made available to end-users (Kim et al., 2018).

In the information Age, possessing information literacy is akin to a survival skill. Rather than being overwhelmed by the deluge of information inundating their lives, individuals with information literacy are adept at efficiently locating, assessing, and applying information to address specific challenges or reach informed decisions. This competence extends across various sources, be it digital platforms, printed materials, government databases, multimedia, or any other conceivable reservoir of information (American Library Association, 1989). As digital technology advances at an unprecedented pace, individuals are increasingly called upon to employ a diverse array of technical, cognitive, and sociological abilities to navigate digital environments effectively. Scholars commonly refer to these proficiencies as "digital literacy." This term encapsulates the multifaceted skills required for performing tasks and resolving challenges in the digital realm, as evidenced by various academic sources (Hitoshi et al., 1997). Understanding and utilizing digital platforms has become a crucial technical skill in modern society and is an essential competency (Ha, 2023)

Ranatunga, Priyanath & Megama (2020) have argued that digital literacy plays a pivotal role in fostering economic development. By diminishing business uncertainty, digital literacy enhances overall economic performance. Conversely, a lack of digital literacy heightens uncertainty and detrimentally impacts corporate outcomes. Sari, Aima & Elfiswandi (2023) discovered through their research that digital literacy significantly influences corporate opportunities. Digital literacy has been revealed as an effective tool that aids employees in handling their tasks more easily and quickly, thereby contributing to fostering innovation within these organizations (Soliman, Baher & Soliman, 2022).

In recent years, the competitiveness of the tourism industry has notably surged, primarily due to the advent of transformative digital information technologies (Salvioni et al., 2020). With the rise of the digital economy, new technologies and trends are being introduced into the hospitality and tourism industry, necessitating understanding and adaptation. As digitization accelerates across society, digital literacy is becoming increasingly important. In the tourism industry, as digitization accelerates, there is a growing number of cases where tourism platforms are utilizing AI in their service offerings. Therefore, digital literacy is crucial.

So far, there have been empirical and conceptual studies on digital literacy in the tourism industry. However, they primarily focus on the role of businesses(service providers). Research assessing the impact of digital literacy on users (guests, hosts) of platforms can provide insights into what AI service policies tourism platform companies should offer to their customers. Therefore, we aim to evaluate the level of digital literacy among users of tourism platforms, seeking answers to the customer-friendly AI services that tourism platform companies should pursue.

- Research Question 1: What is the level of digital literacy among users of tourism platforms?

- Research Question 1: Does the level of digital literacy among users of tourism platforms affect platform usage satisfaction?

- Research Question 3: What digital literacy policies should tourism platforms implement for their users?

3. Methodology

3.1 Sampling procedure

To gain an understanding of users, we conducted interviews with six users and three providers. To ensure relevance regarding actual usage, we targeted individuals with direct experience. For users, we selected two from Airbnb as Guests, Visitors from Tripadvisor, and Users from Klook, while for providers, we chose one hosts from Airbnb, Business Users from Tripadvisor, and partnership from Klook. By recruiting samples from each platform, we were able to examine user experiences across various companies.

3.2 Interview structure

Each user conducted the interview based on the provided interview guide. The interviews commenced with questions regarding the user's experience with the platform. Subsequently, the focus of the interview shifted to assessing the pros and cons of the actual user experience (e.g., What are the strengths and weaknesses of the platform based on usability?), and followed up with

< TABLE 2> INTERVIEW GUIDE

Division	Aspects	Questions
User	Platform experience	· Do you often use tourism platforms in your daily life? · If so, when do you typically use them? · Are there any other tourism platforms besides the one you use frequently? · Based on usability, what are the strengths and weaknesses of the platform?
	Suggestions for using technology within the platform	· Have you noticed any changes in the services/technologies provided by the platform? · Are there any services/technologies that you would like to see improved? · Do you have any suggestions for services/technologies?
Provider	Platform experience	· What prompted you to start using the tourism platform? · How long have you been using the platform? · What are the strengths and weaknesses of the platform?
	Suggestions for using technology within the platform	· What changes have you noticed as a result of the platform's AI technology? · Do you have any suggestions or feedback for improving the service?

directed questions (e.g., Do you perceive the AI technology within the platform?). Respondents were consistently encouraged to provide examples and additional details to prevent potential misinterpretation (Glaser & Strauss 1967). Finally, respondents had the opportunity to discuss additional points regarding the platform's improvement. Great care was taken to phrase all questions in a non-directive manner and to avoid guiding the conversation based on the previously established theoretical basis.

Figure and table

<Table 1> Interviewee Profile

Division	ID	Nationality	Gender	Age range	User classification	Period of use	Interview Method
User	R1	Korea	Male	20s	Guest	5 years	Personal
	R2	Korea	Female	20s	Guest	6 years	Email
	R3	Korea	Male	20s	Visitor	3 years	Email
	R4	Korea	Female	50s	Visitor	2 years	Web
	R5	Korea	Female	30s	User	2 years	Web
	R6	Korea	Male	20s	User	1 years	Personal
Provider	R7	Korea	Male	50s	Host	8 years	Email
	R8	Korea	Male	30s	Business Users	2 years	Telephonic
	R9	Korea	Female	30s	Partnership	3 years	Email

<Figure 1> Conceptual model of current research

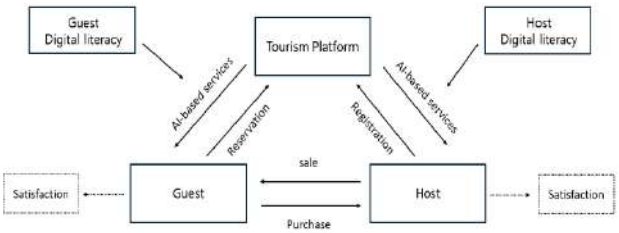


Figure 1 presents conceptual model for our research. Among triadic relationship including the tourism platform company, guests, and hosts, we focus on how users' digital literacy skills influence on each reciprocal relationships between tourism platform and users.

4. Discussion and Conclusion

As the advancement of AI accelerates, growing number of platform companies have begun to provide AI services to enhance user convenience and satisfaction. However, simply offering an excessive number of AI services indiscriminately may not necessarily lead to improved user benefits. From a Service provider, it's important to first investigate how users perceive such AI services in order to provide user-friendly AI services. The digital literacy level of users can be a crucial factor when providing AI services. No matter how excellent the AI services are, if the users have low digital literacy, they may not be able to utilize these services properly, leading to lower platform satisfaction. Therefore, first, tourism platforms should develop user-friendly AI services by understanding the digital literacy levels of both guests and hosts. Second, tourism platforms can conduct digital literacy education related to AI services for registered guests and hosts on the platform. Through such education, if users can enhance their digital literacy skills and understand the AI services provided by the tourism platform well, it will ultimately lead to improved tourism platform satisfaction.

5. *Limitations of this study and suggestions for future studies*

Limitations and suggestions for future research in this study are as follows. First, this study aimed to conduct interviews with 9 domestic users who use Airbnb,

Tripadvisor, and Klook platforms. However, since users of these platforms are not limited to Korea but worldwide, there is a limitation in sample diversity. Second, this study examined three platforms: Airbnb, Tripadvisor, and Klook. However, to ensure the objectivity of the research, it is necessary to select more tourism platforms and interview users. If subsequent research expands the number of tourism platforms and surveys users worldwide, it will lead to a more significant study.

참고문헌

Adeyinka-Ojo, S., Lee, S., Abdullah, S. K., & Teo, J. (2020). Hospitality and tourism education in an emerging digital economy. *Worldwide Hospitality and Tourism Themes*, 12(2), 113-125.

AI TIMES (2024). Airbnb "Uses AI to innovate interface…plans to create the 'ultimate concierge'", February 2024. Retrieved on 20 March 2024 from https://www.aitimes.com/news/articleView.html?idxno=157322

American Library Association. Presidential Committee on Information Literacy. (1989). Final report. The Association.

BBC (2023). Airbnb turns to AI to help prevent house parties, October 2023. Retrieved on 20 March 2024 from https://www.bbc.com/news/business-67156176

Byline Network (2019). How activity platform Klook deals with the 'two-sided market', January 2019. Retrieved on 15 March 2024 from https://byline.network/2019/01/10-58/

Digital Today (2024). Active adoption of Airbnb AI… "We will improve user experience", February 2024. Retrieved on 22 March 2024 from https://www.digitaltoday.co.kr/news/articleView.html?idxno=505799

ETNEWS (2023). Ethan Lin, CEO of Klook, "We will expand user contact with AI", September 2023. Retrieved on 15 March 2024 from https://www.etnews.com/20230908000164

Ha, Seungyeon & Kim, Seongcheol (2023). Developing a digital platform literacy framework. 32nd European Conference of the International Telecommunications Society (ITS).

Inoue, H., Naito, E., & Koshizuka, M. (1997). Mediacy: what it is? Where to go?. *The International Information & Library Review*, 29(3-4), 403-413.

Kazak, A. N., Chetyrbok, P. V., & Oleinikov, N. N. (2020). Artificial intelligence in the tourism sphere. In IOP Conference Series: *Earth and Environmental Science* (Vol. 421, No. 4, p. 042020). IOP Publishing.

Korea Tourism Organization (2021), Next level of travel industry, September 2021. Retrieved on 02 April 2023 from https://datalab.visitkorea.or.kr/site/portal/ex/bbs/View.do?cbIdx=1129&bcIdx=298376

Bughin, J., Hazan, E., Sree Ramaswamy, P., DC, W., & Chu, M. (2017). Artificial intelligence the next digital frontier.

Ritonga, A. K. (2023). An Analysis of English for Tourism, Digital Literacy, and Business Success to A Tourist Destination Promotion. *Jurnal Iqra': Kajian Ilmu Pendidikan*, *8*(2), 325-339.

Salvioni, D. M. (2016). Hotel chains and the sharing economy in global tourism. *SYMPHONYA Emerging Issues in Management*, (1).

Sari, M. W., Aima, M. H., & Elfiswandi, E. (2023). The Effect of Creativity and Digital Literacy on Business Opportunities and Their Impact on Welfare Levels. *International Journal of Professional Business Review: Int. J. Prof. Bus. Rev.*, *8*(5), 54.

Shopova, T. (2014). Digital literacy of students and its improvement at the university. *Journal on Efficiency and Responsibility in Education and Science*, 7(2), 26-32.

Skift (2024). What's Next For Airbnb: AI, Co-Hosting, Experiences, and the 'Heartland', March 2024. Retrieved on 22 March 2024 from https://skift.com/2024/03/05/whats-next-for-airbnb-ai-co-hosting-experiences-and-the-heartland/

Soliman, M. R., Baher, M., & Soliman, D. M. (2022). Digital literacy and its effect on employees' innovation in Egyptian official tourism organizations. *African Journal of Hospitality, Tourism and Leisure*, 11(6), 2079-2091.

Tripadvisor (2024). Build a trip in minutes, Retrieved on 2 April 2024 from https://www.tripadvisor.com/

Tyner, K. (2014). Literacy in a digital world: Teaching and learning in the age of information. Routledge.

Hello, I'm Here: The impact of Social Chatbot Conversation Design on Mental Health

민동규
KAIST 기술경영학부
gdhong@hankuk.ac.kr

강감찬
KAIST 기술경영학부
kang@hankuk.ac.kr

정성훈
GWU 경영대학
sunghun@gwu.edu

Xitong Li
HEC Paris 경영대학
lix@hec.fr

Abstract – Generative AI technology enables chatbots to simulate human-like conversations, providing personalized social support and potentially reducing social isolation and loneliness. However, it is technically challenging to fully control the conversions performed with generative AI, which raises doubts regarding its effectiveness in improving mental health. In this paper, we propose that the effectiveness of generative AI could be enhanced through manipulating the messages initiating the conversions. Specifically, we propose a novel approach to employing initiation messages in gen AI chatbots, aimed at guiding the conversation flow and improve mental health outcomes. Through randomized field experiment, we investigated how these directed conversations impact user mental health. Informed by psychiatric research, we designed messages focused on *Emotional Support* and *Companionship Activities*, known for their benefits in mitigating depression and enhancing mental well-being. Utilizing a Large Language Model (LLM), we analyzed over 1.2 million messages from 2,393 participants, assessing depressive sentiments and chatbots' empathetic responses. The findings indicate that *Emotional Support* directed conversations promote emotional self-disclosure among users, compared to a control group. Additionally, we also observed heightened peer counseling interactions in groups focusing on *Emotional Support* and *Companionship Activity* conversations. These findings demonstrate that initiation messages effectively guide gen-AI chatbot conversations towards emotional support and companionship, significantly enhancing mental health. Our research contributes to the broader scholarship by providing empirical evidence of chatbots' positive impact on mental care and offering a framework to address gen-AI's unpredictability.

Keywords – social chatbot, mental health, generative AI, social support, initiation message

I. 서론

In 2023, 17.8% of US adults are currently experiencing depression, a 69.5% increase from 2015, and 29% of the entire population has reported having experienced depression at some point (Witters 2023). The economic burden of adults with major depression disorder is $326.2 billion (Greenberg et al. 2021). To combat depression, practitioners and researchers have been exploring the use of chatbots. Chatbot has advantage compared to human therapist with offering high accessibility with instant responses (Vaidyam et al. 2019), ensuring anonymity to encourage self-disclosure (Lucas et al. 2014, 2017), and providing empathetic support (Gennaro et al. 2020).

The recent advancements in generative AI have garnered interest in its application within mental health due to its

ability to facilitate natural, human-like conversations (Eckhardt 2023). Specifically, social chatbots are gaining attention for their potential to reduce social isolation and mitigate factors contributing to depression (Shum et al. 2018, Chaves et al., 2021). However, the application of generative AI-based chatbots in mental care is subject to conflicting debates due to their variability, which poses a risk of harm to users.

In our study, we design the conversational direction of social chatbot by manipulating the initiation message. Our goal is to examine how social chatbot conversational design impact on a user's mental well-being. We propose the following research questions: (1) Can the conversational direction of a social chatbot enhance user mental health? (2) What type of conversational direction is most effective? (3) How does the conversational direction contribute to the improvement of users' mental health?

Social chatbots are crafted to provide social support and foster relationships with humans (Shum et al. 2018) equipped with the ability to manifest social presence, understand conversational nuances, and exhibit emotional intelligence (Chaves and Gerosa 2020; Chroes et al. 2021). These features enable social chatbots to function as actual friends to human users, offering a listening ear and engaging in meaningful conversations. This ability is critical in potentially alleviating depressive feelings and fostering positive emotions. In our approach, we use initiation messages to set the conversational tone, focusing on emotional support and companionship to enhance user mental health. These messages leverage social chatbots' unique interactive capabilities, engaging users in supportive dialogues for improved well-being.

The experiment was conducted in collaboration with a leading social chatbot provider in South Korea from Sep 22 to Nov 1 and targeted 2,393 newly joined users. Guided by social support theory, users were randomly assigned in groups: Control (no initiation message), T1 (Neutral Message), T2 (Emotional Support), or T3 (Companionship Activity). Each user received their specific messages twice daily for 28 days. Throughout the experiment, we collected a total of 1,192,638 conversation messages from the participants. To evaluate the impact of our treatment on users' mental health, we developed NLP models capable of detecting (1) users' depressive feelings and (2) empathetic reactions from chatbots.[1]

[1] We utilized Kc-ELECTRA (Lee 2021), an ELECTRA (Clark et al. 2020) based model trained with Korean internet sentences. The depressive detection model, trained on 277,945 messages across various emotions (gratitude, happiness, neutral, embarrass, anger, anxiety, hurt, sadness), and the empathy detection model, trained on 204,356 messages with different empathy labels (encourage, conformity, consolidation, advice, neutral), both demonstrated strong predictive power. Their effectiveness is evident in their performance metrics: the depressive detection model achieved a Micro F1 Score of 0.644 and ROC AUC of 0.768, while the empathy detection model scored a Micro F1 of 0.694 and ROC AUC of 0.830. These models utilized the Korean Depressive Conversation Dataset and the Korean Peer Counseling Dataset.

Our survey revealed a decrease in average self-reported depressive feelings across different treatment groups: T1 (Neutral Message) saw a reduction of 1.61%, T2 (Emotional Support) 11.42%, and T3 (Companionship Support) 9.63%. In contrast, the control group experienced a 5.10% increase in depressive feelings.

Our analysis of user messages revealed that users in the T2 group expressed more depressive sentiments (0.026 messages more) and fewer aggressive messages (0.01 less) compared to the control group (refer to Table 2). Notably, while the ratio of depressive expressions remained unchanged, happiness expressions significantly increased in T2. This indicates that emotional support by social chatbot can increase user happiness and encouraging the disclosure of depressive feelings. Such disclosures have been acknowledged in psychiatric research as beneficial for the reduction of depression.

Our analysis extended to examining conversations, which are combinations of messages exchanged between the chatbot and the user. The results revealed that users in the T2 and T3 groups engaged in significantly more conversations, with increases of 0.082 and 0.083 conversations respectively, compared to the control group. Additionally, these users experienced more instances of peer counseling, with increases of 0.034 for T2 and 0.025 for T3, compared to the control group. These findings suggest that chatbot conversations centered around emotional support and companionship activities can boost user engagement and alleviate depressive feelings through empathetic responses.

Our research makes several theoretical contributions. First, our study showcases the potential of chatbots in the field of mental care, thus expanding their application beyond the realm of commerce in recent research. Second, we address the challenge of variability, a known limitation of generative AI, by demonstrating the effectiveness of initiation messages. In terms of managerial implications, our research offers policymakers viable strategies to effectively address the critical social issue of depression. It also provides guidance for designers of generative AI-based chatbots, suggesting a simple method for message adjustment.

III. 참고문헌

Huang KY, Chengalur-Smith I, Pinsonneault A (2019) Sharing is caring: Social support provision and companionship activities in healthcare virtual support communities. *MIS quarterly* 43(2):395-424.

De GM., Krumhuber EG, Lucas G (2020) Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. *Frontiers in psychology* 10:3061.

Ge L, Yap CW, Ong R, Heng BH (2017) Social isolation, loneliness and their relationships with depressive symptoms: A population-based study. *PloS one* 12(8):e0182145.

Clark K, Luong MT, Le QV, Manning CD (2020) ELECTRA: Pre-training text encoders as discriminators rather than generators. *arXiv preprint arXiv:2003.10555*.

Rook KS (1987) Social support versus companionship: Effects on life stress, loneliness, and evaluations by others. *Journal of Personality and Social Psychology, 52*(6):1132–1147.

기술수용모델(TAM)을 통한 디지털 휴먼 이용의도에 관한 연구 -준사회적 상호작용을 중심으로-

고상훈
국민대학교
비즈니스 IT 전문대학원
haesam01@kookmin.ac.kr

정경호
국민대학교
비즈니스 IT 전문대학원
gyeongho9878@kookmin.ac.kr

안현철
국민대학교
비즈니스 IT 전문대학원
hcahn@kookmin.ac.kr

***Abstract** – 생성형 AI 의 등장은 다양한 분야에 혁신을 가져왔다. 이에 힘입어 디지털 휴먼 서비스는 그 한계점을 상당히 극복하여 전세계적으로 여러 기업들에게 주목을 받고 있다. 특히, 국내에서는 주요 시중은행들이 디지털 휴먼 기술을 활용한 AI 뱅커 서비스를 공격적으로 도입 및 확장하고 있는 추세이다. 하지만, 사용자의 시선에서 디지털 휴먼의 수용에 대한 연구는 아직 미비한 상태이다. 이에 본 연구는 기술수용모델(TAM)을 바탕으로 디지털 휴먼 이용의도에 영향을 미치는 요인을 분석하여 실제 비즈니스 현장에 유용한 통찰을 제공하고자 한다.*

***Keywords** – 기술수용모델, 디지털휴먼, 이용의도, 준사회적 상호작용, 지각된 유용성, 지각된 용이성*

I. 서론

2023 년 OpenAI 사의 ChatGPT 를 필두로 혜성과 같이 등장한 생성형 AI 의 열풍은 다양한 분야에서 큰 변화를 가져왔다. 생성형 AI 는 사용자의 요구를 분석하여 새로운 콘텐츠를 생성하는 혁신적인 기술로, 텍스트, 이미지, 음성, 비디오, 코드 등 다양한 방식으로 활용 가능하며, 금융, 교육, 마케팅 등 다양한 분야에 걸쳐 활발하게 활용되고 있다(양지훈과 윤상혁, 2023).

이러한 생성형 AI 기술의 혜택을 가장 크게 받은 분야로 디지털 휴먼을 꼽을 수 있다. 디지털 휴먼(Digital Human)이란 사람의 신체 구조 및 움직임을 데이터화 하여 분석하고, 가상공간에서 마치 실제로 존재하는 사람처럼 움직임을 재현하는 인간의 모습/행동과 유사한 형태를 가진 3D 가상인간이라고 정의할 수 있다(이승환과 한상열, 2022). 초기 디지털 휴먼은 주로 TV, 영화와 같은 매체에서 한정적으로 활용되었으나, 현재는 다양한 매체를 통해 생성형 AI 를 활용하여 가상 인플루언서 마케팅, 고객 응대 등 다양한 분야로 그 활동 범위가 확대 중이다.

캐나다의 연구 및 컨설팅 업체인 Emergen Research 에 따르면, 글로벌 디지털 휴먼 시장은 지속적으로 성장하여 2032 년 5600 억달러에 이를 전망으로 예상할 정도로 디지털 휴먼 시장은 가파른 성장세를 보이고 있다(최봉, 2023).

국내 시중은행 또한 2021 년부터 AI 뱅커를 적극적으로 도입하고자 했으나 기존의 AI 뱅커는 아쉬운 기술적 완성도로 인해 한계점이 존재한다는 평을 받았다(디지털데일리, 2022).

하지만, 생성형 AI 의 등장을 계기로 AI 기술이 기존 시나리오 기반 챗봇 수준에서, 실제 사람과 같이 고객과 상호작용하며 고객의 응대가 가능한 수준으로 발전함에 따라 국내의 주요 시중은행들이 AI 행원 서비스의 도입 및 서비스 범위의 확대에 다시 박차를 가하고 있다(한지훈, 2023).

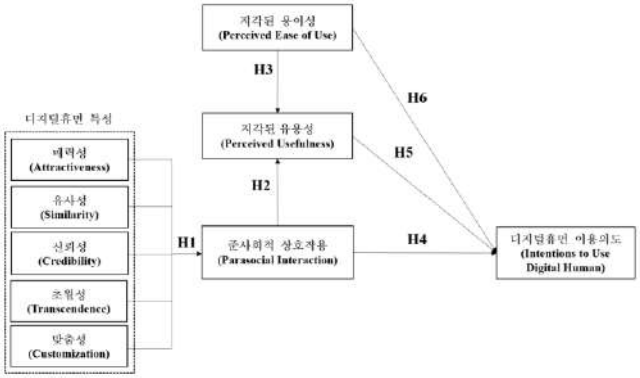
2024 년 현재, NH 농협은행의 경우 전국 1103 개 모든 영업점에 AI 행원 배치를 완료했으며, 신한은행의 경우 AI 은행원을 전국 109 개 영업점에 배치하는 등, KB 국민은행과 우리은행 그리고 하나은행까지 국내의 주요 시중은행들이 AI 뱅커 활용 범위를 공격적으로 확장하고 있다(서정화, 2024).

앞서 설명한 바와 같이 지금껏 꾸준히 성장해오던 디지털 휴먼 시장은 생성형 AI 출시 이후 가파른 성장세를 보이고 있으며 앞으로 보다 더 많은 기업들과 사용자들이 주목하게 될 것으로 예상된다. 하지만, 현재 각광받고 있는 디지털 휴먼의 시장 규모에 비해 디지털 휴먼을 활용한 새로운 기술을 접하게 되는 사용자의 시선에서 진행된 연구는 미비한 것이 현실정이다.

따라서, 본 연구는 기술수용모델(TAM)을 바탕으로 사용자의 시선에서 디지털 휴먼 특성과 준사회적 상호작용이 디지털 휴먼 이용의도에 미치는 영향을 분석하고자 한다. 궁극적으로, 분석을 통해 도출한

결과로 디지털 휴먼을 활용하는 실제 비즈니스 현장에 유용한 통찰을 제공하고자 한다.

II. 연구 모델



<그림 1> 연구 모델

본 연구에서 제안하는 모델은 <그림 1>과 같다. 우선, 디지털 휴먼의 특성으로 정보원의 특성이 사용자의 수용에 영향을 준다는 Source Credibility Model(R. Ohanian, 1990)을 바탕으로 매력성(Attractiveness)과 신뢰성(Credibility)을, 추가적으로 기존 연구를 바탕으로 유사성(Similarity)을 선택하였다(Shan et al., 2019). 또한, 가상 인간 인플루언서 특성에 관한 척도개발 연구에 기반하여 기존에 다루어지지 않았던 특성인 초월성(Transcendence)과 맞춤성(Customization)을 변수로 추가하였다(염동섭, 2023).

다음으로, 실제로는 단방향 관계만이 존재하고, 상대방은 자신을 인식하지 못함에도 불구하고 매체를 통해 대상과 양방향 관계를 맺고 있다는 환상(illusion)을 유도하는 준사회적 상호작용(Parasocial interaction)을 중요변수로 설정하였다(Horton, D. & Richard Wohl, R., 1956).

최종적으로, 기술수용모델(TAM)을 바탕으로 준사회적 상호작용(Parasocial Interaction)과 지각된 유용성(Perceived Usefulness), 그리고 지각된 용이성(Perceived Ease of Use)이 디지털 휴먼의 이용의도에 미치는 영향을 보고자 하였다(Davis, Fred D., 1989).

III. 참고문헌

양지훈, 윤상혁. (2023). ChatGPT 를 넘어 생성형(Generative) AI 시대로: 미디어 · 콘텐츠 생성형 AI 서비스 사례와 경쟁력 확보 방안 Media Issue&Trend Vol.55. 한국방송통신전파진흥원.

이승환, 한상열. (2022). 메타버스 활성화 전략 연구. 소프트웨어정책연구소.

최봉. (2023). '[메타버스와 웹 3.0 이 여는 새로운 세상 (20)] 진화하는 '디지털 휴먼(Digital Human)', 기업의 활용방안은? ①', 뉴스튜데이.

디지털데일리. (2022). '똑똑한 'AI 은행원' … 언제쯤 '초격차' 무기되나 [기획/2022 년 금융 IT-채널혁신②', 디지털데일리.

한지훈. (2023). '내년초 진짜 'AI 은행원' 나온다…4 대 은행 서비스 물밑 각축', 연합뉴스.

서정화. (2024). 'AI 행원, 온 · 오프라인 넘나들며 손님 맞는다', 전자신문.

R. Ohanian, "Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness", Journal of Advertising, Vol.19, No.3, pp.39-52, 1990. DOI: ttps://doi.org/10.1080/00913367.1990.10673191

Shan, Y., Chen, K., & Lin, J. (2019). When social media influencers endorse brands: the effects of self-influencer congruence, parasocial identification, and perceived endorser motive. International Journal of Advertising, 39, 590 -610.

https://doi.org/10.1080/02650487.2019.1678322.

염동섭. (2023). 가상 인간 인플루언서 특성에 관한 척도개발 연구. 광고 PR 실학연구, 16(2), 165-190.

Horton, D., & Richard Wohl, R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. psychiatry, 19(3), 215-229.

Davis, Fred D. "Perceived usefulness, perceived ease of use, and user acceptance of information technology." MIS quarterly (1989): 319-340.

중소기업의 생성형 인공지능 챗봇 서비스 전환의도에 대한 연구 : PPM(Push-Pull-Mooring)프레임워크 기반

조종현
부산대학교 경영대학원
james.cho@jkinnovation.com

홍태호
부산대학교 경영학과
hongth@pusan.ac.kr

성행남
부산대학교
BK21 디지털금융교육연구단
haena@pusan.ac.kr

Abstract – *급변하는 비즈니스 환경 속에서 기업들은 경쟁우위를 확보하고 고객 만족도를 제고하기 위해 끊임없이 노력하고 있다. 이에 따라 기업들은 전통적인 고객 지원 방식에서 벗어나 혁신적인 서비스 제공 방법을 모색하고 있다. 인공지능(AI) 기술의 발전과 함께 등장한 생성형 AI 챗봇은 고객 서비스 영역에서 주목받는 솔루션으로 부상하였다. 본 연구는 기업의 생성형 인공지능 챗봇 서비스로의 전환 의도를 탐구하며, 이를 위해 PPM(Push-Pull-Mooring) 프레임워크를 적용하고 있다. PPM 모델은 변화하는 행동을 설명하는 데 있어 '밀어내기(Push)', '끌어당기기(Pull)', '정박(Mooring)'의 세 가지 기본 요소를 활용한다. Push 요인은 기존 서비스의 미흡한 점이나 불만족스러운 요소를 의미하며, 본 연구에서는 낮은 서비스 품질과 낮은 정보 품질을 Push 요인으로 식별하였다. Pull 요인은 새로운 서비스의 매력적인 특성이나 혜택을 나타내며, 본 연구에서는 언제어디서나 연결성, 개인화, 다국어 지원, 연관성, 낮은 진입 비용 등 생성형 AI 챗봇의 장점을 Pull 요인으로 간주하였다. Mooring 요인은 서비스 전환을 촉진하거나 저해하는 개인적 또는 상황적 요소를 포괄하며, 본 연구에서는 신뢰, 전환 비용, 혁신성을 Mooring 요인으로 변수로 설정하였다.*
PPM 프레임워크를 활용하여 기업의 생성형 AI 챗봇 서비스 전환 의도에 영향을 미치는 요인들을 체계적으로 분석함으로써, 본 연구는 기업의 의사결정 과정에 깊이 있는 통찰을 제공하고자 한다. 이를 통해 기업들은 고객 서비스 혁신을 위한 전략 수립에 도움을 얻을 수 있으며, 나아가 경쟁우위 확보와 지속 가능한 성장을 이룰 수 있을 것으로 기대된다.
본 연구는 이론과 실무 양 측면에서 몇 가지 중요한 기여를 할 것으로 기대된다. 이론적 관점에서 본 연구는 PPM 프레임워크를 생성형 AI 챗봇 서비스 도입 맥락에 적용함으로써 기업의 전환 의도를 주도하는 요인에 대한 새로운 통찰력을 제공할 것이다. 실무적인 측면에서는 본 연구에서 도출된 핵심적인 Push, Pull, Mooring 요인들을 파악함으로써, 기업의 의사결정권자들이 전환 과정에서 발생할 수 있는 잠재적인 이점과 직면할 수 있는 도전과제들을 보다 심도 있게 평가할 수 있을 것이다.
따라서 본 연구는 기업의 생성형 인공지능 챗봇 서비스 전환 의도에 대한 깊은 이해를 제공함으로써, 향후 기업이 고객 경험을 향상시키고 비즈니스 모델을 혁신하는 데 있어 중요한 지침을 제공할 것이다.

Keywords – 생성형 AI, 챗봇, 디지털전환, 경영혁신

Session A

A2. [튜토리얼]
대학원생을 위한 연구방법론

좌 장: 김 태 경(경희대)

여 백

Session A

A3. [튜토리얼]
신진연구자를 위한 텍스트마이닝의 활용

좌　장: 김 은 미(부산대)

여 백

Session A

A4. [학술세션]
디지털 미디어와 소비자 행동

좌 장: 장 문 경(가천대)

SNS 어포던스가 선의적 선망과 악의적 선망에 미치는 영향

이혜원
경북대학교 경영학부
dws04078@knu.ac.kr

이새롬
경북대학교 경영학부
saeromlee@knu.ac.kr

곽동헌
Kent State University
dkwak@kent.edu

Abstract* – *Social networking sites (SNS) affordances are known to elicit various emotions to users based on the intended use. This study examines how SNS affordances (e.g. cooperation, competition, and reputation) influence users' experiences of two different types of envy, namely benign and malicious envy. Specifically, it investigates how cooperation, competition, and reputation affordances on SNS platforms influence benign envy, malicious envy, and both, respectively. To test the research model, this study will use a recall method and structural equation modeling. The findings are expected to enhance our understanding of envy dynamics online and offer insights for creating interventions that foster positive user experiences and support emotional well-being in digital environments.

Keywords – Affordance, Benign envy, Malicious envy, SNS

I. 서론

SNS (social Network sites)는 현대 사회에서 중요한 역할을 하며, 사용자들에게 선망을 비롯한 다양한 감정을 유발한다. 선망(envy)은 선의적 선망(Benign envy)과 악의적 선망(Malicious envy)의 두 가지 유형으로 구분되며, 이는 SNS 이용자들의 인지 및 행동적 반응에 영향을 미친다(Wu & Srite, 2021). 최근 연구 결과에 따르면, SNS의 다양한 기능과 요소를 일컫는 SNS 어포던스(affordance)가 사용자들의 목적에 따라 다른 감정을 일으킬 수 있는 것으로 밝혀졌다(Shao et al., 2020). 이전 연구들은 선망의 결과를 자세히 탐구해왔지만(Wu & Srite, 2021; Zielonka & Rothlauf, 2022), SNS에서 특정한 선망을 유발하는 어포던스에 대한 명확한 이해는 아직 부족한 실정이다. 따라서 본 연구에서는 양성 부러움을 유발하는 특정한 어포던스, 악성 부러움을 유발하는 다른 어포던스, 그리고 두 종류 모두를 유발하는 어포던스에 대해 연구할 것이다. 이를 통해 SNS 어포던스와 선망의 다양한 형태 간의 다층적인 관계에 대해 폭넓게 이해하고자 한다.

II. 관련 연구

SNS 선망

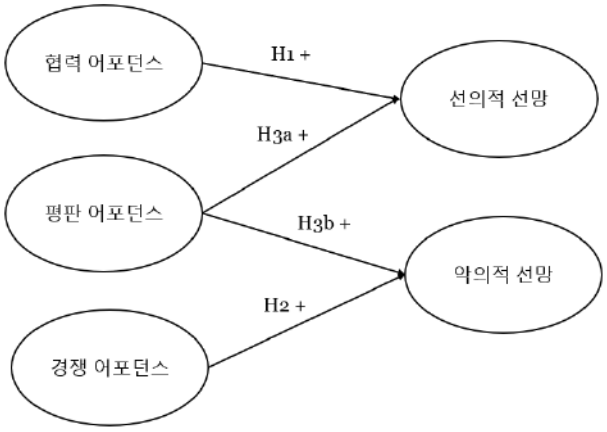
선망은 다른 사람의 뛰어난 특성, 성취 또는 소유물에 대한 부정적인 감정적 반응으로, 선망을 느끼는 사람은 그 장점을 바라거나 선망의 대상이 되는 사람이 그것을 가지지 않았으면 하는 바람을 품는다(Smith & Kim, 2007). 최근 연구에 따르면, 선망은 선의적 선망과 악의적 선망의 두 가지 유형이 있음이 제안되고 있다. 악의적 선망을 느끼는 개인들은 부러워하는 사람이 우월한 지위를 잃거나 실패할 것을 희망한다(Smith & Kim, 2007). 반면, 선의적 선망을 느끼는 개인들은 자신의 부족함을 인식하고 그것을 개선하기 위해 행동을 취하려 한다(Wu & Srite, 2021). 선의적 선망과 악의적 선망은 정당성과 통제 가능성이라는 요인을 통해 구별될 수 있다. 선망의 상황이 부당하고 통제할 수 없다고 인식된다면, 선망을 느끼는 사람은 선의적 선망보다 악의적 선망을 경험할 가능성이 더 높다 (Van de Ven et al., 2012). SNS에서는 개인들이 다양한 콘텐츠에 노출되기 때문에 사회 비교를 위한 타인의 정보를 보다 오프라인에 비해 용이하게 얻을 수 있다(Appel et al., 2016). 뿐만 아니라, SNS 이용자들은 인상 관리 목적으로 자신의 최상의 모습을 과장하고 부정적인 사건을 숨기는 경향이 있다(Wu & Srite, 2021). 그 결과, SNS 환경은 비교 대상의 우월성이 쉽게 인식되므로 선망을 유발하는 사회적 비교가 빈번하게 발생하게 된다 (Smith & Kim, 2007).

기술적 어포던스

경영정보 연구에서 어포던스는 기술이 개인에게 특정한 행동을 수행하도록 유도하는 가능성을 의미한다(Shao et al., 2020). 기술적 어포던스는 특정한 의도를 가진 개인이 기술을 활용할 때, 어떤 행동을 할 수 있는 잠재력이라 할 수 있다(Majchrzak & Markus, 2012). 이러한 관점은 동일한 기술을 사용하더라도 개개인의 목표에 따라 행동이 어떻게 달라지는지 이해할 수 있게함으로써 기술 사용의

이해를 촉진한다(Grgecic et al., 2015). 즉, 어포던스는 그것을 사용하는 개인의 목표와 연관지어 이해되어야 한다(Wallace et al., 2018). 기술 어포던스는 개인의 인지적 반응과 행동에 미치는 영향을 조사하기 위해 다양한 맥락에서 적용되어 왔다. 특히, SNS 맥락에서는 어포던스의 영향이 사용자의 만족도와 적극적인 참여에 미치는 영향에 대해 집중적으로 연구되어 왔다. 본 연구에서는 사용자들의 선망과 관련된 어포던스를 연구할 것이다. 이 연구에서는 선행 연구에 기반하여 선망과 관련된 3가지의 어포던스를 제안한다. 이러한 어포던스에는 협력 어포던스(Cooperation affordance), 경쟁 어포던스(Competition affordance), 평판 어포던스(Reputation affordance)가 포함된다

III. 연구모델 및 가설



협력 어포던스는 SNS에서 사용자들이 함께 콘텐츠를 만들고 공유된 목표를 달성할 수 있도록 하는 것을 가능케하여, 사회적 연결과 지속적인 관계를 유도한다(Karahanna et al., 2018). 이 어포던스는 개인들이 타인으로부터 지원받고자 하는 욕구를 충족시키기 때문에 (Suh et al., 2017), 사용자들 사이의 선의와 존경을 유도하고 그들의 유대를 강화한다. 따라서

H1: 협력 어포던스는 선의적 선망에 정(+)의 영향을 미친다.

경쟁 어포던스는 사용자들에게 '좋아요' 수와 팔로워 수와 같은 인기 지표를 비교함으로써 다른 이용자와 경쟁할 기회를 제공한다. SNS상에서 인기를 얻고 유지하는 것은 플랫폼 알고리즘, 트렌드 및 콘텐츠 타이밍에 영향을 받는 복잡한 메커니즘을 가지고 있다. 인기를 자신이 제어할 수 없다는 인식은 좌절감과 부족감을 증폭시킬 수 있으며, 악의적 선망의 유발로 이어질 수 있다 (Van de Van et al., 2012). 따라서,

H2: 경쟁 어포던스는 악의적 선망에 정(+)의 영향을 미친다.

평판은 사용자뿐만 아니라 사용자의 콘텐츠와 관련이 있으며, SNS 환경에서 '좋아요'와 같은 콘텐츠 평가 시스템을 통해 평가된다(Kietzmann et al., 2011). 평판 어포던스는 자신의 콘텐츠에 대한 통찰력을 제공하며 이로 인해 사용자들이 우월한 타인과의 비교를 통해 상대적인 위치를 평가하게 하는 상향사회비교를 촉진한다. 이러한 상향사회비교는 SNS에서 선의적 선망 및 악의적 선망을 유발하는 요소로 알려져 있다 (Wu & Srite, 2021). 따라서,

H3a: 평판 어포던스는 선의적 선망에 정(+)의 영향을 미친다.
H3b: 평판 어포던스는 악의적 선망에 정(+)의 영향을 미친다.

IV. 예상 분석 방법

선망 유발을 위해서는 회상 방법(Recall method)을 사용할 예정이다. 데이터 분석을 위해서는 두 단계의 접근 방식을 실시할 예정이다. 먼저, 측정 모형을 평가하기 위해 확증적 요인 분석 (CFA)을 실시할 예정이다. 그후, 구조 방정식 모델링 (SEM)을 활용하여 연구 가설을 검증할 예정이다.

V. 예상 기여점

본 연구는 SNS에서 어떤 어포던스가 선의적 선망과 악의적 선망을 유발하는지를 연구했다. SNS의 기술적 특성이 선망에 미치는 영향에 대해 연구가 진행되어 왔음에도 불구하고(Krasnova et al., 2015; Wallace et al., 2018), SNS 플랫폼 내의 특정한 어포던스가 각각 선의적 선망과 악의적 선망 혹은 두 유형의 선망을 모두 유발하는지를 이해하는 데 있어 연구적 공백이 남아 있었다.

어포던스의 독특한 역할을 탐구함으로써, 이 연구는 SNS 특성과 두 가지 전혀 다른 유형의 선망 간의 복잡한 상호작용에 대한 이해를 깊이 있게 하고자 한다. 이러한 세밀한 이해는 디지털 환경에서 발생하는 선망의 개념을 더욱 정교하게 이해하는 데 도움을 주며, 사용자들 사이에서 다양한 감정적 반응을 이끌어내는 특정 메커니즘에 대한 통찰력을 제공한다.

Ⅵ. 참고문헌

Appel, H., Gerlach, A. L., & Crusius, J. (2016). The interplay between Facebook use, social comparison, envy, and depression. *Current Opinion in Psychology*, *9*, 44-49.

Grgecic, D., Holten, R., & Rosenkranz, C. (2015). The impact of functional affordances and symbolic expressions on the formation of beliefs. *Journal of the Association for Information Systems*, *16*(7), 2.

Karahanna, E., Xu, S. X., Xu, Y., & Zhang, N. (2018). The needs–affordances–features perspective for the use of social media. *MIS Quarterly*, *42*(3), 737-A23.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons*, *54*(3), 241-251.

Majchrzak, A., Faraj, S., Kane, G. C., & Azad, B. (2013). The contradictory influence of social media affordances on online communal knowledge sharing. *Journal of Computer-Mediated Communication*, *19*(1), 38-55.

Shao, Z., Zhang, L., Chen, K., & Zhang, C. (2020). Examining user satisfaction and stickiness in social networking sites from a technology affordance lens: uncovering the moderating effect of user experience. *Industrial Management & Data Systems*, *120*(7), 1331-1360.

Smith, R. H., & Kim, S. H. (2007). Comprehending envy. *Psychological bulletin*, *133*(1), 46-64.

Suh, A., Cheung, C. M., Ahuja, M., & Wagner, C. (2017). Gamification in the workplace: The central role of the aesthetic experience. *Journal of Management Information Systems*, *34*(1), 268-305.

이하 생략

OTT 플랫폼 고착화와 구독 지속 의도에 관한 연구

이서현
경북대학교 경상대학
suny8362@gmail.com

이새롬
경북대학교 경상대학
saeromlee@knu.ac.kr

Abstract – 다양한 OTT 스트리밍 서비스가 등장하며 콘텐츠 업계의 경쟁은 강화되고 있다. 본 연구는 자극-조직-반응(Stimulus-Organism-Response: S-O-R) 모델을 통하여 OTT 구독 지속 의도에 미치는 영향을 파악하고자 하였다. 연예인 특성이 자극으로써 준사회적 상호작용을 불러오고 OTT 이용과 충족 요소를 통해 고착화와 구독 지속 의도에 영향을 미치는 연구모형을 설정하고 검증하였다.

Keywords – 연예인 신뢰도. 준사회적 상호작용, OTT, S-O-R 모델

I. 서론

기술의 발전을 통해 다양한 OTT 스트리밍 서비스가 등장하며 콘텐츠 업계의 경쟁은 날이 갈수록 강화되고 있다. 특히 전 세계적으로 팬데믹을 겪으면서 실내에 머무르는 시간이 길어짐으로 인해 OTT 이용자의 수가 급격히 증가하기도 하였다 (김선미 등, 2023). 이러한 상황에서 OTT 가 중요한 의의를 가지는 이유는 많은 이용자들을 끌어들이고 있음은 물론 사람들의 콘텐츠 시청 방식에 큰 변화를 주고 있기 때문이다.

본 연구에서는 OTT 의 지속 사용에 영향을 미치는 요인을 자극-조직-반응(Stimulus-Organism-Response: S-O-R) 모델을 통하여 제시한다. S-O-R 모델에서 자극에 해당하는 OTT 콘텐츠의 특성으로 OTT 콘텐츠 등장 연예인 특성을 제시하였다. 또한 해당 연예인의 특성이 시청자와 준사회적 관계를 형성하여 해당 자극에 대한 반응으로써 구독 지속 의도와 고착화를 형성하는 과정에 대한 메커니즘을 설명하고자 한다.

II. 본론

2.1 이론적 배경

OTT(Over The Top) 서비스는 인터넷 망을 이용하여 여러 단말기를 통해 동영상 콘텐츠를 제공하는 서비스를 의미한다 (정용국 & 장위, 2020). OTT 는 몰아보기 등 새로운 콘텐츠 시청 형식을 만들어내며 사람들의 미디어 소비 습관을 바꾸고 있다. 이러한 변화는 시청자를 적극적 미디어 사용자가 되게 하며, 따라서 최근 연구에서는 다양한 OTT 서비스 이용자의 이용 행태에 초점을 맞춘 연구가 많이 진행되고 있다 (Tefertiller, 2017). 하지만, OTT 콘텐츠의 어떤 자극을 통하여 어떠한 조직 반응이 생겨나고 어떤 반응이 나타나는지에 대한 체계적인 연구는 아직 부족한 실정이다.

S-O-R 모델은 1974 년 Mehrabian 과 Russell 에 의하여 제시되었다. 해당 모델은 자극, 조직 및 반응이라는 세 가지 기본 요소를 포함한다(Tang et al., 2015). 환경적 자극이 사람의 내부 상태인 조직으로 나타나고, 이는 행동적 반응의 결과를 불러올 수 있다는 것을 보여주는 모델이다. 이 모델에서의 자극은 개인에게 영향을 미칠 수 있는 환경적 요인이며, 조직은 자극에 의해 유발된 개인의 정서적, 인지적 상태를 나타낸다(Olfat et al., 2019). 마지막으로 반응은 자극으로 인해 유발된 결과를 의미한다(Suh & Prophet, 2018). 본 연구에서는 OTT 환경에서의 자극으로 콘텐츠에 등장하는 연예인 신뢰도를 제시한다. 조직으로는 준사회적 상호작용과 시간 때우기, 즐거움을 제안하며 이로 인한 반응으로써 고착화와 OTT 구독 지속 의도가 나타난다고 본다.

준사회적 상호작용은 미디어에 등장하는 대상에게 상상적이고 일방적 관계를 형성하는 상호작용으로 정의된다(Horton & Wohl, 1956). 이는 단순하고 일시적인 미디어 소비를 넘어서는 개념이다. 장기적이고 지속적이며 긍정적인 일방적 감정을 특징으로 하며(류은아, 2020), 따라서 사람들의 행동에 크게 영향을 줄 수 있기에 미디어 소비에 있어서 중요하게 연구되어야 하는 부분이다.

이용과 충족 이론은 개인이 미디어를 선택하고 소비함으로써 어떠한 필요나 욕구를 충족시킨다는 것을 설명하는 이론이다(Severin & Tankard, 1997). 해당 이론을 적용한 기존 연구들에서는 주로 즐거움, 정보 탐색, 편리성, 자기표현 등을 연구변수로 사용하였다. 본 연구는 OTT 플랫폼 사용 상황에서 나타날 수 있는 이용과 충족 이론 변수로 시간 때우기와 즐거움을 제안하고자 한다.

2.2 가설

본 연구에서는 S-O-R 모델과 준사회적 상호작용, 이용과 충족 이론을 중심으로 OTT 구독 지속 의도에

영향을 주는 변수들에 대한 가설을 다음과 같이 설정하였다.

가설 1: 연예인 매력도는 준사회적 상호작용에 긍정적 영향을 미칠 것이다.

가설 2: 연예인 진실성은 준사회적 상호작용에 긍정적 영향을 미칠 것이다.

가설 3: 연예인 전문성은 준사회적 상호작용에 긍정적 영향을 미칠 것이다.

가설 4: 준사회적 상호작용은 즐거움에 긍정적 영향을 미칠 것이다.

가설 5: 준사회적 상호작용은 시간 때우기에 긍정적 영향을 미칠 것이다.

가설 6: 즐거움이 높을수록 구독 지속 의도가 커질 것이다.

가설 7: 즐거움이 높을수록 고착화가 커질 것이다.

가설 8: 시간 때우기가 높을수록 구독 지속 의도가 커질 것이다.

가설 9: 시간 때우기가 높을수록 고착화가 커질 것이다.

2.3 데이터 수집 및 분석

본 연구모형 검증을 위하여 설문을 통하여 데이터를 수집하였다. 선행연구를 참고하여 본 연구에 맞게 설문지를 구성하고, 전문 설문조사 업체를 통하여 250 부의 응답을 수집하여 데이터 분석에 활용하였다.

SmartPLS4 프로그램을 사용하여 구조방정식 모형으로 가설을 검증한 결과, 모든 가설이 지지되는 것으로 확인되었다. 이러한 연구의 전반적인 결과는 연예인 매력도, 연예인 진실성, 연예인 전문성으로 나타나는 연예인 신뢰도가 준사회적 상호작용에 긍정적으로 영향을 미쳐서 즐거움, 시간 때우기를 통해 고착화와 구독 지속 의도에 긍정적인 영향을 준다는 것을 시사한다.

III. 결론

본 연구는 S-O-R 모델을 통하여 OTT 구독 지속 의도에 영향을 미치는 요인들을 검증하고자 하였다. 본 연구의 결과는 연예인 신뢰도가 높은 연예인이 OTT 플랫폼의 고착화와 구독 지속 의도를 높일 수 있다는 것을 보여주며 향후 연구자들과 실무자들에게 도움이 될 수 있다. 먼저, 본 연구는 S-O-R 모델을 적용하여 OTT 콘텐츠에서 자극이 되는 연예인의 특성이 어떻게 결과적으로 구독 지속 의도와 고착화에 영향을 미치는지에 대한 이해를 높인다. 또한 OTT 콘텐츠 연예인 특성이 시청자의 행동에 미치는 영향에 대해 더 깊이 이해함으로써 OTT 플랫폼이나 콘텐츠 제작사가 더 효과적으로 콘텐츠를 제작하고 사용자 경험을 개선할 수 있도록 도울 수 있다.

IV. 참고문헌

김선미, 이상원, & 손현정. (2023). 동영상 OTT 플랫폼 사업 전략과 정책 방향: 전문가 델파이 분석. *방송통신연구*, 43-73.

류은아. (2020). SNS 플랫폼에 따른 소셜미디어 인플루언서 평판의 마케팅 커뮤니케이션 효과: 준사회적 상호작용과 동일시의 매개효과를 중심으로: Marketing communication effect of social media influencer reputation according to SNS platform: Focus on the mediating effect of the parasocial interaction and identification.

정용국, & 장위. (2020). 구독형 OTT 서비스 특성이 이용자 만족과 지속 사용 의도에 미치는 영향: 넷플릭스 이용자를 대상으로. *한국콘텐츠학회논문지*, 20(12), 123-135.

Horton, D., & Richard Wohl, R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. psychiatry, 19(3), 215-229.

Olfat, M., Tabarsa, G. A., Ahmadi, S., & Shokouhyar, S. (2019). Disclosing the bright side of SNs in the workplace: A comparison between ESNs and public SNs in benefiting the organizations. Journal of Enterprise Information Management, 32(3), 390-412.

Severin, W. J., & Tankard, J. W. (1997). Communication theories: Origins, methods, and uses in the mass media (pp. 300-310). New York: Longman.

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. Computers in Human behavior, 86, 77-90.

Tang, J., Zhang, P., & Wu, P. F. (2015). Categorizing consumer behavioral responses and artifact design features: The case of online advertising. Information Systems Frontiers, 17, 513-532.

Tefertiller, A. (2017). Moviegoing in the Netflix age: Gratifications, planned behavior, and theatrical attendance. Communication & Society, 30(4), 27-43.

숏폼 영상의 특성이 중독에 미치는 영향 및 결과 : 상대과정이론 관점

한민정
경북대학교 경영학부
hmj9656@gmail.com

이새롬
경북대학교 경영학부
saeromlee@knu.ac.kr

Abstract – 숏폼 영상은 최근 디지털 미디어 소비의 급격한 증가와 함께 중요한 플랫폼으로 부상하고 있다. 이러한 현상은 짧은 시간 동안에도 효과적으로 메시지를 전달하고 관객의 주의를 끌며 높은 사용자 참여율을 유도하는 효과적인 수단으로 인식되고 있다. 그러나 숏폼 영상 사용에서 중요한 문제 중 하나는 사용자 중독입니다. 숏폼 영상은 사용자의 인지된 즐거움과 금단 증상을 활성화하여 중독에 영향을 미칠 수 있습니다. 이러한 현상은 OPT(Opponent Process Theory) 관점에서 설명될 수 있으며, 사용자는 긍정적 감정을 유지하거나 부정적 감정을 줄이기 위해 숏폼 영상을 반복적으로 사용한다. 이로 인해 중독으로 이어질 수 있다. 또한, 중독은 심리적 웰빙과 부정적인 연관성이 있다. 이번 연구는 OPT 관점에서 숏폼 영상의 특성이 중독에 미치는 영향과 중독이 심리적 웰빙이 미치는 영향을 논의하고자 하며, 해당 현상에 대한 새로운 이해를 제공할 것으로 기대된다.

Keywords – 숏폼 영상 중독, 상대과정 이론, 심리적 웰빙

I. 서론

최근 디지털 미디어 소비의 급격한 증가와 함께 숏폼 영상이 중요한 플랫폼으로 떠오르고 있다. 숏폼 영상은 빠르게 변화하는 사람들의 라이프스타일에 응하고(Wang 2020) 짧은 기간에 청중에게 훌륭한 콘텐츠를 제공한다(Wright 2017). 이러한 현상은 짧은 시간 동안에도 효과적으로 메시지를 전달하고 관객의 주의를 끌며 높은 사용자 참여율을 유도하는 효과적인 수단으로 인식되고 있다. 숏폼 영상의 성공적인 활용은 시각적, 감각적 효과를 강조하며 동시에 간결하고 명료한 이야기 구조를 갖추는데, 이는 집중력 부족과 정보 과부하와 같은 시대적 특성과 부합한다. 더불어, 소셜 미디어 플랫폼의 급격한 성장과 이용자들의 미디어 소비 패턴 변화는 숏폼 영상의 인기를 증폭시키는 요인 중 하나로 작용하고 있다. 대표적인 숏폼 동영상 플랫폼인 틱톡은 전 세계적으로 약 16억 회 다운로드 되었다(Yurieff, 2019). 그러나 숏폼 영상 사용에서 가장 두드러진 문제 중 하나는 사용자 중독으로, 이는 사람들의 업무, 생활, 건강에 악영향을 미칠 수 있다(Andreassen and Pallesen, 2014).

이번 연구는 상대성과정이론(Opponent Process Theory: OPT) 관점에서 숏폼 영상의 상대적인 감정을 통해 숏폼 영상의 특성이 중독에 미치는 영향과 숏폼 영상 중독이 심리적 웰빙에 미치는 영향을 논의하고자 한다.

II. 관련 연구

숏폼 영상 중독

숏폼 영상은 사용자가 앱 플랫폼을 통해 온라인 숏폼 영상을 제작하고 공유 및 협업 기능을 실현하는 웹 2.0 시대의 떠오르는 소셜 미디어이다(Kuss and Griffiths, 2017). 그러나 2019년 인터넷 동향 보고서에 따르면 모바일 단편 동영상으로 인해 사용자가 스마트폰에 더 많은 시간을 소비하게 되었으며(Dai, Tai, Ni 2021), 중국에서 단편 동영상 앱의 일일 평균 사용량이 6억 시간으로 증가했다(Meeker 2020). 짧은 기간, 풍부한 콘텐츠, 개인화된 추천, 몰입형 경험을 갖춘 짧은 형식의 비디오는 사용자를 더욱 쉽게 중독시킨다(Anderson 2020; Huang, Hu, and Chen 2022; Meng and Leung 2021). 일부 연구에서는 숏폼 영상이 사람들의 시간을 빼앗고 주의를 산만하게 만들어 본질적인 일에 집중하기 어렵게 만드는 '시간 블랙홀'이라고 주장한다(Zhao 2021). 독특한 특성을 지닌 새롭고 널리 퍼진 소셜 미디어 유형인 짧은 형식의 동영상은 중독을 일으키기 쉽고 사용자에게 부정적인 영향을 미칠 수 있다. 따라서 숏폼 영상 중독은 큰 관심을 끄는 중요한 주제가 되었으며 여러 분야에서 수많은 연구를 진행했다(Hasan, Jha 및 Liu 2018; Huang, Hu 및 Chen 2022; Liu, Ni 및 Niu 2021; Ye et al 2022; Zhang, Wu, Liu 2019).

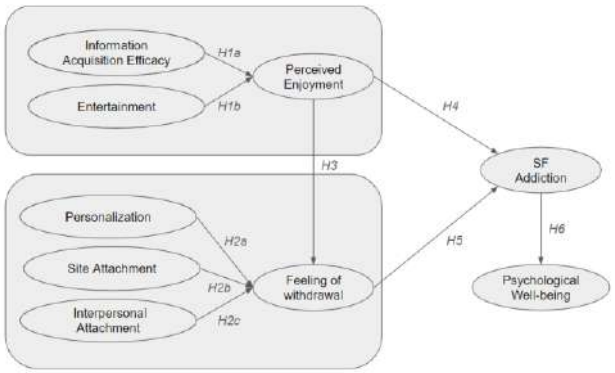
상대과정이론

중독 및 중독과 관련된 내성을 조사하기 위한 대안적이지만 핵심적인 이론적 틀은 반대 과정 이론(Solomon and Corbit, 1974)에서 찾을 수 있다. Solomon & Corbit(1974)은 일단 시작된 쾌락, 정서 또는 감정 상태는 쾌락 감정의 강도를 감소시키는 메커니즘을 통해 중추 신경계에 의해 자동으로 조절된다고 가정했다. 상대 과정 이론은 두 가지 과정으로 정의된다. a-과정에는 정서적 또는 쾌락적 습관화가 포함되고 b-과정에는 정서적 또는 쾌락적 철수가 포함된다. b-과정은 a-과정이 종료된 후에 나타나며, 반복 노출에 따라 더 커진다(Solomon 1980). 이 관점은 중독, 아동 애착, 사랑과 애정, 스릴을 추구하는 행동, 두려움과 회피의 역학, 우울증과 행복감, 비만 등 다양한 심리학 연구 영역에서 다양한 인간 행동을 이해하는데 적용되어 왔으며(Solomon and Corbit, 1974) 특히, OPT 관점은 특정 기술 시스템(예: 숏폼 영상)이 중독성이 강한 이유를 설명하고 있다(Turel and Serenko, 2012). 긍정적 강화와 부정적 강화를 기반으로 한 OPT(Opponent Process Theory) 관점에서, 사용자는 긍정적인 감정을 유지하거나 부정적인 감정을 줄이기 위해 짧은 형식의 비디오와 상호 작용을 반복해야 한다. 이러한 사용 패턴이 지속될 경우, 중독으로 이어질 수 있다.

심리적 웰빙

심리적 웰빙 이론에 따르면, 개인의 심리적 건강은 삶의 특정 측면에서 긍정적인 기능을 하는 데 달려 있다. 개인은 긍정적인 관계를 유지하고 환경을 조절하며, 자기 수용과 과거를 인정하며, 목표와 의미를 가지며, 개인적 발전과 자기 결정 능력을 가져야 한다(Özen, 2005). 이러한 이유로 심리적 안녕, 행복, 발달 사이에는 잠재적인 긴장이 존재합니다(Ryff and Singer, 1998). "심리적 웰빙"은 이전에 우울증, 불안, 분노와 같은 부정적인 심리적 특성이 없는 것으로 명시되었다. 그러나 이 용어에 대한 관점은 시간이 지남에 따라 바뀌었다(Yılmaz, 2013). 심리적 웰빙은 긍정적인 이해 및 긍정적인 사회적 관계와 관련이 있다(Akın, 2009). 즉, 심리적 웰빙은 자신이 경험하는 긍정적인 감정에 다른 사람이 영향을 미치도록 허용하지 않고 자신의 만족감을 표현하는 개인의 능력으로 정의할 수 있다(Diener et al., 2003). 심리적 웰빙은 스트레스와 기타 정신적 문제가 없는 건강한 생리 외에도 자기 수용, 타인과의 긍정적인 관계, 자율성, 환경 지배, 삶의 목적, 개인적 성장과 밀접한 관련이 있습니다(Ryff, 1989).

III. 연구모델 및 가설



<그림 1> 연구모델

짧은 시간에 정보를 효율적으로 습득할 수 있는 정보습득 효율성과 소셜 미디어 사용을 예측하는 중요한 요인인 오락(Chen and Leung, 2016; Oh and Syn, 2015; Yu and Oh, 2018)은 사용자에게 만족감을 주고 시간을 보내는데 도움이 되므로 (Gan and Li, 2018) 인지된 즐거움인 긍정적인 감정을 가지게 된다. 따라서 다음과 같은 가설을 제안한다.

H1a: 정보습득 효율성은 인지된 즐거움에 정(+)의 영향을 미친다

H1b: 오락(Entertainment)은 인지된 즐거움에 정(+)의 영향을 미친다.

그리고 고객이 더 많은 편의성, 더 낮은 비용 또는 기타 혜택을 누릴 수 있도록 제품이나 서비스의 일부 기능을 맞춤화하는 것으로 정의된 개인화는 개인화 프로세스에 대한 고객의 참여가 소셜 미디어에 대한 기능적 의존도에 긍정적인 영향을 미친다는 것을 확인했다(Thongpapanl and Ashraf, 2011). Wan et al. (2017)). 또한 사용자의 정서적 유대감을 의미하는 사이트 애착((Kim et al., 2016b)은 웹사이트(쇼핑 및 여행 사이트)에 대한 사용자 충성도에 영향을 미치는 중요한 요소임을 설명한다고 입증되었다(Jiang and Dong, 2008). 그룹 구성원간의 유대, 그룹 응징력에 대한 대인관계 개념, 개인이 서로에 대한 매력으로 정의되는 대인관계 애착(Sassenberg, 2002)은 온라인 우정을 유지하는데 도움이 된다. 따라서 다음과 같은 가설을 제안한다.

H2a: 개인화는 금단 증상에 정(+)의 영향을 미친다

H2b: 사이트 애착은 금단 증상에 정(+)의 영향을 미친다

H2c: 대인관계 애착은 금단 증상에 정(+)의 영향을 미친다.

숏폼 영상은 사용자의 인지된 즐거움과 금단 증상을 활성화시켜 중독에 영향을 미치는 것으로 나타났다. OPT는 정서가 반대되는 패턴, 즉 사용자가 긍정적인 감정을 유지하거나 부정적인 감정을 줄이기 위해 자극을 반복적으로 활성화해야 한다고 설명한다(Solomon and Corbit, 1978). OPT에 따르면, 긍정적 강화와 부정적 강화 과정은 밀접한 관련이 있다(Lee et al., 2014). 긍정적 강화가 부정적 강화보다 선행되며, 금단 증상의 발생은 본질적으로 긍정적 경험이 감소됨에 따라 부정적 경험이 상대적으로 강화될 때 나타난다. 다시 말해, 숏폼 영상이 중단되거나 감소될 때, 긍정적 경험이 감소하면서 부정적 경험의 심리적 및 생리적 반응이 더욱 강화되는 것으로 나타났다(Greenfield, 2012; Solomon and Corbit, 1974). 따라서 다음과 같은 가설을 제안한다.

H3: 인지된 즐거움은 금단 증상에 정(+)의 영향을 미친다.

H4: 인지된 즐거움은 숏폼 영상 중독에 정(+)의 영향을 미친다.

H5: 금단 증상에 숏폼 영상 중독에 정(+)의 영향을 미친다.

인터넷 중독과 같은 다른 관련 형태의 중독 행동에 대한 연구에서는 심리적 웰빙과 중독 행동 사이의 부정적인 연관성이 지속적으로 보고되고 있다(Cardak, 2013 , Casale et al., 2015). 심리적 웰빙은 주로 개인의 삶의 목적, 개인적 성장, 타인과 긍정적인 관계를 맺을 수 있는 능력에 중점을 둔다(Choi & Lim, 2016). 일부 개인은 중독에 빠질 때 심리적 웰빙을 잃는 경향이 있지만, 다른 사람들은 중독을 통해 일시적인 쾌락이나 안정감을 얻는 것으로 인해 심리적 웰빙을 강화할 수도 있다. 또한, 중독은 자기 통제력을 상실하고 심리적 스트레스를 증가시킬 수 있으며, 이는 심리적 웰빙을 저하시키는 요인으로 작용할 수 있다. 따라서 심리적 웰빙과 문제가 있는 인터넷 사용 사이의 부정적인 연관성은 이전 연구에서 입증되었다(Casale et al., 2015). 따라서 다름과 같은 가설을 제안한다.

H6: 숏폼 영상 중독은 심리적 웰빙에 부(-)의 영향을 미친다.

Ⅳ. 예상 분석 방법

온라인 설문조사로 데이터를 수집할 예정이다. 구조 방정식 모델링 (SEM)을 활용하여 연구 가설을 검증할 예정이다.

Ⅴ. 예상 기여점

이번 연구는 숏폼 영상의 특성이 중독에 미치는 영향과 심리적 웰빙이 숏폼 영상 중독을 조절하는데 어떤 역할을 하는지에 대해 탐구한다. 특히, 숏폼 영상의 상대적인 감정을 중심으로 한 OPT 관점을 적용하여 이를 분석한다. OPT 관점을 통해 숏폼 영상의 중독적 특성을 이해하는데 새로운 이론적 접근을 제시한다. OPT를 통해 숏폼 영상 소비의 감정적 측면을 보다 깊이 있게 이해하고 중독 현상의 메커니즘을 설명한다. 숏폼 영상의 상대적 감정을 이해함으로써 중독을 예방하고 사용자의 심리적 웰빙을 증진시키는 방안을 모색할 수 있을 것이다. 그리고 숏폼 영상의 중독적 특성과 심리적 웰빙 간의 관계를 조사하여 해당 분야에 대한 새로운 지식을 확장한다. 이는 미디어 중독 및 온라인 건강에 대한 이해를 높이고 이를 관리하는데 도움이 될 것이다.

따라서, 이 연구는 숏폼 영상 중독에 대한 이해를 높이고 예방 및 관리에 기여함으로써 학문적, 실용적, 지식적 관점에서 중요한 기여를 할 것으로 예상된다

Ⅵ. 참고문헌

Tian, X., Bi, X., & Chen, H. (2023). How short-form video features influence addiction behavior? Empirical research from the opponent process theory perspective. Information Technology & People, 36(1), 387-408.

Zhang, X., Wu, Y., & Liu, S. (2019). Exploring short-form video application addiction: Socio-technical and attachment perspectives. Telematics and Informatics, 42, 101243.

Chen, Y., Li, M., Guo, F., & Wang, X. (2023). The effect of short-form video addiction on users' attention. Behaviour & Information Technology, 42(16), 2893-2910.

Ponnusamy, S., Iranmanesh, M., Foroughi, B., & Hyun, S. S. (2020). Drivers and outcomes of Instagram Addiction: Psychological well-being as moderator. Computers in human behavior, 107, 106294

이하 생략

Keep an Eye on Me: The Role of Online Communities as a Commitment for Healthier Behaviors

김가윤
한국과학기술원 기술경영학부
kgyoon@kaist.ac.kr

민동규
한국과학기술원 기술경영학부
donggyu.min@kaist.ac.kr

김영인
한국과학기술원 기술경영학부
ykim3@kaist.ac.kr

이철호
한국과학기술원 기술경영학부
irontigerlee@kaist.ac.kr

Abstract – *Individuals join online weight-loss communities (OWCs) to push themselves into an audience that observes their self-control. Despite being an irrational act that involves contributing resources such as time, energy, and attention, and incurs costs, we define the voluntary choice to stand before an audience as "audience commitment." This paper uses data from an online healthcare platform to explore which individuals engage with community features as a form of audience commitment and assesses its impact on health behaviors. To address endogeneity problems, we employ an instrumental variable approach and coarsened exact matching. The study reveals that individuals with hyperbolic preferences tend to utilize the community more as a form of commitment and that community has a positive and significant effect on health behaviors. Such commitment proves more effective when influenced by a strong network, offline-based relationships, and larger community size. We provide insights into the potential underlying mechanisms of how communities help alleviate self-control problems, from before to after individuals join them. This research contributes to the existing literature on the role of commitment in mobile environments.*

Key Terms – *mHealth, online weight-loss communities (OWCs), commitment, self-control*

리뷰 작성자의 문화적 배경요인이 리뷰 유용성에 미치는 영향

김태이
부산대학교 경영대학
kimtaeyi@pusan.ac.kr

홍태호
부산대학교 경영대학
hongth@pusan.ac.kr

Abstract 온라인 리뷰는 리뷰어 정보와 리뷰 정보로 구성되었으며, 소비자들은 리뷰어의 프로파일 정보, 평점, 그리고 텍스트를 포함하는 리뷰 정보를 통해 유용한 정보를 확보할 수 있다(야오즈옌 등 2020). 기업들은 유용한 정보를 구매자들에게 제공하기 위해 유용성 투표 시스템을 도입하였으며 소비자들은 이를 통해 신뢰할 수 있는 정보를 얻으며 구매 결정을 손쉽게 할 수 있게 되었다. 리뷰 유용성은 작성 리뷰에 대한 유용한 정도를 평가하는 것으로 유용성이 높은 리뷰는 제품이나 서비스 효과적으로 설명하고 있다고 볼 수 있다. 리뷰 유용성에 미치는 영향요인으로는 평점, 리뷰길이, 리뷰 내용 등으로 선행연구를 통해 많은 역할을 한다는 것이 입증되었으며 특히 리뷰의 텍스트는 리뷰 유용성에 가장 큰 영향력을 보이는 것으로 나타났다(Liu & Park, 2015). 텍스트 리뷰는 텍스트에서 직접적으로 나타나지 않는 감성 및 감정 뿐만 아니라 제품과 서비스의 다양한 속성도 포함되어 있어 유용한 리뷰를 찾아낼 수 있는 중요한 역할을 한다. 텍스트는 온라인 리뷰 컨텐츠에서 제품을 이해할 수 있는 중요한 요인으로 구매자는 제품을 평가할 때 리뷰 텍스트를 의존하며 판매자도 리뷰 텍스트를 통해 소비자를 이해할 수 있다(Zhu et al., 2017). 한편, 최근 이커머스 기업들이 플랫폼 사업을 확장함에 따라 국경의 경계가 점차 모호해지고 문화적 다양성을 고려한 현지화 전략 등이 기업의 글로벌 전략에 중요한 요소로 부상하고 있다. 미국의 아마존과 한국의 쿠팡은 플랫폼 사업을 확장하면서 다양한 국가의 소비자들에게 상품을 판매하고 있으며 해외배송, 번역서비스 등을 제공하면서 구매의 장벽으로 존재했던 부분들을 해소해 판매 접근성을 높였다. 그러나 구매 경험이 많지 않은 해외 소비자들은 제품이나 서비스에 대한 불확실성으로 자신의 구매의사 결정을 위해 정확하고 상세한 리뷰를 찾게 되며, 이러한 리뷰는 제품에 대한 기대치를 설정해 구매 결정 과정에서 발생할 수 있는 불확실성을 감소시키는 데 도움을 준다. 해외에서 물건을 구매하는 소비자에게 온라인 리뷰는 구매 결정에 필수적인 정보를 제공하며, 리뷰의 내용이 구매자의 선택과 만족도에 직접적으로 영향을 미치는 중요한 요소라 할 수 있다. 또한, 판매자는 이러한 리뷰를 통해 구매 경험의 차이를 파악하고, 소비자의 행동과 선호도를 분석할 수 있다. 리뷰를 통해 구매자의 문화적 다양성을 이해하고 적절한 대응을 하는 것은 이커머스 기업이 글로벌 시장에서 다양한 국가의 소비자들과 효과적으로 소통하고 소비자의 요구를 충족시키기 위한 필수적 요소로 간주될 수 있다. Hong et al., (2016)은 리뷰 작성 및 의사결정 과정에서 문화적 차이가 정보 전달에 영향을 미칠 수 있음을 언급하며, 리뷰 텍스트를 통해 특정 문화 내에서의 가치와 행동 경향성을 살펴볼 수 있다고 하였으며 그 외 많은 선행연구에서 온라인 리뷰에서 문화 간의 차이가 나타난다는 연구를 수행바있다(Wang et al., 2019; Fang et al., 2013; Jia, 2020; Oh et al., 2020). 이에 따라 본 연구는 온라인 리뷰 텍스트를 중심으로 문화이론을 적용하여 구매자의 행동의 차이를 살펴보고 리뷰 유용성을 분석하고자 한다. 본 연구는 다양한 배경과 특성을 가진 리뷰를 광범위하게 확보할 수 있는 **Amazon.com** 의 리뷰 데이터를 이용해 **Hofstede** 의 문화차원이론을 바탕으로 리뷰와 리뷰작성자의 특성에 따른 리뷰 유용성을 측정한다. 일반적으로 문화적 차이나 문화적 다양성 연구는 국가 단위로 설명되거나 연구가 진행되고 있으나 모든 국가의 사람들이 동질적인 특성을 갖지 않으며 한 국가 내에서도 사람들이 저마다 다양한 문화적 스펙트럼을 가지고 있기 때문에(De Mooij, 2011) 문화적 차이를 개인 수준으로 측정하고자 한다. 본 연구는 문화적 배경이 리뷰 유용성에 미치는 영향을 탐색하기 위해 선행연구에서 제시한 리뷰 특성과 리뷰어의 특성을 기반으로 모형을 구축한다. 리뷰의 특성은 감정, 감성, 평점 등의 정보를 이용하며 리뷰어의 특성으로는 전문성, 경험, 평판을 사용한다. 본 연구는 유용하고 신뢰할 수 있는 리뷰를 선별해 소비자는 더 나은 구매 결정을 내릴 수 있는 방안을 모색하고 판매자 측면에서는 마케팅 활용방안으로 사용할 수 있도록 방향성을 제시하고자 한다.

Keywords – 리뷰유용성, 리뷰특성, 리뷰어 특성, 문화적 배경요인

Session B

B1. [학술세션]
디지털 데이터를 활용한 소비자 인사이트 분석

좌　장: 이 형 용(한성대)

Exploring the impact of informational factors in online reviews on review helpfulness

야오즈옌
부산대학교 BK21 디지털금융교육연구단
yaoziyan@pusan.ac.kr

홍태호
부산대학교 경영학과
Hongth@pusan.ac.kr

김종원
동의대학교 경영정보학과
jokim@deu.ac.kr

김은미
부산대학교 경영연구소
keunmi100@pusan.ac.kr

Abstract – *The review helpfulness serves as an evaluative indicator, enabling the distinction of beneficial product reviews by reflecting the information quality of online reviews. Reviews deemed highly helpful can provide online consumers with valuable information more promptly, reducing uncertainties regarding products or services and facilitating purchasing decisions. Prior research has examined the impact of various factors related to the content of online reviews on individual review helpfulness. However, there is a gap in the analysis of the relationship between informational elements of reviews (text, images, videos) and review helpfulness. While the content of online reviews, such as text length and product images, has been shown to influence review helpfulness, further validation and clarification of the non-linear relationship between these content factors and review helpfulness are necessary. Furthermore, given the acknowledged variations among consumers based on product types highlighted in prior research, it is imperative to validate the moderating effect of product types.*

This study collected and analyzed data from online reviews posted on JD.COM, a leading online shopping platform in China. Through thorough analysis, we confirmed a non-linear relationship between review length and the number of review images with review helpfulness, characterized by an inverted U-shaped curve. Additionally, the presence of review videos was found to impact review helpfulness significantly, and the information richness of reviews positively influenced review helpfulness. Moreover, the study investigated the moderating effect of product types by examining reviews of various products, and identifying differences in the impact of informational factors on review helpfulness based on product types. This study examined the impact of textual, image, and video content in online reviews, and the information richness reflecting these three types of content, on review helpfulness. Through analysis, it elucidated the influence and significance of various content components in shaping online reviews and their influence in aiding information processing for online shopping. These findings deepen our understanding of the role of online review structures in evaluating review helpfulness, assist platforms in formulating more effective strategies for review generation and recommendation, and provide guidance for reviewers to write reviews that are more beneficial to potential consumers. Furthermore, this research contributes to the academic literature on information systems and e-commerce by empirically investigating consumers' information processing, review helpfulness, and product types, enhancing scholarly discourse in these fields.

Key Terms – *Review helpfulness, Information richness, Product type, Inverted U-shape.*

Unveiling the Impact of Fake Reviews on Consumer Behavior: An Integrated Approach Using Topic Modeling and Explainable AI

Chernyaeva, Olga
Pusan National University
College of Business Administration

missleka@pusan.ac.kr

Hong, Taeho
Pusan National University
College of Business Administration

hongth@pusan.ac.kr

Kim, Jong-weon
Dong-Eui University
College of Management Information System

jokim@deu.ac.kr

Abstract

The digital marketplace's reliance on online reviews presents a paradox. In contrast, genuine reviews facilitate informed purchasing decisions, fake reviews threaten the integrity of these interactions, misleading consumers and distorting market dynamics. This study aims to address the challenges posed by fake reviews through the lens of detection and its subsequent impact on consumer purchase intentions. By integrating Explainable Artificial Intelligence (XAI) and topic modeling techniques, this study aims to unveil the distinguishing features of fake versus real reviews and examine how these characteristics influence the likelihood of purchase.

Purpose: The core objective of this study is to enhance the accuracy and transparency of fake review detection processes and to understand the effect of review authenticity on consumer behavior towards purchasing. Applying XAI to topic modeling predictions, this research seeks to elucidate the rationale behind model decisions, thus creating more interpretable and trustworthy detection mechanisms. It also investigates the dynamics between the authenticity of reviews and consumers' purchasing intentions, offering a comprehensive analysis of the power of online reviews.

Methodology: This research utilizes the Amazon.com dataset for online reviews, applying topic modeling to identify an optimal number of topics and categorize reviews based on topic probabilities. The study then predicts fake reviews by analyzing whether a product was purchased. SHapley Additive exPlanations (SHAP) values are employed to determine the features most predictive of a review's authenticity.

In the subsequent phase, the same dataset is analyzed to predict consumer purchase intentions, using the derived topics and the classification of reviews as fake or real as inputs. Again, SHAP values are used to identify the key features influencing purchase decisions. This streamlined approach allows for a focused analysis of the relationship between review characteristics, their authenticity, and their impact on consumer behavior.

Contributions to Academia and Practice: Through its methodological rigor, this study advances the theoretical foundations of fake review detection and its implications for consumer purchase behavior. By integrating topic modeling with XAI techniques, specifically SHAP values, the research illuminates the complex decision-making processes underlying predictive models. These insights not only contribute to the development of more accurate and transparent detection systems but also provide e-commerce platforms with actionable strategies to mitigate the influence of fake reviews. The examination of features impacting purchase intentions further equips businesses with the knowledge to enhance the authenticity of online review ecosystems, fostering a more trustworthy environment for consumers.

Implications: The application of SHAP values in analyzing both fake review detection and purchase intention predictions underscores the potential of XAI in enhancing model interpretability across various domains. For academia, this methodology offers a replicable framework for future research in online fraud detection and consumer psychology. Practically, it empowers e-commerce stakeholders to prioritize transparency and reliability, ultimately leading to improved consumer satisfaction and loyalty. Bridging theoretical exploration with practical application, this study significantly contributes to ethical practices in digital commerce. In

conclusion, by meticulously analyzing the Amazon.com dataset through advanced topic modeling and Explainable Artificial Intelligence (XAI) techniques, this research addresses critical challenges in distinguishing fake reviews and understanding their impact on consumer purchase decisions. The strategic use of SHAP values to interpret model predictions offers a ground-breaking approach to enhancing the transparency and trustworthiness of an AI model designed for detecting online fake reviews. As the digital marketplace continues to evolve, the findings of this study mark an important step in promoting ethical consumer engagement and maintaining the integrity of online review platforms.

Keywords:

Fake Review Detection, Topic Modeling, Explainable Artificial Intelligence, Online Review Analysis, SHAP.

SNS 분석 기반 팝업스토어 고객 경험 분석 및 몰입개념을 통한 성공요인 분석

황지선
아주대학교 경영대학
bibionidae@ajou.ac.kr

황은혜
아주대학교 경영대학
heh3800@ajou.ac.kr

강주영
아주대학교 경영대학
jykang@ajou.ac.kr

Abstract *– 유통업계의 D2C(Direct to Consumer)로의 전환으로 인하여 대표적인 오프라인 마케팅 수단으로 떠오른 팝업 스토어(Pop-up Store)는 기록적인 수치를 기록하며 성행하고 있다. 이러한 상황에 비해 팝업스토어 마케팅의 소비자반응을 분석하고 앞으로의 팝업스토어의 방향을 제시하는 연구와 기업문화는 부족한 상황이다. 본 연구는 팝업 스토어의 소비자반응을 SNS 텍스트 데이터의 감성분석과 빈도분석 통하여 분석하고 데이터의 결과를 기반으로 몰입개념을 적용하여 팝업스토어의 성공요인을 분석하였다. 선행연구를 참고하여 팝업 스토어 유형을 판매형, 노출형, 고객 획득형, 시장 테스트형의 마케팅 목적에 따라 구분하고, 각 집단별로 팝업 스토어의 목표성, 수행성, 자기보상, 방문 구조, 참여성, 집단성의 성공요인을 확인하였다. 본 연구의 분석 결과를 바탕으로 팝업 스토어의 소비자 인식을 확인하고, 팝업 스토어의 성공요인을 도출하여 앞으로의 팝업 스토어 마케팅의 방향을 제시할 수 있을 것으로 사료된다. 향후 후속 연구를 통해 효과적인 팝업 스토어 마케팅 전략의 방향과 팝업스토어의 소비자 반응 데이터 분석에 대한 논의가 이루어지길 기대한다.*

Keywords *– 몰입, 텍스트 마이닝, 팝업스토어, SNS 분석*

A Study on Factors Affecting Academic Achievement of College Students: Focusing on Fuzzy set Qualitative Comparative Analysis

권승인
한성대학교 스마트융합컨설팅학과
zpizgi100@hansung.ac.kr

박영식
한성대학교 경영학과
youngsik.park@hansung.ac.kr

박정민
한국발명진흥원
paspark1@gmail.com

이형용
한성대학교 경영학부
leemit@hansung.ac.kr

Abstract – *This study attempted to identify the importance of variables (Feature Importance) by applying the SMOTE technique to the dataset used in Yılmaz & Sekeroglu(2019)'s "Higher Education Students Performance Evaluation" and then applying the algorithm of supervised learning classification analysis. In addition, a complementary study with quantitative comparative methodology was conducted using fuzzy set qualitative comparative analysis (FsQCA). For the classification analysis, we used Decision Trees, Support Vector Machines, Logistic Regression, Gradient Boosting Machine (GBM), and Random Forests. Since the algorithm with the best performance was Random Forest, the variable importance ranking was calculated through Random Forest. Python 3.11.5 version was used as the implementation language and environment, and R-4.3.1 version and Rstudio 2023 were used as the implementation environment. As a result of the analysis, looking at the importance of the variables up to 5th place, the importance was confirmed in the order of 'Last semester's cumulative grade point average', 'Scholarship type', 'Total income', 'Expected graduation cumulative grade point average', and 'Number of siblings'. Using the relevant variables as causal conditions, the minimum simple solution model for the FsQCA analysis results is Last semester's cumulative grade point average*~Scholarship type*~Total income + Last semester's cumulative grade point average*Scholarship type*Number of siblings + Last semester's cumulative grade point average*~Total income*Expected graduation cumulative grade point average → GRADE. The significance of this study is that it attempted to identify effective variables in the field of education through complementary research using both quantitative and qualitative comparative methodologies. Meanwhile, according to prior research, the influence of the student's family background and the desire for academic achievement have a relatively large influence in Korea, whereas the study group, Cyprus university students, are largely influenced by self-efficacy, such as the individual's desire for academic achievement. appear.*

Key Terms – *Academic Achievement, Decision Tree, Random Forest, FsQCA, SMOTE*

Session B

B2. [산업체세션]
취·창업을 위한 산업체 동향

좌 장: 한 동 욱(전주대)

©YoungLimWon Soft Lab

Low-code 앱 개발 SaaS "플렉스튜디오"를 활용한 AI 앱 개발하기

(주)영림원소프트랩

2024.05 ver.

Flextudio

Contents

01 회사 소개

02 로우코드 노코드 특징

03 어플리케이션 개발

04 플렉스튜디오 소개

회사소개

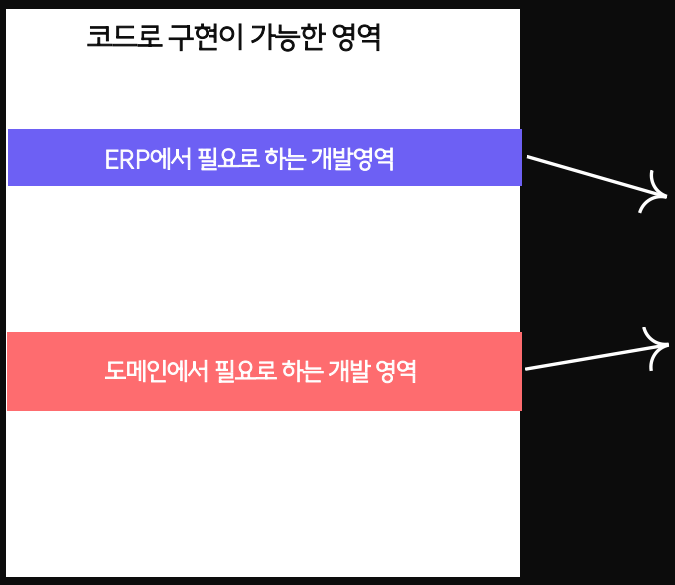
영림원소프트랩

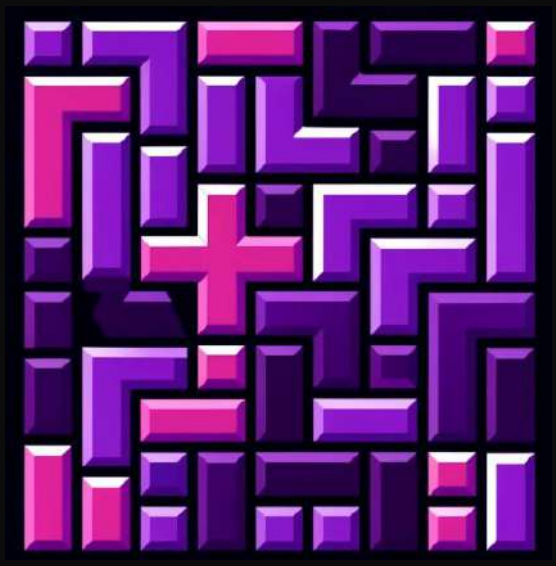
ERP

30년

4만 여개 프로그램

로우코드의 필연





로우코드의 필연

로우코드 특징

구축형

학습곡선 높음

복잡한 로직

기능풍부-유연함

B2B

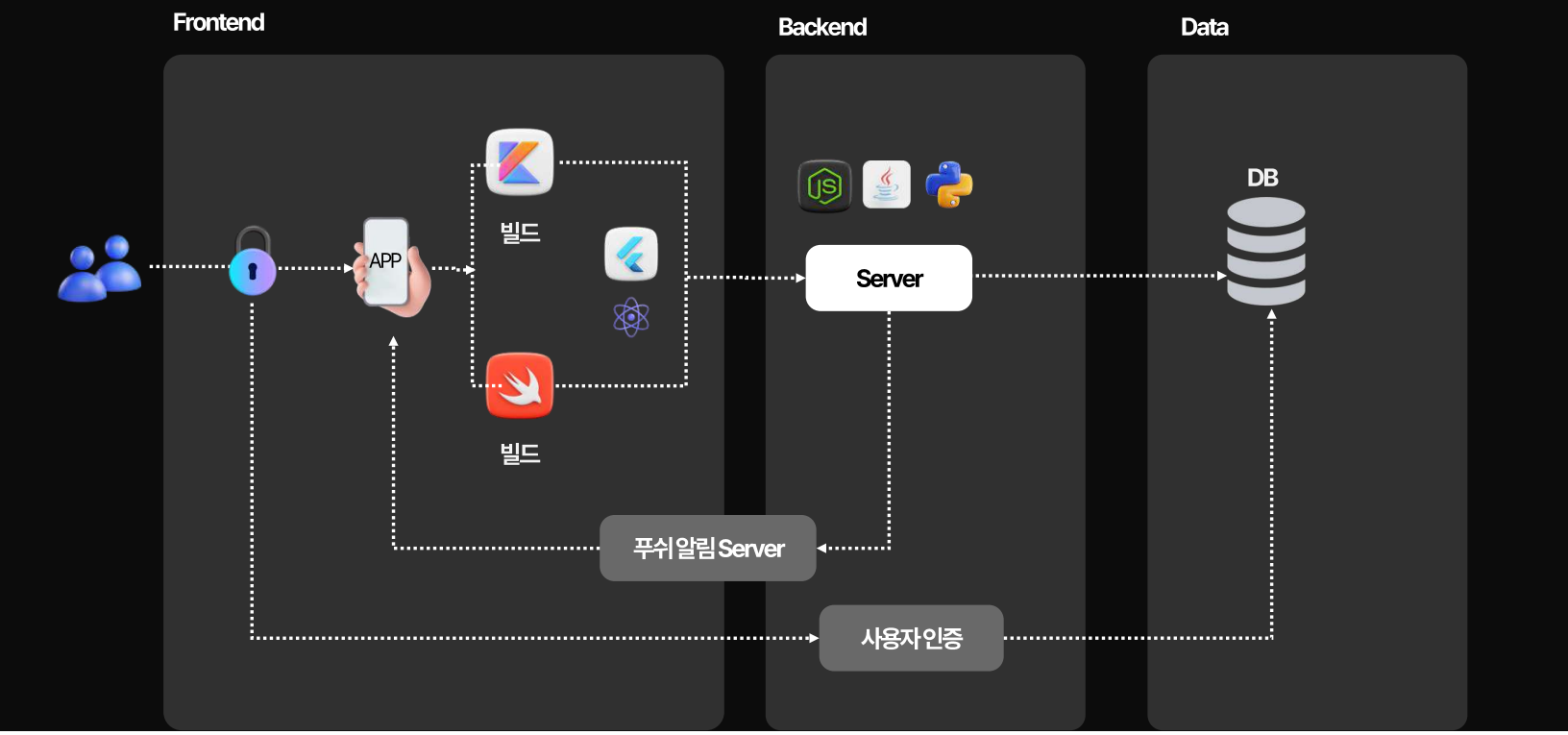
높은비용

로우코드 특징 노코드 특징

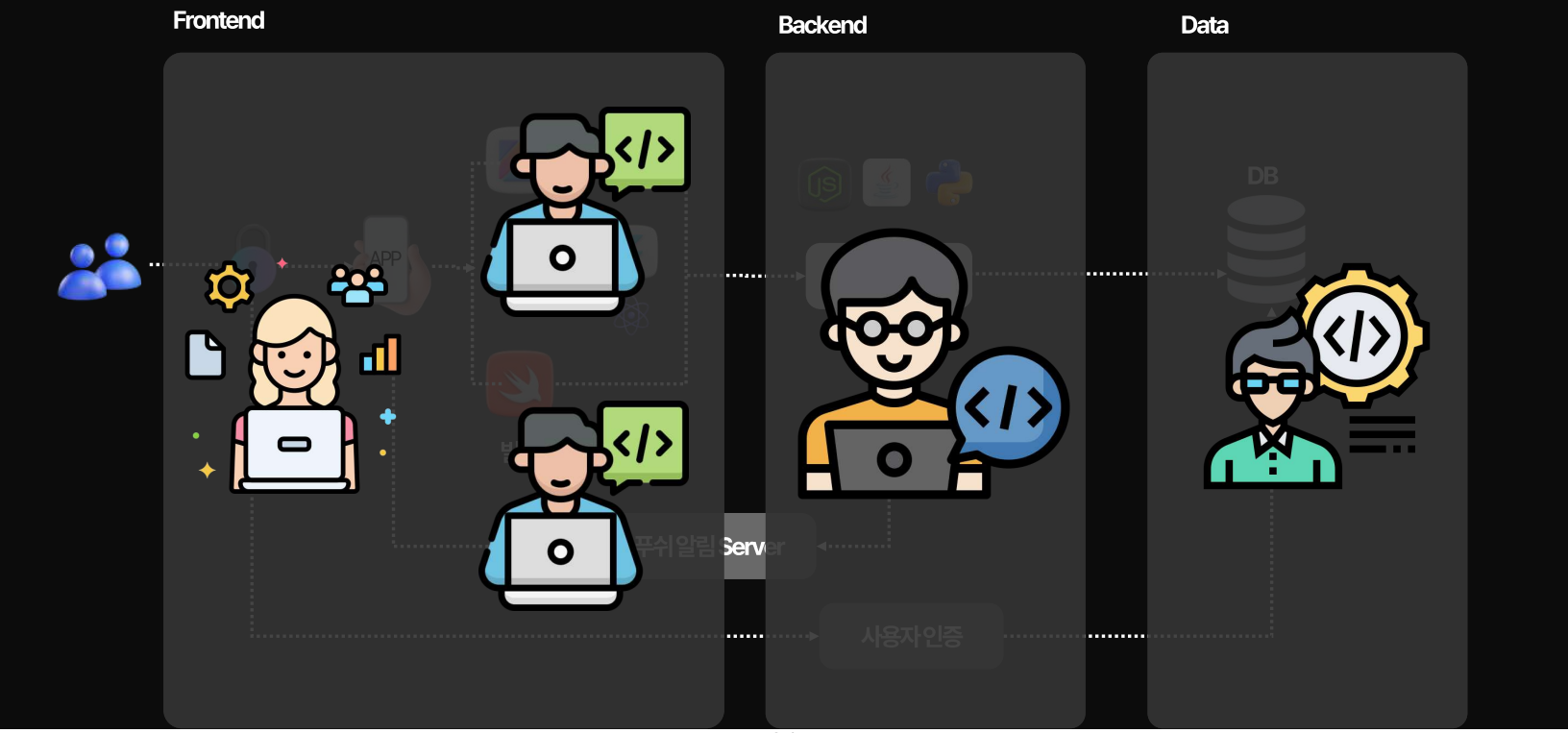
로우코드	노코드
구축형	SaaS
학습곡선 높음	낮음
복잡한 로직	단순한 기능
기능풍부-유연함	코딩지식 없이
B2B	비(시민)개발자
높은비용	구독형-상대적 저렴

노코드 로코드 플랫폼의 기능 세분화

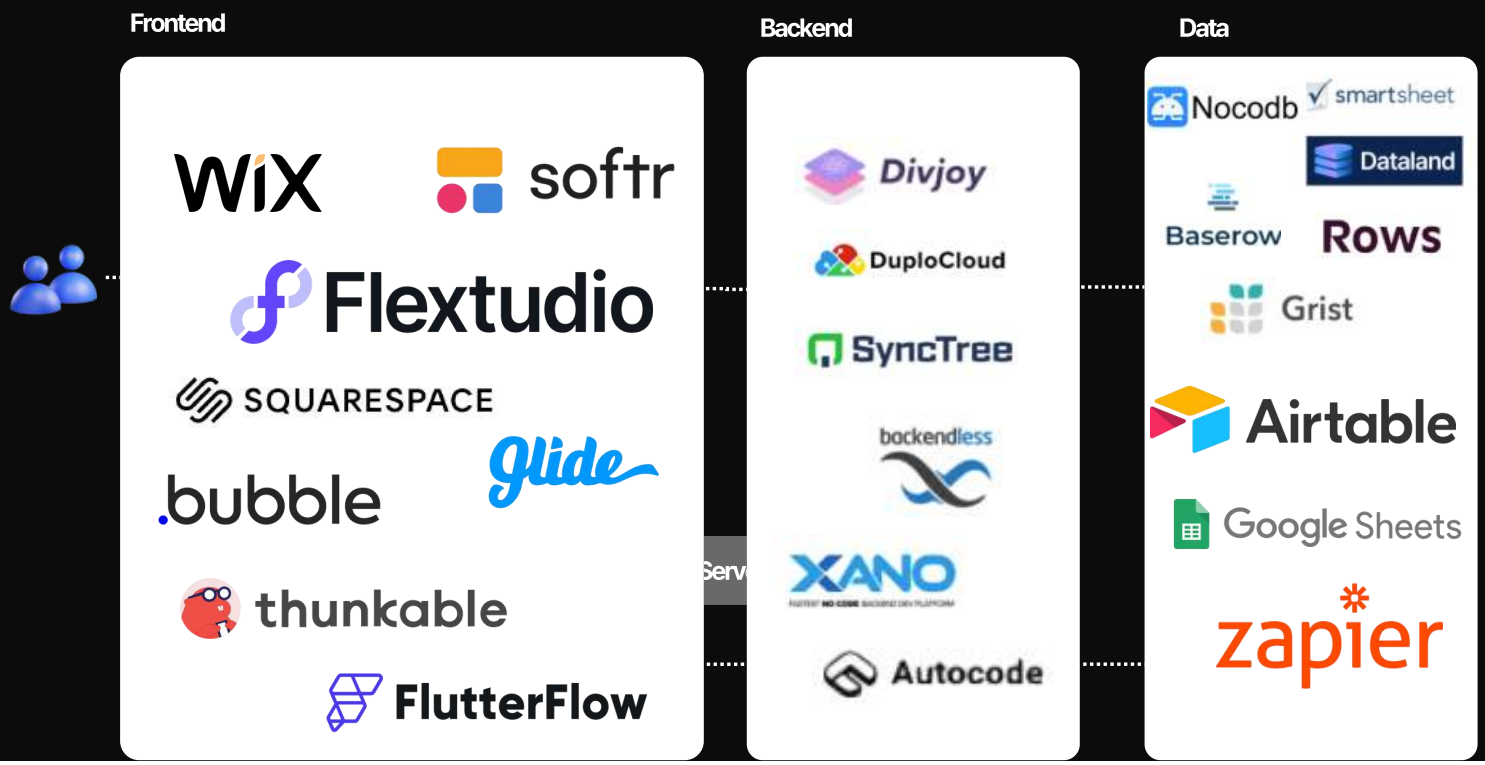
어플리케이션의 개발 방법



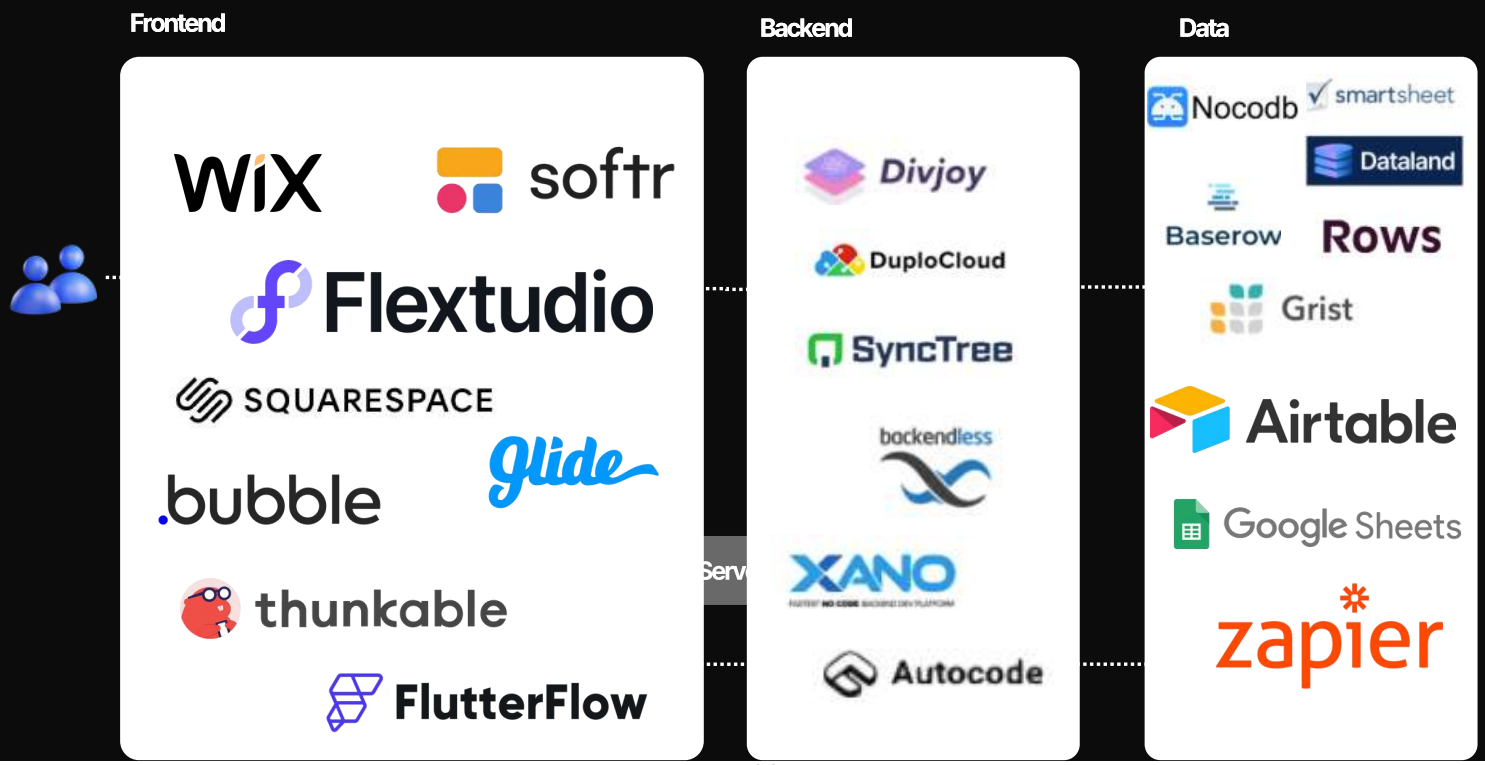
어플리케이션의 개발 방법



어플리케이션의 개발 방법



어플리케이션의 개발 방법



어플리케이션의 개발 방법

Frontend

WIX
softr
Flextudio
SQUARESPACE
Glide
bubble
thunkable
FlutterFlow

Backend

Divjoy
DuploCloud
SyncTree
backendless
XANO
Autocode

Data

Nocodb
smartsheet
Dataland
Baserow
Rows
Grist
Airtable
Google Sheets
zapier

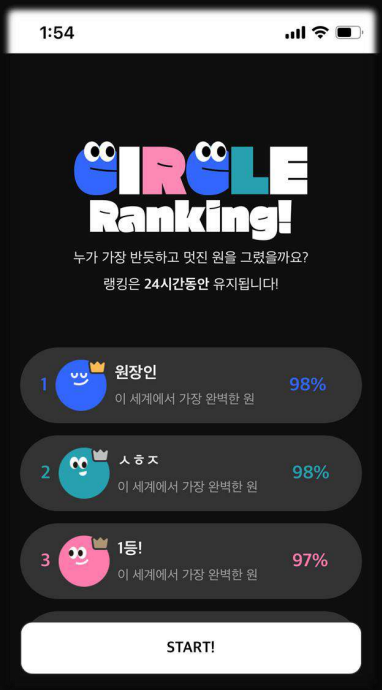
AI

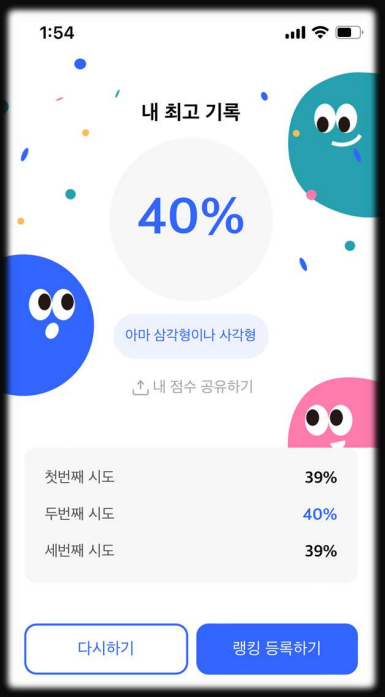
ChatGPT
Gemini
upstage
GitHub Copilot
DALL·E
Jasper
copy.ai
colossyan
synthesia
CLOVA X

Flextudio 이벤트

GOOD CIRCLE
구겨지지 않은 반듯한 원 그리기

체험해보기





모바일 앱 개발 SaaS 플렉스튜디오 →

기획부터 배포까지,
번거로운 앱 제작은 이제 그만!

Flextudio

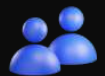
MAKE APPLICATION EASIER

26

어플리케이션 개발, 몇 명이 필요할까?

	Project Manager	UI/UX Designer	Mobile Developer	Backend Developer	QA Engineer	DevOps Engineer
역할과 책임	프로젝트 처음부터 끝까지 관리 팀 Task 분배 일진행률 확인 . . .	목업(Mock-up) 및 스케치 작성 유저 flow 개발 시각적 요소 디자인 및 개발 . . .	목업 및 스케치를 코드로 변환 코드의 품질과 유지보수성 확인 모바일 앱 디버깅 . . .	앱 설계 및 로직 구현 API 개발 DB 및 클라우드 관리 . . .	요구사항 반영여부 확인 테스트 케이스 및 결과 문서화 개발자를 위한 피드백 제공 . . .	개발팀 및 운영팀과의 협의 조율 Integration과 delivery를 자동화하기 위한 툴 설계 . . .
7년차 연봉	5,834만원	4,481만원	5,984만원	5,586만원	4,570만원	5,926만원
주급	1,012,413원	777,619원	1,038,440원	969,377원	793,062원	1,028,376원

최소 6명



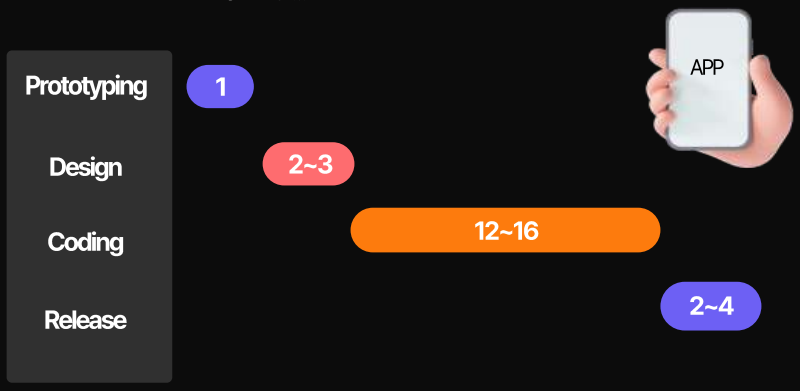
*연봉정보: 채용 정보에 포함된 직무 별 요구 경력(최저, 최대)과 연봉(최저, 최대)을 바탕으로 추정한 예상 데이터(출처 원티드)
*주급계산은 4대보험(9.4%)을 공제한 비용을 바탕으로 계산되었으며, 리소스 산정을 위해 편의로 계산되었음

어플리케이션 개발, 얼마나 걸릴까?

소요 개발기간(총 17~24주)

Duration Week

*앱 개발 난이도에 따라 기간은 상이할 수 있음



직무별 필요 업무 투입 기간

Resources

*24주를 기준으로 함

Project Manger: 12주

UI/UX Designer: 5주

Mobile Developer: 12주

Backend Developer: 8주

QA Engineer: 2주

DevOps Engineer: 2주

어플리케이션 개발, 비용은 어떻게 될까?

3억989만6223원

필요인력(7명) * 소요 개발기간(24주)

외주 개발..?

✓ 추후 유지보수에 대한 어려움

✓ 문제 발생 시, 책임소재가 비교적 불명확

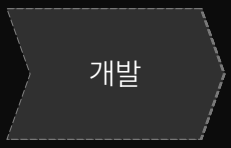
✓ 필요한(필요하게 된) 모듈, 기능에 따른 계속되는 추가 과금

.

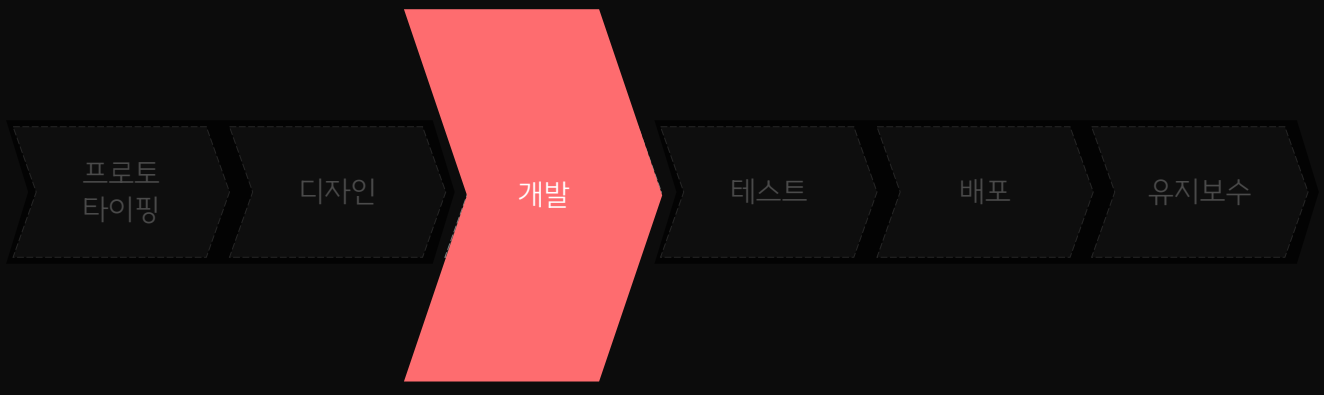
.

.

Flextudio 소개

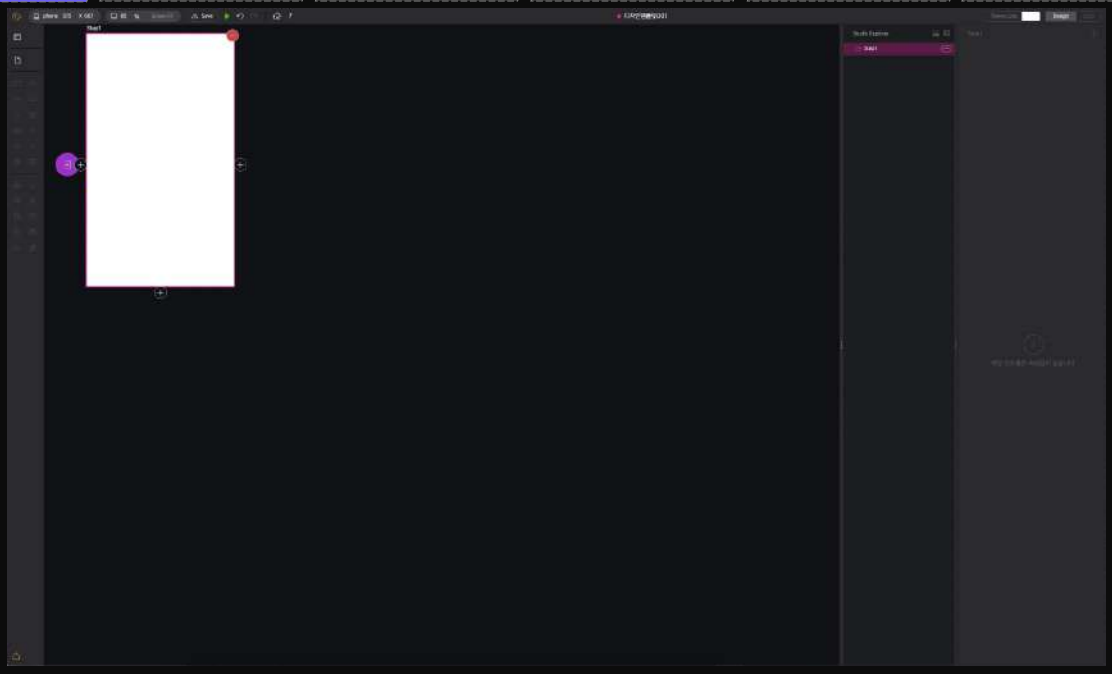


Flextudio 소개



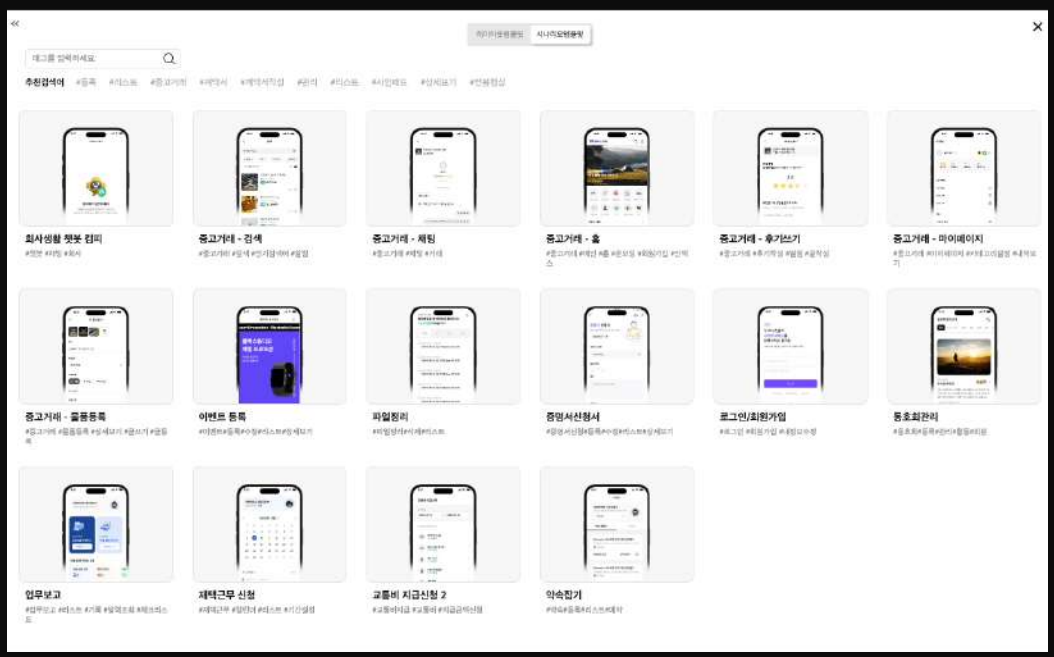
Flextudio 소개

프로토 타이핑 | 디자인 | 개발 | 테스트 | 배포 | 유지보수



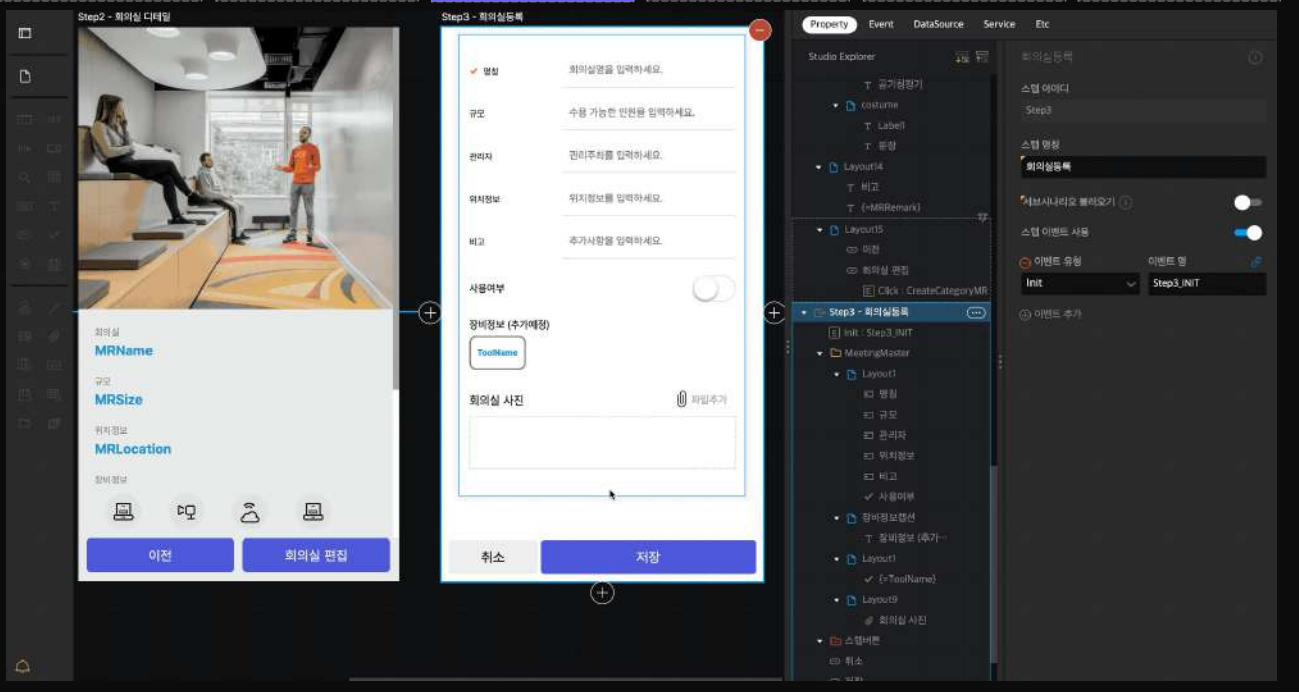
Flextudio 소개

프로토 타이핑 | 디자인 | 개발 | 테스트 | 배포 | 유지보수



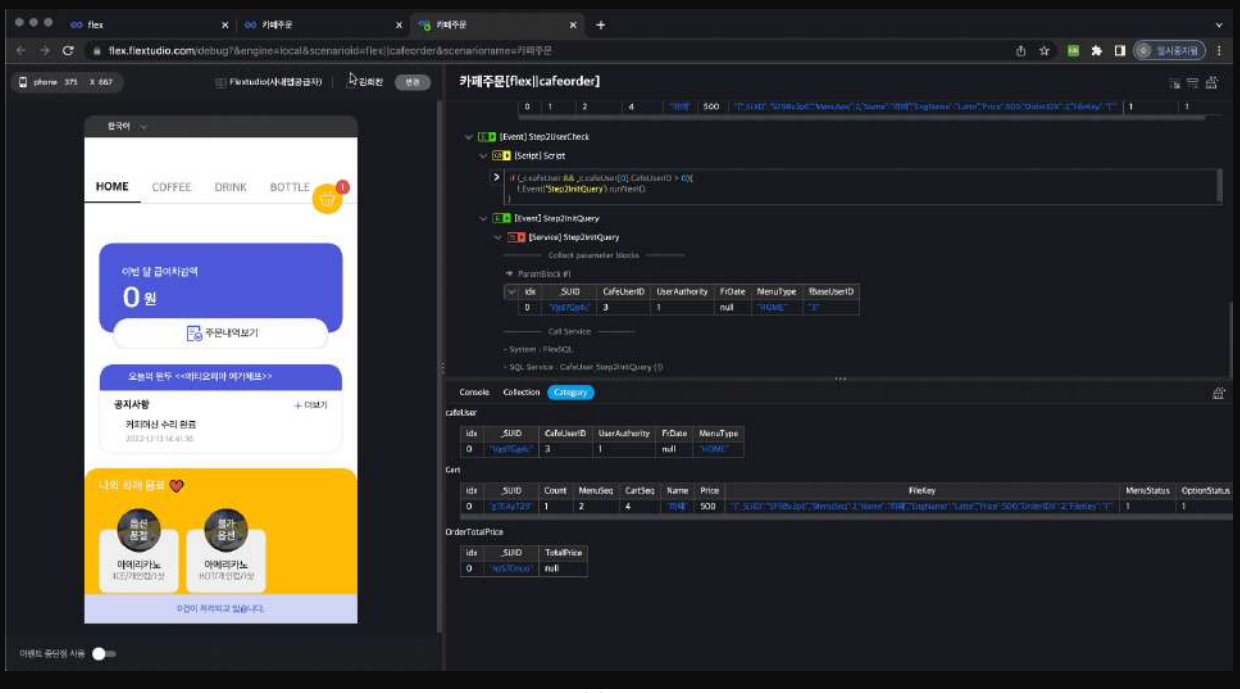
Flextudio 소개

프로토 타이핑 | 디자인 | 개발 | 테스트 | 배포 | 유지보수



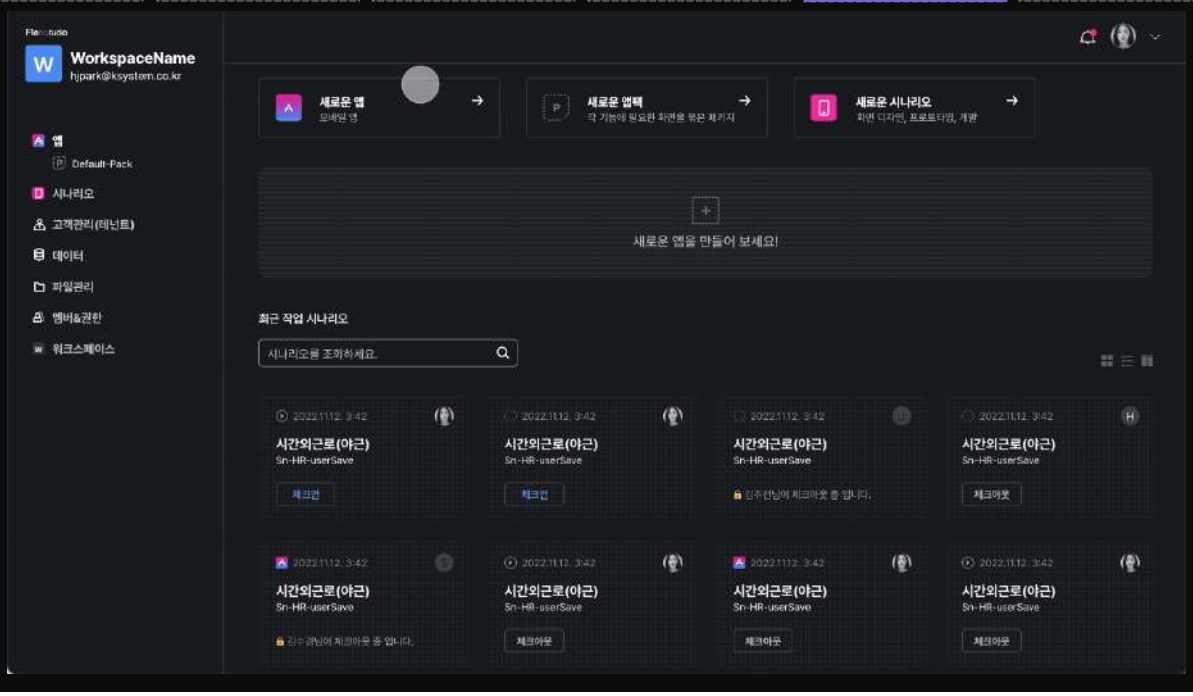
Flextudio 소개

프로토 타이핑 | 디자인 | 개발 | 테스트 | 배포 | 유지보수



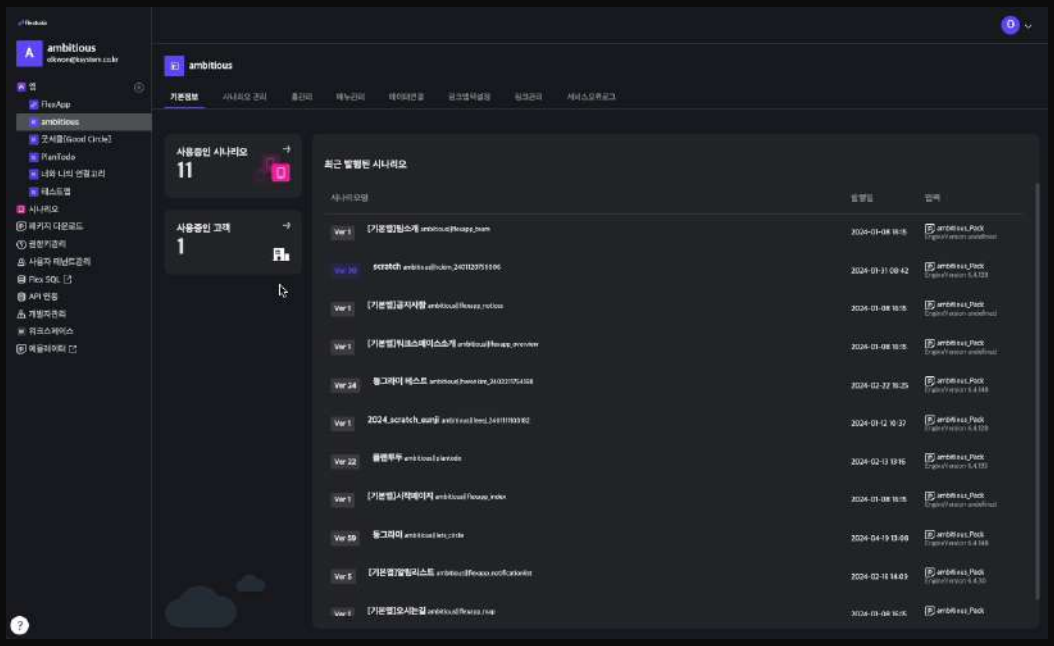
Flextudio 소개

프로토 타이핑 | 디자인 | 개발 | 테스트 | **배포** | 유지보수



Flextudio 소개

프로토 타이핑 | 디자인 | 개발 | 테스트 | 배포 | **유지보수**



어플리케이션 개발, 몇 명이 필요할까?

	Project Manager	UI/UX Designer	Mobile Developer	Backend Developer	QA Engineer	DevOps Engineer
역할과 책임	프로젝트 처음부터 끝까지 관리 팀 Task 분배 일진행률 확인 . . .	목업(Mock-up) 및 스케치 작성 유저 flow 개발 시각적 요소 디자인 및 개발 . . .	목업 및 스케치를 코드로 변환 코드의 품질과 유지보수성 확인 모바일 앱 디버깅 . . .	앱 설계 및 로직 구현 API 개발 DB 및 클라우드 관리 . . .	요구사항 반영여부 확인 테스트 케이스 및 결과 문서화 개발자를 위한 피드백 제공 . . .	개발팀 및 운영팀과의 협의 조율 Integration과 delivery를 자동화하기 위한 툴 설계 . . .
7년차 연봉	5,834만원	4,481만원	5,984만원	5,586만원	4,570만원	5,926만원
주급	1,012,413원	777,619원	1,038,440원	969,377원	793,062원	1,028,376원

최대 3명

*연봉정보: 채용 정보에 포함된 직무 별 요구 경력(최저, 최대)과 연봉(최저, 최대)을 바탕으로 추정한 예상 데이터(출처 원티드)
*주급계산은 4대보험(9.4%)을 공제한 비용을 바탕으로 계산되었으며, 리소스 산정을 위해 편의로 계산되었음

어플리케이션 개발, 비용은 어떻게 될까?

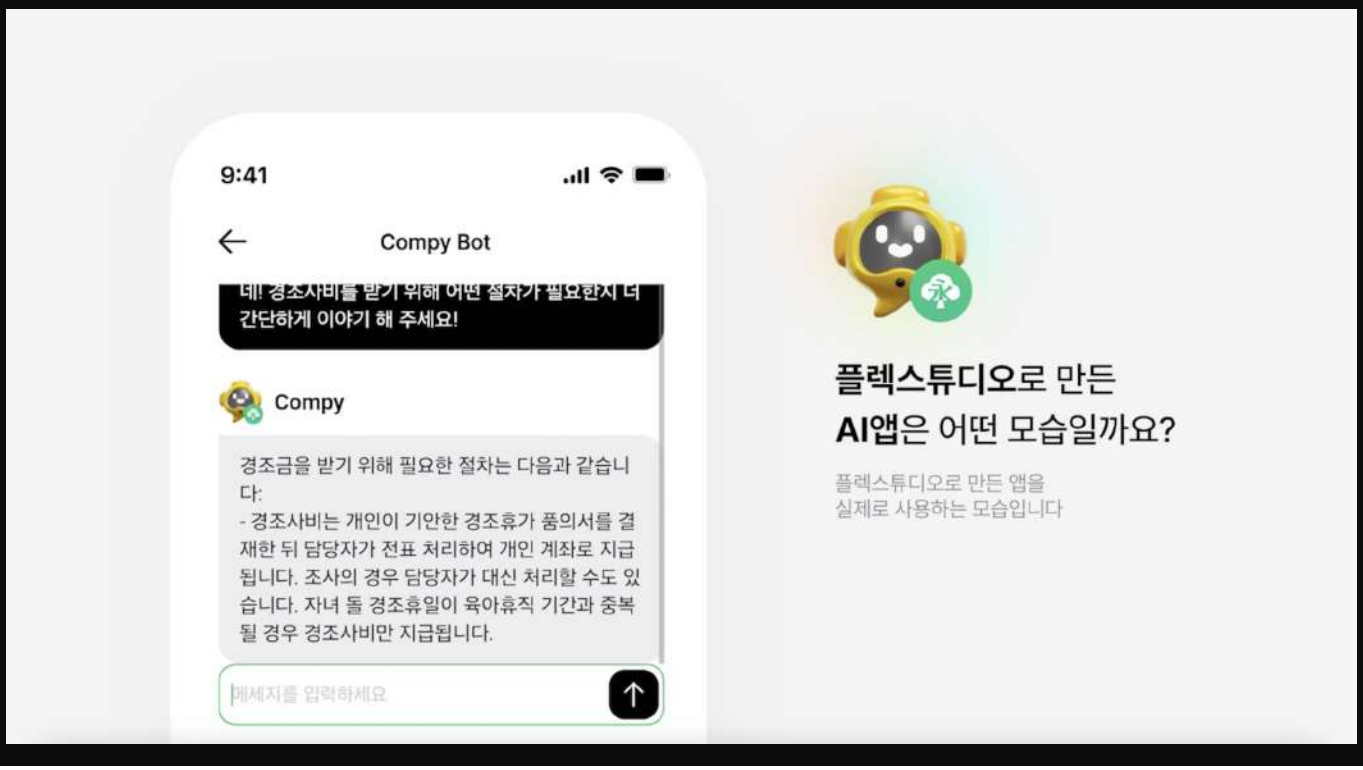
~~3억 989만6223원~~

~~필요인력(6명) * 소요 개발기간(24주)~~

1억 6,010만6223원

필요인력(3명) * 소요 개발기간(14주)

Chat GPT 활용하기 데모



감사합니다

www.flextudio.com

flextudio@ksystem.co.kr

xEV 전기차 모터풀리 외관검사 불량분석 및 인공지능 시스템

조영우
㈜엠티데이타
ywcho@mtdata.co.kr

김재경
㈜엠티데이타
qkqh970605@mtdata.co.kr

장마가
㈜엠티데이타
mgjang @mtdata.co.kr

Abstract *– 2023 년, 글로벌 전기차 시장은 예상을 뛰어넘는 성장을 이뤄냈다. 이는 주요 국가들의 전기차 보조금 정책 조정 및 고금리 여파에도 불구하고, 견조한 수요를 유지하며 성장했음을 나타낸다. 실제로, 전 세계적으로 등록된 전기차량 수는 전년 대비 약 40% 증가한 약 1242.7 만대로 추산된다. 특히 현대자동차그룹은 아이오닉 5/6, EV6, Niro, Kona 를 주력으로 세계적인 판매량을 기록하여 전년 대비 13.7%의 성장률을 달성했다. 이러한 성과는 전기차 시장이 지속적으로 성장하고 있으며, 이에 대한 수요가 높아지고 있음을 시사한다.*

이러한 시장 동향 속에서, 전기차 제조업체들은 효율적인 생산과 품질 관리에 대한 관심을 높이고 있다. 특히, 전기차 모터풀리 외관검사 및 불량분석은 제품 품질을 보장하기 위해 중요한 작업이다. 그러나 인력을 사용한 수작업으로 진행되는 이러한 작업은 인적 오류나 검수 시간의 증가 등의 문제를 야기할 수 있다.

이에 본 연구에서는 인공지능 기술을 활용하여 전기차 모터풀리 외관검사 시스템을 개발하고자 한다. 이를 위해 YOLOv4 알고리즘을 선택하여 객체 인식 모델을 구축하고, 현장에서 필요한 데이터를 수집하여 학습 데이터셋을 구성하였다. 이를 통해 실시간 객체 탐지 및 불량 부품 식별 시스템을 구현하고, 이를 현장에 즉시 적용하여 검수 시간을 단축하고 정확도를 향상시키는데 성공하였다.

본 연구의 최종 결과는 전기차 제조업체들에게 높은 효율성과 품질 보증을 제공함으로써, 전기차 시장의 성장과 함께 제품 생산과 품질 관리를 지원하는데 기여할 것으로 기대된다

Key Terms *– 객체인식, 모터풀리, 불량분석, 외관검사, 컴퓨터 비전*

Acknowledgement (사사)

이 논문은 2023 AI 융합 지역특화산업 지원사업에서 (경남 자동차부품산업 AI 융합 지원) 수행된 연구이다.

국내 법률 대상 다국어 번역을 위한 인공지능 학습 모델 연구

이정기
㈜타임게이트
jk.lee@time-gate.com

정회창
㈜엠티데이타
jeonghoechan@gmail.com

정민기
㈜엠티데이타
mingijeon83@gmail.com

Abstract *– 본 연구는 국내 거주 외국인이 우리 사회에서 다양한 국적과 언어를 사용하는 현실을 고려하여, 이들의 법률적 권리와 의무를 보장하기 위해 국내 거주 외국인 중 통계 인구수가 가장 높은 5개 국가의 언어를 대상으로 하여, 다국어 번역 기계 학습용 데이터를 구축하고 번역 모델을 개발하는 것을 목표로 한다. 기존에 MTPE(Machine Translation Post Editing)을 사용하여 번역을 시도하였으나, 본 연구에서는 Mbart, Mt5 인공지능 모델을 사용하여 실험을 통해 성능이 좋은 모델을 사용 했다. 본 연구의 결과는 다음과 같다. BLEU 지표를 활용하여 Mbart 모델이 다국어 법률 해석 분야에서 작지만 좋은 성능을 내어, 이 서비스가 법률 분야에서 언어 간 이해를 촉진하고 국제적인 법률 업무를 보다 효율적으로 수행하는데에 기여할 것으로 기대 된다.*

Key Terms *– 다국어번역, 법률적 권리, 번역 모델, 법률 서비스*

블록체인 기반 연합학습 리소스 공유 시스템

김명길
㈜스마트엠투엠
myeongkil@smartm2m.co.kr

권민호
㈜스마트엠투엠
alsgh458@smartm2m.co.kr

이현희
㈜스마트엠투엠
hyeonhui@smartm2m.co.kr

김요한
㈜스마트엠투엠
yohan@smartm2m.co.kr

Abstract *– 오늘날 인공지능 기술의 눈부신 발전과 클라우드 컴퓨팅의 광범위한 확산은 대규모 데이터의 처리와 저장 능력을 필수적인 요소로 만들었다. 이에 대응하고자 연합학습이 대두되었지만, 현대의 디지털 환경에서 요구되는 투명성과 신뢰성을 담보하기에는 여전히 한계가 존재한다. 이러한 한계를 극복하기 위해, 본 논문에서는 블록체인을 활용하여 연합학습의 투명성과 신뢰성을 강화한 리소스 공유 시스템을 제안한다. 본 시스템은 컴퓨팅 자원을 보유한 클라이언트들이 자원을 네트워크에 공유함으로써 데이터 처리와 학습 과정을 분산시키고, 이를 통해 경제적 및 환경적 비효율성을 줄이며 인공지능 학습의 접근성을 높일 수 있도록 설계되었다. 또한, 블록체인 환경에서 각 참여자에게 적절한 보상을 제공하여 능동적인 참여를 유도하고, 이를 통해 인공지능 학습의 접근성을 높이고 연구 개발을 촉진하는 선순환 구조를 조성한다. 본 논문은 연합학습과 블록체인 기술의 결합을 통한 새로운 가능성을 탐색하여, 제안 시스템의 구조와 동작방식을 체계적으로 설명한다. 또한, 연구 결과를 요약하고 앞으로의 연구 방향에 대해 논의하여, 이 분야의 지속적인 발전을 위한 기반을 마련한다.*

Key Terms *– 인공지능, 연합학습, 블록체인*

Session B

B3. [학술세션]
블록체인 기술의 응용

좌 장: 김 호 원 (부산대)

블록체인 기반 ESG 정보 공시 플랫폼 구축 방안 연구

최하눌
계명대학교
hanoolchoi@kmu.ac.kr

Abstract *– 본 연구는 ESG 자체 수준 진단과 외부 기관 평가 등을 통해 글로벌 시장 변화에 대응할 수 있는 블록체인 기반의 ESG 정보 공시 플랫폼 구축 방안을 제안하고자 한다. 사회 전반에서 지속 가능한 경영 및 사회적 책임에 대한 요구가 증가하고 있으며 투자자들과 소비자들은 기업들의 ESG 노력에 대한 높은 관심을 보인다. 각국 정부는 다양한 법 제도를 통해 환경과 인권, 사회적 책임에 대한 기업의 의무를 강화하고 있다. 최근 연구에 따르면 ESG 경영 지원을 위해 기업에 필요한 ESG 정보의 체계적인 관리와 이해관계자와 정보 공유가 중요해지고 있다. 이에 본 연구는 블록체인 기반으로 기업의 GRI, SASB, SDGs 등 ESG 평가 지표에서 요구하는 정보 자산을 관리하고 이를 공시하여 외부 기관 평가나 실사에 대응할 수 있는 시스템을 제안하고자 한다. 블록체인은 데이터의 무결성과 투명성, 탈중앙화된 거래 및 정보 공유 등의 장점으로 보안성과 신뢰성이 높아 이를 기반으로 정보의 공시가 이루어진다면 국제 시장에서의 ESG 노력에 대한 공정한 평가 및 공급망 실사에 효과적으로 대응할 수 있을 것이다.*

또한 본 연구는 기업의 ESG 평가를 위한 정보 자산과 평가 결과 등을 NFT 로 발행하여 이를 블록체인을 통해 효과적으로 유통할 수 있는 시스템을 제안하고자 한다. NFT 는 대체불가능 토큰으로 NFT가 대표하는 고유한 자산의 소유권이 블록체인을 통해 증명될 수 있다. 이런 관점에서 NFT를 통해 각각의 ESG 정보가 고유하게 식별되고 추적될 수 있어 신뢰성이 높아지며, 투자자와 이해관계자들에게 투명한 정보를 제공할 수 있다. 따라서 ESG 현황 및 평가 등을 포함한 정보 자산을 대표하는 NFT는 위변조할 수 없으며 고유 정보의 소유자가 명확하므로 공시 수단으로 적합한 도구라고 할 수 있다. ESG NFT 가 시장을 통해 거래되는 것이 주된 목적이 아니라도 블록체인 기반의 NFT는 소비자들과 투자자들의 의사결정 시 신뢰할 수 있는 기업 ESG 현황 및 평가 정보 제공 수단이 될 수 있다. 공시 플랫폼을 통해 기업은 ESG 관련 정보 자산의 효과적인 관리가 가능하며 투자자들과 소비자들은 기업의 ESG 노력에 대한 신뢰성 있는 정보를 제공받을 수 있을 것으로 기대한다.

Key Terms *– 블록체인; ESG; NFT; ESG 정보 공시; ESG NFT*

블록체인 기반 배터리 여권 기술 동향 분석

이상명	김재현	조재한	정한호	김호원
부산대학교	*부산대학교*	*부산대학교*	*부산대학교*	*부산대학교*
sangmyeong @islab.re.kr	jaehyun @islab.re.kr	jaehan @islab.re.kr	hanho @islab.re.kr	howonkim @gmail.com

Abstract* – *최근 지속적으로 성장하는 전기차 시장과 함께 배터리 시장에도 관심이 쏠리고 있다. 전기차 시장의 성공은 당연히 친환경과 이어져야 한다. 하지만 전기차 수가 증가함에 따라 환경 유해 물질을 포함한 수많은 폐배터리의 발생은 환경을 위협하는 존재로 여겨진다. EU 는 친환경 및 순환경제 구축을 목표로 새로운 배터리 규정과 함께 배터리 여권 제도를 도입할 것으로 밝혔으며, 전 세계적으로 배터리 여권 관련 기술과 연구들이 등장하고 있다. 본 논문에서는 배터리 여권과 4 차 산업혁명의 대표 기술 중 하나인 블록체인을 적용한 블록체인 기반 배터리 여권의 핵심 기술과 블록체인 기반 배터리 여권의 기술 동향을 소개한다. 나아가, 블록체인 기반 배터리 여권의 한계점과 연구 필요성에 대해 기술한다.

Keywords* - *공급망, 배터리 여권, 블록체인, 순환경제

사사표기

이 논문은 2024 년도
정부(과학기술정보통신부)의 재원으로
정보통신기획평가원의 지원을 받아 수행된
연구임(No.2018-0-00264, IoT 융합형 블록체인
플랫폼 보안 원천 기술 연구, 100%)

I. 서론

전기차 산업의 성장에 따라 배터리의 수요가 증가하여 원자재 및 공급망 확보는 배터리 산업의 경쟁력과 직결된다. SNE 리서치는 글로벌 배터리 재활용 시장이 연 평균 17%씩 증가해, 2040 년 2089 억 달러에 이를 것으로 전망했으며, 폐배터리의 발생량은 2025 년 338GWh, 2035 년 1329GWh, 2040 년 3339GWh 로 연평균 33% 증가할 것으로 전망했다. 전기차의 폐배터리는 환경 유해 물질을 포함하고 있어 앞으로 발생할 폐배터리의 재사용성을 높이지 못하면 친환경이 목적인 전기차가 오히려 환경을 해치는 격이 되기에 폐배터리의 재활용은 환경 오염 예방을 위한 큰 과제로 여겨진다. 이에 따라 EU 는 배터리의 지속가능성을 향상시키고, 배터리 원재료의 재활용성을 높이는 것을 목표로 하는 배터리 규정을 발표했다.

본 논문은 배터리 규정의 주요 내용 중 하나인 배터리 여권에 블록체인 기술을 적용한 블록체인 기반 배터리 여권의 기술 동향에 대해 설명한다. 나아가 해당 기술의 한계점과 추가적인 연구 및 개발의 필요성에 대해 제시한다.

본 논문의 구성은 다음과 같다. 2 장에서 배터리 여권 기술과 핵심 기술인 블록체인 기술 소개와 함께, 3 장에서 블록체인 기반 배터리 여권 기술 동향에 대해 소개한다. 마지막으로, 4 장에서 블록체인 기반 배터리 여권의 한계점 및 연구 필요성에 대해 기술하고 본 논문을 마무리한다.

II. 배터리 여권 및 핵심 기술 소개

1. 배터리 여권 기술

배터리 제조와 유통 과정에는 다양한 공급자들이 존재하고 각 과정은 복잡하며, 공개되지 않아 제 3 자들은 해당 정보에 대해 알기 어렵다.

이를 해결하기 위한 기술인 배터리 여권은 경제, 환경, 사회적으로 지속가능한 배터리 가치사슬 구축을 위해 결성된 글로벌 배터리 연맹(Global battery Alliance, GBA)에 의해 제안되었다. 가치사슬은 제품과 서비스를 제공하기 위해 여러 조직 및 기업의 활동이 연결된 과정을 의미한다. 배터리 여권의 목적은 가치사슬 내의 모든 이해관계자가 배터리의 정보와 이력을 공유하여 가치사슬의 투명성을 제공하도록 하는 것이다. 배터리 제조부터 유통, 사용, 재사용까지의 모든 이력을 투명하게 관리함으로써 그 목적을 달성할 수 있다. 나아가, 효율적인 배터리 이력 관리를 위해 블록체인을 적용하는 연구와 개발이 활발히 이뤄질 것으로 기대된다.

2. 블록체인

공급망 추적 및 이력 관리를 통한 신뢰할 수 있는 배터리 여권을 위해서는 가치사슬의 투명성, 데이터의 불변성과 무결성이 보장되어야 한다. 중앙집중형 데이터베이스는 데이터가 하나의 중앙 데이

터베이스에 집중되어 있어, 장애 혹은 해킹과 같은 상황이 발생할 경우 매우 취약한 구조를 가진다. 반면, 분산 원장 기술은 분산된 구조로 인해 중앙 집중형 데이터베이스에 비해 높은 효율성, 투명성, 안정성을 제공한다.

블록체인은 분산 원장 기술 중 하나로, 거래내역을 포함한 블록들을 해시화하여 연결한다. 블록체인 네트워크는 탈중앙화된 시스템에서 사용자간 쉽고 자유로운 거래를 가능하게 한다. 블록체인은 암호화 기술을 기반으로, 데이터의 무결성을 보장하고 블록체인에 기록된 데이터가 참가자들에게 공개되기 때문에 투명성을 보장할 수 있다. 뿐만 아니라, 배터리 이력 정보에 따라 자동으로 실행되는 조건을 정의하여 프로그래밍 코드로 작성된 스마트 계약(Smart Contract)을 활용하여 보다 정확하고 효율적인 이력 관리가 가능하다. 블록체인을 활용하면 배터리의 이력을 투명하게 관리하고, 상태를 추적하여 유지보수를 수행할 수 있다.

III. 블록체인 기반 배터리 여권 기술 동향

최근 EU 의 배터리 규정에 따라, 배터리 여권을 위한 연구가 활발히 이루어지고 있다. 본 장에서는 블록체인 기반 배터리 여권 기술 동향에 대해 설명한다.

Ford 와 Volvo 가 콩고 민주 공화국의 전기차 배터리의 핵심 원자재 중 하나인 코발트 공급을 추적하기 위해 블록체인 프로젝트를 진행했다 (Cyrine Soufi et al, 2023). 블록체인을 이용한 코발트 공급망은 아동 노동과 같은 비윤리적인 문제를 투명하게 파악할 수 있게 함으로써, 코발트 공급망의 신뢰도를 높일 수 있다.

Audi, BMW 와 같은 많은 배터리 관련 업체들이 컨소시엄을 구성하여 'Battery Pass' 프로젝트를 진행 중이다. Battery Pass 는 블록체인과 같은 분산 원장 기술을 활용하여 전기차 배터리에 관한 정보를 저장하고 관리하는 배터리 여권 프로젝트다. 배터리 이력 정보를 블록체인에 저장하여 위변조를 방지할 뿐 아니라, 배터리의 출처, 생산 과정, 수명 주기 등을 신뢰할 수 있게 제공하도록 프로젝트를 진행하고 있다. 하지만, Battery Pass 의 Technical Guidance 는 현재와 미래의 글로벌 가치 사슬을 위해 공통된 표준이 필요하며, 전 세계적으로 조직화된 관리가 부족함을 지적했다.

Sara Saberi 가 제안한 Blockchain technology and its relationships to sustainable supply chain management 에서는 블록체인 기술을 통한 지속 가능한 공급망의 이점을 설명함과 동시에, 조직 내 장벽, 조직 간 장벽, 시스템 관련 장벽, 외부 장벽 등을 고려해야 함을 설명했다.

현재, 동적인 전기차 배터리 공급망을 따라 투명성과 보안을 향상시키는 새로운 형태의 데이터 공유가 필요하다. 또한, 배터리 여권이 표준의 부재와 투명성 해결에 해결책으로 지목되었지만, 효율적인 재사용에 방해가 되는 표준과 배터리 설계 등의 대한 연구와 논의가 부족한 상황이다(Jacob Lohmer et al, 2023).

IV. 결론

배터리 여권에 블록체인 기술을 적용하면 크게 두 가지 이점을 얻을 수 있다.

첫번째 이점으로는, 공급망에서 얻는 정보를 신뢰할 수 있다. 이는 블록체인에 기록된 데이터의 위변조가 매우 어렵기 때문이다. 또한, 블록체인의 스마트 계약을 활용하여 거래 과정을 자동화함으로써 중개자 등 사람의 실수로 인한 문제를 줄일 수 있다. 신뢰할 수 있는 데이터는 사용자의 재사용 결정 여부에 큰 영향을 줄 수 있으며, 이는 친환경과 직결되는 중요한 이점이라고 할 수 있다.

두 번째 이점으로는, 신속한 공급망 추적이다. 블록체인 상에 기록된 투명하고 신뢰할 수 있는 데이터는 공급망의 효율성을 높이고, 발생한 문제에 대해 신속한 추적과 대응이 가능하다. 이로써 배터리 가치사슬의 지속 가능한 확장을 가능하게 할 수 있다.

하지만, 배터리 여권에 블록체인 기술을 적용할 경우, 앞선 장점 뿐 아니라 해결해야 할 단점 역시 존재한다.

배터리의 국가 간 이동은 대규모 데이터 처리로 인한 블록체인의 확장성 문제를 야기할 수 있다. 나아가, 상호 운용성과 표준화의 부재는 이 문제를 더욱 악화시킬 수 있다. 앞서 3 장에서 언급했듯, 표준화된 데이터 규격과 프로토콜이 필수적이다. 즉, 전 세계적으로 상호 운용 가능한 시스템이 필요하다.

지속가능한 순환경제를 위해 블록체인 기반 배터리 여권의 이러한 한계점을 개선하기 위해서는 국가, 기술, 경제, 환경 등 다양한 측면에서의 추가적인 연구와 개발을 위한 노력이 필요하다.

V. 참고문헌

The Battery Pass consortium. (2023). Battery Passport Content Guidance. Retrieved April 1, 2024, from https://thebatterypass.eu/assets/images/content-guidance/pdf/2023_Battery_Passport_Content_Guidance.pdf

The Battery Pass consortium. (2024). Battery Passport Technical Guidance. Retrieved April 8, 2024, from https://thebatterypass.eu/assets/images/technical-guidance/pdf/2024_BatteryPassport_Technical_Guidance.pdf

Soufi, C., Mesbahi, T., & Samet, A. (2023, October). Digital Battery Passport as an Enabler of Environmental Impact Assessment in Electric Vehicle Applications. In *2023 IEEE Vehicle Power and Propulsion Conference (VPPC)* (pp. 1-6). IEEE.

Bernards, N., Campbell-Verduyn, M., & Rodima-Taylor, D. (2022). The veil of transparency: Blockchain and sustainability governance in global supply chains. *Environment and Planning C: Politics and Space*, 23996544221142763.

Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. (2019). Blockchain technology and its relationships to sustainable supply chain management. *International journal of production research*, *57*(7), 2117-2135.

da Silva, E. R., Lohmer, J., Rohla, M., & Angelis, J. (2023). Unleashing the circular economy in the electric vehicle battery supply chain: A case study on data sharing and blockchain potential. *Resources, Conservation and Recycling*, *193*, 106969.

Wu, H., Jiang, S., & Cao, J. (2023). High-efficiency blockchain-based supply chain traceability. *IEEE Transactions on Intelligent Transportation Systems*, *24*(4), 3748-3758.

Berger, K., Baumgartner, R. J., Weinzerl, M., Bachler, J., Preston, K., & Schöggl, J. P. (2023). Data requirements and availabilities for a digital battery passport–A value chain actor perspective. *Cleaner Production Letters*, *4*, 100032

Walden, J., Steinbrecher, A., & Marinkovic, M. (2021). Digital product passports as enabler of the circular economy. *Chemie Ingenieur Technik*, *93*(11), 1717-1727

Centobelli, P., Cerchione, R., Del Vecchio, P., Oropallo, E., & Secundo, G. (2022). Blockchain technology for bridging trust, traceability and transparency in circular supply chain. *Information & Management*, *59*(7), 103508.

데이터 양자화를 통한 FPGA 기반 이진 신경망 경량 설계

윤동욱
부산대학교
donk1879@gmail.com

강은세
부산대학교
saintmansiu@gmail.com

김호원
부산대학교
howonkim@gmail.com

Abstract – *DNN(Deep Neural Network) 기술은 경량 환경에서 강력한 성능을 가지는 것으로 입증되며 모바일이나 웨어러블 기기 등의 IoT(Internet of Things) 분야로 확장되었다. 이러한 IoT 디바이스들은 한정된 자원과 전력 소비를 최소화 하기 때문에 ASIC 이나 FPGA 를 활용하는 것이 유리하며, 향후 고신뢰성을 보장하는 BNN 가속기를 구현하기 위해 하드웨어의 자원량 증가에 대비하여 저면적 및 저전력으로 구현해야 한다. 이에 본 연구에서는 FPGA 를 기반으로 한 저면적 및 저전력 BNN 가속기를 구현하였으며, 구현 결과 기존 32bit 가속기와 비교하여 유사한 정확도를 유지 하면서도 자원 사용량에서 BRAM 의 경우 약 62%, DSP 는 약 75% 감소하였다.*

Keywords – *BNN, FPGA, 신경망, 양자화*

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 융합보안핵심인재양성사업의 연구 결과로 수행되었음(IITP-2023-2022-0-01201)
또한, IDEC 에서 EDA Tool 을 지원받아 수행하였습니다.

I. 서론

DNN(Deep Neural Network)는 이미지 및 음성인식, 자연어처리 비디오 분석과 같은 다양한 작업에서 매우 강력한 것으로 입증되면서, 모바일, 헬스케어, 가전과 같은 IoT(Internet of Things) 분야로 확장되고 있다. 하지만, 이러한 IoT 디바이스는 자원이 한정적이고, 전력 소비를 최소화하여 데이터 처리 및 제어와 같은 작업을 수행하는 것이 특징이다. 이에 특정 목적으로 설계된 ASIC(Application Specific Integrated Circuit)이나 FPGA(Field Programmable Gate Array)와 같은 하드웨어를 통해 구현하는 것이 유리하며, 또한 많은 연구가 이루어지고 있다. 이에 본 연구에서는 FPGA 기반으로 BNN(Binarized Neural Networks) 모델에 대한 저면적 및 저전력 프로세서를 구현하며, 가속기 내부 Data Width 를 16bit 로 양자화를 통하여 메모리 사용량을 대폭 줄이고, 추론 시 기존 32bit 가속기과 유사한 정확도를 보장한다.

II. 본론

2.1 배경지식

기존 CNN(Convolutional Neural Network)은 네트워크 가중치(Weight), 특징 맵(Feature Map)을 부동 소수점(Floating Point)로 표현하여, 계산 복잡성과 메모리 사용량을 많이 필요로 하는 반면, BNN 은 부동 소수점이 아닌 이진수(-1 또는 +1)로 표현이 가능하여 모바일 및 임베디드 시스템과 같은 자원이 제한된 환경에서 구현하기 적합하다.

이진화(Binarization) 기법으로 Stochastic 과 Deterministic 두가지 방법이 있다. 이는 학습 시 가중치와 활성화를 모두 +1 또는 -1 로 제한할 수 있으며, 이로 인해 하드웨어 관점에서 저장용량, 면적, 연속속도 등 다양한 이점을 가질 수 있다.

$$x_b = \begin{cases} +1, & with\ p = \max(0, \min\left(1, \frac{x+1}{2}\right)) \\ -1, & with\ \ 1-p \end{cases} \quad (1)$$

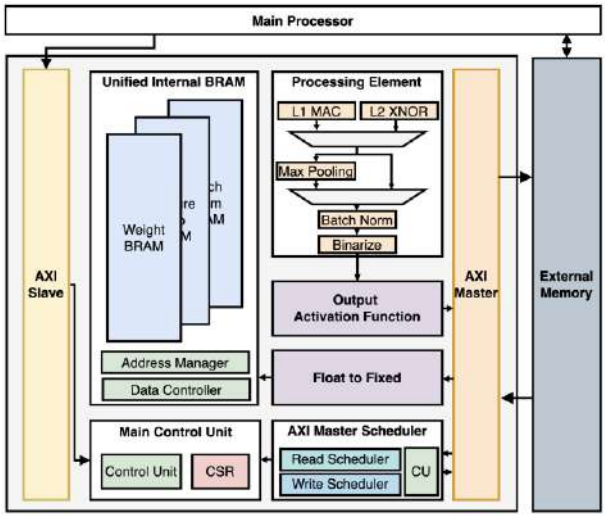
[수식 1]은 Stochastic 방법으로 랜덤으로 생성된 비트스트림에서 “1”의 비율을 확률적으로 이진 값(-1 또는 +1)을 결정한다. 이를 통해 정보를 유연하게 표현할 수 있는 장점은 있지만, 랜덤 비트를 생성하는 난수 발생기를 필요로 하여 이를 통한 하드웨어 자원을 필요로 하는 단점이 있다.

$$x_b = Sign(x) = \begin{cases} +1, & if\ x \geq 0 \\ -1, & otherwise \end{cases} \quad (2)$$

[수식 2]은 Deterministic 한 방법으로 입력 값에 대해 0 보다 크거나 같을 경우 +1 로 이진화 되며, 입력 값이 0 보다 작을 경우 -1 로 이진화가 결정된다. 본 연구에서도 Deterministic 방식을 사용하며, 하드웨어에서 “-1”을 표현하기 위해 2bit 가 필요하여 -1 을 “0”으로 치환하여 1bit 의 크기를 가질 수 있도록 표현하였다. 따라서 본 구현은 +1 과 0 으로 치환하여 이진화를 표현한다.

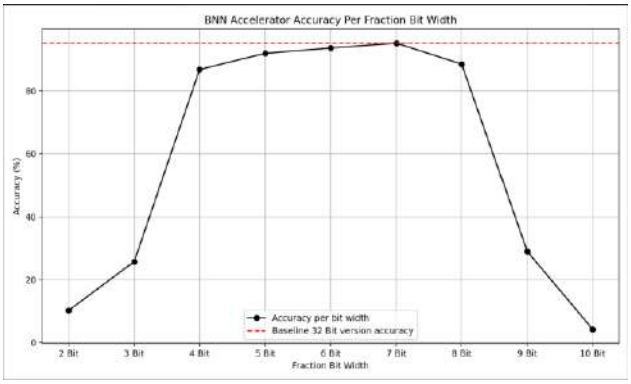
2.2 제안구조 및 실험

본 연구에서는 MNIST 데이터 셋을 사용하여 내부 연산을 고정 소수점 기반으로 Data Width 가 16bit 인 BNN 가속기를 구현하였으며, 제안하는 구조는 [그림 1]과 같다.



<그림 1> BNN 가속기 전체 구조

제안된 가속기는 AXI Interface 를 통해 입출력 데이터가 전달되며, 이를 제어하기 위한 스케줄러와 제어 로직을 포함하고 있다. 가속기 내부는 고정 소수점 기반 연산을 수행하여, 입력 부동 소수점을 고정 소수점으로 변환하는 모듈과 Batch Normalization, Weight, Feature map 값을 저장하기 위한 통합 BRAM(Block RAM)과 각 주소와 데이터를 제어하기 위한 Address Manager 와 Data Controller 가 존재한다. 또한, 각 Layer 연산 MAC(Multiply Accumulation)와 MAX Pooling, Batch Normalization, Binarize 와 같은 연산을 수행하기 위한 Processing Element 로 구성되어 있다.

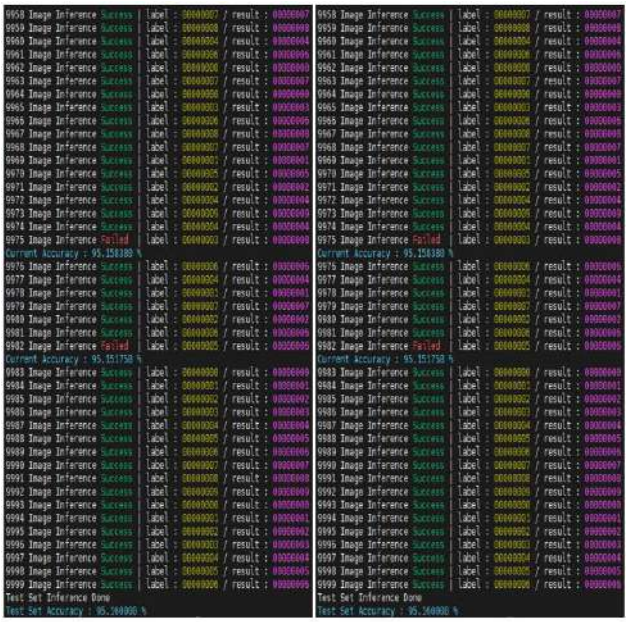


<그림 2> 소수 비트 크기별 추론 정확도

본 연구는 향후 부채널 공격 대응 기술 적용하여 고신뢰성을 보장하는 BNN 가속기 구현하기 위해 [그림 2]와 같은 실험을 진행했다. 부채널 공격 대응 기술로 전력분석 대응 기술인 Masking 또는 Hiding 기법 또는 오류주입 대응 기술인 Integrity Check, Float Representation 등의 기법 중 적용 시에 하드웨어 자원 증가가 발생한다. 이를 대비하기 위해 내부 Data Width 를 기존 32bit 에서 16bit 로 양자화를 수행했다. [그림 2]의 실험은 16bit 고정 소수점에서 소수 비트 크기를 조절하여 그에 따른 정확도를 보여주는 실험이다. 그 결과로, 소수 비트가 7 비트일때 기존 32bit BNN 가속기와 가장 유사한 정확도를 유지하였다.

2.3 구현 및 결과

제안된 BNN 가속기는 Xilinx Zynq Ultrascale+ XCZU7EV MPSoC FPGA 칩이 내장된 ZCU104 보드를 사용하였다. 내부 PS(Processing System) 영역에 Petalinux 를 빌드하여 Model 의 Weight, Image Data 와 같은 파라미터를 외부 메모리에 저장한 후, 가속기 내부 CSR(Control Status Register)에 파라미터와 start 신호를 입력한다. 가속기 내부에서는 CSR 값에 따라 외부 메모리에서 데이터를 읽어와 추론을 시작하고, 최종 추론 결과를 다시 외부 메모리에 저장한다.



<그림 3> (좌) 32BIT (우) 16BIT MNIST 테스트 셋 추론 결과

[그림 3]은 MNIST 데이터 셋에 10,000 장의 테스트 셋을 이용하여 추론한 결과로 16bit 양자화를 거쳤음에도 불구하고, 거의 동일한 테스트 결과 및 정확도를 얻은 것을 확인할 수 있다.

<표 1>. 하드웨어 자원 사용량 비교표

	LUT	F/F	BRAM	DSP
Ver. 32bit	12208 (5.30%)	12876 (2.79%)	14 (4.49%)	8 (0.48%)
Ver. 16bit	11034 (4.79%)	13081 (2.84%)	5 (1.60%)	2 (0.12%)

또한 [표 1] BNN 구현에 따른 합성 결과를 나타내며, 기존 32bit 모델과 16bit 모델의 하드웨어 자원 사용량을 나타낸다. 16bit 양자화를 통해 내부 연산에 사용되는 DSP 사용 개수가 8 개에서 2 개로 75% 감소되었으며, BRAM 메모리 사용량이 기존 14 개에서 5 개로 62% 감소한 결과를 도출했다.

III. 결론

본 연구는 향후 부채널 공격 대응 기술 적용으로 하드웨어 자원량 증가에 대비하여 내부 Data Width 를 16bit 양자화를 적용하여 면적 최적화를 수행했다. 온 보드 검증 결과 MNIST 데이터 셋 10,000 장 기준으로 테스트 이미지에 대한 반복 추론한 결과 95.16%로 기존과 정확도를 거의 동일한 추론 얻을 수 있었다. 향후 대응 기술 적용에 있어 기존 32bit 가속기 대비 연산기 및 필요 연산 비트로 인한 면적 증가량이 감소할 뿐더러, 동일한 추론 정확도를 보장하기 때문에 하드웨어 자원 사용량 즉, 면적을 효율적으로 사용할 수 있을 것으로 기대된다.

IV. 참고문헌

Courbariaux, Matthieu, et al. (2016). “Binarized neural networks: Training deep neural networks with weights and activations constrained to +1 or -1”. arXiv:1602. 02830.

Yichi Zhang, Junhao Pan, Xinheng Liu, Hongzheng Chen, et al. (2020). “FracBNN: Accurate and FPGA-Efficient Binary Neural Networks with Fractional Activations”. arXiv:2012. 12206.

Goubin, Louis. (2001). “A sound method for switching between boolean and arithmetic masking”. Cryptographic Hardware and Embedded Systems-CHES2001: Third International Workshop Paris, France, May 14-16, 2001 Proceedings 3.

Gilbert Goodwill, Benjamin Jun, Josh Jaffe, and Pankaj Rohatgi. (2011). “A testing methodology for side-channel resistance validation”. NIST non-invasive attack t

Hong, S., Frigo, P., Kaya, Y., Giuffrida, C., & Dumitras, T. (2019). Terminal brain damage : Exposing the graceless degradation in deep neural networks under hardware fault attacks. In 28th USENIX Security Symposium(USENIX Security 19) (pp.497-514).

Breier, J., Hou, X., Jap, D., Ma, L., Bhasin, S., & Liu, Y. (2018). Deeplaser: Practical fault attack on deep neural networks. arXiv preprint arXiv:1806. 05859.

블록체인 기반 LLM 생성 컨텐츠 필터링 및 모니터링 시스템

김재현
부산대학교
jaehyun
@islab.re.kr

이상명
부산대학교
sangmyeong
@islab.re.kr

조재한
부산대학교
jaehan
@islab.re.kr

정한호
부산대학교
hanho
@islab.re.kr

김호원
부산대학교
howonkim
@gmail.com

초록 – 인터넷과 생성 콘텐츠의 발전은 정보 접근성을 증가시키고 사회적 상호작용을 증진시키는 데 기여했지만, 악의적인 콘텐츠 생성 문제는 현대 사회에서 점점 더 중요한 문제로 부상하고 있으며, 적극적인 대응이 필요하다. 본 논문에서는 블록체인 기반 LLM 생성 컨텐츠 필터링 및 모니터링 시스템을 설계하여 기술적 해결책을 제시한다. 제안하는 시스템은 사용자 쿼리와 LLM 응답에 대한 검사와 블록체인 기술을 활용하여 이력을 관리하고 데이터의 무결성과 개인정보 암호화를 보장한다. 또한, DID 기반 사용자 인증 및 접근 제어를 통해 서비스 이용을 관리한다.

Keywords – AI Contents Filtering, Blockchain, Decentralized Identity, Large Language Model

블록체인 기반 및 플랫폼 분야 핵심기술 개발 및 미래 혁신인재 양성(Development of Core Technology for Blockchain Platform and Cultivating Next Generation Creative Talent)

I. 서론

인터넷과 생성 콘텐츠의 발전은 정보 접근성을 증가시키고 사회적 상호작용을 증진시키는 데 큰 기여를 했다. 특히, 자연어 생성 모델을 사용한 대화형 인공지능 시스템은 콘텐츠 생성의 새로운 지평을 열었지만 악의적인 콘텐츠를 생성할 수 있다는 점은 현대 사회에서 점점 더 중요한 문제로 부상하고 있으며, 적극적인 대응이 필요하다. 이러한 문제에 대응하기 위해 컨텐츠 필터링 및 모니터링 시스템이 필요하다. 생성된 콘텐츠의 부적절한 내용을 감지하고 차단함으로써 온라인 공간에서의 다양성과 포용성을 증진시키는 데 기여할 수 있으며 추후 학습에 사용되는 데이터 수집을 위한 방법으로 활용될 수 있다.

본 논문의 목표는 블록체인 기반 LLM(Large Language Model) 생성 컨텐츠 필터링 및 모니터링 시스템의 설계를 통해 기술적 해결책을 제시하는 것이다.

II. 관련연구

2.1 블록체인

블록체인은 1) 데이터 검증 및 저장을 위한 체인 구조, 2) 데이터 생성 및 업데이트를 위한 분산 합의 알고리즘, 3) 데이터 전송 및 액세스 보안을 보장하기 위한 암호화 기술, 그리고 4) 데이터 프로그래밍 및 운영을 위한 자동화된 스마트 계약을 통해 분산화, 무결성, 불변성, 내결함성, 투명성 기능을 갖춘 기술이다(Guo et al.,2022; Leng et al.,2022).

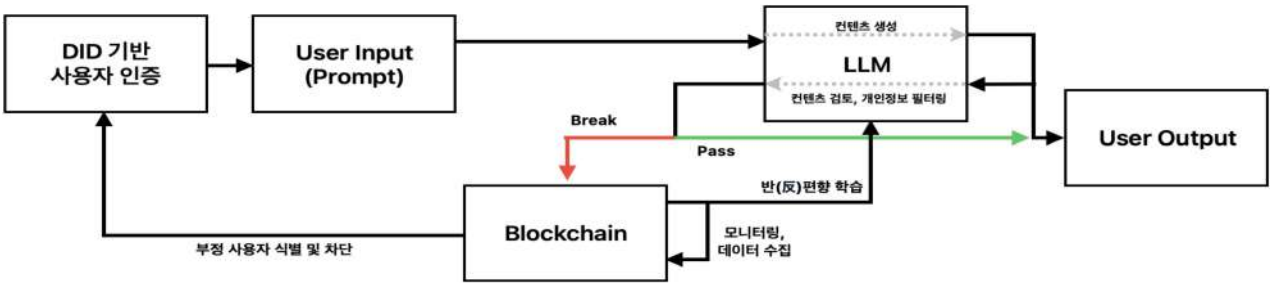
블록체인을 사용하면 중앙집중식 데이터베이스에 의존하는 솔루션은 서비스 거부(Denial of Service) 공격과 같은 보안 문제와 단일 실패 지점에 취약하다는 문제점도 해결할 수 있다(Wang et al., 2018).

이러한 특징을 바탕으로 블록체인은 빅 데이터 관리(수집 및 저장), 빅 데이터 개인 정보 보호 및 보안을 위한 연구가 활발하다(Deepa et al., 2022). 대표적으로 공급망 분야에서 데이터 및 정보 사용에 대한 투명성, 추적성 및 신뢰를 촉진하기 위해 블록체인 기술을 사용하고 있다(Centobelli et al., 2022; Chauhan et al., 2022).

Ethereum 의 창시자 Vitalik Buterin 은 블록체인과 AI 에 대한 연구가 활발해지며 두 영역이 교차하는 사용 사례가 늘어나고 있고 인공지능과 블록체인 기술의 융합에서 오는 시너지 효과가 있음을 시사했다(Vitalik, 2024).

2.2 LLM

LLM 은 엄청난 양의 데이터를 학습하기 때문에 폭발물이나 약물 제조 방법과 같이 쉽게 접근할 수 없는 위험한 정보를 필연적으로 포함하게 된다(Gupta et al., 2023). 따라서 사용자가 이 정보에 액세스하지 못하도록 각 회사는 가드레일을 설치한다(Sami, 2023). 하지만 LLM 개발자가 설치한 안전 가드레일을 회피하는 데 사용할 수 있는 방법인 "탈옥" 기술을 통해 위험한 정보에 의도적으로 접근할 수 있고, 모델의 윤리적 제약을 우회하여 악의적인 정보를 유출하기 위해 악용할 수 있다(Anil et al., 2024).



<그림 1> 전체 시스템 구조

III. 시스템 구조

3.1 전체 시스템 구조

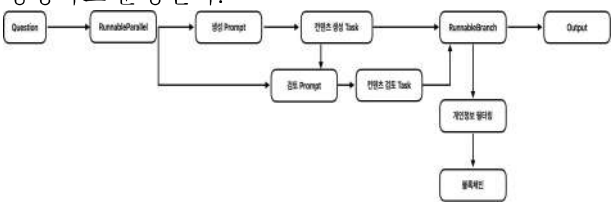
본 논문에서 제안하는 시스템은 사용자의 쿼리와 LLM 의 응답에 대한 엄격한 검사와 투명한 기록 과정을 포함한다. <그림 1>에는 본 논문에서 제안하는 시스템의 흐름이 나타난다. 사용자는 가입단계에서 DID 인증을 통해 신원을 확인한 후 LLM 과 상호작용할 수 있다.

인증이 완료된 사용자의 쿼리는 LLM 을 통해 대한 답변을 생성하고, 생성된 답변에 대한 부정적인 내용을 검토한다. 검토 결과가 부정적인 경우, 해당 쿼리와 답변에 대해 DID(Decentralized Identity) 정보를 바탕으로 개인정보가 포함된 문장을 블라인드 처리되어 블록체인에 기록된다. 검토가 완료되면 사용자는 생성된 컨텐츠를 확인할 수 있다.

3.2 시스템 동작 방식

3.2.1 LLM

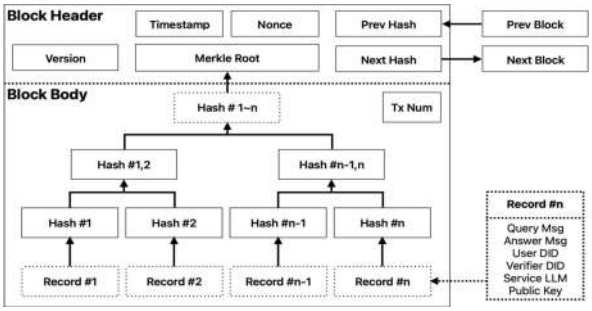
본 논문에서 LLM 은 사용자의 쿼리와 관련된 정보를 분석하고, 그에 따른 적절한 응답을 생성하는데 사용된다. LLM 모델에 대한 공격을 통해 부적절한 내용이 포함하는 있는 경우가 있기 때문에 대답하기 전에 자체적으로 컨텐츠를 검토한다. 대부분의 공격은 만약 부적절한 내용이 포함되어 있다면 블록체인에 개인정보가 기록되지 않도록 하기위해 답변에 포함된 개인정보에 대한 필터링을 추가적으로 진행한다. 이 과정에서 LLMOps 를 사용하여 LLM 파이프라인을 생성하고 운영한다.



<그림 2> LLM 파이프라인

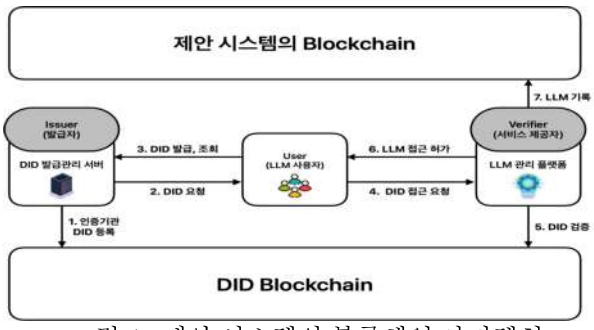
<그림 2>에 해당 시스템에서 사용되는 LLM 파이프라인이 나타나 있다. 파이프라인을 살펴보면 부정적인 컨텐츠를 포함한 답변일 경우, 개인정보 필터링을 위한 추가 비용이 지출되며 답변까지 받아볼 수 없음을 시사한다. 이러한 접근 방식은 사용자가 LLM 이 부정적인 컨텐츠를 생성하지 않도록 장려하는 효과를 가진다.

3.2.2 블록체인



<그림 3> 블록체인에 기록되는 블록 구조

시스템은 데이터의 무결성과 개인정보 암호화를 위해 블록체인 기술을 활용한다. 사용자 부정적인 쿼리와 LLM 의 응답은 개인정보가 블라인드 처리된 상태로 암호화되어 저장되며, 데이터는 <그림 3> 형태로 블록체인에 기록된다. 블록에는 암호화된 쿼리와 답변 메시지, DID 관련 정보, 사용한 LLM 모델과 암호화에 사용된 공개키가 저장된다.



<그림 4> 제안 시스템의 블록체인 아키텍처

제안하는 시스템은 <그림 4>와 같이 DID 를 기반으로 한 사용자 인증 및 접근 제어를 통해 서비스에 대한 접근을 관리할 수 있는 방법을 제공한다. 사용자는 DID 정보를 이용하여 LLM 플랫폼에

접속하게 되고, 전달받은 DID 정보를 DID 블록체인에 검증하여 LLM 에 대한 접근을 허가 받는다. 또한 악의적인 행위에 대한 이력이 블록체인에 저장되어 해당 사용자의 신뢰도를 평가하는 데 활용되어 서비스 제공자는 악의적인 사용자가 서비스를 이용하는 것을 차단할 수 있다.

IV. 결론

제안하는 시스템은 디지털 콘텐츠 생성과 정에서 AI 가 유해하거나 불법적인 콘텐츠 등을 생성하는 문제에 대한 대안을 제공한다. 시스템은 사용자 쿼리와 LLM 응답의 부정적 의도를 검사하여 다양성과 포용성을 강화하여 사용자들에게 안전하고 포용적인 환경을 제공하고, 데이터 암호화와 블록체인 기술을 활용하여 데이터의 무결성과 개인정보 보호를 강화한다. 수집된 데이터는 공격을 방어하기 위해 사용되고, 데이터 공유 및 거래가 가능하며, 악의적인 사용자의 행위 이력은 블록체인에 저장되어 서비스 제공자는 악의적인 사용자의 서비스 이용을 실시간으로 차단할 수 있다.

본 연구는 블록체인 기반 LLM 생성 컨텐츠 필터링 및 모니터링 시스템의 개념적 시스템을 제안하였으나, 시스템의 구체적인 설계 및 프로토타입 구현이 향후 필요하다.

V. 참고문헌

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy. IEEE Access.

Chauhan, C., Parida, V., & Dhir, A. (2022). Linking circular economy and digitalisation technologies: A systematic literature review of past achievements and future promises. Technological Forecasting and Social Change, 177, 121508.

Sami, M., Sami, A., & Barclay, P. (2023, October). A case study of fairness in generated images of Large Language Models for Software Engineering tasks. In 2023 IEEE International Conference on Software Maintenance and Evolution (ICSME) (pp. 391-396). IEEE.

Deepa, N., Pham, Q. V., Nguyen, D. C., Bhattacharya, S., Prabadevi, B., Gadekallu, T. R., ... & Pathirana, P. N. (2022). A survey on blockchain for big data: Approaches, opportunities, and future directions. Future Generation Computer Systems, 131, 209-226.

Leng, J., Zhou, M., Zhao, J. L., Huang, Y., & Bian, Y. (2020). Blockchain security: A survey of techniques and research directions. IEEE Transactions on Services Computing, 15(4), 2490-2510.

Anil, C., Durmus, E., Sharma, M., Benton, J., Kundu, S., Batson, J., ... & Duvenaud, D. Many-shot Jailbreaking.

Centobelli, P., Cerchione, R., Del Vecchio, P., Oropallo, E., & Secundo, G. (2022). Blockchain technology for bridging trust, traceability and transparency in circular supply chain. Information & Management, 59(7), 103508.

Guo, H., & Yu, X. (2022). A survey on blockchain technology and its security. Blockchain: research and applications, 3(2), 100067.

Wang, S., Zhang, Y., & Zhang, Y. (2018). A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems. Ieee Access, 6, 38437-38450.

Vitalik Buterin(2024, January 30) The promise and challenges of crypto + AI applications. https://vitalik.eth.limo/general/2024/01/30/cryptoai.html

전문가 혼합 신경망 적용을 통한 객체 탐지 트랜스포머 성능 개선 방안 연구

박지성
스마트엠투엠
parkjiseung@smartm2m.co.kr

김용수
스마트엠투엠
yongsu@smartm2m.co.kr

강효은
스마트엠투엠
hyoeun@smartm2m.co.kr

김호원
스마트엠투엠
howonkim@gmail.com

***Abstract* – 본 연구에서는 전문가 혼합 신경망(Mixture of Experts, MoE) (Albert Q. Jiang, 2024)의 적용이 객체 탐지 트랜스포머(DEtection TRansformer, DETR) (Nicolas Carion, 2020) 기반 모델의 성능 향상에 유의미한 영향을 미친다는 결과를 제시한다.**

객체 탐지 트랜스포머에 대한 전문가 혼합 신경망 적용을 위해 아래의 방법을 사용하였다.

1. **객체 탐지 트랜스포머의 피드포워드 신경망(Feed-Forward Networks, FFNs)을 4 개의 피드포워드 신경망(즉, 전문가 신경망)과 1 개의 라우터로 구성된 전문가 혼합 신경망으로 대체하였다.**
2. **전문가 혼합 신경망의 라우터에 객체 쿼리(Object Queries) (Nicolas Carion, 2020)에 대한 마스크(Mask)를 적용하여 각 객체 쿼리가 선택된 전문가 신경망의 가중치를 병렬적으로 적용 받을 수 있도록 하였다.**

본 실험에서는 객체 탐지 트랜스포머의 느린 수렴 속도를 개선한 변형 객체 탐지 트랜스포머 (Deformable DEtection TRansformer, Deformable DETR) (Xizhou Zhu, 2021)을 기반 모델로 사용하였고, COCO 데이터셋을 사용하여 학습과 평가를 진행하였다.

COCO val 데이터셋에 대한 성능 평가 결과, 전문가 혼합 신경망의 적용은 +2.5 mAP, +2.8 AP-medium, +4.3 AP-large 의 성능 향상을 가져왔다.

***Keywords* – 전문가 혼합 신경망(Mixture of Experts, MoE), 객체 탐지 트랜스포머(DEtection TRansformer, DETR), 변형 객체 탐지 트랜스포머(Deformable DEtection TRansformer, Deformable DETR)**

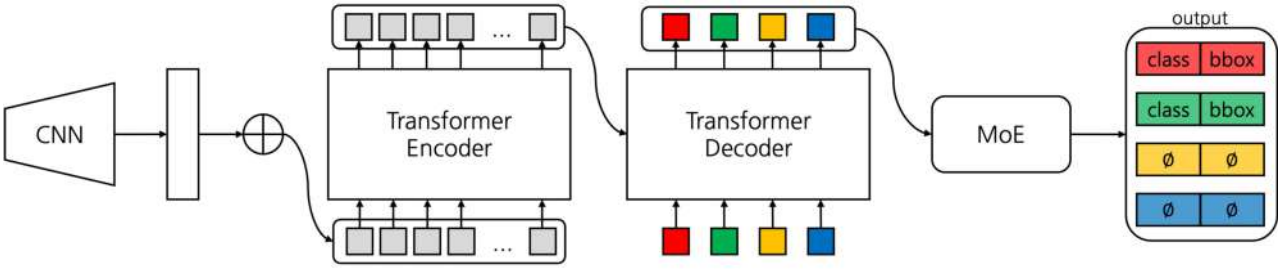
이 논문은 2024 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00431, (3 세부) 지능형 디지털 트윈 연합 지원 개방형 서비스 플랫폼 및 저작도구 기술 개발)

I. 서론 및 관련 연구

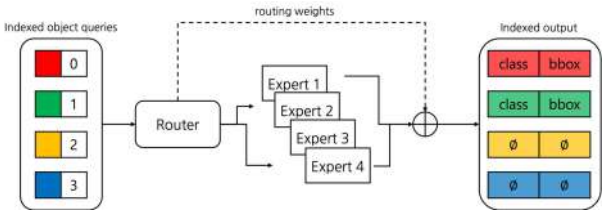
DETR 은 복잡한 사전 처리 단계 없이 이미지 내 객체를 직접 탐지하는 End-to-End 모델로, 객체 탐지 분야에서 혁신적인 발전을 이루었다. 하지만 느린 수렴 속도, Negative Examples 에 대한 편향성 등 모델의 성능과 효율성을 제한하는 여러 문제점들이 존재하고 있다. 이와 관련하여 최근 연구에서는 Deformable DETR(Xizhou Zhu, 2021), Group DETR(Qiang Chen, 2023), Co-DETR(Zhuofan Zon, 2023)과 같은 다양한 변형 모델이 등장하였다. 해당 모델들은 Sparse Spatial Sampling 기법, Group Object Queries 에 대한 One-to-Many 매칭 기법 등을 적용하여 DETR 의 수렴 속도를 개선하였고, Aux Head 를 통해 Negative Examples 에 대한 편향성을 완화하여 DETR 의 성능을 개선하였다. 이 중 Co-DETR 은 COCO test-dev, COCO val 등 대부분의 객체 탐지 리더보드에서 State-Of-The-Art 모델로 자리잡고 있다.

한편, 최근 Large Language Models(LLM)에 대한 연구가 활발해지면서, Transformer 기반 모델의 성능을 향상시키는 다양한 기법들이 등장하고 있다. 그 중 Mixture of Experts(MoE)는 Transformer 의 Feed-Forward Networks(FFNs)를 여러 개의 FFNs(즉, 전문가)와 1 개의 라우터로 구성된 MoE Networks 로 대체하여, 모델이 다양한 전문가들의 조합을 통해 특정 작업을 더 효과적으로 처리할 수 있도록 하는 기법이다. 해당 기법은 Transformer 기반 모델의 유연성과 확장성을 향상시키며, 복잡한 문제를 해결하는 데 있어서 잠재력을 입증하였다.

본 연구에서는 MoE 기법을 Transformer 기반 객체 탐지 모델에 적용하여, 모델이 다양한 상황과 객체 유형에 대해 더 정밀한 예측을 할 수 있도록 성능을



<그림 1> MoE DETR 구조



<그림 2> MoE 라우팅

향상시키고자 한다. 더 나아가 Transformer 기반 객체 탐지 모델의 새로운 발전 방향을 제시하여, 객체 탐지 분야의 미래 연구에 기여하고자 한다.

II. 방법 및 결과

1) 전문가 혼합 신경망

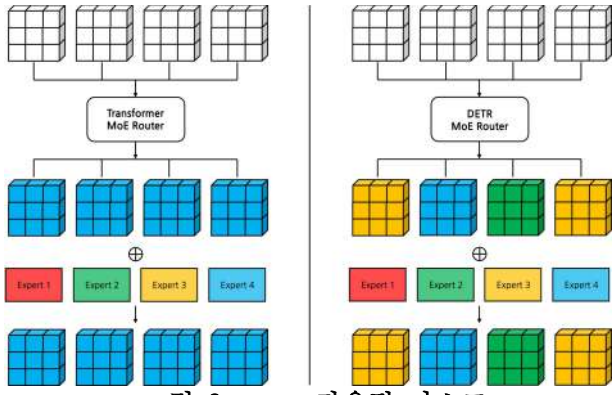
Transformer 의 Decoder 만 사용하는 LLM 의 구조와 달리, DETR 은 Transformer 의 Encoder 와 Decoder 를 모두 사용하는 구조이다. 따라서 LLM 에 적용되었던 MoE 방법과 유사하게 DETR 의 Decoder 의 FFNs 에만 MoE 를 적용하여, DETR 의 MoE 가 LLM 의 MoE 와 동일한 기능을 수행할 수 있도록 하였다.

본 연구에서는 MoE 의 Experts 를 각각 작은 객체, 중간 객체, 큰 객체, 불완전 객체(일부분만 존재하거나 해상도 품질이 낮은 객체)에 대한 전문가로 가정하여 총 4 개의 Experts 를 가지도록 설계하였다. MoE 가 적용된 DETR 의 구조는 <그림 1>과 같다.

2) 라우팅 마스크

DETR 은 일반적인 Transformer 와 다르게 입력 토큰(즉, 객체 쿼리)에 대한 Attention 연산을 독립적으로 수행하여 N 개의 객체 쿼리에 대해 중복되지 않는 N 개의 객체를 예측한다. 따라서 모든 입력 토큰을 한 번에 처리하는 LLM 의 MoE 와 다르게, <그림 2>와 같이 각 입력 토큰을 독립적으로 처리할 수 있도록 라우팅을 수행해야 한다.

이를 위해, 본 연구에서는 각 객체 쿼리가 선택된 전문가의 가중치를 독립적으로 적용 받을 수 있도록 <그림 3>와 같이 라우팅 마스크에 객체 쿼리 인덱스에 대한 차원을 추가하여 확장하였다.



<그림 3> MoE 라우팅 마스크

결과적으로, DETR 에 대한 MoE 와 수정된 라우팅 마스크의 적용은 <표 1>, <그림 4>와 같이 +2.5 mAP, +2.8 AP-medium, +4.3 AP-large 의 유의미한 성능 향상을 이끌었다.

III. 결론

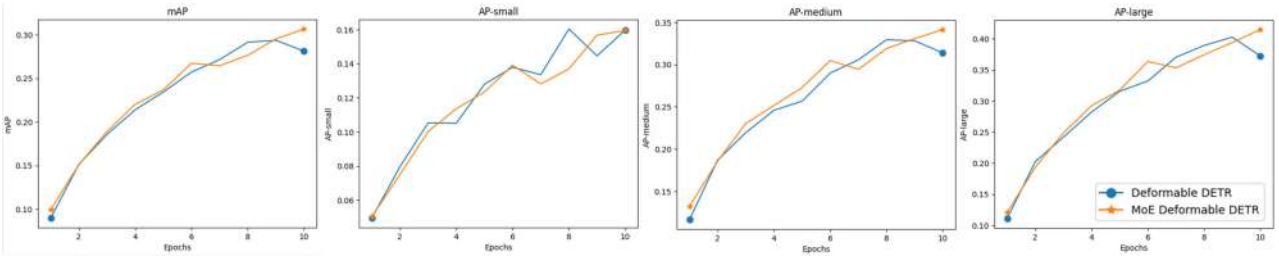
DETR 기반 모델의 느린 수렴 속도와 추론 속도, 작은 객체 탐지에 대한 낮은 정확성 등은 지속적인 연구를 통한 개선이 필요한 문제이다. 본 연구는 LLM 에 적용된 Transformer 기반 성능 향상 기법 중 MoE 기법을 DETR 기반 모델에 적용하여 유의미한 성능 향상이 있음을 보였고, 이를 통해 DETR 기반 모델의 성능 개선 방안에 대한 새로운 관점을 제시한다.

하지만 본 연구에서는 10 epochs 까지의 학습 결과를 토대로 분석하였기 때문에, DETR 모델이 최적화된 시점에서의 결과를 제시하지 못한다. 또한 <표 1>을 보면, 작은 객체에 대해 가정과 달리 -0.4 AP-small 의 더 낮은 성능을 보여 MoE 의 적용이 DETR 의 모든 성능을 향상시키지 못한다는 점을 시사한다.

그러나 후속 연구를 통해 본 연구에서 적용한 MoE 뿐만 아니라, LLM 분야에서 활용되고 있는 다양한 기법들을 DETR 기반 모델에 적합하게 조정하고 통합한다면, 전체적인 모델 성능 개선에 대한 비약적인 도약이 있을 것으로 기대한다.

<표 1> Deformable DETR 에 대한 MoE 적용 여부에 따른 성능 평가 결과 (COCO val 데이터셋)

모델	mAP	AP-small	AP-medium	AP-Large
Deformable DETR	28.1	**16.0**	31.3	37.2
MoE Deformable DETR	**30.6**	15.6	**34.1**	**41.5**



<그림 4> Deformable DETR 에 대한 MoE 적용 여부에 따른 성능 평가 결과 (COCO val 데이터셋)

Ⅳ. 참고문헌

Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, Sergey Zagoruyko(2020). End-to-End Object Detection with Transformers(ECCV 2020). Facebook AI.

Xizhou Zhu, Weijie Su, Lewei Lu, Bin Li, Xiaogang Wang, Jifeng Dai. Deformable DETR: Deformable Transformers for End-to-End Object Detection(ICLR 2021), SenseTime Research.

Qiang Chen, Xiaokang Chen, Jian Wang, Shan Zhang, Kun Yao, Haocheng Feng, Junyu Han, Errui Ding, Gang Zeng, Jingdong Wang. Group DETR: Fast DETR Training with Group-Wise One-to-Many Assignment(ICCV 2023). Baidu VIS.

Zhuofan Zong, Guanglu Song, Yu Liu. DETRs with Collaborative Hybrid Assignments Training(ICCV 2023). SenseTime Research.

Barret Zoph, Irwan Bello, Sameer Kumar, Nan Du, Yanping Huang, Jeff Dean, Noam Shazeer, William Fedus (2022). ST-MoE: Designing Stable And Transferable Sparse Expert Models. Google Brain.

Albert Q. Jiang, Alexandre Sablayrolles, Antoine Roux, Arthur Mensch, Blanche Savary, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Emma Bou Hanna, Florian Bressand, Gianna Lengyel, Guillaume Bour, Guillaume Lample, Lélio Renard Lavaud, Lucile Saulnier, Marie-Anne Lachaux, Pierre Stock, Sandeep Subramanian, Sophia Yang, Szymon Antoniak, Teven Le Scao, Théophile Gervet, Thibaut Lavril, Thomas Wang, Timothée Lacroix, William El Sayed (2024). Mixtral of Experts. MistralAI.

Session B

B4. [학술세션]
비지도 학습을 통한 이상 탐지 기술의 혁신

좌　장: 이 철 호(KAIST)

멀티 모달 데이터를 위한 Flow-based Generative Models 기반 비지도 학습 다변량 시계열 이상 탐지 알고리즘

하은겸
연세대학교 산업공학과
eungyeom_ha@yonsei.ac.kr

김우주
연세대학교 산업공학과
wkim@yonsei.ac.kr

Abstract *– 다변량 시계열 이상 탐지를 위한 완전히 정상적인 데이터셋 준비의 내재된 어려움을 고려할 때, MTGFlow 의 비지도 학습 방법 제안은 본 분야에서 중대한 발전을 나타낸다. MTGFlow 는 동적 그래프와 개체 인식 정규화 흐름을 활용하여 다변량 시계열 데이터의 복잡한 상호 의존성과 다양한 특성을 탐색한다. 그래프 구조를 통해 개체 간의 동적 관계를 학습하고, 개체 인식 정규화 흐름으로 개별 개체의 고유 분포를 포착함으로써, MTGFlow 는 이상 탐지를 위한 세밀한 밀도 추정을 달성하는 정교한 방법을 제공한다.*

이 기반 위에, 우리는 Log Embedding 을 도입하여 MTGFlow 모델의 능력을 멀티모달 데이터로 확장시키는 개선을 제안한다. 이 수정은 모델이 데이터의 로그 속성을 더 잘 이해하고 표현할 수 있게 하여, 데이터가 다양한 규모와 분포를 가로지르는 멀티모달 컨텍스트에서 특히 중요하다. Log Embedding 을 MTGFlow 에 통합함으로써, 모델은 데이터의 미묘한 부분을 더 효과적으로 포착할 수 있는 능력을 풍부하게 하여, 다양한 시나리오에서 더 정확한 이상 탐지를 가능하게 한다.

Log Embedding 향상을 통해 강화된 MTGFlow 모델은 다중 도메인에서의 이상 탐지 성능에서 우월한 성능을 보여준다. 데이터의 로그 속성과 동적 그래프 기반 관계, 개체별 분포를 활용함으로써, 모델은 다변량 시계열 내의 정상 및 비정상 패턴을 더욱 미묘하게 이해한다. 여러 공개 데이터셋에 대해 수행된 기초 모델들과의 비교 실험은 수정된 MTGFlow 모델의 향상된 성능을 보여준다. 결과는 이상 탐지 정확도에서의 상당한 개선을 나타내며, Log Embedding 추가의 효과를 입증한다.
이 작업은 복잡한 이상 탐지 작업을 처리하는 데 있어 비지도 학습 접근법의 잠재력을 강조할 뿐만 아니라, 멀티모달 데이터에 적응할 수 있는 MTGFlow 의 다양한 가능성을 강조한다. Log Embedding 을 MTGFlow 에 통합하는 것은 다양한 도메인에서 이 모델의 추가적인 탐색 및 적용을 위한 길을 열며, 다변량 시계열 데이터의 이상 탐지 도전 과제에 대한 강력한 해결책을 제공한다.

Key Terms – Multivariate Time Series Anomaly Detection, MTGFlow, Unsupervised Learning, Dynamic Graphs Object-aware Normalizing Flows

"이 논문은 국토교통부의 스마트시티 혁신인재 육성사업으로 지원되었습니다."

III. 참고문헌

[1] Zhou, Qihang, et al. "Detecting multivariate time series anomalies with zero known label." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. No. 4. 2023.

[2] Pang, Guansong, et al. "Deep learning for anomaly detection: A review." ACM computing surveys (CSUR) 54.2 (2021): 1-38.

[3] Rudolph, Marco, Bastian Wandt, and Bodo Rosenhahn. "Same same but differnet: Semi-supervised defect detection with normalizing flows." Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2021.

[4] Veličković, Petar, et al. "Graph attention networks." arXiv preprint arXiv:1710.10903 (2017).

트랜스포머 기반 오토인코더를 활용한 전류 센서 데이터의 비지도 학습 이상탐지

박종호
국민대학교 일반대학원
데이터사이언스학과
jjong2961@kookmin.ac.kr

이제혁
국민대학교 경영대학
AI 빅데이터융합경영학과
jehyuk.lee@kookmin.ac.kr

***Abstract** – 전기를 생산하고 사용하는 다양한 산업 분야에서 회전기계는 필수적인 역할을 수행한다. 특히, 풍력, 수력, 원자력 발전 및 터빈과 같은 중요한 전력 설비에서 사용되는 이러한 기계들은 전력 공급의 안정성과 효율성에 직접적으로 영향을 미친다. 따라서 회전기계의 안정적인 운영과 신뢰성 있는 관리는 지속가능한 전력 공급 체계의 핵심 요소이다.*

본 논문은 Transformer based Autoencoder 모델을 사용하여, 지속적으로 수집되는 전류 센서 데이터에서 레이블이 없는 상태로 훈련하는 Unsupervised Learning을 통해 이상치(Anomaly)를 식별하는 방법을 제안한다. 이를 실제 수집된 데이터에 적용한 결과, LSTM 기반의 autoencoder 기반 방법보다 성능이 우수함을 확인하였다. 이를 통해 본 논문은 센서 데이터의 시간적인 패턴과 복잡한 관계를 더욱 효과적으로 분석하고, 이상치를 보다 효율적으로 탐지할 수 있는 방법을 제안한다.

***Keywords** – Anomaly Detection, Autoencoder, Sensor Data, Transformer, Unsupervised Learning*

I. 서론

최근 IoT 기술의 발전으로 기계 시설물에서 발생하는 다양한 데이터를 실시간으로 수집하고 분석하는 것이 가능 해졌다. 이러한 변화는 기계 시설물의 고장을 사전에 예지하고, 필요한 조치를 취함으로써 유지보수 비용을 절감하고 시설물의 안정성을 높이는 새로운 기회를 제공하고 있다(Cook et al., 2019).

본 논문에서는 전류 센서 데이터를 기반으로, 기계 시설물의 결함을 정밀하게 탐지할 수 있는 비지도 학습 방법론을 제안하고 탐구한다. 현대 산업 시스템에서 기계 시설물의 고장이나 결함은 예기치 않은 중단, 생산성 손실, 심지어 안전사고로 이어질 수 있는 만큼 이러한 결함을 조기에 탐지하고 예방하는 것은 매우 중요하다(안동주 외, 2022). 전통적인 고장 진단 방법은 대부분 지도 학습에 의존하며, 높은 수준의 전문 지식과 상당한 양의 레이블링된 데이터를 필요로 한다(Sikorska et al., 2011). 그러나 실제 산업 현장에서는 레이블링된 데이터를 얻기 어렵거나 비용이 많이 드는 경우가 일반적이다(Zhou et al., 2015). 이에 본 논문은 비지도 학습 기반의 Transformer-based Autoencoder 모델을 활용하여 이상 탐지 기법을 제안한다. 자연어 처리(NLP)분야에서 두각을 나타낸 Transformer 모델을 기계 시설물의 고장 진단 분야에 적용함으로써, 기존 방식에서 간과되었던 복잡한 시계열 데이터의 내재된 패턴을 포착한다(Zhang et al., 2020). Autoencoder 는 이상이 있는 데이터를 입력으로 받았을 때, 데이터를 복원하는 과정에서 Reconstruction Error 가 발생한다. 이를 통해 정상적인 작동 상태와 이상 상태 사이의 차이를 구분해낼 수 있는 잠재력을 지닌다(Sakurada & Yairi, 2014).

본 논문에서는 기계 시설물의 이상 탐지를 위한 비교 연구로, LSTM Autoencoder 모델을 도입하여 성능을 분석하였다. LSTM Autoencoder 는 시퀀스 데이터의 장기적인 의존성을 학습할 수 있는 능력으로 인해, 이전 연구에서 시계열 데이터의 이상 탐지에 널리 사용되어 왔다(Malhotra et al., 2016). 이러한 배경 하에, 본 연구는 제안한 Transformer 기반의 Autoencoder 모델과 LSTM Autoencoder 모델 간의 성능을 직접 비교하여 두 모델의 장단점 및 적용 가능성을 평가하고자 한다.

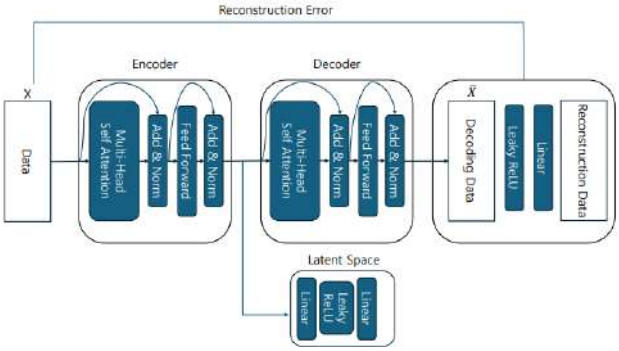
II. 제안 방법론

그림 1 은 본 논문에서 제안하는 Transformer 기반의 Autoencoder 모델의 Architecture 이다. 기존 Transformer 의 모델 구조에서 encoder, decoder 부분을 Autoencoder 의 구조로 변환하였으며, 이는 encoder 단계에서 데이터의 특징을 추출하여 잠재 공간에

매핑한 후 다시 decoder 를 통해 복원하여 하나의 새로운 심층 벡터로 재구성된다. 새롭게 재구성된 심층 벡터는 아래의 수식을 통해 Reconstruction Error 를 계산한다.

$$E = \frac{1}{2}\sum_{x}^{t} \|x - \hat{x}\|^2$$

이후, Reconstruction Error 함수를 통해 평가 데이터셋에 대한 모델의 재구성 오차를 계산하고, 오차가 임계값보다 큰 경우 데이터를 '이상'으로 판단한다.



<그림 1> TRANSFORMER AUTOENCODER 구조

본 논문에서 임계값은 Reconstruction Error 의 평균 절대 오차 MAE(Mean Absolute Error)를 사용하였으며, MAE loss 누적분포 함수의 99%에 해당하는 지점을 임계값으로 선택하였다(Zhou et al., 2017). 이후, 임계값을 기준으로 평가 데이터셋으로 예측 레이블링을 형성하고 실제 레이블과 비교하여 Confusion Matrix 를 통해 해당 모델의 성능을 평가하였다.

III. 실험 및 결과

본 연구에서는 AI Hub 에서 제공되는 "기계시설물 고장 예지 센서 데이터"(AI-Hub., 2021)를 사용하였다. 이는 대전도시철도공사에 설치된 2.2kW 급 실제 모터에서 수집된 데이터이며, 회전체 불균형 샘플을 사용하였다. 학습 데이터로는 전류 센서 데이터를 18000 초 동안 수집된 전류값 (R 상, T 상, S 상)의 RMS(root mean square)를 1 초 단위로 계산하여 사용하였다. 모델을 학습할 때에는 레이블링이 되어있지 않은 상태에서의 데이터만을 사용하였으며, 레이블의 예측이 잘 되었는지 확인하기 위한 평가 데이터로는 레이블링이 표시된 정상상태 1500 개의 샘플과 회전체 불평형 상태 500 개의 샘플을 결합하여 사용하였다.

제안하는 방법론의 효과성을 검증하기 위해 LSTM-Autoencoder 를 활용하였을 때와 성능을 비교하였다. 성능 평가 지표로 Accuracy, Precision, Recall, F1-Score 를 사용하였다. [표 1]은 이 실험의 결과를 나타내고 있으며, 본 논문에서 제안한 방법이 LSTM-Autoencoder 보다 더 좋은 성능을 보였다. 특히 높은 재현율(Recall)을 포함한 나머지 지표에서 우수한 성능 지표를 달성함으로써 레이블링이 되어있지 않거나 데이터의 비정상 클래스의 불균형으로 인한 문제를 완화하고 실제 적용 환경에서 이상탐지의 성능을 개선하는데 기여할 것으로 기대된다.

<표 1> 모델 성능 평가지표

Model	Precision	Recall	Accuracy	F1-score
Transformer Autoencoder	0.7982	0.9970	0.8725	**0.8866**
LSTM-Autoencoder	0.7990	0.8500	0.7990	**0.8088**

IV. 참고문헌

안동주, 신재광, 이수안 (2022), 전동기 기계시설물 고장 분류를 위한 이미지 인코딩 기반 경량화된 딥러닝 모델, *전자공학회논문지*, 59(5), 57-63.

장현준, 김광주, & 김상훈. (2022). 기계 시설물 고장 여부 및 고장 범주 분류 기법 비교. *한국통신학회 학술대회논문집*, 942-943.

AI-Hub, "Mechanical facility failure prediction sensor", Retrieved March 10, 2024, from https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=238

Cook, A. A., Mısırlı, G., & Fan, Z. (2019). Anomaly detection for IoT time-series data: A survey. IEEE Internet of Things Journal, 7(7), 6481-6494.

Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). LSTM-based encoder-decoder for multi-sensor anomaly detection. arXiv preprint arXiv:1607.00148.

Sakurada, M., & Yairi, T. (2014, December). Anomaly detection using autoencoders with nonlinear dimensionality reduction. In Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis, 4-11.

Sikorska, J. Z., Hodkiewicz, M., & Ma, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. Mechanical systems and signal processing, 25(5), 1803-1836.

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into deep learning. arXiv preprint arXiv:2106.11342.

Zhou, C., & Paffenroth, R. C. (2017, August). Anomaly detection with robust deep autoencoders. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 665-674.

Zhou, K., Liu, T., & Zhou, L. (2015, August). Industry 4.0: Towards future industrial opportunities and challenges. In 2015 12th International conference on fuzzy systems and knowledge discovery (FSKD) (pp. 2147-2152). IEEE.

Efficient AD의 손실함수를 활용한 비지도 이상 탐지 성능 개선

김명섭
한양대학교 인공지능융합학과
kimas1215@hanyang.ac.kr

오홍규
한양대학교 인공지능융합학과
omar0701@hanyang.ac.kr

이승재
한양대학교 인공지능융합학과
rkqnfdltkrhk @hanyang.ac.kr

고병진
한양대학교 스마트융합공학부
byungjinko @hanyang.ac.kr

윤종완
한양대학교 지능형로봇학과
jongwanyoon @hanyang.ac.kr

박태준
한양대학교 로봇공학과
taejoon @hanyang.ac.kr

Abstract *– 지식 증류의 최근 발전은 비지도 이상 탐지라는 도전적인 과제에 상당한 가능성을 보여주었다. 그 중 교사-학생 모델 기반의 역 증류(Reverse Distillation) 모델은 교사 인코더, 학생 디코더와 훈련 가능한 압축 모듈인 OCBE 모듈로 구성되어 있다. 하지만 추론 시 이상 정보가 들어왔을 때 OCBE 모듈에서 이상 정보를 압축시켜 그대로 학생 디코더로 넣어주기 때문에 학생 디코더도 이상 정보를 일부 복원해버리게 되는 문제가 있다. 이 문제를 해결하기 위해 Efficient AD에서 제안한 방법인 하드 피쳐 로스와 페널티 로스의 아이디어를 역 증류 모델에 도입하였다. 산업 이상 감지 데이터 셋 중 하나인 MVTec AD 데이터셋 중 5개 클래스에서 성능을 평가한 결과, 제안된 모델의 Detection AU-ROC의 평균은 98.5%로, 기존 역 증류 모델의 97.9%보다 0.6% 향상됐다. Segmentation AU-ROC의 평균은 기존 역 증류 모델의 97.7%에서 0.3% 증가한 98.0%를, Segmentation AU-PRO는 94.2%에서 1% 증가한 95.2%를 기록했다. 이 결과는 추가된 손실 함수들이 모델 성능을 향상시키는 데 기여했음을 보여준다.*

Key Terms *– 교사-학생 모델, 딥러닝, 비지도 학습, 이상 탐지, 지식 증류*

특징 어댑터를 사용한 비지도 이상 탐지를 위한 Reverse Distillation 모델의 편향 완화

오홍규
한양대학교 인공지능융합학과
omar0701@hanyang.ac.kr

김명섭
한양대학교 인공지능융합학과
kimas1215@hanyang.ac.kr

이승재
한양대학교 인공지능융합학과
rkqnfdltkrhk@hanyang.ac.kr

윤종완
한양대학교 지능형로봇학과
jongwanyoon@hanyang.ac.kr

고병진
한양대학교 스마트융합공학부
byungjinko@hanyang.ac.kr

박태준
한양대학교 로봇공학과
taejoon@hanyang.ac.kr

Abstract *– 이상 탐지는 제조 및 의료 분야에서 중요한 응용 프로그램이다. 최신 이상 탐지 기법인 Reverse Distillation 모델은 ImageNet 데이터셋으로 사전 학습된 모듈을 통해 입력 이미지의 특징을 학습한다. 그러나, ImageNet 과 학습 이미지의 이미지 분포가 다른 경우에는 모델이 ImageNet 으로 편향된 특징으로 학습될 수 있고, 이로 인해 이상 탐지 성능이 저하될 수 있다. 본 논문에서는 이를 해결하기 위해 Reverse Distillation 에 특징 어댑터를 도입한 Reverse Adaptive Distillation 모델을 제시한다. 여기서 특징 어댑터는 편향된 특징 T를 T' 로 조정함으로써 편향을 완화시켜준다. Reverse Adaptive Distillation 은 Reverse Distillation 과 비교하여 이상 탐지 정확도를 나타내는 Sample-AUROC 가 MVTec 데이터셋의 5 개 class 에서 평균적으로 0.6% 증가한 98.4 %, VisA 데이터셋의 5 개 class 에서는 평균적으로 1.0% 증가해 98.0%를 나타내었다. 이와 같은 Reverse Adaptive Distillation 의 높은 성능은 실제 현장에서 이상 탐지 실용 가능성이 높다는 것을 입증한다.*

Key Terms *– 딥러닝, 비지도 학습, 이상 탐지, 지식 증류, 특징 어댑터*

Session C

C1. [학술세션]
언어모델의 최적화와 응용

좌 장: 강 주 영(아주대)

한국어 특화 대규모 언어 모델의 토크나이저 심층 분석

최창우 김용수 강효은 김호원
스마트엠투엠 *스마트엠투엠* *스마트엠투엠* *스마트엠투엠*
changwoo@smartm2m.co.kr yongsu@smartm2m.co.kr hyoeun@smartm2m.co.kr howonkim@gmail.com

Abstract* – *본 논문은 한국어에 특화된 대규모 언어 모델(Large Language Models, LLM) 개발의 중요성을 탐구하고, 한국어 자연어 처리 기술의 발전에 기여하는 주요 방법론 중 하나인 Vocab Expansion 과 Pre-Training 의 접근 방법을 중점적으로 다룬다. 본 논문은 한국어 LLM 의 토크나이저와 Vocab 의 최적화가 모델의 이해력과 생성 능력을 향상시킬 수 있음을 심층적으로 분석한다. 연구 결과, 적절한 Vocab Expansion 과 Pre-Training 을 거친 모델이 다양한 한국어 벤치마크에서 우수한 성능을 나타냄을 확인하였다. 이러한 결과는 한국어 뿐만 아니라 다른 비영어권 언어들에 대한 언어 모델 개발에도 중요한 시사점을 제공한다.

Keywords* – *Artificial Intelligence, Korean Language Processing, Large Language Models, Pre-Training, Vocab Expansion

이 논문은 2024 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00431, (3 세부) 지능형 디지털 트윈 연합 지원 개방형 서비스 플랫폼 및 저작도구 기술 개발)

I. 서론

인공지능 기술의 비약적인 발전은 현대 사회의 다양한 분야에 혁신적인 변화를 가져왔다. 이 중에서도 대규모 언어 모델(Large Language Models, LLM)의 출현은 자연어 처리(Natural Language Processing, NLP) 분야에서 특히 주목할 만한 성과를 이루어내며, 이는 인간과 기계 간의 소통 방식에 근본적인 전환점을 마련하였다. LLM 은 방대한 양의 텍스트 데이터를 학습하여 언어의 복잡성과 미묘함을 이해하고, 이를 바탕으로 텍스트 생성, 번역, 요약 등 다양한 언어 관련 작업을 수행할 수 있는 능력을 갖추고 있다. 그러나 이러한 진보에도 불구하고, 대부분의 LLM 연구와 개발은 영어와 같은 소수의 주요 언어에 집중되어 왔다. 이는 비영어권 언어들이 직면한 한계점을 드러내는 것으로, 특히 풍부한 문법 구조와 고유한 어휘 체계를 갖고 있는 한국어의 경우에 LLM 을 훈련시키고 최적화하는데 어려움이 존재한다.

한국어 기반의 LLM 을 개발하는 과정에서 직면하는 공통적인 문제점은 '자원의 부족' 이다. 영어 데이터세트의 경우 대량의 고품질 데이터가 공개되어 누구나 사용가능한 상황이지만, 한국어의 경우에는 많은 기관에서 수집 및 생성하고 있지만 영어에 비해 현저히 부족하며 Huggingface 에 공개된 데이터세트의 경우에도 영어 데이터를 번역 프로그램을 통해 번역하여 결과를 집계한 것 또는 성능이 입증된 언어 모델 (e.g. ChatGPT)을 사용하여 데이터를 수집한 경우가 다수이다.

'자원의 부족' 을 해결하기 위해서 지속적인 데이터 수집 및 생성이 필요한 것도 사실이나, 현실적인 상황에서 적용 가능한 방안은 성능이 입증된 LLM 에 대하여 Vocab Expansion 을 수행하고 Pre-Training 을 해주는 것이며, 이는 현 시점 한국어에 특화된 언어모델을 개발하는 가장 빠른 방법이라고 할 수 있다. 점차 한국어 기반 언어모델 개발의 한계성을 직면하고 Vocab Expansion 의 중요성이 부각되면서, 많은 국내 연구진들이 영어권에서 개발된 LLM 에 대하여 Vocab Expansion 수행 후 Pre-Training 을 통해 한국어 기반의 Foundation Model 을 Hugging Face 에 배포하고 있는 상황이다.

본 논문에서는 한국어 특화 대규모 언어 모델의 토크나이저에 대한 심층적인 분석을 통해, 언어 모델의 이해력과 생성 능력을 향상시키는 데 기여하는 요소들을 탐구하고자 한다. 이를 통해 비영어권 언어의 자연어 처리 기술 발전에 기여하고, 다양한 언어 환경에서의 인공지능 활용 가능성을 넓히는 데 목적이 있다.

II. Vocab 과 토크나이저

Vocab (Vocabulary) 이란 단어 또는 토큰의 집합을 의미한다. 이 집합은 언어 모델이 이해하고 처리할 수 있는 모든 단어 또는 토큰을 포함하는데, 각 토큰은 고유한 인덱스에 매핑되어 있어서 모델이 텍스트를 숫자 데이터로 변환할 때 이 인덱스를 사용한다. 그리고 토크나이저는 텍스트를 작은 단위인 토큰으로 분리하는 역할을 한다. 이 토큰은 단어일 수도 있고, 단어의 일부 또는 여러 단어의 조합일 수도 있다. 예를 들어, 문장 "나는 학교에

간다."를 ["나는", "학교", "에"," 간다"]와 같은 토큰의 시퀀스로 분리할 수 있다.

토크나이저는 텍스트를 모델이 처리할 수 있는 데이터 형식으로 변환하는 과정에서 중추적인 역할을 한다. 특히, 한국어의 경우 어절과 문장 구조가 영어와는 다르게 작동하기 때문에, 한국어 특화 토크나이저의 개발은 모델이 언어의 미묘한 뉘앙스를 정확히 파악하는 데 필수적인 요소이다. 올바르게 설계된 토크나이저는 텍스트의 의미를 모델에 정확히 전달하도록 돕고, 이는 모델의 성능에 직접적으로 영향을 미치게 된다. 표 1 에서 모델별로 사용된 토크나이저가 상이하기 때문에, 다른 토크나이징 결과가 나타나는 것을 확인할 수 있는데, 이는 모델의 추론속도와 정확도 측면에서 중요한 역할을 한다. 따라서, Vocab 과 토크나이저는 자연어 처리에서 텍스트를 숫자로 변환하는 과정에서 서로 밀접하게 연관되어 작동하며 적절한 Vocab 과 토크나이저 선택은 효과적인 언어 모델을 구축하는데 필수적인 역할을 한다고 볼 수 있다.

<표 1> TOKENIZING RESULT

기본 문장 : "화창한 날씨에는 바깥으로 나가 산책을 하거나, 가까운 공원에서 책을 읽는 것도 좋을 것 같아요."	
모델	토크나이징 결과
Llama-2-7b	['화', '창', '한', '_', '날', '<0xEC>', '<0x94>', '<0xA8>', '에', '는', '_', '바', '<0xEA>', '<0xB9>', '<0xA5>', '으', '로', '_', '나', '가', '_', '산', '<0xEC>', '<0xB1>', '<0x85>', '을', '_', '하', '거', '나', ',', '_', '가', '까', '운', '_', '공', '원', '에', '서', '_', '<0xEC>', '<0xB1>', '<0x85>', '을', '_', '읽', '는', '_', '것', '도', '_', '좋', '을', '_', '것', '_', '같', '아', '요', '.']
Llama-2-ko-7b	['_화', '창', '한', '_날씨', '에는', '_바', '깥', '으로', '_나가', '_산', '책을', '_하', '거나', ',', '_가까운', '_공', '원에서', '_책', '을', '_읽', '는', '_것도', '_좋을', '_것', '_같아요', '.']
EEVE-Korean-10.8b	['_화', '창', '한', '_날씨', '에는', '_바', '깥', '으로', '_나가', '_산', '책을', '_하거나', ',', '_가', '까', '운', '_공', '원에서', '_책을', '_읽는', '_것도', '_좋을', '_것', '_같아요', '.']

III. Vocab Expansion 과 Pre-Training

한국어는 고유의 문법적 구조와 어휘 사용 패턴을 가지고 있어, 이에 최적화된 언어 모델이 필요하다. 하지만, 현재 널리 사용되는 대부분의 LLM 은 주로 영어를 중심으로 개발되어왔다. 이로 인해 한국어에서 주로 사용되는 토큰들이 Vocab 에 적절히 반영되지 않아, 언어 모델이 한국어 데이터를 처리할 때 정확도와 효율성이 크게 떨어질 수 있다. 따라서 대량의 고품질 한국어 데이터를 수집하여 적절한 Vocab 을 구성하고 한국어만을 위한 Foundation Model 을 자체 개발하는 것이 필수적이라 볼 수 있지만, Foundation Model 을 개발하는 과정은 막대한 양의 데이터와 계산 자원을 요구하기 때문에 단기간에 한국어에 특화된 좋은 Foundation Model 을 개발하는 것은 쉽지 않다. 처음부터 막대한 양의 고품질 한국어 데이터를 학습하여 Foundation Model 을 구축하는 것이 어렵다면, 선택할 수 있는 부분은 기존에 성능에 입증된 Foundation Model 에 대하여 Vocab Expansion 을 통해 한국어에 적합한 Vocab 을 구성하고 Pre-Training 을 통해 변경된 Vocab 에 대해서 학습을 시켜주는 것이다.

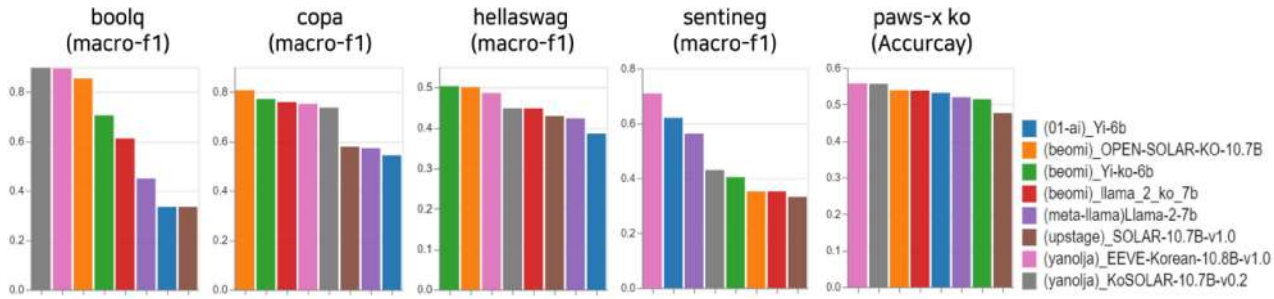
Vocab Expansion 은 기존 언어 모델의 어휘 범위를 확장하는 과정으로, 모델이 더 넓은 범위의 텍스트를 이해하고 생성할 수 있도록 돕는다. 이 과정은 특히, 비영어권 언어에 대한 모델을 개발할 때 중요하며, 한국어의 경우 자주 사용되는 표현, 전문 용어 등을 포괄적으로 수용할 수 있도록 돕는다.

Pre-training 은 대규모 데이터세트를 사용하여 언어 모델을 사전에 학습시키는 과정을 의미한다. 주요 목적은 언어 모델이 언어의 구조와 문맥을 이해하도록 하는 것이며, 이를 통해 모델은 단어, 구, 문장의 관계를 배우고, 언어 사용에서 나타나는 다양한 패턴과 규칙을 학습하게 된다. Pre-Training 을 적절히 수행할 경우 언어 모델은 미세조정(Fine-Tuning)을 거치지 않더라도 여러 언어 작업에서 기본적인 성능을 발휘할 수 있게 된다.

IV. Vocab Expansion 유용성

현재 많은 국내 연구진들이 해당방식으로 좋은 Foundation Model 을 개발하고 있으며, 해당 Foundation Model 을 통해 Downstream Task 에 대해서도 적절한 언어 모델을 개발할 수 있는 상황이다. 본 논문에서는 Vocab Expansion 과 Pre-Training 의 유용성을 분석하고자 한국어 벤치마크를 일부 사용하여 각 언어모델의 성능을 평가하였다. 사용된 벤치마크는 표 2 와 같으며, 평가에 사용된 모델은 표 3 과 같다.

8 개의 대상 모델 중 3 가지(No. 1,2,3)는 Base Model 이며, 5 가지(No. 4,5,6,7,8)는 Base Model 을 대상으로 Vocab Expansion 과 Pretraining 이 수행된 모델이다. 'ko' 가 붙어있는 모델명의 경우 후자에 속하는데, 해당 모델들은 5 가지 벤치마크 중 한국어 텍스트의 문맥 이해 성능을 판단하는데 사용되는 boolq, copa, hellaswag 에서 모두 상위권을



<그림 1> PERFORMANCE VISUALIZATION

차지하였다. 즉, Vocab Expansion 과 Pre-Training 과정이 한국어에 특화된 모델 개발에 유용하다는 것을 입증한다.

<표 2> KOREAN BENCHMARK

Cetegory	Content
boolq	문맥을 파악하여 질문에 대한 답변
copa	문맥을 이해하고 질문에 대한 정답을 선택
hellaswag	문맥을 이해하고 다음 문장을 예측
sentineg	감성 분석
Paws-x ko	특정 단어의 문맥상 의미 동일 여부 판단

<표 3> MODEL CANDIDATES

No	Category	Basemodel	Ref (Huggingface)
1	Llama-2-7b	-	Meta
2	SOLAR-10.7b	-	upstage
3	Yi-6b	-	01-ai
4	llama-2-ko-7b	Llama-2-7b	beomi
5	OPEN-SOLAR-KO-10.7b	SOLAR-10.7b	beomi
6	Yi-ko-6b	Yi-6b	beomi
7	KoSoLAR-10.7b	SOLAR-10.7b	yanolja
8	EEVE-Korean-10.8b	SOLAR-10.7b	yanolja

<표 4> Evaluation Results

No	Boolq (macro-f1)	Copa (macro-f1)	Hellaswag (macro-f1)	Sentineg (macro-f1)	Paws-x ko (acc)
1	0.4502	0.5724	0.4231	*0.5625 (3)*	0.5200
2	0.3343	0.5786	0.4292	0.3316	0.4765
3	0.3343	0.5432	0.3856	*0.6196 (2)*	0.5315
4	0.6121	*0.7598 (3)*	0.4480	0.3517	0.5380
5	*0.8547 (3)*	*0.8076 (1)*	*0.5011 (2)*	0.3517	*0.5390 (3)*
6	0.7060	*0.7716 (2)*	*0.5031 (1)*	0.4044	0.5145
7	*0.8971 (1)*	0.7365	0.4481	0.4294	*0.5565 (2)*
8	*0.8948 (2)*	0.7516	*0.4860 (3)*	*0.7086 (1)*	*0.5575 (1)*

V. 결론

본 논문은 한국어 특화 대규모 언어 모델(Large Language Models, LLM) 개발의 중요성과 그 과정에서 직면하는 주요 도전 과제를 탐구하였다. 특히, 한국어 데이터의 부족 문제와 이를 해결하기 위한 Vocab Expansion 과 Pre-Training 의 중요성에 초점을 맞췄다. 연구 결과, Vocab Expansion 과 Pre-Training 은 한국어에 최적화된 언어 모델을 구축하는데 핵심적인 방법론으로 나타났으며, 이를 통해 얻어진 모델은 다양한 한국어 벤치마크에서 우수한 성능을 보였다. 이는 비영어권 언어에 대한 언어 모델 개발에 있어 Vocab Expansion 과 Pre-Training 이 효과적인 전략임을 시사한다. 또한, 이러한 접근 방식은 한국어 뿐만 아니라 다른 비영어권 언어에 대한 언어 모델 개발에도 적용될 수 있을 것으로 기대된다. 본 연구는 향후 다양한 언어 환경에서의 인공지능 활용 가능성을 넓히는 데 기여할 것으로 전망된다.

VI. 참고문헌

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.

Kim, D., Park, C., Kim, S., Lee, W., Song, W., Kim, Y., ... & Kim, S. (2023). Solar 10.7 b: Scaling large language models with simple yet effective depth up-scaling. arXiv preprint arXiv:2312.15166.

Kim, S., Choi, S., & Jeong, M. (2024). Efficient and Effective Vocabulary Expansion Towards Multilingual Large Language Models. *arXiv preprint arXiv:2402.14714*.

Jang, M., Kim, D., Kwon, D. S., & Davis, E. (2022, October). Kobest: Korean balanced evaluation of significant tasks. In Proceedings of the 29th International Conference on Computational Linguistics (pp. 3697-3708).

Text-to-Gloss를 위한 사전학습 언어모델 단어사전 제어 연구

이명진
연세대학교 산업공학과
myeongjin216@yonsei.ac.kr

김우영
연세대학교 산업공학과
timothy@yonsei.ac.kr

김태용
연세대학교 지능융합협동과정
kasamdi5@yonsei.ac.kr

손의성
연세대학교 산업공학과
esohn@yonsei.ac.kr

김우주
연세대학교 산업공학과
wkim@yonsei.ac.kr

Abstract *– 수어는 농아인들의 의사소통에 중요한 역할을 하는 시각언어로, 손의 모양과 움직임을 나타내는 수지 요소와 얼굴 표지, 입 표지, 몸 표지와 같은 비수지 요소로 표현된다. 그러나 음성언어와 수어 간 번역 연구는 음성언어 간 번역 연구에 비해 부족한 상황이다. 특히, 수어 관련 연구는 수어의 특성을 고려한 새로운 접근이 필요하다. 본 연구에서는 파이프라인 방식의 수어 생성에서 필수적으로 요구되는 수어 표현인 글로스(Gloss)를 도출하기 위해 음성언어를 수어의 형태소인 글로스로 변환하는 Text-to-Gloss 방법을 제안한다. 이를 위해 사전 학습된 언어 모델에서 단어사전을 확장하거나 교체하는 방식의 접근을 제안하였으며 이 방식들의 Text-to-Gloss 성능의 차이를 분석하였다. 우리의 접근은 파이프라인 방식의 수어 생성에서 정확도를 향상시켜 농아인들의 의사소통을 개선하는데 도움이 될 것으로 기대된다.*

Key Terms *– 기계번역, 수어, 글로스, 수어 생성*

LLM 과 RAG 에 기반한 대화형 매칭 에이전트 프롬프트 엔지니어링 과정과 최적화 이슈

이경전
경희대학교 경영대학 & 빅데이터응용학과
klee@khu.ac.kr

윤이지
경희대학교 빅데이터응용학과
chokogangdo1015@khu.ac.kr

이수영
경희대학교 빅데이터응용학과
qasd132@khu.ac.kr

정백
경희대학교 산학협력단
ylbaek@khu.ac.kr

안은정
서울여자대학교 경영학과
dksdmswjd62@swu.ac.kr

심민준
경희대학교 빅데이터응용학과
minjunkevin@khu.ac.kr

정규윤
경희대학교 빅데이터응용학과
jrbdbs@khu.ac.kr

옥근우
경희대학교 빅데이터응용학과
corn2719@khu.ac.kr

김담
하렉스인포텍
yedam.kim@ubplatform.net

박동주
하렉스인포텍
dongju.park@ubplatform.net

Abstract – *기업이나 정부 등 사회 조직에서 일반적인 상식과 지식을 가진 사람을 채용하게 되면, 그 채용된 사람은 해당 조직의 규칙과 전문 지식을 교육받아 습득하게 된 이후에 업무에 투입되는 데, 이와 마찬가지로, 일반적인 상식과 지식을 어느 정도 갖춘 LLM (Large Language Model)과 같은 초거대 생성AI가 발전한다 하더라도, 그 AI 시스템이 특정 조직과 목적을 위해 일할 때는, 해당 조직의 규칙을 습득하고 이에 기반하여 일을 할 수 있어야 하는데, 이러한 것을 가능하게 하는 과정이 프롬프트 엔지니어링이다. 그 직무가 매칭(Matching)일 경우, 매칭 에이전트는 시시각각 변하는 사용자들의 요구(Request)와 상황(Situation) 정보를 반영할 필요가 있는데, 이는 현재, RAG(Retrieval Augmented Generation)기술에 기반해야 한다. 본 연구는, 다양한 분야에서 매칭을 하는 에이전트를 위한 핵심 알고리듬을 설계하고 개발하기 위한 프롬프트 엔지니어링 과정과 최적화 이슈를 소개한다. 발전하고 있는 LLM과 RAG 기술을 활용하면, 사람과 사람, 사람과 상품, 사람과 비즈니스 등을 모두 이어줄 수 있는, 말 그대로 특정 카테고리의 구분 없이 상대를 매칭해주는 인공지능 시스템을 개발할 수 있을 것이라고 예상된다. 이 시스템은, 실시간으로 조건에 맞는 상대를 매칭하는데, 사용자가 매칭을 요구하는 대화를 입력했을 때, 해당 조건에 맞는 다른 사용자를 추천해주고, 추천된 상대와 연결하고, 남녀매칭이나 중고거래뿐만 아니라 특정 카테고리에 속하기 어려운 요구들도 매칭하는 것을 목표로 한다.*

본 연구에서, LLM은 OpenAI의 GPT-4-Turbo를, RAG를 위한 embedding에는 OpenAI의 "text-embedding-3-large" 모델을 사용하여 다음과 같이 프롬프트 엔지니어링을 진행하고 있다. 모든 입력 대화가 매칭을 원한다는 가정 하에, 사용자가 입력한 대화를 입력값으로하는 RAG를 수행하여 유사한 대화들을 N개 추출한다. 유사한 대화들임을 판단하는 기준은 L2거리로, FAISS 라이브러리를 이용하여 Embedding을 통해 벡터값으로 저장되어 있던 타 사용자들의 대화들과 사용자 대화의 벡터값을 비교함으로써 구할 수 있다. 추출된 N개의 대화와 사용자가 입력한 대화를 함께 프롬프트에 넣고, 사용자의 요구사항에 가장 부합하는 타 사용자의 대화를 매칭하는 답변을 생성한다. 만약, 매칭을 찾지 못할 경우 차선이 될 수 있는 매칭 대화를 연결해주되 차선임을 알리는 답변도 함께 생성하도록 설정하였다. 그럼에도 적절한 매칭이 어려운 경우에는, 매칭되지 못했음을 안내하고 다른 매칭으로 재시도해달라는 답변을 생성한다.

매칭 알고리듬의 성능을 높이기 위한 RAG의 개선과 최적화 실험도 진행하였다. 먼저, RAG 검색을 다양화하여 적절한 문장을 검색하도록 하였다. RAG 검색 시 사용자의 입력 문장을 그대로 검색하는 것뿐만 아니라 형태소 단위로 나누어 검색한다. 이와 더불어, 사용자 입력 문장으로부터 명사를 추출하

여 해당 단어에 대한 일반적인 텍스트 검색도 수행한다. 이렇듯 검색 방식을 다양화하는 이유는, 전통적인 키워드 검색 방식을 사용하거나 chunk 단위로 검색할 경우에 더 잘 검색되는 대화들이 있기 때문이다. 가령, '제가 쓰던 축구공 팔아요.'라는 대화는 RAG 시 '축구공'보다 '팔아요'라는 의미에 더욱 집중하여 어떠한 제품을 판매하는 대화들을 추출하지만, 해당 문장을 형태소 단위로 분리 시 '축구공'과 더 관련된 문장들을 추출하게 된다. 프롬프트 역시 단순히 사용자의 요구에 따라 매칭된 타 사용자를 출력해달라 하는 것보다 매칭 이유까지 출력하도록 할 때, 그리고 사용자가 어떠한 물건을 사고 싶어할 때 사용자가 원하는 물건과 타 사용자가 판매 중인 물건을 함께 출력해보라 작성할 경우 더 적절하고 정확히 매칭하는 양상을 발견하였다.

매칭 시 응답 속도를 높이는 방법으로서, 유사 대화들과 사용자의 입력 대화를 프롬프트에 넣고 매칭을 요청하는 과정에서, 매칭된 상대의 활동명만을 먼저 출력한 뒤 답변을 생성하는 '필터링' 방법을 시도하였다. 이 방법의 유효성을 검증하기 위해, 판매 상점 및 취미 관련 1,102개의 데이터를 활용하여 RAG를 이용한 문장 필터링 유무에 따른 LLM의 응답 속도를 비교 분석하기 위해 30번의 실험 후, 평균 응답 속도를 측정하고, 독립 표본 T-검정과 단순 회귀분석을 실시하였다. 필터링 미적용에서의 평균 응답 속도는 12.59초인데, 필터링 적용시 경우 평균 응답 속도는 10.85초로 감소하였고, 독립 표본 T-검정의 결과, 필터링 적용 전후의 응답 속도 차이는 p-value 0.0015로 매우 유의미한 것으로 나타났다. 또한 회귀분석을 통해 계산된 p-value는 0.002로, 필터링이 응답 속도에 유의한 영향을 미친다는 결과를 재확인했다. 회귀분석에서 도출된 R-squared 값은 0.735로, 필터링 적용의 설명력이 높음을 시사하며, 이는 필터링이 LLM 응답 속도에 영향을 끼치는 주요 요인임을 나타낸다. 본 실험결과는, RAG를 이용한 문장 필터링과 같은 최적화 기법이 LLM 응답 속도를 향상시켜서, 인공지능 시스템의 효율성과 사용자 경험을 개선하는 데 기여할 수 있음을 보여준다. 이러한 매칭 알고리듬과 이에 기반한 매칭 에이전트 서비스를 개발하는 과정은 다양한 LLM 대안을 놓고, 속도와 비용, 매칭 성과를 고려하여 최선의 조합을 취사 선택하고, 또 RAG 자체의 성과를 최대하기 위한 알고리듬을 실험하는 과정을 필요로 한다. 실험결과로 나온 지식들은 프롬프트로 반영되며, 그것은 절차적 알고리듬의 형태로 표현되기보다는 LLM에게 명령하는 프롬프트에 선언적으로 반영하여, 매칭 알고리듬에서의 코드와 지식의 독립성을 높이는 노력을 계속할 예정이다.

매칭에이전트는 거래 상대방을 찾는 데에 그치지 않고, 둘간의 채팅, 만남, 거래 완결까지 진행할 필요도 있다. 이 과정을 완전히 대화형 인터페이스(CUI: Conversational User Interface)로 구현할 수도 있겠지만, 결제 등의 프로세스는 사용자의 명확한 의사 표시와 행동에 근거할 필요가 있으므로, 이를 대화형이 아닌 사용자 클릭 등의 GUI(Graphical User Interface)로 개발하는 한편, 사용자의 의도 파악 알고리듬이 경량화, 내재화되어 비용과 처리시간이 최소화되는 동시에 정확성을 담보하게 된다면, 완전한 대화형 서비스로 구현하는 것을 준비중이다.

매칭 알고리듬은 다양한 분야에서 사용자 요구를 파악하고 원하는 조건의 상대를 찾아주는 매칭 Agent 시스템에서 유연하게 사용되는 것을 목표로 한다. 특정 도메인에서의 매칭에서 사용될 경우, RAG 검색 방식이나 프롬프트에 포함된 변수들을 파라미터화 하여 특화하여 사용할 수 있도록 개발중이다. 개발된 알고리듬은 현재 ㈜하렉스인포텍의 매칭 에이전트 Jarvis Just (JJ)에 채용되어, 중부시장 등 전통시장에 먼저 적용되고 있다.

본 연구는 처음에는 오직 순수한 LLM기반의 이른바 No Code 프로그래밍의 형태로 구현하는 새로운 에이전트 개발 방식을 목표로 출발하였다. 그런데, 중요한 태스크(Task)는 그저 상식과 일반 지식을 가진 LLM의 채용으로 해결되지 않는다는 점을 개발 과정에서 절감하였다. 더구나 다양한 태스크에 대해서 비용과 시간이 많이 드는 LLM API(Application Programming Interface)를 외부에서 일률적으로 부르는 것은, 실험실 연구가 아닌 실제 산업 현장에 적용하는 연구에는 비현실적이었다. 결국 LLM 그 자체의 최적화, RAG 방법 그 자체의 최적화 등 각 프로세스 단위의 최적화가 필요한 동시에, 여러 다양한 LLM API와 자체 AI 모델에 대한 Make-or-Buy Decision을 가설에 근거한 실험을 통해 채택하고 조합해야 하는 최적화 과정이라는 점을 확인할 수 있었다.

또한, 그러한 최적화의 결과가 프롬프트 상에서의 절차적 코드 형태로 표현되는 것은 시스템 유지보수 관점과 에이전트의 특정 태스크 지식을 관리한다는 측면에서 모두 비효율적이므로, 프롬프트 엔지니어링 과정에서 코드와 지식을 어떻게 분리함으로써, 최적화를 달성할 수 있는가라는 새로운 이슈를 제기한다는 점을 발견하였다.

Key words- Large Language Model, Retrieval-Augmented Generation, Matching Intelligence

텍스트 마이닝을 활용한 반려동물 관련 연구의 동향 연구

이동현
아주대학교 경영대학
960ldh910@gmail.com

강주영
아주대학교 경영대학
jykang@ajou.ac.kr

이한솔
아주대학교 경영대학
lhs15@ajou.ac.kr

***Abstract** – 최근 반려동물 양육인구 비율이 지속적으로 증가함에 따라 반려동물 연관 산업이 급격하게 성장하고 있다. 이에 본 연구에서는 반려동물 관련 연구의 동향을 분석하여 관련 산업체에서 활용할 수 있는 기초자료를 제시하고자 한다. 국가과학기술지식정보서비스(NTIS)의 특허, 과제, 연구보고서 자료를 활용하여 동향분석을 실시하였다. Python 을 이용하여 텍스트 마이닝을 실시하고 불용어 사전에 불용어를 추가한 후 상위 10 개 키워드만을 추출한 결과, 반려동물 관련 특허의 경우 '사료', '균주', '진단', '추출', '함유', '건강', '모니터링', '수제', '락토바실러스', '관절염' 순으로 단어가 추출되었고, 과제의 경우 '개발', '창업', '질병', '진단', '디딤돌', '복지', '중소기업', '의약품', '형질', '산학' 순으로 키워드가 추출되었다. 연구보고서의 경우 '사료', '진단', '치료', '질병', '건강', '바이러스', '면역', '항생제', '모니터링', '예방' 순으로 키워드가 추출되었다. 본 연구를 바탕으로 국내 업체들이 산업 내 연구 동향을 파악하고 경쟁력을 강화할 수 있을 것으로 기대한다.*

Keywords** – **반려동물, 반려동물산업, 반려동물특허, 반려동물연구동향, 펫테크

사사표기 (논문 지원기관이 있는 경우, 여기에 표기)

이 논문은 2024 년 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0020632, 2022 년 산업혁신인재성장지원사업)

Session C

C2. [특별세션]
AI 신뢰성의 과거 및 현재 이슈와 미래 전망

좌 장: 허 용 석(영남대)

여 백

Session C

C3. [학술세션]
디지털 금융 혁신: AI와 딥러닝

좌 장: 고 지 훈(부산대)

Anchoring bias and the cross-section of stock return patterns

김동훈
KAIST
kdh9406@kaist.ac.kr

고지훈
부산대학교
jihoongoh@pusan.ac.kr

Abstract – This study explores the impact of anchoring bias on the daily tug-of-war in financial markets, specifically examining the reversal dynamics between overnight and daytime periods in relation to the 52-week high price. After analyzing the interplay between anchoring bias and the abnormal intensity of the tug-of-war (AB_NR), our findings reveal a significant AB_NR premium for stocks far from their 52-week high prices. This sheds light on the important role of psychological barriers faced by noise investors during overnight periods in shaping return dynamics. Furthermore, our research will also be helpful in explaining the return patterns in the cryptocurrency market where investors exhibit behavioral biases.

Key Terms – Overnight return, Daytime reversal, Arbitrageurs, Anchoring bias, 52-week high

The dynamics of financial markets, including the effects of major economic news and information released outside regular trading hours, have received significant attention in the extant literature. Specifically, the concept of reversal between overnight and daytime periods has recently been the focus of considerable research (e.g., Berkman et al., 2012; Hendershott et al., 2020; Bogousslavsky, 2021; Cheema et al., 2022). Lou et al. (2019) find that a tug-of-war occurs between two distinct groups—individual investors and institutional investors—resulting in return reversals between overnight and daytime periods. Akbas et al. (2022) document that an intense daily tug-of-war leads to the undervaluation of a stock because arbitrageurs often disregard the portion of overnight gains associated with fundamental changes and consistently exhibit overcorrection behavior in response to consecutive positive overnight returns. They propose the measure AB_NR, the abnormal intensity of the tug-of-war, which captures the abnormal frequency of positive overnight returns followed by negative reversals during regular trading hours.

Anchoring bias is also well documented in the literature, and it plays a crucial role in investors' decision-making processes. Several studies suggest that investors rely on anchoring bias, meaning that investors use the 52-week high price as a reference point to evaluate a stock's potential price movement (e.g., Kahneman and Tversky, 1974; George and Hwang, 2004; Birru, 2015). In particular, anchoring bias is strongly observed mainly among individual investors (e.g., Huddart et al., 2009). If stocks are near their 52-week high prices, investors tend to believe that positive news has already influenced prices. This results in a decrease in aggressive bidding and subdued price reactions to positive events. Thus, these stocks tend to be undervalued, leading to a high expected return.

This study examines the influence of anchoring bias on the daily tug-of-war. We conjecture that when stocks are far from their 52-week high prices, overnight noise traders (mainly individual investors) tend to determine that there is enough room for further price adjustments rather than assessing the intrinsic value of a stock. These stocks experience substantial upward pressure from overnight noise traders, who perceive greater potential for price increases. In response, daytime arbitrageurs overcorrect this overvaluation, resulting in relatively undervalued high AB_NR stocks compared with low AB_NR stocks. Conversely, for stocks near their 52-week high prices, overnight noise traders perceive less room for further price movement, leading to reduced pressure on prices. In response, daytime arbitrageurs exhibit less overcorrection behavior, causing high AB_NR stocks to be less undervalued, potentially weakening the return predictability of AB_NR.

Our results indicate that the AB_NR premium is significant among stocks that are far from their 52-week high prices but marginally significant for stocks near their 52-week high prices. Fama and MacBeth (1973) cross-sectional regressions show significant negative coefficients for the interaction between NH and AB_NR. These results are amplified following high-sentiment states, supporting our hypothesis.

This study strengthens the previous research regarding overnight returns and daytime reversals. By exploring the connection to anchoring bias, we offer a novel perspective on the psychological barrier faced by noise investors during overnight periods in shaping return dynamics.

This study sheds light on the role of anchoring bias in the tug-of-war between overnight traders and daytime traders. Our findings reveal that the predictive power of the

intense daily tug-of-war, as measured by AB_NR, is more pronounced among stocks trading far below their 52-week high prices. This suggests that overnight noise traders (mainly individual investors) may be influenced by anchoring bias. Noise traders generate more (less) upward price pressure for positive stock information when the stock prices are far from (near) their 52-week high prices; this is then followed by the overcorrection by arbitrageurs. Our findings provide a fresh perspective on the psychological barriers that investors face when dealing with overnight returns. Furthermore, our research will also be helpful in explaining the return patterns in the cryptocurrency market where investors exhibit behavioral biases.

<Table 1> Fama and MacBeth (1973) regression

This table presents the Fama–Macbeth cross-sectional regression results between stock returns and several variables.

Variable	Model 1	Model 2	Model 3	Model 4
AB_NR	0.189	0.091	0.168	0.107
	(6.52)	(5.83)	(7.50)	(6.63)
AB_PR	–0.066	–0.04	–0.028	–0.039
	(–1.76)	(–1.39)	(–0.98)	(–1.55)
RETOC_M	–0.253	–0.366	–0.483	–0.374
	(–2.53)	(–3.87)	(–6.04)	(–5.32)
RETCO_M	–0.276	–0.217	–0.344	–0.225
	(–4.48)	(–5.04)	(–5.89)	(–4.73)
AB_NR × NH			–0.078	–0.073
			(–4.04)	(–3.83)
AB_PR × NH			0.039	0.035
			(1.95)	(1.77)
NH			0.445	0.042
			(2.93)	(0.33)
BETA		–0.021		–0.020
		(–0.54)		(–0.61)
ME		–0.162		–0.183
		(–2.33)		(–3.10)
BM		0.114		0.087
		(1.49)		(1.19)
MOM		0.188		0.211
		(2.12)		(2.67)
ILLIQ		0.006		0.006
		(0.13)		(0.13)
TURN		0.297		0.276
		(5.37)		(5.30)
IVOL		–0.549		–0.496
		(–5.41)		(–6.82)

III. 참고문헌

Akbas, F., Boehmer, E., Jiang, C., Koch, P.D., 2022. Overnight returns, daytime reversals, and future stock returns. Journal of Financial Economics 145, 850-875

Berkman, H., Koch, P.D., Tuttle, L., Zhang, Y.J., 2012. Paying attention: overnight returns and the hidden cost of buying at the open. Journal of Financial and Quantitative Analysis 47, 715-741

Birru, J., 2015. Confusion of confusions: A test of the disposition effect and momentum. The Review of Financial Studies 28, 1849-1873

Bogousslavsky, V., 2021. The cross-section of intraday and overnight returns. Journal of Financial Economics 141, 172-194

Cheema, M.A., Chiah, M., Man, Y., 2022. Overnight returns, daytime reversals, and future stock returns: Is China different? Pacific-Basin Finance Journal 74, 101809

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607-636

George, T.J., Hwang, C.Y., 2004. The 52-week high and momentum investing. The Journal of Finance 59, 2145-2176

Hendershott, T., Livdan, D., Rösch, D., 2020. Asset pricing: A tale of night and day. Journal of Financial Economics 138, 635-662

Huddart, S., Lang, M., Yetman, M.H., 2009. Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. Management Science 55, 16-31

Lou, D., Polk, C., Skouras, S., 2019. A tug of war: Overnight versus intraday expected returns. Journal of Financial Economics 134, 192-213

Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. Science 185, 1124-1131

ChatGPT 의 지시문 조정을 통한 한국 주식시장 투자 전략 구축

이재욱
부산대학교 경영학과
디지털금융전공
jaaewook@pusan.ac.kr

도영훈
부산대학교 경영학과
디지털금융전공
dudgns342@naver.com

정다훈
부산대학교 경영학과
디지털금융전공
wjdekgns980@gmail.com

김태현
부산대학교 물리학과
tae6709@gmail.com

이현빈
부산대학교 기계공학부
lhb823@pusan.ac.kr

이민혁
부산대학교 경영학과
디지털금융전공
minhyuk.lee@pusan.ac.kr

Abstract– *금융 시장에서 텍스트 마이닝은 대량의 비정형 텍스트 데이터를 분석하여, 시장 동향, 투자자 심리, 소셜 미디어의 영향 등에서 숨겨진 패턴을 파악하기 위해 사용된다. 최근에는 **OpenAI** 사의 **ChatGPT** 와 **Meta** 의 **LLAMA2** 로 대표되는 대규모 언어모델(**Large Language Model, LLM**)이 금융 분야 텍스트 마이닝 방법으로 각광받고 있다. 하지만, 응답의 무작위성과 실제로 사실이 아닌 정보를 사실처럼 생성하여 제공하는 "모델 환각(**Hallucination**)"과 같은 문제가 실제 이용을 어렵게 하는 요인으로 지적되고 있다. 본 연구에서는, 뉴스 식별, 투자종목 선정, 감성 분석에 이어지는 연쇄적인 과정을 **ChatGPT** 를 이용하여 구현하였다. 이 결과에 기반하여, 일일 매매 매도 전략을 구축하여 실제 투자 상황을 시뮬레이팅하여, 결과를 도출했다. 이후, 기존 재무 분야 텍스트 마이닝에 사용되었던 **FinBert** 와 범용 텍스트 마이닝 기법인 **KlueBert** 와의 투자 성과 차이를 관찰했다. 이 연구는 비교적 간단한 조정 방식인 프롬프트 조정을 통해 금융 분야에서의 **LLM** 사용에서 발생할 수 있는 문제들을 해결할 수 있는 실질적인 가이드라인을 제공하며, 나아가 **AI** 기술이 투자 결정 및 전략 개발에 어떻게 기여할 수 있는지에 대한 통찰력을 제공할 것으로 기대된다.*

<Keywords> ***BERT, ChatGPT, LLM, Prompt, Stock Preditcion, Transformer***

I. 서론

텍스트마이닝은 비정형 텍스트 데이터에 담겨있는 패턴이나 상관관계를 찾아내어, 유의미한 정보로 전환하는 일련의 과정이다. 초기 텍스트마이닝은 단순하게 특정 단어나 문장의 빈도를 분석하는 방식으로 이루어졌다. 하지만, 컴퓨터의 하드웨어와 알고리즘이 발전하며, 단순 빈도 분석보다 정교하고 복잡한 의미론 및 구문 분석이나 초기 머신러닝 기법들이 이용되기 시작했다. 2010 년대 후반 트랜스포머 알고리즘이 제안되며, 텍스트의 복잡한 의미를 깊이 이해하고, 숨겨진 문맥 및 문장 간의 관계에 대한 탐색이 가능해지며 자연어처리 및 텍스트마이닝 분야는 빠르게 발전하고 있다.

텍스트마이닝의 발전과 함께 금융 및 투자 분야로의 도입 역시 가속화되고 있다. 특히 거대언어모델(Large Language Model, LLM)과 트랜스포머 알고리즘이 ChatGPT 와 Llama 로 대표되는 서비스를 통해 구현되며, 금융 분야에서 트렌드 파악, 시장 동향 분석, 투자자 심리 분석 등에서 사용될 수 있을 것으로 기대되고 있다. 하지만, LLM 을 통한 텍스트마이닝 진행과, 이를 금융 및 투자 분야에 활용하는 데에는 몇 가지 문제점이 있다. LLM 은

단순한 분석을 넘어 텍스트를 생성하고, 이에 따라 무작위적 답변 생성의 위험에 노출되어 있다. 또한, 현실에 존재하지 않거나 부합하지 않는 답변을 진짜인 것처럼 답하는 환각(Hallucination) 역시 LLM 의 신뢰성을 하락시키는 요소이다. LLM 의 무작위성과 환각 현상을 완화하기 위해서 구조화된 프롬프트의 구성, 개체명 식별(Named Entity Recognition), 검색 증강 시스템 (Retrieval-Augmented Generation) 등이 이용될 수 있다. 본 연구에서는, LLM 의 무작위성과 환각 현상을 최소화하여, 신뢰성 있는 분석을 진행하기 위해, 구조화된 프롬프트 조정을 진행하고, NER 을 이용하여 투자 종목 식별을 진행한다. 이후, 실제 뉴스 헤드라인에 대한 감성 분석을, ChatGPT 를 통해 진행하여, 일일 매수 종목을 선정하고, 이 결과를 한국거래소의 KOSPI 의 가격 데이터를 통해 검증한다. 마지막으로, 기존 BERT 알고리즘들을 통한 감성 분석 결과와 비교하여, 본 연구에서 진행한 ChatGPT 조정의 효율성과 한계를 탐색하고자 한다.

II. 연구 방법

본 연구에서 사용되는 뉴스 헤드라인은 뉴스 빅데이터 분석 서비스 빅카인즈(BIGKinds)에서 수집하였다. 수집 기간은 2023 년 1 월 2 일부터 4 월 30 일까지의 기간으로 설정하였으며, LLM 의 베이스 모델은 ChatGPT 3.5-Turbo, BERT 의 베이스 모델은 서울대학교 Ambient 자연어처리 연구실에서 개발한 KR-FinBert 와 KlueBert 를 사용하였다. 투자 대상 자산은 2022 년 12 월 28 일 기준 KOSPI200 편입 종목으로 선정했다.

ChatGPT 3.5-Turbo 를 이용하여, 전체 뉴스 헤드라인 중 단기 KOSPI 변동에 대해, <그림 1>에서 제시된 프롬프트를 통해, 유의미한 헤드라인만 선별했다. 그 이후, <그림 2>, <그림 3>에서 제시된 프롬프트에 따라, 각 뉴스 헤드라인이 어떤 자산과 관련이 있는지를 식별하고 그 식별 자산에 대한 감성분석을 진행했다. 뉴스 헤드라인 하나당 두 개의 종목을 식별했다.

- prompt = """Forget all previous instructions and processes. You are an expert who has been investing in and advising on the South Korean financial market for a long time. You have a precise understanding of the KOSPI200 stocks and their characteristics, as well as how they have evolved over time. Please accurately carry out the tasks I will instruct. Determine whether the news headlines I provide will have an impact on the short-term price fluctuations of KOSPI200 stocks. Output the result as either "O" for influential or "X" for not influential."‘
  prompt += f"\nNews headline: '{headline}'\nImpact on KOSPI200 short-term price fluctuations: "

<그림 1> 헤드라인 선별 프롬프트

- Forget all previous instructions. Pretend you're a proficient securities analyst in the South Korean stock market, particularly skilled in utilizing text mining to understand market sentiment and generate above-average returns. Your task is to select two stocks from the KOSPI 200 that will be most significantly impacted in the short term by the provided news headlines : '{headline}'. Select two stocks and format your response as {{“Name_Ticker"}} for each, separated by a comma.

<그림 2> 종목 식별 프롬프트

- Based on the headline: '{headline}', and considering the stock {asset}, Please determine whether the presented news will have a positive or negative correlation with short-term price fluctuations. Provide me with either "O" for a positive correlation or "X" for a negative correlation, without any detailed explanation.

<그림 3> 감성분석 프롬프트

ChatGPT3.5-Turbo 와의 비교를 위해, 동일한 기간의 뉴스 헤드라인 데이터를 KR-FinBert 와 KlueBert 를 이용하여 분석했다. 뉴스 선별 및 식별은 KcElectra 의 NER(Named Entity Recognition) 기능을 이용했다. 이후, KR-FinBert 와 KlueBert 를 이용해 뉴스 헤드라인에 대한 감성분석을 진행했다.

III. 참고문헌

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

Fatouros, G., Soldatos, J., Kouroumali, K., Makridis, G., & Kyriazis, D. (2023). Transforming sentiment analysis in the financial domain with ChatGPT. *Machine Learning with Applications*, *14*, 100508.

Kim, J. H. (2023). What if ChatGPT were a quant asset manager. *Finance Research Letters*, *58*, 104580.

Konstantinidis, T., Golan, M., Xu, Y. L., Constantinides, T. G., & Mandic, D. P. (2023, June). Text Mining for Sentiment Analysis in Bond Portfolio Construction. In *2023 24th International Conference on Digital Signal Processing (DSP)* (pp. 1-5). IEEE.

Lopez-Lira, A., & Tang, Y. (2023). Can chatgpt forecast stock price movements? return predictability and large language models. *arXiv preprint arXiv:2304.07619*.

Pelster, M., & Val, J. (2024). Can ChatGPT assist in picking stocks?. *Finance Research Letters*, *59*, 104786.

Wu, R. (2024). Portfolio performance based on LLM news scores and related economical analysis. *Available at SSRN 4709617*.

Zhang, B., Yang, H., Zhou, T., Ali Babar, M., & Liu, X. Y. (2023, November). Enhancing financial sentiment analysis via retrieval augmented large language models. In *Proceedings of the Fourth ACM International Conference on AI in Finance* (pp. 349-356).

Designing Mobile Message to Reduce Loan Repayment Default in Microfinance: Evidence from Field Experiment

Nigusse Gebre
Department of Business and Technology Management, KAIST
GebreN@kaist.ac.kr

Woojin Yang
Department of Business and Technology Management, KAIST
dnwls2669@kaist.ac.kr

Chul Ho Lee
Department of Business and Technology Management, KAIST
irontigerlee@kaist.ac.kr

Abstract – *The study is about how financial nudging with text messages changes the behavior of microfinance borrowers. This study is the first to test the impact of financial nudging on microfinance borrowers. Focuses on how text message reminders reduce repayment default in microfinance. We use timing and information to test how psychological distance determines a consumer's response to text message reminders for loan payments. The results showed a positive relationship; sending text reminders on the previous day to the due date with the amount of money and ten days before the due without the amount of money increases the chances of borrowers' intention to pay in contrast with sending them ten days before the due date with the amount of money of the payment. Also, the borrowers who received ten days before the due without the amount of money show that their intention to pay behavior stayed long in the post-treatment period due to their strong self-control capability. Our paper is the first in IS literature to show the role of performance ambiguity empirically and to associate the ambiguity with antecedents derived from construal-level theory. Besides, it advises that considering the when, where, and how of mobile targeting strategies is vital and designing mobile text message reminders prudently, how microfinance can mitigate default risks. Therefore, policymakers can benefit from developing a practical policy for mobile applications by considering the contextual factors we proposed.*

Key Terms – mobile banking; microfinance; loan repayment default; self-control

동백전 빅데이터를 활용한 인공지능 기반 신흥 상권 도출에 관한 연구

송위창
동아대학교 경영정보학과
2378338@donga.ac.kr

김상진
동아대학교 경영정보학과
skim10@dau.ac.kr

***Abstract** – 본 연구에서는 부산시 동백전 지역화폐 빅데이터를 활용하여 신흥 상권을 도출하고 통계 기반 알고리즘과 딥러닝 기반 알고리즘에 신흥 상권의 카드 승인 건수 및 승인 금액 시계열 데이터를 적용, 모델링하고 학습하여 상권의 성장성을 예측하였다.*
지역 인구 유입 및 유출, 유동 인구의 통계정보를 이용하거나 설문조사 방법을 활용하여 신흥 상권 도출에 대한 연구를 수행할 경우 통계적 오류로 인하여 모집단의 대표성을 충분히 반영하지 못하는 한계점이 발생한다. 반면 빅데이터를 이용한 분석은 실제 소비 활동과 데이터 수집 및 분석 사이에서 발생하는 간극을 최소화하여 상권의 변동성에 민첩하게 대응할 수 있다.
*부산시 동백전 지역화폐 가맹점에서 사용한 카드 거래 내역을 우편번호 기반으로 집계 및 분석작업을 수행하였고, 신흥 상권으로 선정된 지역의 시계열 데이터를 기반으로 통계 기반 예측 알고리즘인 **Auto-ARIMA, Prophet** 과 딥러닝 기반 예측 알고리즘인 **RNN, LSTM, GRU** 를 사용하여 모델링 및 학습으로 상권의 성장을 예측하였으며, 각 모델의 성능을 **MAE, RMSE, R2 Score** 값으로 평가하였다*
*신흥 상권으로 선정된 지역의 시계열 데이터를 통계 기반 예측 알고리즘과 딥러닝 기반 예측 알고리즘으로 모델링하고 학습하여 성장성을 예측한 결과 통계 기반 모델인 **Prophet** 과 **Auto-ARIMA** 의 **R2 Scores** 는 각각 **44.9%, 44.7%**로 **Prophet** 이 약 **0.2%** 정도의 근소한 차이로 우수하였으나 두 모델의 적합도는 낮다. 반면, 딥러닝 기반 모델인 **RNN, LSTM, GRU** 의 성능 평가 지표인 **R2 Scores** 는 각 **21%, 74%, 93%**로 **GRU** 모델의 성능이 가장 우수하였다. 향후 본 연구 결과를 기반으로 빅데이터 및 딥러닝 기반 모델을 기반으로 객관적이며 과학적으로 상권의 트렌드를 분석하여 다양한 시장 참여자들에게 경제활동을 통한 성장의 기회를 제공할 수 있다.*

***Keywords** – 동백전, 빅데이터, 신흥상권, 통계기반모델, 딥러닝기반모델*

I. 서론

분석 대상 빅데이터는 2021 년 4 월 부터 2022 년 3 월 까지 약 1 년간 발행된 부산시 동백전 지역화폐 거래 내역이다. 빅데이터는 고객 정보, 가맹점 정보, 카드 승인 정보로 구분되며 <표 1> 분석 데이터 현황과 같이 카드 승인 건수 약 9,000 만 건, 승인금액 약 2조 원을 대상으로 분석을 수행하였다.

<표 1> 분석대상 데이터

구 분	내 용
데이터명	동백전 카드 승인 내역
분석기간	2021.04 ~ 2022.03
데이터 종류 (전처리 後)	√회원정보 -성별 : F(여성), M(남성) -연령대(10, 20, 30, 40, 50, 60, 70, 80, 90) √가맹점정보 -우편번호(주소), 업종명 √카드승인내역 -승인건수, 승인금액
분석대상 (카드승인)	√ 승인 건수 : 90,992,896 건 √ 승인 금액 : 약 2 조

분석된 상권 정보를 활용하여 상권의 속성이나 트렌드를 면밀하게 파악할 수 있으며, 이를 기반으로 시장 참여자들에게 Insight 를 제공하는데 본 연구의 목적이 있다.
시장 참여자들이 얻을 수 있는 Insight 는 첫 번째, 기업에서는 본 연구의 방법론과 결과를 기반으로 데이터를 분석하여 점포 신설 및 대고객 타깃 마케팅 등 효율적인 영업 전략 수립에 활용할 수 있다. 두 번째, 행정기관에서는 시민 및 소상공인

지원을 위한 행정 정책에 이용할 수 있다. 세 번째 소상공인은 데이터를 기반으로 상권의 속성과 점포(업종) 현황을 과학적으로 분석하여 창업을 통한 영업의 기회를 창출할 수 있다.
4 차 산업혁명의 구성 요소 중 핵심은 데이터이다. 빅데이터 및 인공지능을 기반으로 한 데이터 분석 및 예측 모델은 조직의 통찰력, 대응력, 경쟁력 향상에 기여하여 서비스 품질 및 가치를 높이고 업무의 효율성을 향상할 수 있어 여러 분야에서 활용할 수 있다.

II. 연구결과

동백전 빅데이터를 분**석하여** 신흥 상권으로 선정된 기장군 일광면과 강서구 대저 1 동에 대하여 카드 승인 건수 시계열 데이터를 활용하여 전통적인 통계 기반 모델인 Auto-ARIMA 와 Prophet, Deep Learning 기반 모델인 RNN, LSTM, GRU 를 이용하여 학습한 후 예측한 결과를 가장 예측 성능이 우수한 GRU 모델은 <그림 1>, <그림 2>와 같다.

<그림 1> GRU PREDICT(일광면)

<그림 2> GRU PREDICT(대저 1 동)

모델별 예측 성능은 <표 2>과 같으며 일광면을 기준으로 랭킹 1 위 GRU(일광:92.8%, 대저:89.7%), 랭킹 2 위 LSTM(일광:74.2%, 대저:73.9%), 랭킹 3 위 Prophet(일광:44.9%, 대저:21.5%), 랭킹 4 위 Auto-ARIMA(일광:44.7 측 모델을 수행한 결과 GRU 모델의 성능이 90% 이상으로 가장 우수하였다.

<표 2> 모델 별 성능(R SQUARE) 지표

구분	모델	R Square	랭킹
기장군 일광면 상권	Auto-ARIMA	0.447353	4
	Prophet	0.449313	3
	RNN	0.212716	5
	LSTM	0.742104	2
	GRU	0.928348	1
강서구 대저 1 동 상권	Auto-ARIMA	0.210188	5
	Prophet	0.215478	3
	RNN	0.212415	4
	LSTM	0.738882	2
	GRU	0.897446	1

III. 결론

IT 인프라 및 인공지능 등 다양한 알고리즘의 발전으로 빅데이터를 실시간으로 수집하고 분석하여 데이터 기반 영업 및 마케팅 전략을 수립할 수 있다.
본 연구에서는 빅데이터 분석을 통해 도출된 신흥 상권의 카드 거래 내역 시계열 데이터를 전통적인 통계 기반 알고리즘과 인공지능 딥러닝 기반 알고리즘을 이용하여 모델링하고 학습하여 상권의 성장을 과학적으로 예측하였다.
동백전 시계열 데이터를 학습하고 예측한 결과 전통적인 통계 기반 모델보다 딥러닝 기반 모델이 우수하였으며, 성능 측정 지표인 R2 Scores 를 기준으로 비교한 결과 딥러닝 기반 모델인 GRU 가 90% 이상, LSTM 이 74% 이상으로 측정되어 하이퍼파라미터를 추가 튜닝할 경우 Reproduct 과정을 통해 업무 시스템에 적용이 가능한 수준의 우수한 결과를 도출하였다.
통계정보를 이용하여 상권을 예측할 경우 표본 추출과 통계 이면에 내포된 정보의 왜곡이 발생할 수 있는 반면, 본 연구에서 사용된 동백전 카드 거래 현황 데이터는 1 년간 기록된 전체 모수를 대상으로 수행되어 데이터를 통한 패턴과 규칙을 찾아 수집, 집계, 분석하여 활용함으로 인해 보다 현실적인 신흥 상권에 대한 연구 결과를 도출할 수 있다.

VI. 참고문헌

정연승, & 최순화. (2015). 한국형 스트리트형(문화쇼핑거리) 상권 현황과 발전방안. 한국유통학회 학술대회 발표논문집, 28-37..

강현화. (2011). 글로벌환경하에서의 부산지역상권의 활성화 전략 (Doctoral dissertation, 부경대학교).

김원겸. 지방분권화 정책과 충청권 유통상권 변화전망: 신도시 개발에 따른 대전광역시 상권변화를 중심으로: 신도시 개발에 따른 대전광역시 상권변화를 중심으로. 한국유통학회 학술대회 발표논문집, 2004, 379-393..

Song, K., & Lee, H. (2020). 지역화폐 도입이 지역경제에 미친 영향 분석 (The Effect of Introducing Community Currency on Regional Economy: The Case of Korea). Available at SSRN 3554493.

동욱김, 정윤조, 문엽장, & 주일송. (2021). Development of heatwave probability prediction model based on recurrent neural network: A case of seoul metropolitan government. Journal of the Korean Society of Hazard Mitigation, 21(5), 49-57.

우정사업본부 고시 제 2014-63 호(2014). 국가기초구역 체계로의 우편번호 개편.

백충기(2020). 동남권 주요 상권의 변화 분석. BNK 경제연구원.

부산광역시(2023). 지역화폐 동백전 가맹점 약관집

신동하, 최광호, & 김창복. (2017). RNN 과 LSTM 을 이용한 주가 예측율 향상을 위한 딥러닝 모델. 한국정보기술학회논문지, 15(10), 9-16.

안준호, 박순찬. (2020). 폐업 통계는 내지도 않고, 정부 "창업 늘었다". 조선일보. https://www.chosun.com

[국외 문헌]

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

Chris Kuo/Dr. Dataman(2020). Technical Guide on RNN, LSTM, GRU for Stock Price Prediction. https://medium.com/swlh/a-technical-guide-on-rnn-lstm-gru-for-stock-price-prediction-bce2f7f30346

Fowler Jr, F. J. (2013). Survey research methods. Sage publications.

Facebook Open Source. (2023). https://facebook.github.io/prophet/docs/quick_start.html

Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. Journal of machine learning research, 3(Aug), 115-143.

Google Map. (2023). https://www.google.co.kr/maps/place

Hansun, S. (2013, November). A new approach of moving average method in time series analysis. In 2013 conference on new media studies (CoNMedia) (pp. 1-4). IEEE.

Lyberg, L., & Weisberg, H. (2016). Total survey error: a paradigm for survey methodology. The Sage handbook of survey methodology, 10(9781473957893), n3.

Shi, J., Jain, M., & Narasimhan, G. (2022). Time series forecasting (tsf) using various deep learning models. arXiv preprint arXiv:2204.11115.

Stankeviciene, J., Maknickiene, N., & Maknickas, A. (2017). High-low strategy of portfolio composition using Evolino RNN ensembles. Engineering Economics, 28(2), 162-169.

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. The American Statistician, 72(1), 37-45.

Taylor G.S. (2023). https://alkaline-ml.com/pmdarima/index.html

김치 프리미엄 현상과 가상화폐 2 차 시장 간의 매개효과

조승연
한양대학교 경영대학
shengj0712@gmail.com

임규건(교신저자)
한양대학교 경영대학
gglim@hanyang.ac.kr

Abstract – 블록체인 시장의 지속적인 투기 가능한 환경은 가상화폐 거래소에 유통량을 가져다주는 동시에 많은 위기에 직면해 있다. 우리가 Binance 거래소의 5 년간의 패널 데이터 수집과 사회 외부 환경 변화, 그리고 블록 생성 효율성을 결합하여 김치 프리미엄 현상 가상화폐 거래소의 구체적인 영향을 검증했다. 그리고 김치 프리미엄 현상과 가상화폐 거래소 발전 규모에 큰 영향을 미치는 요인을 탐색하였다. 본 연구를 통해 가상화폐 거래소 발전의 효율성을 높이고 가상화폐 거래 환경의 정보 비대칭성을 줄일 수 있기를 바란다.

Keywords – 가상화폐 거래소, 김치 프리미엄 현상, 매개효과, 패널 데이터

I. 서론

지역, 법률, 환율 등의 이유로 가상화폐 거래소 간 동일 가상화폐의 가격이 다른 현상이 있다. 이러한 가격 차이는 전세계적으로 가상화폐 시장의 특성과 환경에 영향을 받다. 예를 들어, 한 국가에서는 가상화폐 거래에 대한 규제가 엄격하여 거래소의 운영이 제한되어 있을 수 있다. 또한, 환율 변동이나 특정 지역의 수요와 공급에 따라 가격이 상이할 수 있다. 이러한 가격 차이는 사용자들에게 다양한 투자 기회를 제공한다. 일부 사용자들은 이러한 가격 차이를 활용하여 헤지(hedge), 공매도(short sale) 등의 전략을 사용하여 수익을 얻으려고 한다. 이러한 행위는 가상화폐 시장의 투기적인 성격을 반영하며, 거래소의 거래량을 늘리고 이익을 창출할 수 있다. 하지만 일방적인 거래는 과연 거래소한테 긍정적인 영향만 미치는 것인지 심층적인 분석이 필요한다.

기존 연구는 시장의 김치 프리미엄 현상, 투자자들이 차익거래를 하는 다양한 패턴, 거래소의 거래 효율성을 높이기 위한 데이터 조작 등을 잘 보여주었다. 그러나 가상화폐 2 차 시장의 짧은 발전 기간으로 인해 패널 데이터를 사용하는 연구는 매우 드물고, 김치 프리미엄 현상이 가상화폐 거래소에 미치는 부정적인 영향에 대한 연구도 많지 않다. 이러한 부분은 우리의 연구에서 주목할 부분이다. 우리는 가상화폐 시장에서의 김치 프리미엄 현상과 거래소의 거래 효율성에 대한 새로운 관점을 제시하고, 이를 통해 시장의 이해를 더욱 깊이 있게 파헤칠 것이다.

우리의 연구는 거래소한테 운영규모를 증가시키는 방향을 제공하고, 일부 리스크를 회피하도록 돕는 동시에 개인 투자자한테 유용한 투자 정보 참고 방향을 제시하고, 더 효율적인 투자 계획을 세워줄 수 있다.

II. 선행연구

비트코인은 가격 편차가 크며, 이는 ID 가 필요 없는 시장 환경과 관련이 있다(Pieters, 2017). 이에 대한 상세한 연구를 통해 거래소 간의 차익 실현 방법을 분석하고, 이론적으로 차익 실현이 가능하다는 것을 보여주었지만, 실제 운영 시간이 오래 걸리고 개인의 외환 전환이 불편하다는 점을 지적했다(Makarov, 2020). 더욱이, 데이터 세척이 심한 거래소의 경우 데이터 거래량이 25~50 배 과대되었으며, 데이터 공개가 미비한 거래소는 거래 데이터를 1.2533 배 확대한 것으로 나타났다(Pennec, 2021). 이는 거래소 정보의 신뢰성이 감소하고 투자 위험이 증가하는 결과를 보인다. 그 외에도 거래소에는 고위험 요소가 많다. Chohan(2018)의 연구에서 거래소에 나타난 가상화폐 도난 및 거래소 폐쇄 등의 정보가 상세히 기록되어 있음을 시사하며, 이는 가상화폐 시장의 특성과 거래소의 운영에 대한 중요한 인사이트를 제공한다.

김치 프리미엄 현상과 가상화폐 거래소의 기본적인 운영 문제를 확인한 후, 관련 문헌을 대량으로 수집하여 분류 처리하였다. 분류 내용에 따라 연구 모델 구성을 수행했다.

III. 연구 모델

김치 프리미엄 현상 및 가상화폐 거래소에 영향을 미치는 요인을 심도 있게 탐색하고, 김치 프리미엄 현상과 가상화폐 거래소의 관계를 파헤치기 위해 우리가 다음과 같은 모형을 만들었다.

<그림 1> 연구 모델

연구 모델은 <그림 1>과 같이 5 개의 독립 변수와 1 개의 매개 변수를 설정하고 이 모델을 기반으로 데이터를 정리하고 분석한다.

IV. 참고문헌

Chohan U W. The problems of cryptocurrency thefts and exchange shutdowns[J]. Available at SSRN 3131702, 2018.

Le Pennec G, Fiedler I, Ante L. Wash trading at cryptocurrency exchanges[J]. Finance Research Letters, 2021, 43: 101982.

Makarov I, Schoar A. Trading and arbitrage in cryptocurrency markets[J]. Journal of Financial Economics, 2020, 135(2): 293-319.

Pieters G, Vivanco S. Financial regulations and price inconsistencies across Bitcoin markets[J]. Information Economics and Policy, 2017, 39: 1-14.

Session C

C4. [학술세션]
인공지능을 활용한 지식 기반 질의응답과 사용자 경험 최적화

좌 장: 양 성 병(경희대)

라이브커머스 카테고리별 소비자 행태 분석: 실시간 채팅을 활용한 탐색적 연구

강은경
경희대학교 빅데이터응용학과
luckiness1@khu.ac.kr

김진수
경희대학교 빅데이터응용학과
wlstn010203@khu.ac.kr

장하렴
경희대학교 빅데이터응용학과
gkfua00@khu.ac.kr

양성병
경희대학교 빅데이터응용학과
sbyang@khu.ac.kr

Abstract *– 최근 라이브커머스가 급속히 성장하며, 다양한 모바일 디바이스를 활용해 시간적, 공간적 제약을 극복하고 '실시간 채팅'을 통해 고객과의 상호작용을 강화하는 새로운 소비 경향이 나타나고 있다. 이에, 본 연구에서는 라이브커머스 채널의 모든 카테고리에서 소비자들의 실시간 채팅데이터를 수집하고, '이벤트성 채팅' 중에서 '구매인증' 이 많이 나타나는 카테고리(가구/인테리어, 도서, 디지털/가전, 생활/건강, 생활/편의, 스포츠/레저, 식품, 출산/육아, 화장품/미용)와 많이 나타나지 않는 카테고리(패션의류, 패션잡화)에서의 소비자 반응을 비교분석하고 발생원인을 파악하고자 하였다. 구체적으로, (1) KNU 감성분석을 실시하여 감성점수를 도출한 후, 두 그룹 간 감성점수의 양상을 비교하고, (2) 토픽모델링(예: LDA, g-DMR)을 활용하여 소비자 채팅 데이터의 주요 토픽을 식별하고 토픽의 주제변화 양상을 확인하였다. 이후, (3) 머신러닝을 활용하여 변수의 상대적 중요도를 파악하여, 라이브커머스 맥락에서 나타나는 소비자의 태도 및 행동양상에 미치는 영향을 확인하고자 하였다. 본 연구를 통해 라이브커머스 맥락에서의 소비자 행동에 관한 학술적인 시사점을 도출하고, 라이브커머스 플랫폼 운영자 및 마케터에게 카테고리별 맞춤화된 마케팅 전략을 수립하는데 일조하고자 한다.*

Key Terms *– 라이브커머스, 머신러닝, 상대적 중요도 분석, 실시간 채팅, g-DMR 토픽모델링*

RGQA : Reasoning Guideline을 활용한 LLM 기반 KGQA

남정재
연세대학교 산업공학과
njj97@yonsei.ac.kr

이태훈
연세대학교 디지털애널리틱스학과
xogns5037@yonsei.ac.kr

왕수연
연세대학교 산업공학과
suyeon00@yonsei.ac.kr

김우주
연세대학교 산업공학과
wkim@yonsei.ac.kr

Abstract *– 최근 거대 언어 모델은 자연어 처리 분야에서 주목받고 있으나, 거대 언어 모델의 주요한 문제인 환각 문제는 여전히 발생한다. 이러한 문제를 완화시키는 방법 중 외부 지식을 검색 및 증강을 통해 환각을 줄여 답을 생성할 수 있는 RAG 프레임워크가 존재한다. 특히, 우리는 이러한 RAG 프레임워크의 프로세스를 기반으로 지식 그래프를 활용하여 KGQA 작업을 해결하는 방법을 연구한다. 기존 KGQA 연구 중에는 추론 경로를 검색한 뒤, 질문에 대한 답을 구하기 위해 추론 경로 정보를 증강시켜주는 방법이 존재한다. 이러한 기존의 방법은 이웃 relation을 질문 정보만을 갖고 추론 경로를 찾아가는 방식이다. 이는 질문에 대한 넓은 관점 기반의 추론이 부족하기에 추론 경로 구성에 있어 신뢰성이 떨어질 수 있다. 이에 대응하여, 우리는 질문에 대해 넓은 관점을 기반으로 추론할 수 있도록 참고할 Reasoning Guideline을 생성하여 추론 경로를 찾아가는 LLM 기반의 추론 프레임워크 RGQA를 제안한다. 이를 위해 추론 경로를 구성하는 검색 프로세스에 Reasoning Guideline 과 Stacked Path 정보를 증강하여 질문에 대한 추론을 위해 넓은 관점에서의 정보를 제공한다. 또한 텍스트 임베딩 유사도 기반으로 질문에 대한 답을 하기에 유의미한 경로만을 선별한다. 우리는 단계적 추론을 위한 정보와 효과적인 경로 검색으로 구성된 RGQA의 성능 및 확장가능성을 확인한다. 이를 통해 Reasoning Guideline을 활용하여 KGQA에 대한 성능 향상을 입증하였다.*

Key Terms *– 거대 언어 모델, 검색 증강 생성, 거대 언어 모델 프롬프팅, 지식 그래프 질의 응답, 추론 가이드라인*

Acknowledgement (사사)
본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음 (과제번호 ***RS-2023-00244769*****).**

본 연구는 국토교통부의 스마트시티 혁신인재육성 사업으로 수행되었음.

Knowledge-Augmented Prompting Noise 감소 전략 : Large Language Models을 활용한 KGQA의 성능 향상을 중심으로

정세영
연세대학교 산업공학과
tpdud3406@yonsei.ac.kr

김대식
연세대학교 산업공학과
kds8266@yonsei.ac.kr

김우영
연세대학교 산업공학과
timothy@yonsei.ac.kr

정세영
연세대학교 산업공학과
wkim@yonsei.ac.kr

Abstract *– 최근 대규모 언어 모델 (Large Language Models) 의 성능이 크게 향상됨에 따라 다양한 분야에서 생성형 AI 모델을 활용하여 텍스트를 처리하는 연구에 대한 관심이 높아지고 있다. 하지만 이러한 대규모 언어 모델은 텍스트 처리 능력은 우수하지만 검증되지 않은 지식을 생성하여 부정확한 답을 응답하는 환각 (hallucination) 현상이 나타날 수 있다. 대규모 언어 모델의 답변 성능을 향상시키고 환각 현상을 줄이기 위해 최근에는 지식 그래프를 활용하여 대규모 언어 모델의 프롬프트로 활용하는 Knowledge Graph Question Answering 연구가 보고되었다. 이에 본 연구는 Knowledge Graph Question Answering의 답변 성능을 높이기 위해 Retrieval Augmented Generation 시스템을 활용한다. 우리는 전통적인 Retrieval Augmented Generation 시스템에 필터링 과정을 더하여 검색-필터링-생성 구조를 제안한다. 우리가 제안한 필터링 방법은 검색된 지식 중에서 실제로 질문과 관련이 없는 Noise 를 제거하는 방식이다. 우리는 필터링을 통해 Noise가 제거된 지식을 대규모 언어 모델의 프롬프트로 활용함으로써 답변 성능을 향상시켰다. 결과적으로 추가적인 미세조정 없이 검색된 지식을 적절히 선별함으로 KGQA의 답변 성능이 향상될 수 있음을 입증하였다.*

Key Terms *– 대규모 언어 모델, 검색 증강 생성, 지식 그래프, 지식 그래프 질의 응답, 생성형 AI*

Acknowledgement (사사)
본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음(과제번호 ***RS-2023-00244769).***
본 연구는 국토교통부의 스마트시티 혁신인재 육성사업으로 수행된 연구임

인공지능 법령 질의응답시스템을 위한 규범 기반 사례 데이터 생성과 활용에 관한 연구

박예린
경희대학교
빅데이터응용학과
yyam1020@khu.ac.kr

박가림
이화여자대학교
법학전문대학원
garim@ewha.ac.kr

이유봉
한국법제연구원
미래법제본부
ebearth@klri.re.kr

손권상
인하대학교
경영학과
ksohn@inha.ac.kr

***Abstract** – 법령정보는 국민의 일상과 밀접하게 연관되어 있지만, 이해하기 어려운 용어와 판단의 전문성이 요구된다. 특히, 위법행위에 대한 규제를 다루는 법령의 경우 규제 대상자의 법 이해도에 따라 판단의 기준이 상이하게 나타날 수 있다. 이에 법령 접근 편의성을 높이기 위해 챗봇을 비롯한 질의응답 형식의 다양한 서비스가 개발되고 있다. 본 연구에서는 2022 년 5 월부터 시행된 공직자의 이해충돌방지법을 대상으로 규제위반 여부를 질의를 통해 사전에 판단할 수 있는 법령 질의응답시스템 개발을 위한 사례 데이터 생성과 활용에 대한 방법론을 제시하고자 한다. 이를 위해 이해충돌방지법의 법조문 분석을 통해 규범 패턴을 도출하여 가상의 사례 약 31 만 건을 생성하였으며, 해당 데이터를 학습 데이터로 활용하여 법 위반 여부를 판단하는 BERT 기반 분류 모델을 구축하여 이해충돌방지법 위반 여부에 대한 분류 정확도를 검증하였다.*

***Keywords** – 법령질의응답시스템, 이해충돌방지법, AI Compliance, KoBERT*

I. 서론

법령 정보는 사회의 기반 구조를 이루며, 일상에서 겪는 다양한 문제와 직접적으로 연관되어 있다. 특히, 특정 행위에 대한 규제는 위반할 경우 법적 처벌의 위험이 따르기에 이를 사전에 인지하고 규범을 준수하는 것이 요구된다. 그러나 법조문의 복잡성과 전문적인 용어는 일반 국민이 쉽게 접근하고 이해하는 데에 장벽으로 작용해왔다. 이러한 문제를 해결하기 위해 최근에는 인공지능 기술을 활용한 법령 서비스 개발되고 있으며, 이를 통해 사용자가 법령에 대해 쉽고 빠르게 접근하며 법조문에 대한 이해도 역시 높일 수 있다 (김나리&김형중, 2017; 김지현 외, 2012).

특히 공직자의 이해충돌방지법과 같이 복잡한 규제를 포함하는 법령은 더욱 철저한 이해가 필요하며, 규제 위반의 판단 기준 또한 사용자의 법적 이해도에 따라 달라질 수 있다(정신교, 2021). 이 법은 공직자가 직무 수행 중 개인적 이해관계를 우선하지 않도록 규정하고 있으며, 해당 법의 정확한 이해는 공정한 법 집행에 있어 필수적이다. 이러한 문제를 해결하기 위해, 본 연구는 이해충돌방지법 법조문의 규범적 패턴을 분석하고, 이를 바탕으로 대규모의 사례 데이터를 생성하며, BERT 기반 분류 모델을 통해 법 위반 여부를 판단하는 방법론을 검증하였다. 이를 통해 인공지능 기반 법령 질의응답시스템의 효율적인 개발 방향을 제시하고, 복잡한 법령을 일반 시민들에게 보다 쉽게 설명하고 법적 문제에 대한 신속한 판단을 가능하게 하는 기반을 마련하고자 한다.

II. 데이터셋

본 연구에서는 이해충돌방지법에서 정의하는 공직자의 규제 위바 여부 판단을 자동으로 추론하기 위해 법조문을 기반으로 주제어사전을 구축하였다. 주제어사전은 조문에서 추출한 법적용대상자, 사적이해관계자, 의무규범, 금지규범, 허용규범 등 구성요소를 포함하며, 규범의 성격에 따른 정보제공이 가능하다.

다음으로 이해충돌방지법 해석에 필요한 공직자 행위 관련 업무에 대한 법령을 수집하기 위해 동법 제 5 조 제 1 항 제 1 호 내지 제 16 호 및 동법 시행령 제 4 조에 열거된 사적이해관계자의 직무관련성 판단에 기초가 되는 직무행위에 대한 키워드를 설정하였다. 그리고 이를 대상으로 법률전문가의 검토를 거쳐 행위 관련 키워드 93 개를 선정하여 국가법령정보센터를 통해 데이터를 수집하였다. 수집된 데이터는 93 개 키워드 기준 13,797 개의

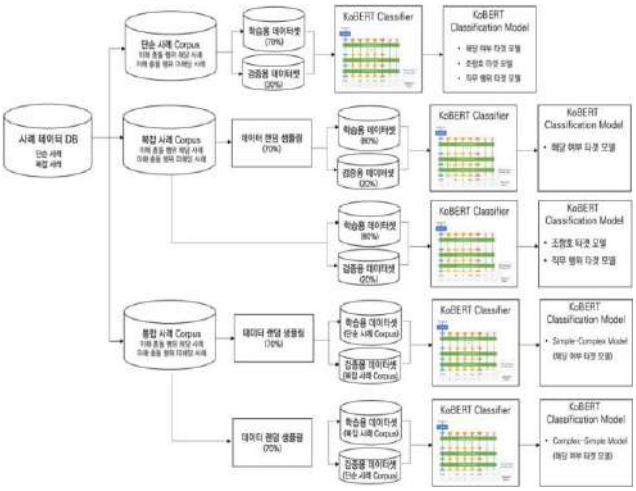
법령 데이터와 33,666 개의 조문 데이터로 구성되었다.

III. 사례 생성

이해충돌방지법에서 가장 중요한 해석 쟁점은 사적이해관계자 해당 여부이다. 따라서 사례데이터 생성의 목적을 사적이해관계자 해당 여부로 설정하였으며, 앞서 수집한 법령 데이터 및 행위 관련 키워드를 기반으로 사례 데이터를 생성하였다. 사례 생성 방법은 사적이해관계자 해당 여부 및 법령에 의한 구성요건 요소를 변수로 설정하여 구조화된 문장의 형태로 생성하였다. 구조화된 사례의 다양성을 포함하고자 문장의 길이와 변수의 구성에 따라 2 문장 이내의 단순사례와 4 문장 이상의 복잡사례로 생성하였다.

IV. 분석 모델링

생성된 사례 데이터를 대상으로 단순 사례 코퍼스 기반의 분류 모델 3 개와 복잡 사례 코퍼스 기반의 분류 모델 3 개, 통합 코퍼스 기반의 분류 모델 2 개 등 총 8 개의 모델을 구축하였으며, 데이터셋은 학습용과 검증용 각 8:2 비율로 분류하여 PYTORCH-KOBERT-MODEL 을 사용하였다<그림 1>.



<그림 1> 사례 데이터 분류모델 프레임워크

III. 참고문헌

김나리, 김형중. (2017). 연관법령 검색을 위한 워드 임베딩 기반 Law2Vec 모형 연구. *디지털콘텐츠학회 논문지*, 18(7), 1419-1425.

김지현, 이종서, 이명진, 김우주, 홍준석. (2012). 법령정보 검색을 위한 생활용어와 법률용어 간의 대응관계 탐색 방법론. *지능정보연구*, 18(3), 137-152.

정신교. (2021). 공직자의 부패방지를 위한 이해충돌방지법의 실효성 확보방안. *한국부패학회보*, 26(3), 27-47.

Lee, J. H., Kim, B. R., & Park, M. Y. (1994). Knowledge-based model management systems. Decision Support System, 3(2), 65-82.

Session D

D1. [학술세션]
AI기반 추천시스템

좌 장: 고 준(전남대)

Distributional Reinforcement Learning for Recommender System Addressing Randomness in Returns

류명진
이화여자대학교 빅데이터분석학협동과정
ryumyoungjin@ewhain.net

강윤철
이화여자대학교 경영대학 경영학과
yckang@ewha.ac.kr

Abstract –
Research in the domain of recommender systems has achieved significant advancements, from traditional methods such as content-based filtering, collaborative filtering, and matrix factorization to sophisticated deep learning-based frameworks. Nonetheless, these conventional approaches often fall short in capturing dynamic interactions between users and systems, as well as adequately accounting for long-term rewards. To address these limitations, there has been an increasing focus on the exploration of deep reinforcement learning (DRL) based recommender systems. The application of DRL is particularly promising within this context, as it enables an agent to adaptively learn from users' changing preferences through real-time feedback. Among various DRL techniques, this study is distinguished by its integration of the actor-critic method with distributional reinforcement learning, specifically through the implementation of Quantile Regression Deep Q-Networks (QR-DQN). An empirical investigation, utilizing real-world datasets in an offline environment, demonstrates that this hybrid approaches outperforms other multi-critic strategies in performance. This enhancement underscores the potential of combining actor-critic frameworks with distributional reinforcement learning principles to refine the efficacy of recommender systems.

Keywords – Deep Reinforcement Learning, Distributional Learning, Recommender System

I. Introduction

Recommender systems have become a crucial part of today's digital landscape, aiding users in finding goods (*e.g.*, products, news, music, services). Over time, these systems have undergone substantial development, moving from classic approaches such as content-based filtering (Lops, P. et al., 2011), collaborative filtering (Schafer, J. B. et al., 2007), hybrid method (Burke, R.,2002), matrix factorization (Koren, Y. et al., 2009), logistic regression (Y. Wang et al., 2016) to advanced deep learning algorithms (Zhang et al., 2019).

However, these conventional methods are typically static in that they often rely heavily on historical interaction to make recommendations. In other words, they may not capture the latest changes in user preferences unless the model is frequently retrained, which can be computationally expensive and time-consuming. The lack of real-time adaptation may lead to continuous recommendation based on outdated preferences, failing to adapt to the user's current interests.

Furthermore, such mentioned approaches tend to discard the long-term rewards since they are trained to maximize the immediate rewards. Focusing merely on instant rewards may result in neglection of long-term effects from the recommendations.

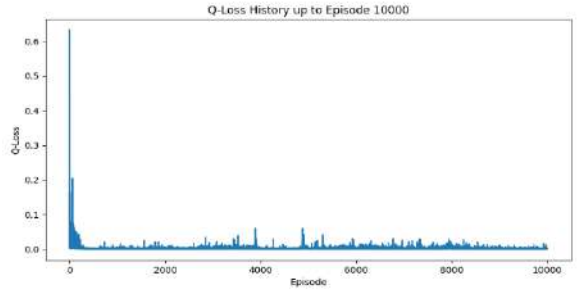
In an attempt to overcome these limitations, deep reinforcement learning (DRL) (K. Arulkumaran et al., 2017) algorithms have been widely utilized in recommender system. A DRL based recommender system typically involves three main stages: (1) State representation (Liu, F. et al., 2020), (2) Reward/cost Formulation, (3) Environment setup. State representation module is often used to better model the complicated interactions between users and items which eventually improves recommendation performance. This module uses multi-layer networks to capture user-item interactions based on historical data, generating a continuous representation of user behavior. Furthermore, the reward/cost signal provided by the environment shows the performance of agent in selecting actions. Thus, designing an adequate reward/cost function is required in DRL based recommender system. An environment can be divided into two settings: offline and online. In an offline approach, the environment consists of a static dataset containing user information. Usually, in an offline setting, the trained model is used to evaluate recommendation performance with the learned policy. Conversely, in online methodologies, algorithms undergo testing through real-

time interactions with actual users, which can be accurate but more costly.

In our study, we combined Deep Deterministic Policy Gradient (DDPG) (Lillicrap, T. P. et al.. 2015) algorithm with Distributional Reinforcement Learning, especially Quantile Regression Deep Q-Networks (QR-DQN) (Dabney et al., 2018). In the existing DDPG algorithm the critic network calculates the state-action value function, also known as the Q-function. DDPG algorithm efficiently merges the benefits of policy gradient methods (Sutton, R. S. et al., 1999) and deep Q-networks (Mnih, V. et al., 2013) by employing the critic network for Q value estimation and the actor network for determining the optimal actions. In this study, we modified the critic network of existing DDPG algorithm by adopting QR-DQN architecture. Instead of estimating a q-value which is a single scalar value, it estimates a q-value distribution. In addition, when calculating a Temporal Difference (TD) Error, instead of using MSE loss function, it uses Quantile Huber Loss function.

II. Methods and Results

In this study, we utilized the MovieLens 1M dataset, allocating 80% as training set and 20% as test set. The reward is set to 1 for recommended items with a rating greater than 3. In order to prevent recommending repeated items, any once-recommended movies are removed from the candidate set. After training our proposed model over 10,000 episodes, we observed that the q loss converged to a value between 0 and 0.1 as shown in Fig.1.



<Fig. 1> Q Loss History up to 10000 episode

For comparison, we employed both multi-critic models and a random model. The model with two critics has an additional critic network that imposes a cost if a rating for recommended items is 2 or lower, and the user's stickiness falls within the bottom 25%. Furthermore, the model with three critics additionally incorporates a critic network that measures another cost based on the popularity of the movies. Specifically, the cost is imposed when the rating for recommended items is lower than 3, and the popularity of the items falls within the top 10%. Lastly, random model without any training process is also used for a comparison.

Precision, NDCG, total rewards, and total costs are used as evaluation metrics. As shown in the Table 1 and 2, it can be observed that the model proposed in this study outperforms the multi-critic models and the random model in all aspects.

<Table 1>
Performance Comparison on Top 5 Recommended Items

	Precision	NDCG	Reward	Cost 1	Cost 2
QR-DDPG	0.5106	0.4981	3084		
2 Critic	0.4502	0.4464	2719	120	
3 Critic	0.4177	0.4104	2523	36	71
Random	0.0323	0.0308	195		

<Table 2>
Performance Comparison on Top 10 Recommended Items

	Precision	NDCG	Reward	Cost 1	Cost 2
QR-DDPG	0.4755	0.4597	5744		
2 Critic	0.4382	0.429	5293	227	
3 Critic	0.4012	0.3943	4846	78	136
Random	0.0327	0.0338	395		

III. References

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, *12*, 331-370.

Dabney, W., Rowland, M., Bellemare, M., & Munos, R. (2018, April). Distributional reinforcement learning with quantile regression. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).

K. Arulkumaran, M. P. Deisenroth, M. Brundage and A. A. Bharath, "Deep Reinforcement Learning: A Brief Survey," in IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 26-38, Nov. 2017, doi: 10.1109/MSP.2017.2743240.

Konda, V., & Tsitsiklis, J. (1999). Actor-critic algorithms. *Advances in neural information processing systems*, *12*.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, *42*(8), 30-37.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Wierstra, D. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*..

Liu, F., Tang, R., Li, X., Zhang, W., Ye, Y., Chen, H., ... & Zhang, Y. (2018). Deep reinforcement learning based recommendation with explicit user-item interactions modeling. *arXiv preprint arXiv:1810.12027*.

Liu, F., Tang, R., Li, X., Zhang, W., Ye, Y., Chen, H., ... & He, X. (2020). State representation modeling for deep reinforcement learning based recommendation. *Knowledge-Based Systems*, *205*, 106170.

Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. *Recommender systems handbook*, 73-105.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web: methods and strategies of web personalization* (pp. 291-324). Berlin, Heidelberg: Springer Berlin Heidelberg.

Schaul, T., Quan, J., Antonoglou, I., & Silver, D. (2015). Prioritized experience replay. *arXiv preprint arXiv:1511.05952*.

Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems*, *12*.

Wu, J., Wang, R., Li, R., Zhang, H., & Hu, X. (2018, October). Multi-critic DDPG method and double experience replay. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 165-171). IEEE.

Y. Wang, D. Feng, D. Li, X. Chen, Y. Zhao and X. Niu, (2016) "A mobile recommendation system based on logistic regression and Gradient Boosting Decision Trees," *International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, 2016, pp. 1896-1902.

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM computing surveys (CSUR)*, *52*(1), 1-38.

The effect of applying Aristotle's rhetoric on persuasive cues in AI's travel recommendation on user trust

Hanna Jeong
Kyung Hee University
jeonggreen6@gmail.com

Seunghun Shin
The Hong Kong Polytechnic University
Seung-hun.shin@polyu.edu.hk

Chulmo Ckoo
Kyung Hee University
helmetgu@khu.ac.kr

Abstract – The emergence of Large Language Models (LLMs) has elevated the importance of ChatGPT (Chat Generative Pre-trained Transformer) and its application in travel planning recommendations. Furthermore, user-friendly interfaces and efficient data processing enhance users' decision-making processes. However, ChatGPT's responses vary based on linguistic cues and question formats, containing a multitude of possibilities. This study aims to interpret AI's persuasive cues from the perspective of Aristotle's rhetorical appeals (Ethos, Pathos, Logos). Additionally, it seeks to examine whether customer user competence moderates the effect of persuasive cues on trust in travel recommendations. The study not only investigates the effective factors of AI persuasive cues but also anticipate that it will be possible to examine the underlying mechanisms of the user's persuasion process in trusting AI.

Keywords – Aristotle's rhetorical appeals, AI persuasive cues, trip planning recommendation, trust, user competence

1. Introduction

The emergence of ChatGPT has made trip planning more convenient in the travel industry by providing users with novel travel information (Kim, Kim, Kim & Park, 2023). Users can now diversify their travel plans by receiving travel recommendations from AI. However, due to the nature of LLMs, ChatGPT tends to provide abstract responses based on context (Sundar, 2023), and user competence in how questions are posed may influence the responses (Callum, 2023). Previous studies have explored recommendation techniques for LLMs (Kirshner, 2024), as well as factors affecting option reduction in travel planning and decision making (Shin, Kim, Lee, Yhee & Koo, 2023). Studies have also examined factors influencing perceived trust in personalized travel recommendations (Ali, Yasar, Ali & Dogan, 2023). However, considering that AI's responses vary based on user competence (Azaria, Azoulay & Reches, 2023), research on cues for travel destination recommendations should be approached from various perspectives. Additionally, Ali et al. (2023) highlighted the need for future research on factors influencing trust in AI-based travel recommendations. Therefore, this study aims to approach trust in travel recommendations from a different perspective by constructing persuasive cues inspired by Aristotle's rhetoric and applying them to AI-powered travel recommendations. Benjamin (1997) stated that persuasion is pervasive in society, while Aristotle defined rhetoric as the ability to see the available means of persuasion, distinguishing three means: logos (logical appeal), pathos (emotional appeal), and ethos (credibility) (Borchers, 2005). Building on this, the study extends persuasive cues from rhetoric to the context of AI, assuming that persuasive cues will positively influence user trust in travel recommendations. Since AI is influenced by user expertise (Azaria et al., 2023), the study assumes that user competence will moderate the effect of recommendations based on user competence type. The goal of this study is to propose tourism-oriented persuasive cues for trust in travel recommendations by applying ChatGPT's persuasive cues from an Aristotelian rhetorical perspective, which is meaningful in

attempting to adopt rhetorical concepts in travel recommendations.

2. Literature Review

2.1. Aristotle's rhetoric theory

Aristotle described rhetoric as the art of persuasion and as a means of engaging in good conversation with others. Murphy (1981) recognized three fundamental components of persuasion: ethos, which refers to credibility; pathos, which involves appealing to emotions; and logos, which pertains to the logical form of an argument. Fogg (1998) asserted that computers have the ability to function as persuasive devices. Research has been conducted on the application of rhetorical appeals in internet e-commerce design and interface-related studies, with Winn & Beck (2002) suggesting that web design elements inherently embody basic persuasive elements to facilitate consumer purchases, particularly when objects are credible, emotionally impactful, and factual. The application of rhetoric has extended to social communication, as evidenced by Bronstein (2013), who noted that the application of rhetorical appeals in online communication during presidential campaigns led to significantly increased support. Additionally, Aristotle's appeals are more persuasive when used together rather than individually, with emotional and credible appeals having a greater impact than appeals to logic (Chu, Deng & Chung, 2014). Thus, Aristotle's rhetoric has been applied over time to various media. Hence, the objective of this study is to examine whether the three appeals of Aristotle (Ethos, Pathos, Logos) also have an effect on user trust as persuasive cues in AI.

2.2. Persuasive cue adoption in AI

Aristotle's rhetoric is comprised of three fundamental components: ethos, which pertains to credibility; pathos, which involves evoking emotions; and logos, which refers to the logical structure. These parts have been expounded upon by Bronstein (2013). This study has reinterpreted these characteristics as ethos (credibility), pathos (social presence), and logos (information quality) for AI persuasive cues. Credibility can be described as the quality of being believable or trustworthy (Fogg & Tseng, 1999). Scholars have explored dimensions of credibility such as trustworthiness, expertise, dynamism, goodwill, likeability, and attractiveness (Choi & Stvilia, 2015; Ohanian, 1990). Fogg & Tseng (1999) propose that trustworthiness and expertise are fundamental elements of perceived credibility, hence establishing the dimensions of credibility in the context of AI as a) Trustworthiness, and b) Expertise. Pathos, in the context of AI, can be utilized as social presence, which pertains to the perception of being in the company of another during human-computer interaction (HCI) (Gefen & Straub, 2003). Pathos, which refers to a sense of human connection and social interaction, can be relevant to AI. Logos, interpreted as the logical discourse used to convey a message, is rational appeal to recipients (Buttle, 1998; Xun & Reynolds, 2010). In AI, logos is applied as information quality, which includes accuracy, format, completeness, and currency of information produced by digital technologies (Setia, 2013). The study examines four elements of information quality: relevancy, reliability, conciseness, and quickness. It specifically considers ChatGPT's tendency to prioritize quickness above timeliness, as highlighted by Niu and Mvondo (2024). Therefore, in this study, we aim to construct the variables of persuasive cues when applied to AI as Credibility, Social presence, and Information quality. Also We intend to adopt the dimensions of each variable and conduct a survey afterward

2.2. User competence

Due to the progress of artificial intelligence (AI), ChatGPT is capable of offering a wide range of responses that are influenced by the structure of user queries (Sundar, 2023). User competency, as described by Marcolin, Compeau, Munro & Huff (2000), refers to the ability of users to effectively employ technology in accordance with their job

tasks. User competence is classified into three categories: breadth of personal knowledge, depth of knowledge, and finesse in creatively applying technology (Munro, Huff, Marcolin & Compeau, 1997). Furthermore, when it comes to end-user computing competence, competence is classified into three categories based on user characteristics, usage characteristics, and application characteristics (Blili, Raymond & Rivard, 1998). Moreover, the competence of users is linked to their level of experience, skill, and knowledge in utilizing smartphones, which in turn leads to user pleasure in utilizing the diverse range of multifunctional features and information available on smartphones (Koo, Chung & Kim, 2015). Thus, user competence has been applied to various technological contexts throughout different eras.

2.3. Trust formation in AI

According to Lewicki & Wiethoff (2000), trust is the belief and willingness of an individual to take action based on the words, actions, and decisions of another person. Trust is the belief a person has in the actions of others, which is influenced by their expectations (Gu, Lee & Suh, 2009). Nevertheless, as technology has progressed, trust has become an essential factor in the interaction between technology and individuals, playing a critical role in the adoption of technology (Johnson, Bardhi & Dunn, 2008). Specifically, when users depend on AI for particular tasks, this dependence can also be regarded as trust (Lankton, Mcknight & Thatcher, 2014). Therefore, studies have been carried out to investigate the way in which humans perceive and embrace technology, specifically recommendation agents in the field of technology (Lankton, et al, 2014). Therefore, the extent to which users accept information based on the plan provided by AI can be regarded as trust in AI (Ali et al, 2023). Therefore, this study aims to investigate whether applying Aristotle's rhetoric to AI as persuasive cues influences trust in travel recommendations.

2.4. Research Model and Hypotheses Development

This study has utilized Aristotle's rhetoric, a foundational framework for persuasive communication (Murphy, 1981), to present persuasive cues. The components of rhetoric encompass Ethos (credibility), Pathos (social presence), and Logos (information quality), which have been employed in the context of AI. The purpose of this study is to investigate the impact of these persuasive cues on trust in travel recommendation. Additionally, the study seeks to investigate whether user competence, categorized by user type (i.e., customer vs. supplier), moderates the relationship between persuasive cues and trust. Based on these discussions, the following hypotheses are proposed:

H1: Credibility has a positive influence on trust in travel recommendation.
H2: Social Presence has a positive influence on trust in travel recommendation.
H3: Information Quality has a positive influence on trust in travel recommendation.
H4a: The relationship between credibility and trust in travel recommendation will be moderated by User Competence.
H4b: The relationship between social presence and trust in travel recommendation will be moderated by User Competence.
H4c: The relationship between information quality and trust in travel recommendation will be moderated by User Competence.

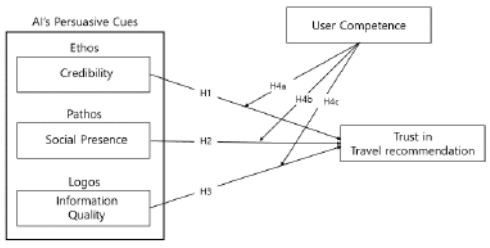


Figure1. Reserch Model

3. Methodology

This study will utilize Partial Least Squares-based Structural Equation Modeling (PLS-SEM) to evaluate the influence of Aristotle's persuasive cues on Trust in travel recommendation. This method is recommended by Hair, Hult, Ringle, & Sarstedt (2017). The research will conduct a survey using the online survey platform Amazon Mturk. Additionally, the study will categorize users into two types: those seeking trip planning information, divided into customers receiving travel recommendations from AI, and those interacting with travel agencies. The aim is to examine whether different user competences among customers moderate the effects of persuasive cues on trust in travel recommendation. To achieve this, Multi-Group Analysis (MGA) will be employed. MGA is a non-parametric method that relies on permutation tests and Henseler's Bootstrap, making it one of the most conservative techniques for evaluating differences between two groups (Henseler, Ringle & Sarstedt, 2016). Furthermore, according to Barreto, Rubio & Campo (2020), there are cases where the validity of results was established through surveys using various types of stimuli. Therefore, this study will provide stimuli for different scenarios of travel recommendation conversations between customers and AI as well as travel agency staff, aiming to measure the moderating effects of competence through MGA.

4. Conclusion

This study is based on Aristotle's rhetoric and investigates the roles of three persuasive cues (i.e., credibility, social presence, information quality). It proposes a new perspective on how these cues influence trust when users utilize AI for travel recommendation. By examining whether persuasive cues are moderated by user competence in influencing trust in travel recommendation, this study aims to provide insights into the effect of user competence on travel recommendation. It is anticipated that this research will contribute to revealing the underlying mechanisms of the user's persuasion process in trusting AI.

References

Ali, F., Yasar, B., Ali, L., & Dogan, S. (2023). Antecedents and consequences of travelers' trust towards personalized travel recommendations offered by ChatGPT. *International Journal of Hospitality Management*, *114*, 103588.

Azaria, A., Azoulay, R., & Reches, S. (2023). ChatGPT is a Remarkable Tool—For Experts. *Data Intelligence*, 1-49. https://doi.org/10.1162/dint_a_00235

Blili, S., Raymond, L., & Rivard, S. (1998). Impact of task uncertainty, end-user involvement, and competence on the success of end-user computing. *Information & Management*, *33*(3), 137-153.

Borchers, T. (2012). *Persuasion in the media age*. Waveland Press.

Bronstein, J. (2013). Like me! Analyzing the 2012 presidential candidates' Facebook pages. *Online Information Review*, *37*(2), 173-192.

Buttle, F. A. (1998). Word of mouth: Understanding and managing referral marketing. *Journal of Strategic Marketing*, *6*(3), 241-254.

Callum Bains. (2023, July 29). *AI prompt engineering: learn how not to ask a chatbot a silly question*. The Guarian. https://www.theguardian.com/technology/2023/jul/29/ai-prompt-engineering-chatbot-questions-art-writing-dalle-midjourney-chatgpt-bard

Choi, W., & Stvilia, B. (2015). Web credibility assessment: Conceptualization, operationalization, variability, and models. *Journal of the Association for Information Science and Technology*, *66*(12), 2399-2414.

Chu, H.-L., Deng, Y.-S., & Chuang, M.-C. (2014). Investigating the persuasiveness of e-commerce product pages within a rhetorical perspective. *International Journal of Business and Management*, *9*(4), 31.

Fogg, B. J. (1998, April). Captology: the study of computers as persuasive technologies. In *CHI 98 Conference Summary on Human Factors in Computing Systems* (p. 385).

Fogg, B. J., & Tseng, H. (1999, May). The elements of computer credibility. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 80-87).

Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What makes users trust a chatbot for customer service? An exploratory interview study. In *Internet Science: 5th International Conference, INSCI 2018, St. Petersburg, Russia, October 24–26, 2018, Proceedings*

5 (pp. 194-208). Springer International Publishing.

Gefen, D., & Straub, D. (2003). Managing user trust in B2C e-services. *e-Service*, *2*(2), 7-24.

Gu, J.-C., Lee, S.-C., & Suh, Y.-H. (2009). Determinants of behavioral intention to mobile banking. *Expert Systems with Applications*, *36*(9), 11605-11616.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017). Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science*, *45*, 616-632.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International marketing review*, *33*(3), 405-431.

Jiménez-Barreto, J., Rubio, N., & Campo, S. (2020). Destination brand authenticity: What an experiential simulacrum! A multigroup analysis of its antecedents and outcomes through official online platforms. *Tourism Management*, *77*, 104022.

Johnson, D. S., Bardhi, F., & Dunn, D. T. (2008). Understanding how technology paradoxes affect customer satisfaction with self-service technology: The role of performance ambiguity and trust in technology. *Psychology & Marketing*, *25*(5), 416-443.

Khan, A. W., & Mishra, A. (2024). AI credibility and consumer-AI experiences: a conceptual framework. *Journal of Service Theory and Practice*, *34*(1), 66-97.

Kim, H., & Niehm, L. S. (2009). The impact of website quality on information quality, value, and loyalty intentions in apparel retailing. *Journal of Interactive Marketing*, *23*(3), 221-233.

Kim, J., Kim, J. H., Kim, C., & Park, J. (2023). Decisions with ChatGPT: Reexamining choice overload in ChatGPT recommendations. *Journal of Retailing and Consumer Services*, *75*, 103494.

Kirshner, S. N. (2024). GPT and CLT: The impact of ChatGPT's level of abstraction on consumer recommendations. *Journal of Retailing and Consumer Services*, *76*, 103580.

Koo, C., Chung, N., & Kim, H.-W. (2015). Examining explorative and exploitative uses of smartphones: A user competence perspective. *Information Technology & People*, *28*(1), 133-162.

Lankton, N., McKnight, D. H., & Thatcher, J. B. (2014). Incorporating trust-in-technology into expectation disconfirmation theory. *The Journal of Strategic Information Systems*, *23*(2), 128-145.

Lewicki, R.J., Wiethoff, C., 2000. Trust, trust development, and trust repair. In: Deutsch, M., Coleman, P.T. (Eds.), The Handbook of conflict resolution: Theory and Practice. San Francisco, CA: Jossey-Bass, pp. 86–107.

Marcolin, B. L., Compeau, D. R., Munro, M. C., & Huff, S. L. (2000). Assessing user competence: Conceptualization and measurement. *Information Systems Research*, *11*(1), 37-60.

Munro, M. C., Huff, S. L., Marcolin, B. L., & Compeau, D. R. (1997). Understanding and measuring user competence. *Information & Management*, *33*(1), 45-57.

Murphy, J. J. (1981). *Rhetoric in the Middle Ages: A history of rhetorical theory from Saint Augustine to the Renaissance* (Vol. 277). Univ of California Press.

Niu, B., & Mvondo, G. F. N. (2024). I Am ChatGPT, the ultimate AI Chatbot! Investigating the determinants of users' loyalty and ethical usage concerns of ChatGPT. *Journal of Retailing and Consumer Services*, *76*, 103562.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, *19*(3), 39-52.

Sundar, S. (2023, Feb 7). ChatGPT 'may make up facts,' OpenAI's chief technology officer says. Business Insider. **https://www.businessinsider.com/chatgpt-may-make-upfacts-openai-cto-mira-murati-says-2023-2**

Setia, P., Setia, P., Venkatesh, V., & Joglekar, S. (2013). Leveraging digital technologies: How information quality leads to localized capabilities and customer service performance. *MIS Quarterly*, 565-590.

Shin, S., Kim, J., Lee, E., Yhee, Y., & Koo, C. (2023). ChatGPT for Trip Planning: The effect of narrowing down options. *Journal of Travel Research*, 00472875231214196.

Steinfield, C. W. (1986). Computer-mediated communication in an organizational setting: Explaining task-related and socioemotional uses. *Annals of the International Communication Association*, *9*(1), 777-804.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5-33.

Winn, W., & Beck, K. (2002). The persuasive power of design elements on an e-commerce web site. *Technical Communication*, *49*(1), 17-35.

Xun, J., & Reynolds, J. (2010). Applying netnography to market research: The case of the online forum. *Journal of Targeting, Measurement and Analysis for Marketing*, *18*, 17-31.

Yang, S.-B., Lee, H., Lee, K., & Koo, C. (2018). The application of Aristotle's rhetorical theory to the sharing economy: An empirical study of Airbnb. *Journal of Travel & Tourism Marketing*, *35*(7), 938-957.

딥러닝 추천시스템: Wide & Deep 모델의 적용

박준석
국민대학교 소프트웨어학부
20191271@kookmin.ac.kr

조현진
국민대학교 소프트웨어학부
mclub4@kookmin.ac.kr

최지훈
국민대학교 소프트웨어학부
jihunchoi @kookmin.ac.kr

이지윤
국민대학교 AI 디자인학과
lgy4494@kookmin.ac.kr

Abstract *– This study explores the implementation and applicability of the Wide & Deep model in a real-world operational environment, starting from a theoretical foundation, through data preparation, to a detailed analysis of the model's structure, learning methodology, and optimization techniques. Specifically, through the operational case study of the platform "e-Um," this research has demonstrated that the Wide & Deep model can accurately reflect users' preferences and behavioral patterns, providing more personalized recommendations. This finding illustrates the model's capability to achieve an effective balance between memorization and generalization within recommendation systems, showcasing superior performance. Based on the results presented in this study, it is expected that recommendation systems based on Wide & Deep algorithms will contribute to the expansion in various fields.*

Key Terms *– 추천 시스템, Wide & Deep 모델*

본 연구는 국민대학교 북악교육인성교육센터의 지원에 의해 이루어진 것임

한국지능정보시스템학회
Korea Intelligent Information Systems Society

Advancing Tourist Experience Through Personalized Menu Recommendations with Auto-Tagging

Kim, M. H., Park, B. H., Lee, H. J., & Lee, J. M. (2024)

Abstract

- This study introduces a novel method for enhancing culinary tourism experiences by providing personalized menu recommendations to travelers.
- Leveraging context-aware clustering and tag-based filtering on a dataset of over 30,000 restaurant reservation entries, we identified distinct dining preferences among users.
- By employing a multi-step recommendation process integrating cluster weights, tag similarity, and past reservation data, our approach significantly outperformed baseline models, achieving over a tenfold increase in recall and mean average precision metrics.
- Our model's focus on culinary tags enables more precise and contextually relevant recommendations compared to traditional menu-based approaches.

Firm’s cloud adoption decision according to environmental policy (ETS)

서정권
카이스트 기술경영학부
sjk9204@kaist.ac.kr

이철호
카이스트 기술경영학부
irontigerlee@kaist.ac.kr

이나래
카이스트 기술경영학부
naree.lee@kaist.ac.kr

Abstract – Amidst the prevailing global concern for environmental preservation due to Earth's pollution, nations worldwide are increasingly formulating environmental policies. Consequently, firms face heightened pressure to mitigate carbon emissions, resulting in additional environmental costs due to regulatory measures. Our research focuses on the decision of cloud adoption by firms as a strategy for carbon reduction, particularly following the implementation of the California Cap and Trade (ETS) regulation in 2012. Employing the Difference-in-Differences (DID) methodology, we meticulously control for various influencing factors to investigate this phenomenon. Our findings reveal a notable increase in cloud adoption among companies in California after the ETS, compared to those in other states. Moreover, this effect is accentuated within the green industry sector and among companies with higher GHG emissions. Through this research, we uncover novel factors influencing cloud adoption decisions, demonstrating the influence of ETS on corporate IT strategies. We provide evidence that environmental policies, such as ETS, influence companies to opt for cloud outsourcing as a means of reducing carbon emissions and mitigating environmental costs.

Keywords – California cap and trade(ETS), GHG emission, Cloud adoption, Green industry, Outsourcing

I. 서론

Currently, there is a heightened global interest in environmental issues. Since the Tokyo Protocol, governments worldwide have agreed to reduce carbon emissions, and more concrete targets and figures have emerged post the Paris Agreement. In 2021, Larry Fink, of the global investment firm BlackRock, declared that they would only invest in companies adhering to ESG (Environmental, Social, Governance) standards. As a result of such global agreements, regulations, and the stance of hedge funds, companies are increasingly compelled to consider the environment.

Our research focuses on corporate decisions regarding environmental factors, specifically the reduction of carbon emissions. In making these decisions, companies may opt not to reduce carbon emissions if the environmental costs are too high, or they may be required to reduce emissions due to regulations. However, most companies lack knowledge, expertise, and specific assets for carbon emission reduction, making it challenging for them to prioritize it.

We have chosen to investigate corporate decisions regarding cloud adoption in situations necessitating carbon emission reduction. While previous studies have explored factors influencing cloud adoption, such as the Technology, Organization, and Environmental (TOE) framework and financial considerations like IT infrastructure size and firm size, as well as security risks, our research aims to empirically demonstrate factors not previously considered.

In cloud adoption literatures, Research on firms' cloud adoption decisions has been steadily progressing in the Information Systems (IS) literature. (Low et al., 2011) examined firms' cloud adoption decisions through the Technology, Organization, and Environment (TOE) framework. (Misra & Mondal, 2011) analyzed cloud adoption from a financial perspective, finding that it is influenced by factors such as IT infrastructure size and firm size. Khalil (2019) identified three main reasons for cloud adoption: the highly competitive market, the need for innovative solutions, and the desire to lower costs. Zhang et al. (2020) studied how situations with high-security risks versus low-security risks affect firms' choices, as well as vendors' pricing and investment decisions. Schneider and Sunyaev (2016) aimed to analyze factors and determinants affecting cloud adoption by reviewing existing literature. They identified various factors related to transaction costs, including site specificity, technical specificity, firm internal IT capability, firm size, coercive pressures, mimetic pressures, and normative pressures.

Cloud adoption involves outsourcing activities for computing power, servers, cooling systems, storage, etc. All of these activities incur electricity costs and emit Scope 2 emissions. Therefore, firms' decisions regarding cloud adoption can be considered as decisions to outsource carbon emissions.

Firms often outsource their emissions to foreign suppliers rather than investing in abatement efforts (Dai ei al., 2021). Bartram et al. (2022) empirically demonstrated the effect of the Emissions Trading System (ETS) in California. They found that financially constrained firms tend to shift emissions and output from California to other states where they have similar underutilized plants. Conversely, unconstrained firms do not make such adjustments.

We examine the impact of environmental regulations on corporate cloud adoption decisions. Specifically, we provide empirical evidence of the influence of the California Cap-and-Trade, implemented in 2012, on cloud adoption by companies in California. The California Cap-and-Trade is a type of Emissions Trading System (ETS), which employs market-based mechanisms to control pollution.

Our findings indicate that since the implementation of the California Cap-and-Trade, companies in California have increased their cloud adoption despite controlling for existing factors outlined in previous literature. We utilized the Difference-in-Difference (DID) method and confirmed the robustness of our results through propensity score matching (PSM). Particularly, companies in green industries exhibit stronger responses due to higher demands for carbon reduction. In conclusion, our research provides evidence that corporate cloud adoption decisions are influenced by environmental regulations.

II. 그림과 표

<표 1> OLS REGRESSION FOR MAIN RESULT

DV : cloud adopted			
Variables	Coef.	Std.	p-value
ETS*CA	0.196	0.065	0.003
California dummy	0.074	0.058	0.207
After ETS	0.704	0.068	0.000
Size	0.025	0.020	0.211
Lev	-0.007	0.144	0.962
Cash flow	0.005	0.003	0.093
Tobin's Q	-0.012	0.014	0.413
Age	-0.001	0.002	0.541
CAPEX	-0.568	0.162	0.001
Servers	0.0001	0.00007	0.115
IT budget	0.046	0.014	0.001
Electricity price industrial	-0.017	0.010	0.094
High level breach	-0.023	0.027	0.401
HHI index	-0.301	0.648	0.643
KZ index	0.001	0.0006	0.120
CEO change	0.064	0.029	0.029
CFO change	0.031	0.028	0.277

III. 참고문헌

Bartram, S. M., K. Hou, and S. Kim (2022). Real effects of climate policy: Financial constraints and spillovers. Journal of Financial Economics 143 (2), 668–696.

Dai, R., R. Duan, H. Liang, and L. Ng (2021). Outsourcing climate change. European Corporate Governance Institute–Finance Working Paper (723).

Dong, M., D. Hirshleifer, S. Richardson, and S. H. Teoh (2006). Does investor misvaluation drive the takeover market? Journal of Finance 61 (2), 725–762.

Khalil, S. (2019). Adopting the cloud: how it affects firm strategy. Journal of Business Strategy 40 (4), 28–35.

Low, C., Y. Chen, and M. Wu (2011). Understanding the determinants of cloud computing adoption. Industrial management & data systems 111 (7), 1006–1023.

Misra, S. C. and A. Mondal (2011). Identification of a company's suitability for the adoption of cloud computing and modelling its corresponding return on investment. Mathematical and Computer modelling 53 (3-4), 504–521.

Park, J., K. Han, and B. Lee (2023). Green cloud? an empirical analysis of cloud computing and energy efficiency. Management Science 69 (3), 1639–1664.

Schneider, S. and A. Sunyaev (2016). Determinant factors of cloud-sourcing decisions: reflecting on the it outsourcing literature in the era of cloud computing. Journal of Information Technology 31, 1–31.

Zhang, Z., G. Nan, and Y. Tan (2020). Cloud services vs. on-premises software: Competition under security risk and product customization. Information Systems Research 31 (3), 848–864.

6

Session D

D2. [특별세션]
안전한 의료데이터 활용

좌 장: 한 현 욱(차의과대학)

의료데이터 안전활용 정책

한국인터넷진흥원 데이터안전활용단 심동욱

2024.05.11. 한국지능정보시스템학회

CONTENTS

Ⅰ. INTRO
Ⅱ. 가명정보
Ⅲ. 보건의료데이터 활용
Ⅳ. 비정형데이터 가명처리
Ⅴ. 개인정보 안심구역
Ⅵ. 마이데이터

Ⅰ Intro

1. 데이터 안전활용 관련 주요 정책 변화(개인정보위)

2020년 데이터3법 이후 개인정보보호법 중심으로 개인정보를 포함한 데이터의 안전 활용 관련 다양한 정책 개발 및 시행

A 가명정보

데이터3법 개정에 따라 '가명정보 처리 특례제도' 2020년8월 시행

『가명정보 활용 확대방안 』 발표('23.7월, 관계부처합동)

B 마이데이터

2023년3월 개인정보 전송요구권 도입(개인정보보호법 개정)

『국가 마이데이터 혁신 추진전략』 발표('23.8.17., 관계부처합동)

C 인공지능

ChatGPT 등 생성형 AI 이슈 부각('23년 초)

『인공지능 시대 안전한 개인정보 활용 정책방향 』 발표('23.8월, 개인정보보호위원회)

2. 데이터 안전활용 관련 주요 정책 변화(보건복지부)

2024년 보건복지부 주요 데이터 정책 : 국가데이터 자원 구축·활용 제고

A 바이오 빅데이터

임상 및 유전체 데이터 등 100만명 규모 빅데이터 구축으로 개인 맞춤의료 등 데이터 기반 정밀의료 실현

* 1단계('24~'28) 77.2만명 우선 추진(총 6천억원 규모, 복지부/과기부/산자부/질병청)

B 건강정보 고속도로

상급종합병원 중심 의료기관 데이터 연계 확대(860->1,003개)를 통해 국민 삶터 속 맞춤형 건강관리 서비스 활성화

* 공공보건서비스, 의료기관·헬스케어사업드에 건강정보 고속도로 플랫폼 및 데이터 활용지원

C 첨단재생의료

기존 임상연구와 별도의 치료목적 세포·유전자치료 등 허용 제도 신설 및 희귀 ·난치 등으로 제한된 연구대상 확대 추진

* 실시기관 지정 : ('23)85개 -> ('24) 115개 / 임상연구승인 : ('21~'23) 37건 -> ('24) 50건

5

Ⅱ 가명정보

1. 양질의 데이터 제공·공유 확대

공공데이터 공공데이터 개방·활용 촉진을 위해 법률 개정을 추진하고 가명정보 활용 성과를 공공기관 평가기준에 신설

- ☑ 공공기관 평가 — 공공기관 데이터 정책평가시 가명처리된 공공데이터 민간 제공 및 데이터기반행정 활용 관련 평가항목 신설 ('24)

AI·자율주행 비정형데이터를 안전하게 가명처리하여 AI학습, 자율주행 기술개발 등에 활용할 수 있는 기준 마련

- ☑ 가명처리 기준 — 비정형데이터(영상, 음성, 텍스트 등)의 가명처리 기준 마련 ('23.下)
- ☑ 합성데이터 — 스타트업 등 민간 데이터 수요를 조사·반영하여 유망분야별 합성데이터셋을 제작하여 제공 ('23.下)

바이오헬스 양질의 보건의료데이터의 민간활용 가속화 추진

- ☑ 공공의료데이터 — 가명처리된 건강보험 데이터의 민간기업 활용촉진을 위한 지침 개정 및 데이터 개방 ('23.下)
- ☑ 유전체정보 — 가명처리가 가능한 유전체 정보 범위를 확대하여 관련 연구 및 제품·서비스 개발 촉진 ('23.下)

2. 가명정보 활용 절차 합리화

절차개선 현실 적합성이 낮은 가명정보 처리·활용 기준을 합리적으로 개선하여 원활한 가명정보 활용 촉진

- ☑ 절차 간소화 — 데이터심의위원회(DRB) 법정기구화 및 데이터심의위원회(DRB)-생명윤리위원회(IRB) 간 중복심의 간소화*
  * DRB 심의 통과시, 7일 이내 'IRB 심의 면제확인서' 발급하는 신속절차 신설

규제 합리화 가명정보 관련 규제완화 · 법적책임 명확화하여 현장의 가명정보 활용 부담 완화

- ☑ 자체결합 — 가명정보 결합전문기관의 자기활용 목적을 위한 데이터 자체결합을 허용하고 부작용에 대한 보완장치 마련

자체결합 허용시 부작용	보완장치
다량의 데이터를 보유한 기관의 제3자 데이터 제공 축소	제3자에게 데이터를 제공한 실적에 비례하여 자체결합 허용
가명정보 결합·반출시 공공성·안전성 약화	가명정보 결합·반출에 대한 적정성 평가시 유사 경쟁기관 참여 의무화 자체결합 목적·실적을 투명하게 공시

- ☑ 법적책임 명확화 — ① 가명정보 처리 과정에서 의도치 않게 특정 개인을 알아볼 수 있는 정보가 생성되었다는 사실만으로는 제재하지 않음
  ② 가명정보를 제공받은 자의 안전조치 미이행 등으로 문제가 발생하였거나 고의로 재식별 행위를 하였을시, 해당 행위자(제공받은자)만 제재

3. 가명정보 활용지원 확대

개인정보 안심구역 데이터 처리의 환경적 안전성을 높임으로써, 개인·가명정보를 보다 유연하게 활용할 수 있는 제도 마련

- ☑ 대상 : 既 운영중인 결합전문기관, 가명정보 활용지원센터, 데이터안심구역 중 추가 안전조치를 갖춘 기관을 우선 지정·시범운영
- ☑ 환경적 안전성 : 제로 트러스트(Zero Trust) 보안모델 적용, 사전·사후적 데이터 처리과정 통제 등 안전조치 마련
  * 아무것도 신뢰하지 않는다는 원칙을 전제로 무조건적으로 신뢰하지 않고 데이터 처리 과정 전체를 검증하는 것을 목표로 하는 보안모델

중소·벤처 지원 가명처리 역량·인력이 부족한 중소기업·스타트업에 대한 지원 강화로 데이터 활용 사각지대 해소

- ☑ 활용지원센터 : '가명정보 활용지원센터' 내실화 및 추가구축
- ☑ 가명처리 솔루션 : 데이터 유형·분야·목적·기능별로 신뢰할 수 있는 가명처리 솔루션 목록 제공
- ☑ 데이터 바우처 : 가명·익명처리 등 데이터 가공 전 과정을 데이터바우처로 지원

전문 인력 가명정보 처리·활용 역량 강화를 위한 가명정보 전문인력 지속 양성을 통해 데이터 활용 생태계 조성

- ☑ 인력양성 : 맞춤형 교육을 통해 가명정보 전문인력 양성(26년까지 4,000명 배출)
- ☑ 전문가풀 : 가명정보 전문가풀 추가 확충(25년까지 100→200명) 및 전문성 제고

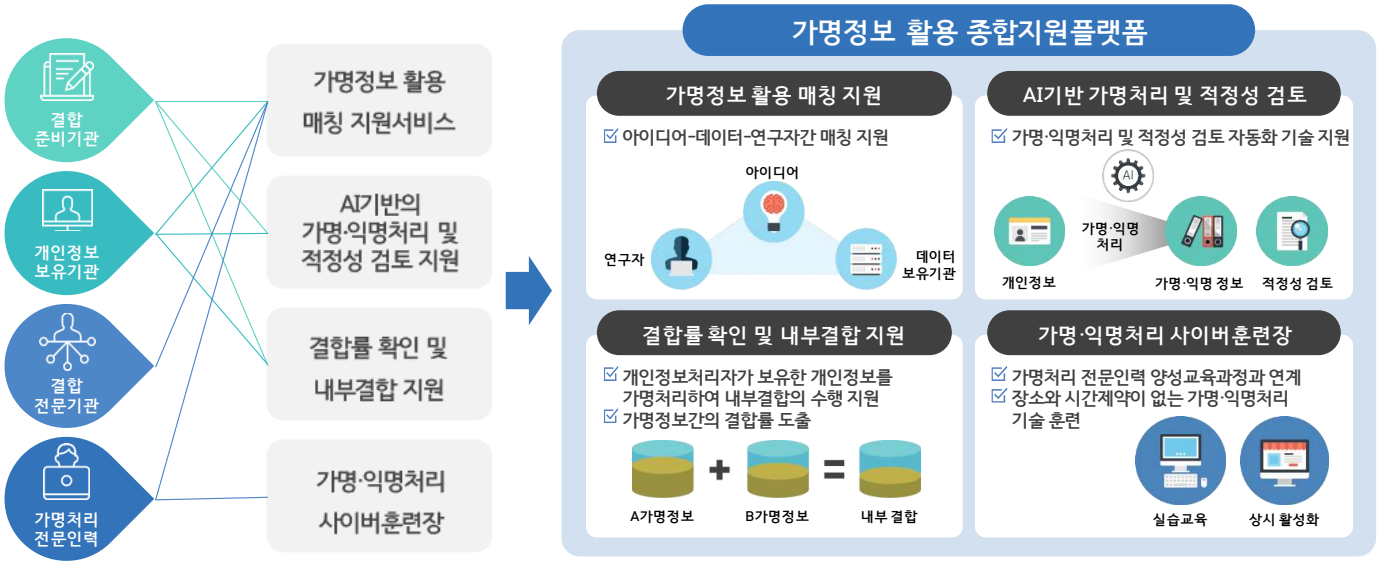
4. 가명정보 활용 절차의 안전관리 강화

R&D, 모니터링 AI개발 등에 필요한 개인정보보호 신기술·솔루션 개발을 지원하고 가명정보 모니터링 및 법률지원 강화

- ☑ 신기술 대응 보호 강화 : AI 등 신기술 프라이버시 수준 진단 및 보호 강화 기술, 자동화된 처리 대응권 지원 기술 R&D 추진
- ☑ 모니터링 : 데이터의 민감성과 처리·결합 과정의 위험성 등을 종합 고려한 모니터링 체계 구축
- ☑ 법률지원 : 지역 가명정보 활용지원센터의 '찾아가는 법률·기술지원 컨설팅' 확대

[참고] 가명정보 활용 종합지원플랫폼 (data.privacy.go.kr)

"데이터 중소기업·새싹기업(스타트업) 등을 대상으로 가명정보의 처리·결합·활용까지 전 과정을 종합적으로 지원하기 위한 온라인 서비스"



11

Ⅲ 보건의료데이터 활용 가이드라인

1 | 보건의료데이터 활용 가이드라인

가명정보 제공 시 법적책임 범위 명확화

※『가명정보 처리 가이드라인』(개인정보위, '24.2.), 『보건의료데이터 활용 가이드라인』(복지부·개인정보위, '24.1.)에 안내

➢ 개인정보를 보호법에서 정한 처리 목적에 따라 가명처리하고 관련 안전조치 등 법률에서 정한 사항을 모두 준수하여 가명정보를 제공한 경우

가명정보 처리·활용 시	가명정보 제3자 제공 시
가명정보 처리·활용 과정에서 **의도치 않게 특정 개인을 알아볼 수 있는 정보가 생성**되었다는 사실만으로 가명정보를 처리한 자 또는 해당 가명정보를 제공한 자를 처벌하지 않음 ※단, 해당 정보의 처리를 즉시 중지하고, 지체없이 회수·파기하여야 함	가명정보를 제공받은 자의 **안전조치 미이행** 등으로 가명정보 유출 등의 문제가 발생하였거나 **고의로 재식별** 행위를 하는 등 그 행위 주체의 위법으로 발생한 사안에 대해서는 **해당 행위자만 제재함**

➢ 가명정보를 제공받은 자가 안전조치 미이행 등으로 가명정보를 유출하였거나 고의로 재식별 행위를 하는 경우, **해당 행위자만 제재함**

13

2 | 보건의료데이터 활용 가이드라인

보건의료데이터 활용 관련 『보건의료데이터 활용 가이드라인』 개정('24.1)

- 보건의료 비정형데이터 가명처리 범위 확대

구분	기 존	개 선
유전체 데이터	널리 알려진 질병의 유전자 변이유무 확인 등 예외적인 경우를 제외하고 가명처리 **유보**	• NGS 기반 유전자 검사를 통해 생성된 표준화된 SAM 및 VCF 등은 메타데이터 및 염기서열에 대한 가명처리
자유입력 데이터	가명처리 **유보**	• 자연어 처리기술 등 정형데이터로 변환 후 식별정보 삭제·대체 등 방법으로 가명처리
음성 데이터	가명처리 **유보**	• 음성인식 기술 이용하여 텍스트 처리한 뒤 자유입력 데이터 가명처리 방법으로 처리

14

‘데이터심의위원회(DRB)’, ‘생명윤리위원회(IRB)’ 절차·제도 개선

- **DRB-IRB 간 중복심의 간소화*** (「IRB 관리안내지침」 개정예정)

  * 데이터심의위원회(DRB) 심의 통과시, ‘DRB 승인서’ 및 ‘데이터활용계획서’ 만으로 7일 이내 ‘IRB 심의면제 확인서’를 발급하는 신속절차 신설

- **IRB 신속심의 절차 신설** (「IRB 관리안내지침」 개정예정)

  * 연구진 변경, 연구기간 연장 등 **연구계획서 변경 내용이 최초 승인목적과 합리적 관련성이 있는 연구일 경우** 신속심의 등을 통해 7일 이내 처리

- DRB 법정기구화, 가명처리 시 IRB 면제, **공용 DRB 설치·운영 근거 마련** 등 추진

  * 「디지털헬스케어법」 제정 추진 中

15

Ⅳ 비정형데이터 가명처리 가이드라인

1 | 비정형데이터 가명처리 기준 마련

추진 배경

- AI 기술 발전과 컴퓨팅 자원 발달로 데이터 활용수요가
  전통적 정형데이터(수치)에서 비정형데이터(이미지, 영상, 음성, 텍스트)로 변화

* 전 세계 데이터 중 이미지, 영상, 음성, 텍스트 등 비정형데이터가 최대 90%를 차지 (IDC, '23)

- 그러나, 기존 「가명정보 처리 가이드라인」은 정형데이터 기준으로 되어있어,
  비정형데이터 고유 특징과 AI 시대 기술발전 등을 반영하지 못하는 한계

『가명정보 처리 가이드라인』 개정(24.2.4.)을 통해 비정형데이터 가명처리 기준 마련

17

2 | 비정형데이터 가명처리 기준 마련

비정형데이터의 가명처리·활용 시 특징

개인식별성 판단의 어려움	개인식별이 가능한 정보와 그렇지 않은 정보의 구분이 상대적이며, 처리 목적·환경 등에 따라 다르게 판단될 수 있음
가명처리 기술의 불완전성	비정형데이터 내 개인식별 위험성이 있는 모든 항목을 완벽하게 탐지하여 처리할 수 있는 기술이 부재(정확도 90~98%)
재식별 공격 위험	AI 및 데이터 복원기술의 발달로, 다른 정보와의 연계·결합 없이도 가명처리된 비정형데이터를 재식별해낼 수 있는 공격 위험성 증가

18

3 | 비정형데이터 가명처리 기준 마련

비정형데이터의 가명처리 기본 원칙

[1] 데이터 처리 목적. 환경, 민감도 등을 종합적으로 고려하여
개인식별 위험성이 있는 정보를 판단하고, 합리적인 처리 방법·수준 설정

- 연구목적에 맞춰 **데이터 자체 훼손을 최소화**하면서, **관리적·환경적 통제 등 다양한 안전성 확보 방안 적용** 가능
- 연구목적 달성에 필수적인 정보항목은 그대로 사용, 그 외 정보항목에 대한 가명처리 수준을 높이거나, **다른 정보 및 SW 반입 제한** 등 충분한 안전조치를 보완하여 활용

19

4 | 비정형데이터 가명처리 기준 마련

비정형데이터의 가명처리 기본 원칙

[2] 가명처리 기술의 한계 등을 보완하기 위해 사전 준비단계(연구 및 기술개발 기획)부터
위험성을 충실히 검토하고 적절한 안전조치를 수행

- 가명처리 기술의 한계 보완을 위해 다음 조치를 이행할 것을 **권고**

❶ 가명처리 기술의 적절성·신뢰성을 확인할 수 있는 근거 작성·보관

❷ 가명처리 기술 적용 이후, 처리 결과에 대해 자체적인 검수 수행

❸ 가명처리 적정성 검토 과정에서 ❶, ❷를 포함하여 점검
(외부전문가 과반 이상 참여 바람직)

20

5 | 비정형데이터 가명처리 기준 마련

비정형데이터의 가명처리 기본 원칙

3 데이터 복원기술 발달 등에 대응하여, 가명처리된 비정형데이터 활용 시 관련 시스템·SW의 접근·사용 제한 등 통제방안 마련

- 원본 복원에 활용될 수 있는 추가정보 분리보관, 복원 SW에 대한 접근권한 제한 등
- AI 개발·활용 과정에서 나타날 수 있는 다양한 위험을 사전에 완벽하게 제거하는 것은 불가능하므로, AI 서비스 제공과정에서도 개인식별 위험 등 정보주체 권익 침해 가능성을 지속 모니터링

21

6 | 비정형데이터 가명처리 기준 마련

비정형데이터의 가명처리 기준 주요 내용

- 가명처리 단계별 고려사항 안내
- 위험성을 사전 진단할 수 있는 '비정형데이터 개인식별 위험성 검토 체크리스트' 제공
- 실제 가명처리를 위한 '비정형데이터 개인식별 위험성 검토 항목별 조치가이드' 제공
- '비정형데이터 가명처리 기술·예시' 안내
- 실제 의료데이터, CCTV 영상, 음성 대화·상담정보 등을 활용하였던 사례를 재구성해 가명처리 전 단계를 상세 안내한 7종의 '주요 가명처리 시나리오 예시' 안내

22

7 | 비정형데이터 가명처리 기준 마련

주요 비정형데이터 가명처리 시나리오 예시

분야	비정형데이터 유형	가명처리 시나리오 예시
의료	이미지, 영상, 텍스트	① 유방암 골밀도 감소 여부 진단 AI 개발
	이미지	② 구강질환 진단 AI 개발
	이미지, 영상	③ 안면골 골절 진단 AI 개발
교통	이미지, 영상	④ 자율주행차 주행 시 비정상 상황인지 AI 개발
	이미지	⑤ 고속도로 다인승전용차 단속 AI 개발
대화·검색	텍스트	⑥ 한국어 대화가 가능한 AI 챗봇 제작
대화·교육	음성, 텍스트	⑦ 콜센터 직원 교육용 가상상담 시나리오 생성 AI 개발

※ 위 시나리오는 **실제사례**를 관련 기업·기관 및 전문가와 **재구성한 것으로 참고용**!

처리자 및 적정성 검토 위원회 판단에 따라 활용분야·상황에 맞게 **자유롭게 변용 가능!**

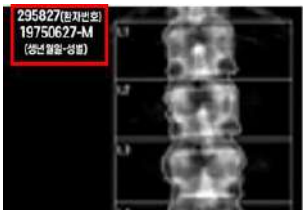
23

8 | 비정형데이터 가명처리 기준 마련

주요 비정형데이터 가명처리 시나리오 예시

(사례1) 유방암·골밀도 감소 여부 진단 AI 개발

<골밀도CT사진>
(환자관련정보 포함)



개인식별 위험성 검토

- 이미지 내 표시된 환자관련정보 * 는 타 정보와 결합되어 분석될 경우 개인식별 위험성이 있음

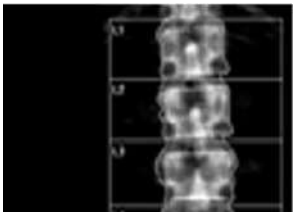
  * DICOM 헤더정보(환자번호, 생년월일, 성별) 표시

- 해당 정보는 연구에 필요하지 않은 정보임

데이터 처리 방안

➜ 블랙마스킹 기법을 통해 환자관련정보 삭제

(블랙마스킹 처리)



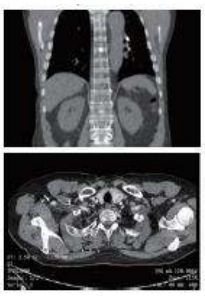
24

9 | 비정형데이터 가명처리 기준 마련

주요 비정형데이터 가명처리 시나리오 예시

(사례1) 유방암·골밀도 감소 여부 진단 AI 개발

<흉부 CT사진>



개인식별 위험성 검토

- 흉부 CT사진만으로는 **개인식별 위험성 거의 없음**
- 개인당 200장씩 촬영된 CT사진이 활용되어 **3차원 재건 기술**을 통해 신체형상의 입체적 복원이 가능
- 입체적 복원 시 특이한 외형·흉터 등이 있는 극히 일부 환자의 경우 낮은 확률로 **개인식별 위험성** 존재
- 클라우드 기반 **폐쇄연구분석환경을 이용**하고 인가되지 않은 데이터·SW 반입을 철저히 통제하고 있어 **3차원 재건기술 적용 불가**

데이터 처리 방안

➔ 재건에 의한 개인식별 위험성이 존재하나, 환경적 통제로 인해 해당 위험이 발생할 가능성이 없으므로 별도 가명처리 없이 그대로 활용 가능

(그대로 활용)

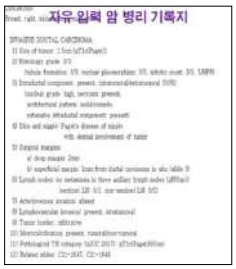
25

10 | 비정형데이터 가명처리 기준 마련

주요 비정형데이터 가명처리 시나리오 예시

(사례1) 유방암·골밀도 감소 여부 진단 AI 개발

<병리기록 텍스트>



개인식별 위험성 검토

- 암병리기록지에는 연구에 불필요한 다양한 개인 식별가능정보가 **정제되지 않은 형태**로 존재하여 **개인식별 위험이 존재**

데이터 처리 방안

➔ 자연어처리 기술을 통해 **정형데이터로 변환**한 후에 활용하고 개인식별 위험성이 있는 항목이 있으면 가명처리 수행

➔ 자연어처리 기술 및 텍스트데이터 가명처리 기술의 정확도가 100%가 아니므로, 정형데이터 변환 후 추가 **전수검사 등을 통해 보완**

(정형데이터 변환 후 활용)

26

11 | 비정형데이터 가명처리 기준 마련

주요 비정형데이터 가명처리 시나리오 예시

(사례2) 구강질환 진단 AI 개발

<구강 촬영사진>

개인식별 위험성 검토

- **구강사진 자체로는 개인식별 위험성 거의 없음**
- **충치 영역 외 부분은 연구에 필요 없음**
- **구강사진에 대한 메타데이터(이름, 나이 등)는 구강사진과 결합되어 개인식별 위험성 존재**

데이터 처리 방안

→ 연구에 필요한 충치 영역은 그대로 활용하고, **연구에 필요 없는** 그 외 영역은 **블러링 처리**

※ 블러링 수준은 현재 복원기술 발전수준 및 데이터 처리 환경(타 정보·복원기술 접근성) 등을 고려하여 설정

→ 메타데이터는 **연구에 필요 없어 삭제**

(충치부분: 그대로 활용)
(그 외: 블러링 처리)

27

V 개인정보 안심구역

1 | 개인정보 안심구역 소개 및 운영계획

'개인정보 안심구역' 이란?

- 연구자·스타트업들이 개인정보(특히, 개인정보)를
  보다 유연하고 탄력적으로 활용할 수 있는 공간

- 제로트러스트 원칙을 기반으로 데이터 활용 환경의 안전성을 높이되
  기존에 사실상 제한되던 다양한 데이터 처리가 가능토록 한 제도

  * 데이터 활용 환경·절차·기준은 개인정보위가 직접 설계·검증하여 안전

'23.12월, 통계청, 국립암센터 지정 완료

29

2 | 개인정보 안심구역 소개 및 운영계획

주요 기능

1. 가명정보의 유연한 활용

- **가명처리 수준**을 적정수준으로 완화 → 데이터 품질 훼손 최소화
- CI (Connecting Information) 일부값 등 **다양한 결합키** 사용 허용 → 결합률 극대화
- **비정형(영상, 이미지, 텍스트 등) 빅데이터** 가명처리시 샘플링 검사후 활용 → 시간·비용 절약

30

3 | 개인정보 안심구역 소개 및 운영계획

주요 기능

2. 가명정보의 장기활용 및 재사용

- AI 개발 등 지속적·반복적 연구목적 넓게 인정 → 가명정보 장기간 보관 가능(최대 5년)
  * 단, 개인정보 안심구역 내 보관·활용 전제
- 개인정보 안심구역에 보관된 가명정보는 제3자가 재사용 신청하여 활용 가능
  (원 보유기관과 협의 지원)

31

4 | 개인정보 안심구역 소개 및 운영계획

주요 기능

3. PET 실증

- PET 기술은 기존 제도 적용이 모호하거나 안전성 검증체계가 없어서 활용 부진
  → 전문심의위원회의 심의·검증 하에 PET를 적용한 개인정보 처리 폭넓게 허용

* PET (Privacy Enhancing Technology)
: 가명·익명처리 기술, 동형암호, 합성데이터, 차분프라이버시 기술 등 프라이버시를 향상시킬 수 있는 다양한 개인정보 보호·활용 기술을 통칭

32

5 | 개인정보 안심구역 소개 및 운영계획

주요 기능

※개인정보 안심구역 전문심의위원회

- 연구분야, 처리기술 등에 맞게 개인정보위가 국내 최고전문가로 구성
- 가명정보 결합시 기관별 반출심사 없이 통합하여 반출 심사(원스톱 심사)
- 전문심의위원회 운영 비용은 개인정보위가 지원

33

6 | 개인정보 안심구역 소개 및 운영계획

개인정보 안심구역 시범운영 기관

[국립암센터] 보건의료분야 연구자·기업의 가명정보 활용·결합 연구 수행 지원

- (기존 제약이 많았던 보건의료데이터 활용 연구 지원)
  - 샘플 수가 적어서 개인식별 위험이 크다는 이유로 연구자체가 불가능했던 경우
  - 희귀암·민감상병 관련 연구에서 과도한 가명처리로 데이터 품질이 크게 훼손되어 연구결과를 신뢰할 수 없었던 경우 등
- (보건의료분야 PET 실증연구 지원)
  - 동형암호, 합성데이터 등 주요 PET 기술의 보건의료 분야 실증사례 연구 도출

34

7 | 개인정보 안심구역 소개 및 운영계획

개인정보 안심구역 시범운영 기관

[국립암센터] 보건의료분야 연구자·기업의 가명정보 활용·결합 연구 수행 지원

- (대용량 의료 영상·이미지 데이터 샘플링 검사 후 활용)

- 가명처리 적정성에 대한 샘플링 검사 허용으로 전수 검사 시간·비용 대폭 절감 가능

< 예시: 폐암 CT 데이터를 활용한 진단 및 예후 예측 인공지능 연구 >

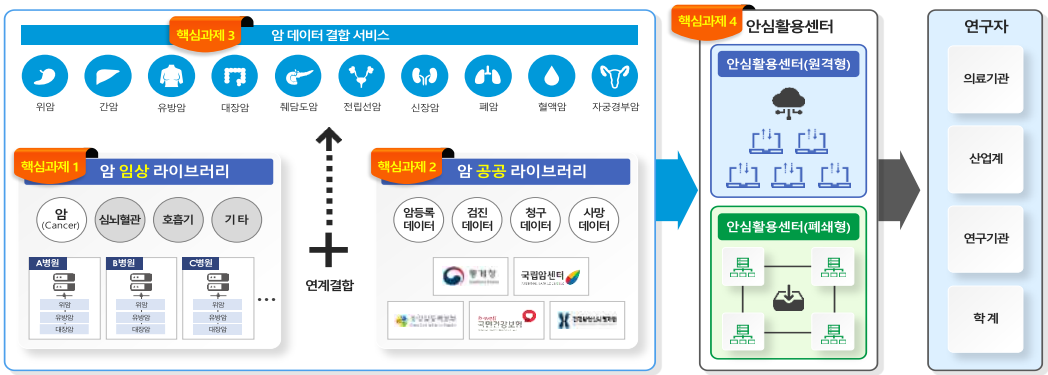
35

8 | 의료데이터 안심활용센터

의료데이터 안심활용센터

- 연구자가 공공데이터와 병원의 임상데이터를 연구 목적으로 안전하게 활용하기 위해 마련된 물리·기술·관리적 보안 대책을 갖춘 폐쇄분석 공간

- 한국보건의료정보원(중앙센터, 서울), 계명대학교동산의료원(대구), 원주세브란스기독병원(원주), 충남대학교병원(충남), 국립암센터



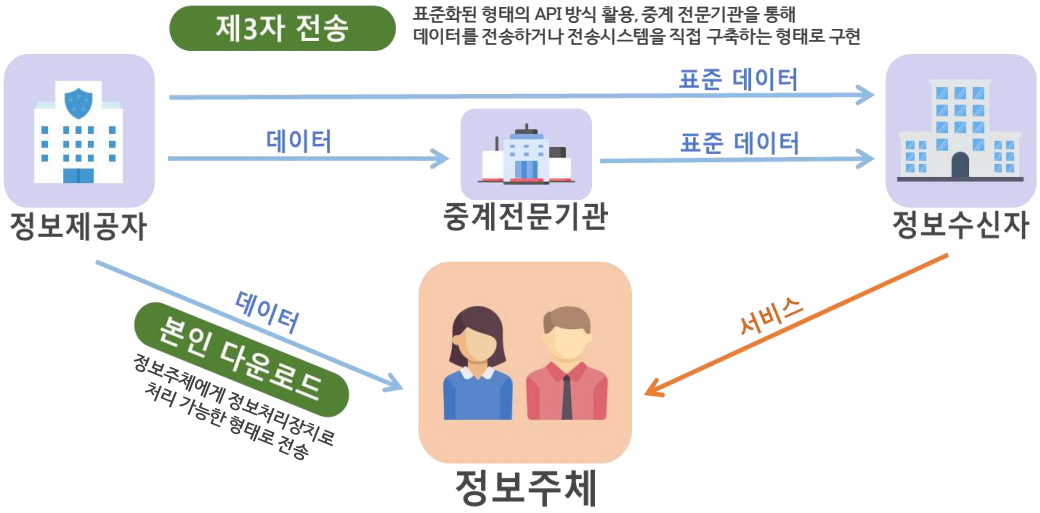
36

Ⅵ 마이데이터

1. 마이데이터 개요

마이데이터(MyData)

- 정보주체가 자신이 원하는 곳으로 개인정보를 이동시켜 본인이 원하는 서비스에 활용되도록 하는 제도



2. 마이데이터 추진방향

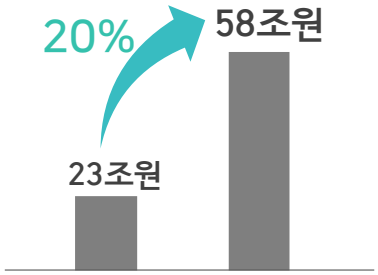
마이데이터로 선도하는 디지털 대전환 시대

국민 체감 서비스

국민이 체감할 수 있는 선도서비스

30개 지정

데이터 시장 규모



마이데이터 기업

신기술, 서비스 분야 마이데이터 기업

500개 발굴

3. 정보주체 관점에서 중점 추진부문선정 및 선도서비스발굴

마이데이터 10대 중점 추진부문 선정

- 국민들의 수요가 많고 제도 도입 효과를 체감할 수 있는 부문을 선정하여 제도 우선 시행
- 2025년 제도 시행 초기에는 중점부문부터 마이데이터를 도입하되, 이후 시장 상황에 따라 부문 및 분야 단계적 확대

기존부문

보건의료	통신 · 인터넷 서비스	교통	교육	유통
• 의료(병·의원) • 의약품(약국) • 웨어러블기기 건강정보	• 무선통신 • 플랫폼 · 포털서비스	• 철도 • 항공 • 자동차임대	• 초·중·고등 교육 • 평생교육 • 온라인 교습	• 온라인 쇼핑 • 대형마트

신규부문

에너지	고용노동	부동산	복지	여가
• 전기 • 가스 • 수도	• 고용알선 · 인력 공급	• 부동산 임대 · 공급	• 사회복지 서비스	• 숙박 • 여행

3. 정보주체 관점에서 중점 추진부문 선정 및 선도서비스 발굴

선도 서비스 발굴

- 중점부문을 중심으로 다양한 **혁신 서비스를 발굴**하고, 마이데이터 인프라를 통해 **산업간 데이터 융합 · 연계 확대**

전 세대(공통)	청소년기(10대)	청년기(20~30대)	장년기(40~50대)	노년기(60대 이상)
· 합리적인 통신요금 관리 · 의료기관 간 진료정보 전송	· 맞춤형 학습, 교육멘토링 · 병원진료, 봉사활동 증명	· 맞춤형 일자리 추천 · 청년 맞춤 대출 · 청약	· 만성 건강질환 관리 · 에너지·공공요금 통합관리	· 독거노인 위급 알림 · 요양내역 알림 서비스

4. 국민이 신뢰할 수 있는 마이데이터 프라이버시 보호 구현

권리보장 원칙

- 필요 정보 최소 수집
- 전송목적 범위 내 활용
- 데이터 활용 목적 명확히 제시

데이터 관리

- 데이터의 정확성·완전성·최신성 유지
- 지체 없는 정보 전송
- 자체 보유정보와 구분하여 관리

국민이 신뢰하는 마이데이터

보안 · 식별 메커니즘

- 부당 전송 유도행위 방지
- 안전한 식별 · 인증체계 마련
- 전송보안 가이드라인 제시

금지 의무

- 개인정보 보호의무 위반 시 과징금, 시정명령 등 제재
- 프라이버시 신고센터 운영

5. 플랫폼을 통해 국민의 투명한 마이데이터 권리행사 지원

마이데이터 지원 플랫폼 구축

- 전송요구, 전송이력 확인, 전송중단, 파기 등 정보주체의 전송요구 권리를 원스톱으로 지원하는 '마이데이터 지원 플랫폼' 구축

마이데이터 지원 플랫폼

정보주체 권리 보장

- 모든 개인정보 전송이력(전송기업, 데이터 항목 및 날짜 등)을 플랫폼에서 한 번에 확인하고 관리
- 전송이력 확인 후 원치 않는 전송을 즉시 중단하거나 기존 전송 데이터의 파기를 요청할 수 있도록 지원

전송 요구 → 전송 이력 확인 → 전송 중단 → 파기

디지털 카탈로그

- 마이데이터 서비스 및 전송 가능한 데이터를 한눈에 파악하고 혜택을 누릴 수 있도록 디지털 카탈로그 제공

프라이버시 신고

- 부당한 전송유도(다크패턴 등) 및 전송요구 거부, 전송 오류 등 국민 불편에 대한 신고센터 운영

※위법사항 확인 시 신속히 조사·처벌 등 후속조치

6. 시장 수용성, 기업 부담 등을 감안하여 단계적·점진적 확대

전송의무자 및 전송대상 정보항목 범위

- 제도 도입 초기에는 **수요가 높은 부문**을 중심으로 도입하고, 제도 도입취지 · 수용성 · 준비부담 등을 감안하여 **단계적 · 점진적 확대**

전송의무자

빅테크 등 대기업 및 중점부문 관련 공공기관

전송대상 정보항목

(1차) 기존 부문 공개

교육, 정보통신, 유통, 교통, 여가

단계적·점진적 범위 확대

전송의무자

매출액, 개인정보 보유규모, 처리능력, 산업별 특성 등을 조사하여 정보제공자 기준 결정

※ISMS, 정보보호 공시의무 대상자 등 기준 참조

전송대상 정보항목

(2차) 신규 부문 공개

보건의료, 복지, 에너지, 고용노동, 부동산

정보항목 추가 · 구체화, 삭제 · 축소 등 민간 의견을 받아 검토한 후 조정하여 전송대상 정보범위 결정

7. 민간 시장의 역동성이 충분히 확보될 수 있도록 제도정책 운용

개인정보관리 전문기관

특수부문 사업 전문기관

개인정보의 안전한 처리에 필요한 시설 · 기술요건은 면밀히하되,
불필요한 진입규제는 최소화하고 필요한 부문*만 허가 · 지정제 운영

*정보자체의 민감도, 유·노출 시 파급력 및 위험도 등 고려(의료, 미성년 교육정보 등)

지정요건

- 기술수준 및 전문성
- 안전성 확보조치
- 재정능력
- 사업계획 적정성 등

개인정보위 또는 관계 중앙 행정기관의 장이 지정

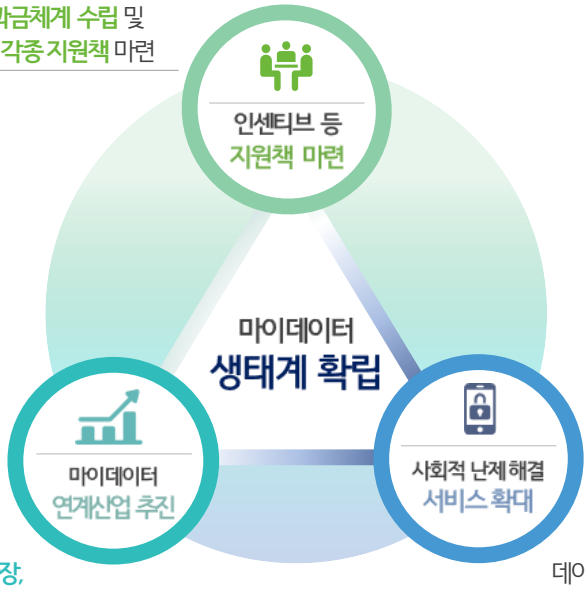
분야별 중계 전문기관

정보제공자로부터 **데이터를 받아 표준화된 형태로**
데이터를 변환하여 수신자에게 전송하는 등
전송요구권 행사를 지원하는 인프라 기관 지정

분야별 지정	민간 · 공공기관 제한없이 개인정보위 또는 중앙행정기관 장이 지정 ※ 중계집중식, 아카이브식 등 다양하게 검토
표준참조 모델	중계 전문기관 준비 지원을 위해 표준 참조 중계모델 마련
사업분리	특정서비스 몰아주기, 데이터 독과점 등 공정경쟁 저해 가능성을 감안하여 중계기관과 사업기관 분리

8. 지속적으로 확산하는 마이데이터 생태계 확립

기업 참여유인 제고를 위한 **과금체계 수립** 및
설비투자에 대한 금융지원 등 **각종 지원책** 마련



생태계 활성화를 위한 **API 인프라 확장,**
클라우드, SW 아키텍처, 데이터 마이닝 확산 추진

데이터 활용으로 안전 · 복지 · 범죄예방, · 재난 ·
환경 등 **사회적 난제를 해결**하는 **정부 · 공공 서비스 확대**

9. 단계별 추진전략

준비기 ('23~'24년)
성공적인 제도 안착을 위한 인프라 구축 집중

주요내용

- (법제도) 전송요구권, 다운로드권, 열람·확인권, 전송·활용 중단권 등 권리 법제화
- (플랫폼) 마이데이터 전 이용과정을 기술적으로 지원하는 온라인 플랫폼 구축
- (표준화) 데이터 형식·전송규격의 공통규칙을 만드는 표준화 추진

도약기 ('24~'25년)
국민이 지지하는 지속가능한 시장기반 구현

주요내용

- (선도서비스) 마이데이터 편익을 체감할 수 있는 선도서비스 발굴 및 홍보 강화
- (참여유인) 과금체계 수립, 설비 투자에 대한 지원책 마련, 중계시스템 지원 확대
- (프라이버시 보호) 식별·인증체계, 보안체계, 사후재제 등 메커니즘 보강

확산기 ('25~'27년)
서비스 혁신 생태계 고도화

주요내용

- (API 확산) 데이터와 서비스 융복합을 위한 API 경제인프라 확장
- (혁신기업) 클라우드 및 서비스형 SW 전문기업 지원
- (사회 프로젝트) 안전·복지·범죄예방·재난·환경 등 사회적 난제 해결

감사합니다
THANK YOU

Q&A

질의응답

의료기기 네트워크 환경의 사이버보안 위협분석 기술

2024.05.11.
정일안 이사
㈜휴네시온

HUNESION

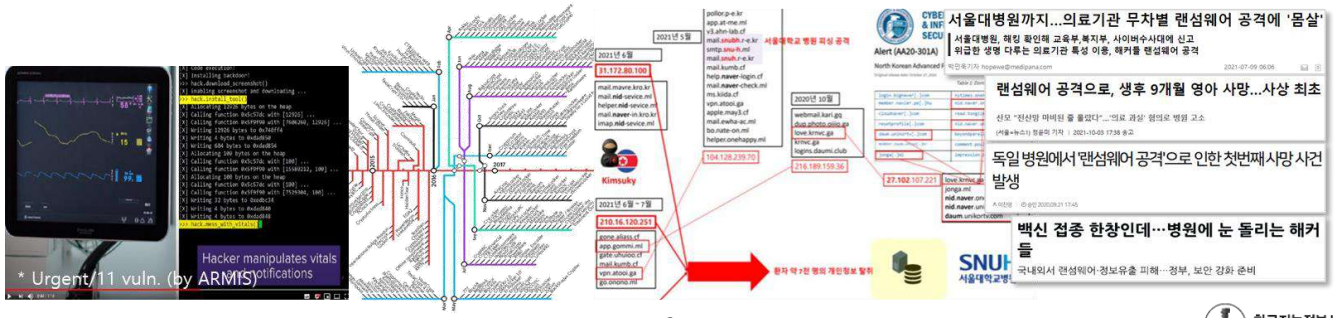
Copyright©2024 HUNESION Co.Ltd. All rights reserved. (무단 전재·재배포 금지)

한국지능정보시스템학회
Korea Intelligent Information Systems Society

목 차

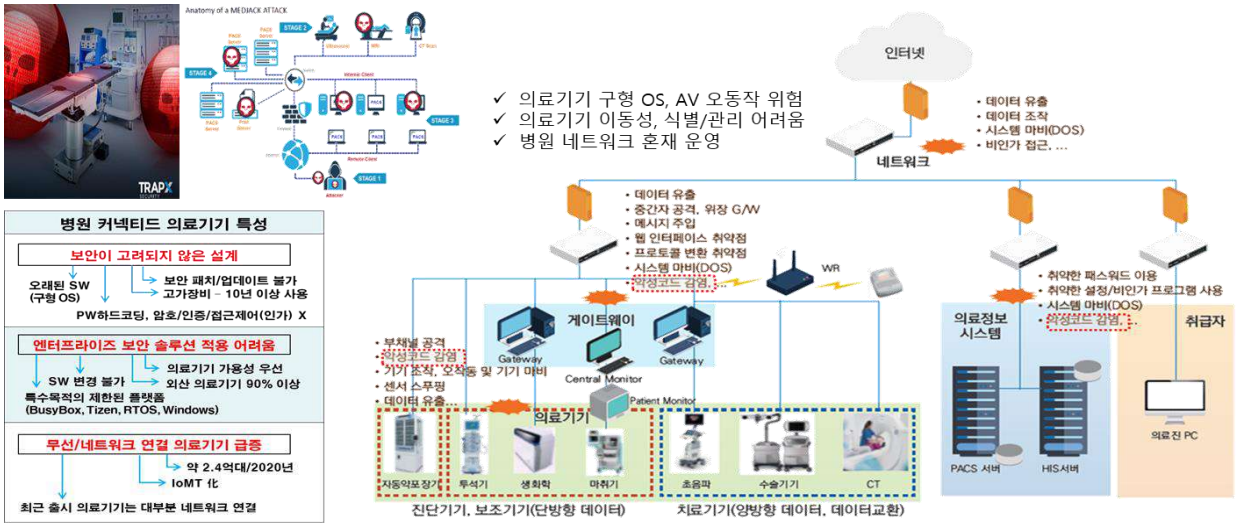
01. 의료 분야 사이버보안 동향
02. 의료 네트워크 트래픽 분석 연구
03. 병원망 트래픽 이상행위 분석 사례

- 사이버보안 위협은 IT, OT/ICS, IoT 영역 뿐만 아니라,
  커넥티드 의료기기 (IoMT; Internet of Medical Things) 대상으로 확대
  - ✓ [Claroty XIoT Report] 2022년 상반기에 발견된 XIoT (Extended Internet of Things) 취약성
    - XIoT Vuln. : IT 16.47%, OT 65.33%, IoT Only 15.13%, IoMT 3.08% (하반기 2%)
  - ✓ [PaloAlto Report] 2022년 20만개 주입 펌프 중에서 75% 제품이 보안에 취약함
    - 주입 펌프의 52% (CVSS 9.8, CVE-2019-12255; VxWorks BOF in TCP component)
- 병원 대상 랜섬웨어 및 해킹 발생 사례
  - ✓ WannaCry, Petya, Ryuk, Conti, Sodinokibi, REvil 등 다수 랜섬웨어 발생 추세
  - ✓ **(국외)** 미국 스프링힐 병원 ('19.07, 영아 사망), 독일 뒤셀도르프 대학병원 ('20.9, 치료 환자 이송 사망)
  - ✓ **(상급종합병원)** 서울대병원 해킹/랜섬웨어 피해 ('21.7, 병원 방문자 7천명 정보 유출)
  - ✓ **(의원급)** 부산 소재 여성의원, 서울/경기 피부과, 강남 성형외과 해킹/랜섬웨어 피해 등



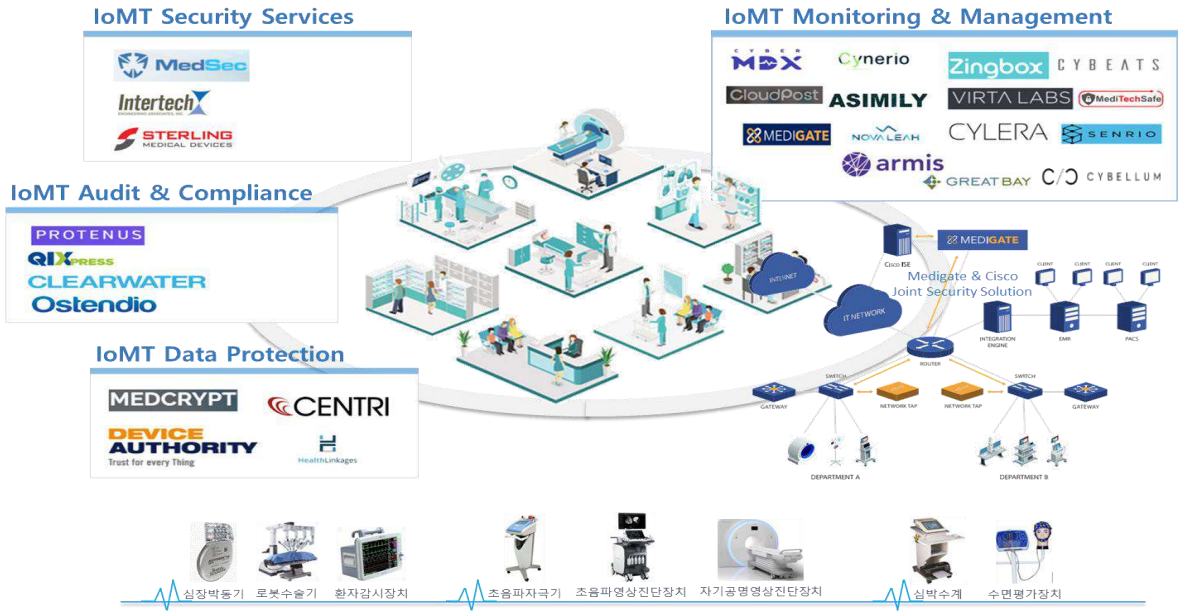
01 의료 분야 사이버보안 동향 - 커넥티드 의료기기 보안 동향 KIISS 2024

- MEDJACK (2015~2017, by TrapX Report)
  - ✓ 의료기기 악성코드 감염 전파, 백도어 생성, Botnet 구축, 탐지 시스템 우회 등
- 커넥티드 의료기기 사이버보안 (2021, ETRI Report)
  - ✓ 보안 취약 요인 : 보안 未고려 설계, 기존 보안솔루션 적용 어려움, 네트워크 연결 의료기기 급증
  - ✓ 랜섬웨어 동작 단계 : 단말 침투 및 권한 획득(Exploit), 동작(Ransomware), 확산(Spreading)

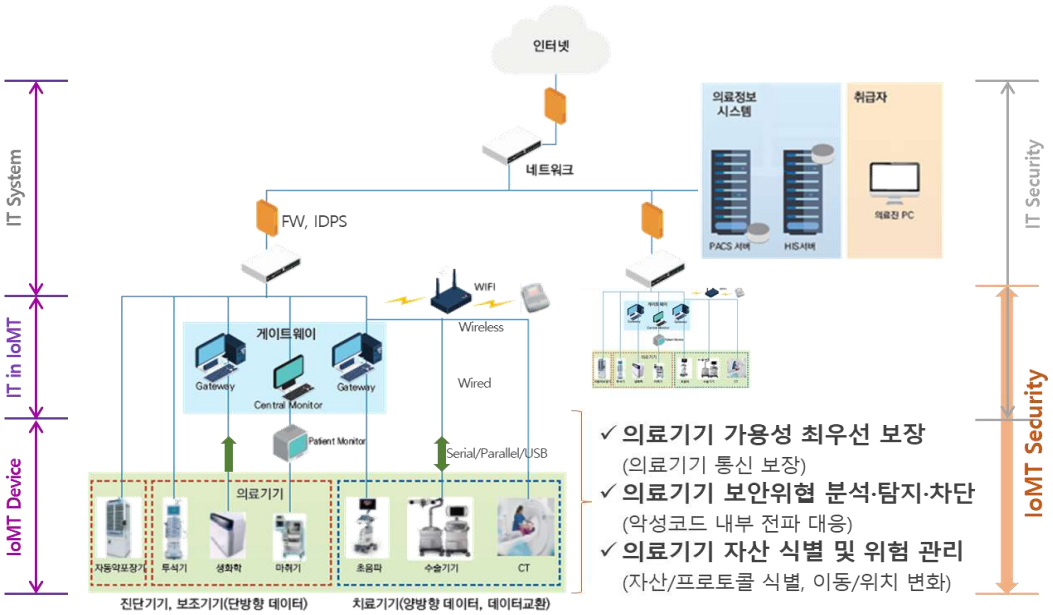


※ 출처 : ETRI 커넥티드 의료기기 해킹 및 랜섬웨어 대응기술 동향 (2021)

- 의료기기 보안 국외 기술 및 솔루션 동향
  - ✓ Medigate : 병원 네트워크의 의료기기 보안과 자산 관리 (device type, network comm. 등) ⊂ **Claroty**
  - ✓ Zingbox : 의료기기 네트워크 행위에 대한 딥인사이트 분석 (vendor, model, serial number 등) ⊂ **PaloAlto**
  - ✓ 기존 보안 솔루션 : 병원 네트워크 행위 분석, 네트워크 장비와 연동하여 대응 (Cisco, Fortinet 등)

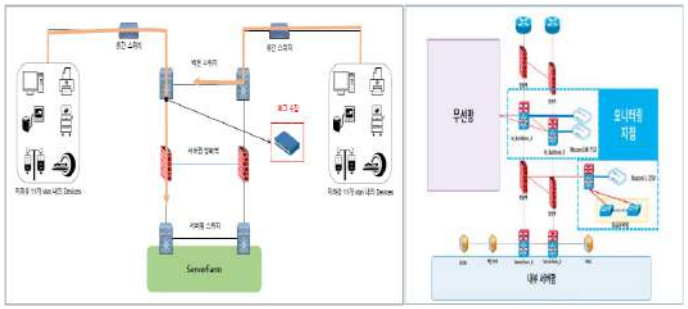


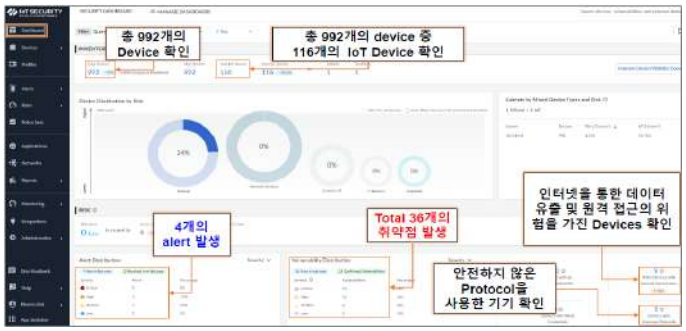
- 병원 의료기기 네트워크 개념도 및 보안 고려사항
  - ✓ 커넥티드 의료기기(IoMT) 네트워크 통신 가용성 보장
  - ✓ 의료기기 네트워크 보안위협 분석-탐지-차단
  - ✓ 의료기기 자산 식별 및 위험 관리



※ 그림 출처 : KISA 스마트의료 보안 가이드(2019)

- 의료장비 보안 취약점 분석을 위한 POC 사례
  - ✓ 목적 : 병원망 대상 각종 자산(의료장비 포함)의 가시화 및 통신 현황 분석
  - ✓ PaloAlto/Nozomi : 미관리 IoMT 확인(IP, Asset), 취약점 점검, 통신 토폴로지 모니터링 등

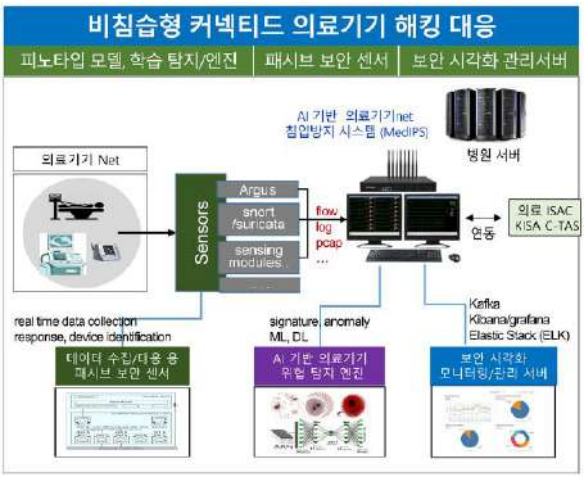


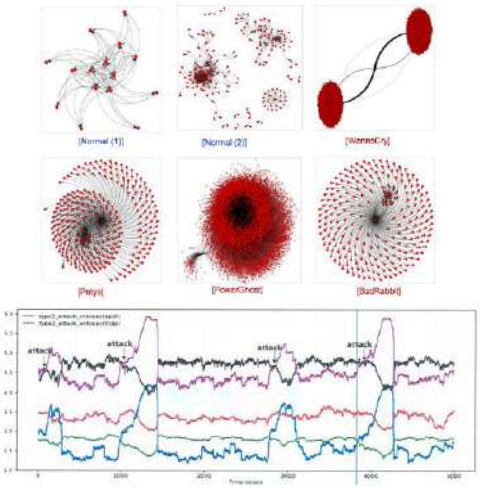


2-3) Dashboard : Medical 확인 (DICOM-Workstation : 의료영상 디스플레이 기기)

※ 출처 : 의료기관 개인정보보호 & 정보보안 컨퍼런스(MPIS 2023)

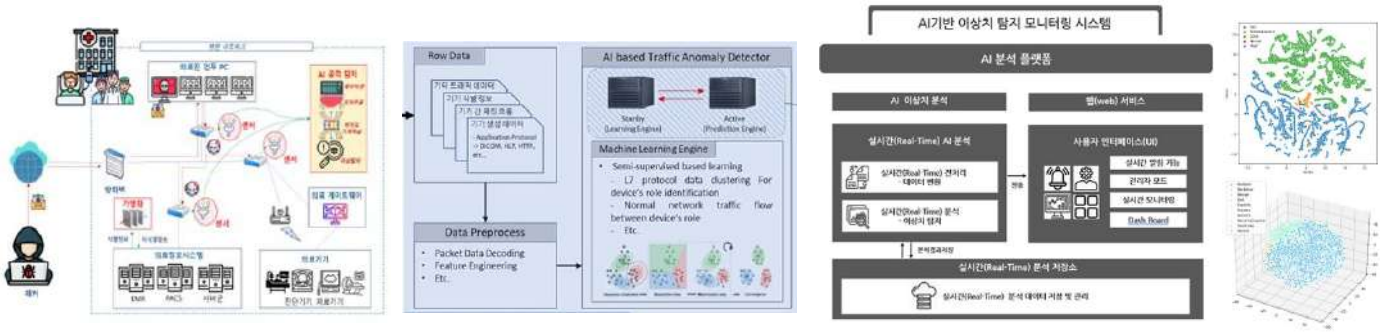
- 커넥티드 의료기기 해킹대응 기술 연구개발
  - ✓ 안전한 의료-헬스케어 서비스를 위한 커넥티드 의료기기 해킹대응 핵심기술 개발
    - 주관연구기관(ETRI), 공동연구기관(휴네시온, 건국대학교병원, 스마트의료보안포럼)
  - ✓ AI 기반 커넥티드 의료기기 이상징후 분석/학습/탐지 연구
    - 의료기기 네트워크 행위의 트래픽 Flow 특성화·피처화 및 학습·탐지
    - 주기성(Periodicity), 친숙성(Familiarity), 엔트로피(Entropy), Flow의 통계적·환경적 특성
    - 의료기기 랜섬웨어 감염에 따른 의료 네트워크 행위 시각화, 랜섬웨어 공격 발생 엔트로피 이상변화 추이





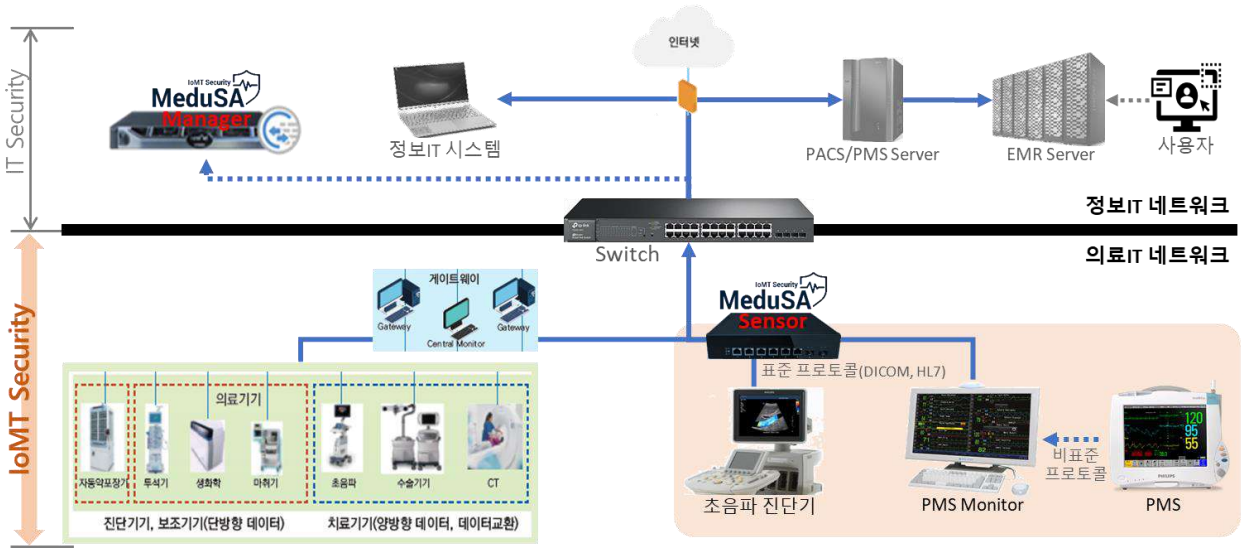
※ 출처 : 커넥티드 의료기기 해킹 및 랜섬웨어 대응기술 동향(ETRI Magazine, 2023)

- AI 기반 의료 네트워크 이상 탐지 기술 연구개발
  - ✓ AIM : AI 기반 차세대 보안 정보관리기법적용 Cognitive Intelligence 및 Secure-오픈 프레임워크 (S-OFW) 기술 개발[2단계, 2023~2026]
    - 주관연구기관 : 차의과학대학교 정보의학연구소
    - 공동연구기관 : 휴네시온, 엘티포, 미소정보기술
  - ✓ AI 기반 네트워킹 행태 적응형 정보보안 플랫폼 기술 개발
    - 의료 네트워크 트래픽의 Metadata 정보 수집 및 분석 기술
    - S-OFW 프레임워크 및 AI 운영 기술
    - AI 기반 의료 네트워크 이상행위 탐지 및 XAI 기반 탐지 근거 분석 시각화 기술

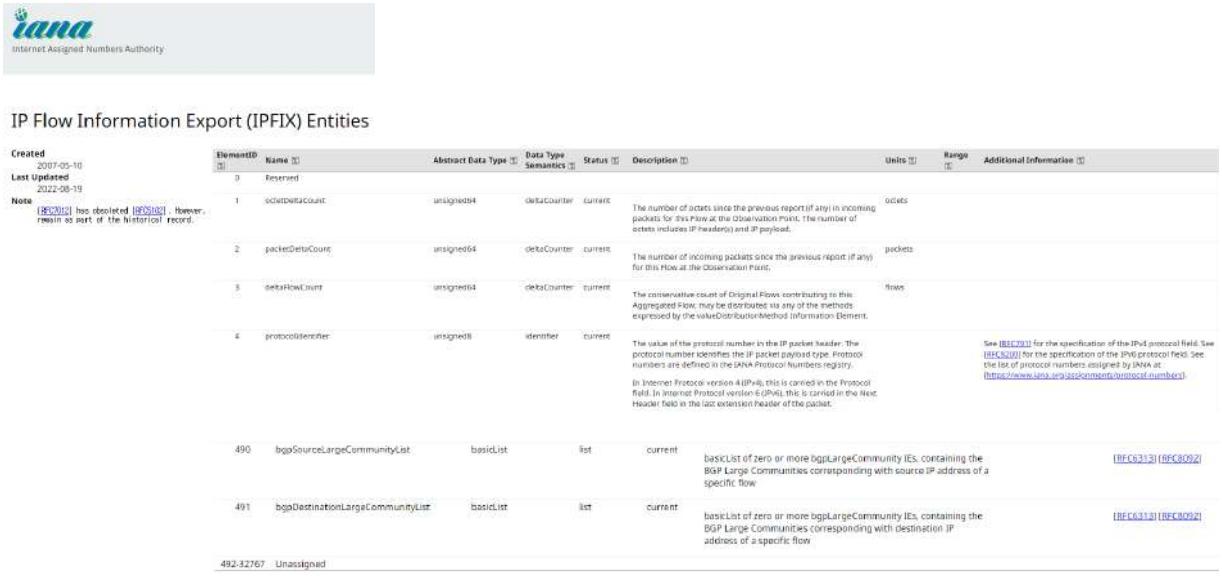


01 의료 분야 사이버보안 동향 - 국내 의료 보안 연구개발 동향 KIISS 2024

- ㈜휴네시온, 커넥티드 의료기기(IoMT) 해킹 대응을 위한 **보안 센서** 및 **보안 서버** 기술
  - ✓ IoMT Security 프로젝트명 : **MedúSA** (Medical únified Security Architecture)
  - ✓ **보안 센서** : 네트워크 기반 의료 자산 식별, 의료 통신 보장/분석, 보안위협 탐지/차단 등
  - ✓ **보안 서버** : 보안 모니터링, 보안 시각화, 보안 관리(의료기기 자산, 보안 센서, 보안 정책) 등



- IP Flow Information Export (IPFIX) Entities
  - ✓ IANA, RFC 7011, RFC 7012
  - ✓ Element IDs : 491 (2022-08-19)
  - ✓ IPFIXcol2 : IPFIX flow data collector for NetFlow v5/v9



iana
Internet Assigned Numbers Authority

IP Flow Information Export (IPFIX) Entities

Created
2007-05-10
Last Updated
2022-08-19
Note
[RFC2012] has obsoleted [RFC5102]. However, remain as part of the historical record.

ElementID	Name	Abstract Data Type	Data Type Semantics	Status	Description	Units	Range	Additional Informatise
0	Reserved							
1	octetDeltaCount	unsigned64	deltaCounter	current	The number of octets since the previous report (if any) in incoming packets for this Flow at the Observation Point. The number of octets includes IP header(s) and IP payload.	octets		
2	packetDeltaCount	unsigned64	deltaCounter	current	The number of incoming packets since the previous report (if any) for this Flow at the Observation Point.	packets		
3	deltaFlowCount	unsigned64	deltaCounter	current	The conservative count of Original Flows contributing to this Aggregated Flow; may be distributed via any of the methods expressed by the valueDistributionMethod Information Element.	flows		
4	protocolIdentifier	unsigned8	identifier	current	The value of the protocol number in the IP packet header. The protocol number identifies the IP packet payload type. Protocol numbers are defined in the IANA Protocol Numbers registry. In Internet Protocol version 4 (IPv4), this is carried in the Protocol field. In Internet Protocol version 6 (IPv6), this is carried in the Next Header field in the last extension header of the packet.			See [RFC791] for the specification of the IPv4 protocol field. See [RFC8200] for the specification of the IPv6 protocol field. See the list of protocol numbers assigned by IANA at [https://www.iana.org/assignments/protocol-numbers].
490	bgpSourceLargeCommunityList	basicList	list	current	basicList of zero or more bgpLargeCommunity IEs, containing the BGP Large Communities corresponding with source IP address of a specific flow			[RFC6313] [RFC8092]
491	bgpDestinationLargeCommunityList	basicList	list	current	basicList of zero or more bgpLargeCommunity IEs, containing the BGP Large Communities corresponding with destination IP address of a specific flow			[RFC6313] [RFC8092]
492-32767	Unassigned							

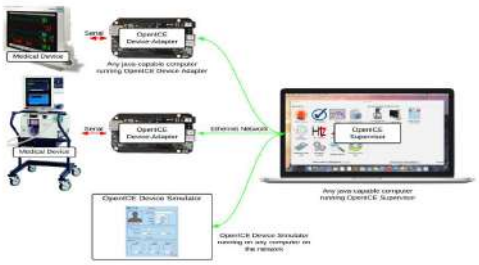
* https://www.iana.org/assignments/ipfix/ipfix.xhtml

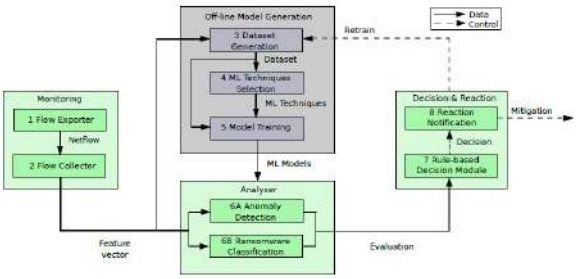
02 의료 네트워크 트래픽 분석 연구 - 의료 Network Flow 분석 KIISS 2024

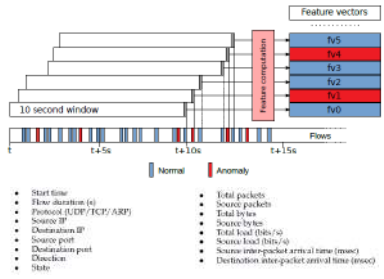
- IPFIX 기반 Network Flow 분석 정보 예시
  - ❖ **Traffic Volume Statistics**
    - Source/Destination IP별 송수신 bytes/packets
    - 프로토콜별 송수신 전체 bytes/packets
    - 송수신 평균 bytes
    - Source/Destination IP별 Flow 수
  - ❖ **Traffic Flow Statistics**
    - 각 Source IP에서 시작된 Connection 수
    - 각 Source/Destination IP 사이의 Connection 수
    - 각 Connection Time 및 유지 시간
    - 주고 받은 byte 기준 Top Source/Destination IP 주소
    - 주고 받은 byte 기준 Top 포트/프로토콜
  - ❖ **Network Performance Statistics**
    - 네트워크 대기시간(Network latency)
    - 왕복 시간(Round trip time)
    - 대역폭 사용(Bandwidth utilization)
    - 패킷 손실률(Packet loss rate)

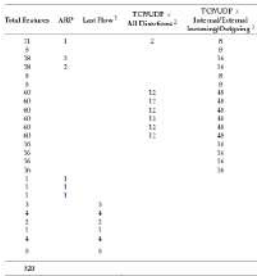
```
RA_FIELD_SPECIFIER
srcid seq stime ltime dur sstime sltime sdur
dstime dltime ddur srng drng trans flgs avgdur
stddev mindur maxdur saddr dir daddr proto
sport dport sco dco stos dtos sdsb ddsb sttl
dttl shops dhops sipid dipid pkts spkts dpkts
bytes sbytes dbytes appbytes sappbytes
dappbytes load sload dload rate srate drate
loss sloss dloss ploss sploss dploss senc denc
smac dmac smpls dmpls svlan dvlan svid dvid
svpri dvpri sintpkt dintpkt sintpktact
dintpktact sintpktidl dintpktidl sintpktmax
sintpktmin dintpktmax dintpktmin sintpktactmax
sintpktactmin dintpktactmax dintpktactmin
sintpktidlmax sintpktidlmin dintpktidlmax
dintpktidlmin jit sjit djit jitact sjitact
djitact jitidl sjitidl djitidl state deldur
delstime delltime dspkts ddpkts dsbytes ddbytes
pdspkts pddpkts pdsbytes pddbytes suser:1500
duser:1500 tcpext swin dwin jdelay ldelay bins
binnum stcpb dtcpb tcprtt synack ackdat inode
smaxsz sminsz dmaxsz dminsz

RA_FIELD_SPECIFIER
stime dur proto saddr sport dir daddr dport
state stos dtos pkts spkts bytes sbytes load
sload sintpkt dintpkt label
```

- OpenICE (Open-Source Integrated Clinical Environment) 플랫폼
  - ✓ 의료기기 시뮬레이터 기반 시험 환경 구성, 정상 트래픽 및 4종의 랜섬웨어 감염 트래픽 수집
  - ✓ ICE Dataset : Clean, WannaCry, Petya, BadRabbit, PowerGhost
  - ✓ 학습 feature : 단일 플로우 기반 16개 추출, Time Window 기반 플로우 집합에서 520차원 생성





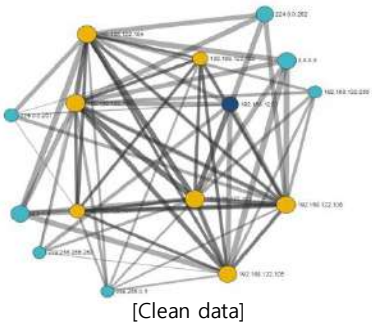




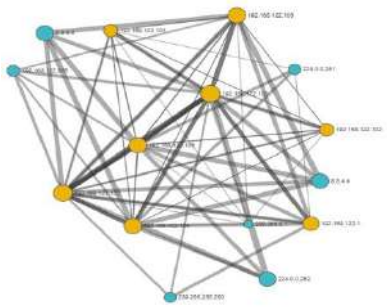
* https://www.openice.info

02 의료 네트워크 트래픽 분석 연구 - 네트워크 트래픽 분석 플랫폼
KIISS 2024

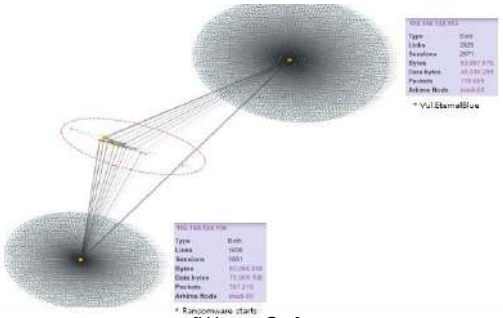
- OpenICE 데이터셋의 랜섬웨어 행위에 의한 Connection 형태 분석
  - ✓ 큰 방사형 2개가 대칭/비대칭 연결 : WannaCry, PowerGhost
  - ✓ 다각형 다수 연결 : Petya, BadRabbit



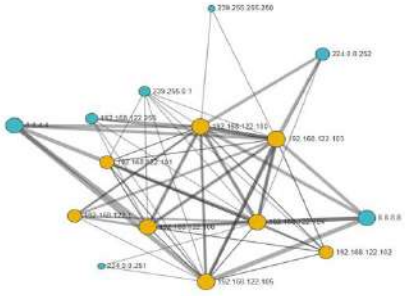
[Clean data]



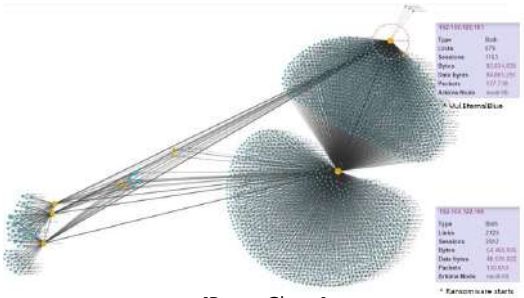
[Petya]



[WannaCry]

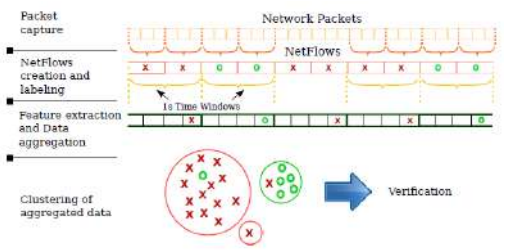


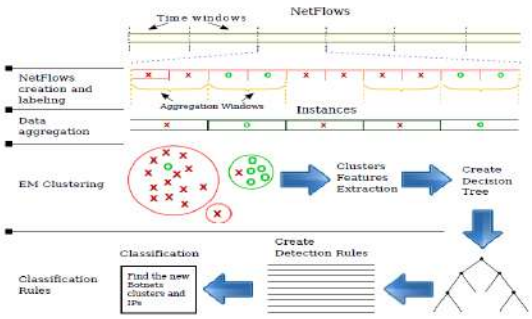
[BadRabbit]

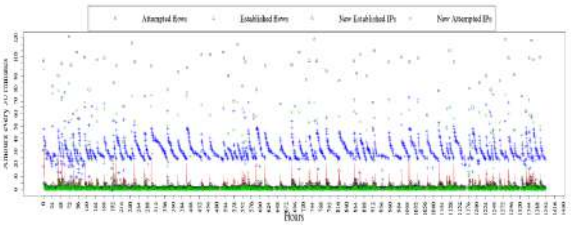


[PowerGhost]

- 봇넷 행위 분석 및 탐지 연구
  - ✓ SimDetect : Bot 간 유사성 탐지 방법
  - ✓ BClus : Botnet 행위 클러스터링 방법
  - ✓ Botnet C&C 채널 통신 행위 분석 : UDP, TCP, HTTP
  - ✓ CCDetector : Botnet 행위 모델 및 Markov Chains







Duration	Small Size			Medium Size			Big Size		
	Short	Medium	Long	Short	Medium	Long	Short	Medium	Long
Not enough data	1	2	3	4	5	6	7	8	9
Strongly periodic	a	b	c	d	e	f	g	h	i
Weakly periodic	A	B	C	D	E	F	G	H	I
Not periodic	r	s	t	u	v	w	x	y	z

```
Label:From-Botnet-V42-TCP-CC53-HTTP-Not-Encrypted
> 4-tuple : 147.32.84.165-31.192.109.167-80-tcp
> state model : 44DdddddddddfddddddddddddddddddddiddddddddDDddddddDddd
```

※ 출처 : Identifying, Modeling and Detecting Botnet Behaviors in the Network

- Network Flow/Packet 분석 기반 의료기기 자산 식별
  - ✓ 의료기기 통신 표준 프로토콜 분석
    - DICOM (Ultrasonic) : A_ASSOCIATE_RQ
    - HL7 (PMS) : ID_String
  - ✓ "1.2.840.10008.5.1.4.1.1.2" 형식의 OID 정보
    - DICOM 데이터 모델에서 사용하는 SOP Class UID로 보임
    - CT Image Storage (SOP Class Name에서 확인)

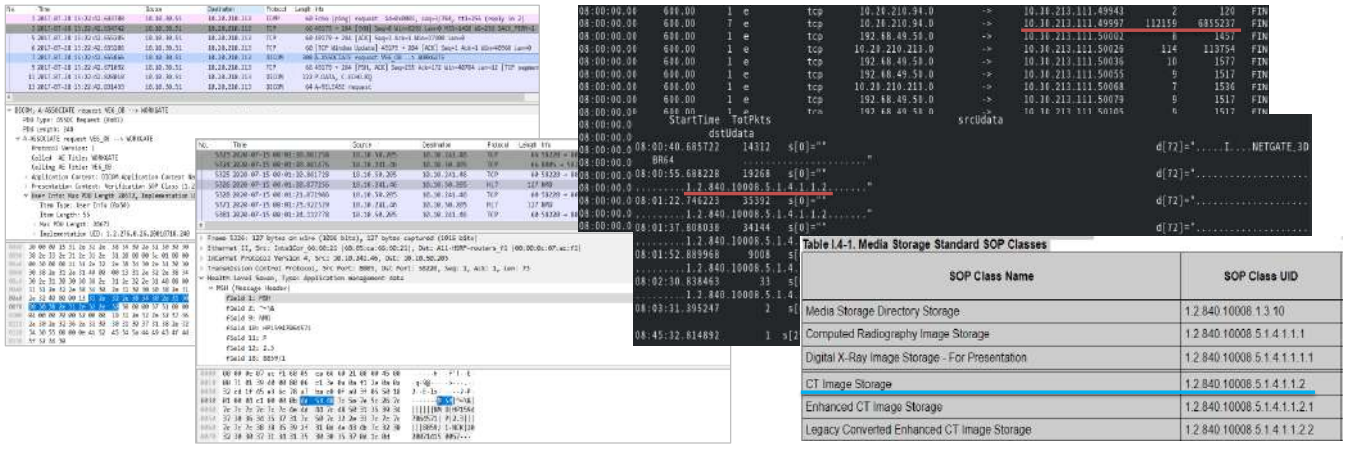
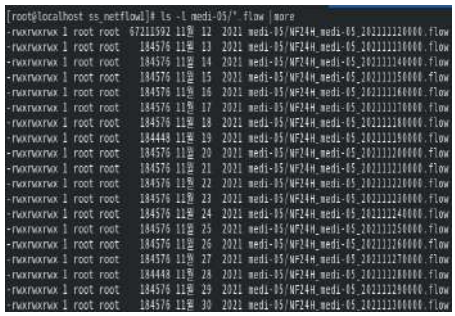
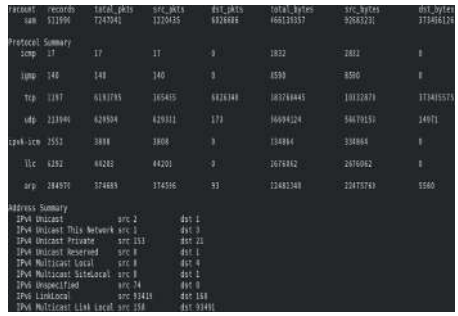


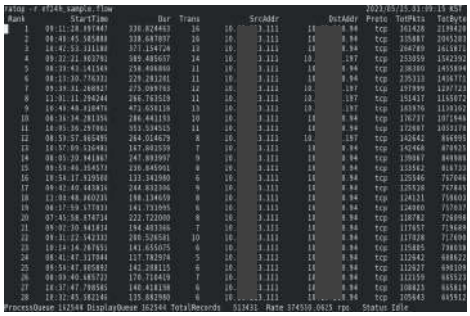
Table I.4-1. Media Storage Standard SOP Classes

SOP Class Name	SOP Class UID
Media Storage Directory Storage	1.2.840.10008.1.3.10
Computed Radiography Image Storage	1.2.840.10008.5.1.4.1.1.1
Digital X-Ray Image Storage - For Presentation	1.2.840.10008.5.1.4.1.1.1.1
CT Image Storage	1.2.840.10008.5.1.4.1.1.2
Enhanced CT Image Storage	1.2.840.10008.5.1.4.1.1.2.1
Legacy Converted Enhanced CT Image Storage	1.2.840.10008.5.1.4.1.1.2.2

- 병원 실망 환경의 Network Flow 기반 Metadata 수집 및 기본 분석
  - ✓ 수집 기간 : 2021.11.12 ~ 2023.05.02 (18개월+)
  - ✓ 수집 대상 : 의료기기 2대 (EKG, CT)
  - ✓ Metadata 수집량 (text data format)
    - 수집 파일 개수 : 1,062개 (센서#1 531개, 센서#2 531개)
    - 수집 파일 크기 : 67GB (센서#1 36GB, 센서#2 31GB)
  - ✓ Metadata 기본 분석 정보
    - Protocol summary : TCP (85.47%) >> UDP (8.69%) > ARP (5.17%)
    - TCP(26) : 10.20.0.0 -> 10.30.xxx.111.0 (99.71%)
    - UDP(15) : 10.30.0.0 -> 224.0.0.xxx:5353 (24.67%), 224.0.0.xxx:5355 (14.74%), IPv6 Pkts (24.58%)
    - ARP(2,585) : 10.30.xxx.158 -> 10.30.xxx.135 (9.22%), 10.30.xxx.1 -> 10.30.xxx.2~254 (0.19%~0.38%)

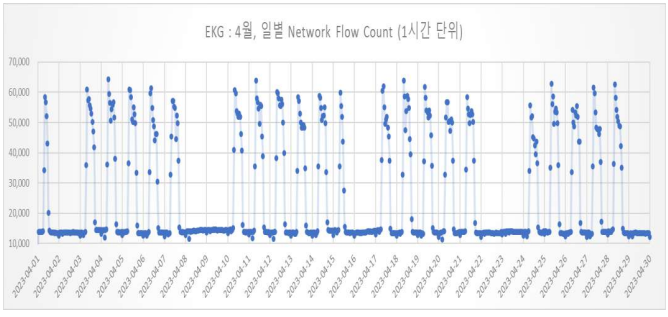


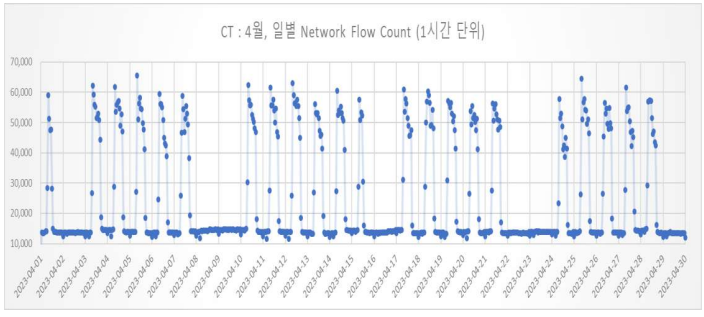


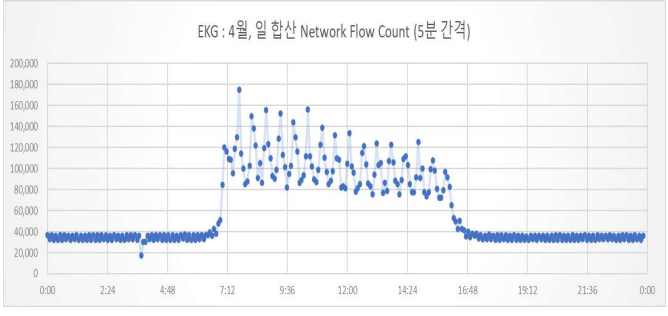


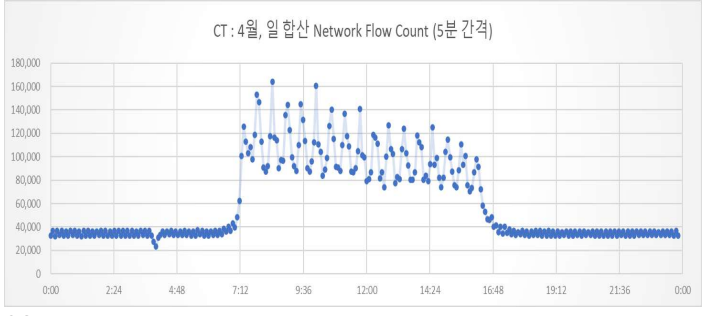
03 병원망 트래픽 이상행위 분석 사례 - 실망 트래픽 수집 및 분석 KIISS 2024

- Network Flow Count 분석 결과 요약
  - ✓ EKG 및 CT 대상 : 1개월 기간, 일별 & 일합산 Network Flow Count

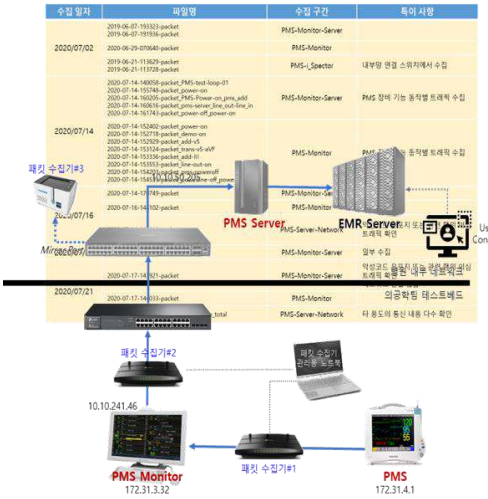






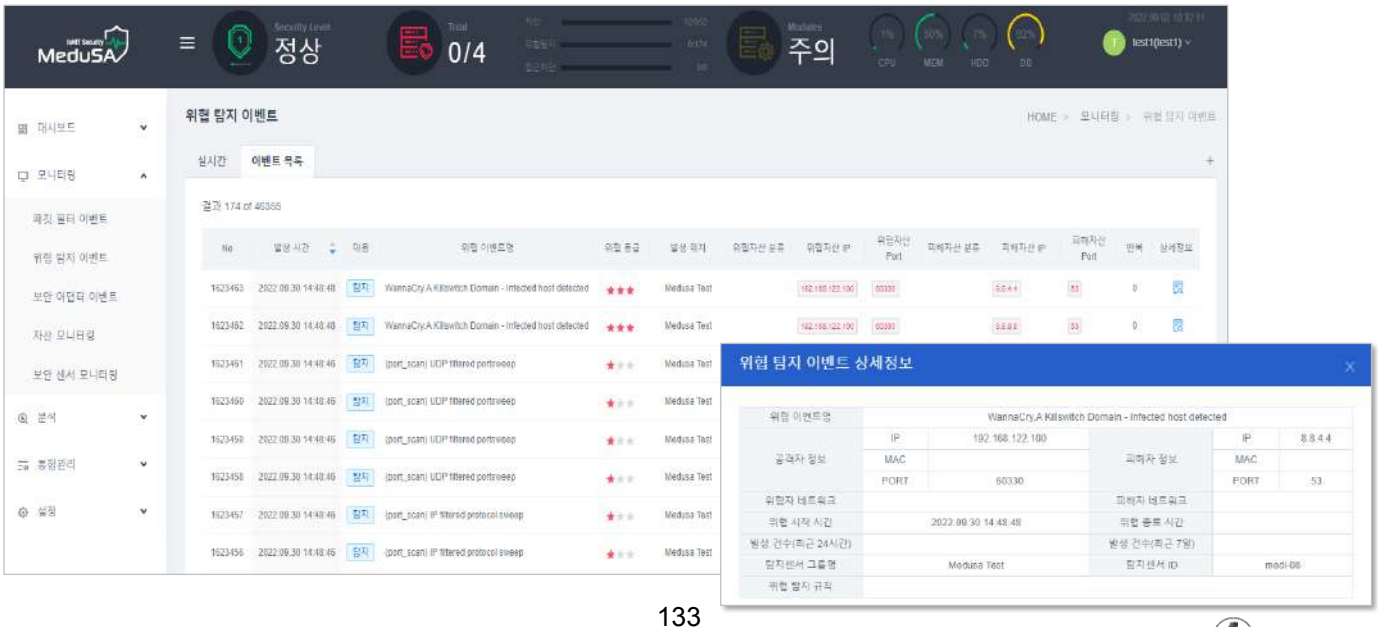


- (랜섬웨어 전파) 병원 실망 환경에서, 감염 대상 검사기 서버의 트래픽 분석 결과
  - ✓ 내시경 Monitor로부터 SMB 취약점 악용한 공격 발생 확인(Ransomware)
  - ✓ SMB 관련 포트 : 137(UDP), 138(UDP), 139(TCP), 445(TCP)
    - 445/TCP, EthernalBlue Tool (WannaCry, Petya)



03 병원망 트래픽 이상행위 분석 사례 - 실망 시범 운영 분석 KIISS 2024

- (랜섬웨어 전파) 시그니처 기반 네트워크 공격 탐지 결과
  - ✓ 공격탐지 이벤트명 : WannaCry.A Killswitch Domain
  - ✓ 감염된 시스템 행위 : src IP (192.168.122.100) → dst IP (8.8.4.4, 8.8.8.8)
    - DNS Query (외부 통신), www.iuqerfsodp9ifjaposdfjhgosurijfaewrwergwea.com



HUNESION

Copyright©2024 HUNESION Co. Ltd. All rights reserved.(무단 전재·재배포 금지)

자연어처리 기반 우울증 중증도 진단 분류 모델 개발

차의과학대학교 정보의학연구소

김명관 연구교수

Contents

자연어처리 기반 우울증 중증도 진단 분류 모델 개발

Ⅰ 개요

1. 배경 및 필요성

Ⅱ 방법

1. 데이터 구축
2. 도구 및 방법

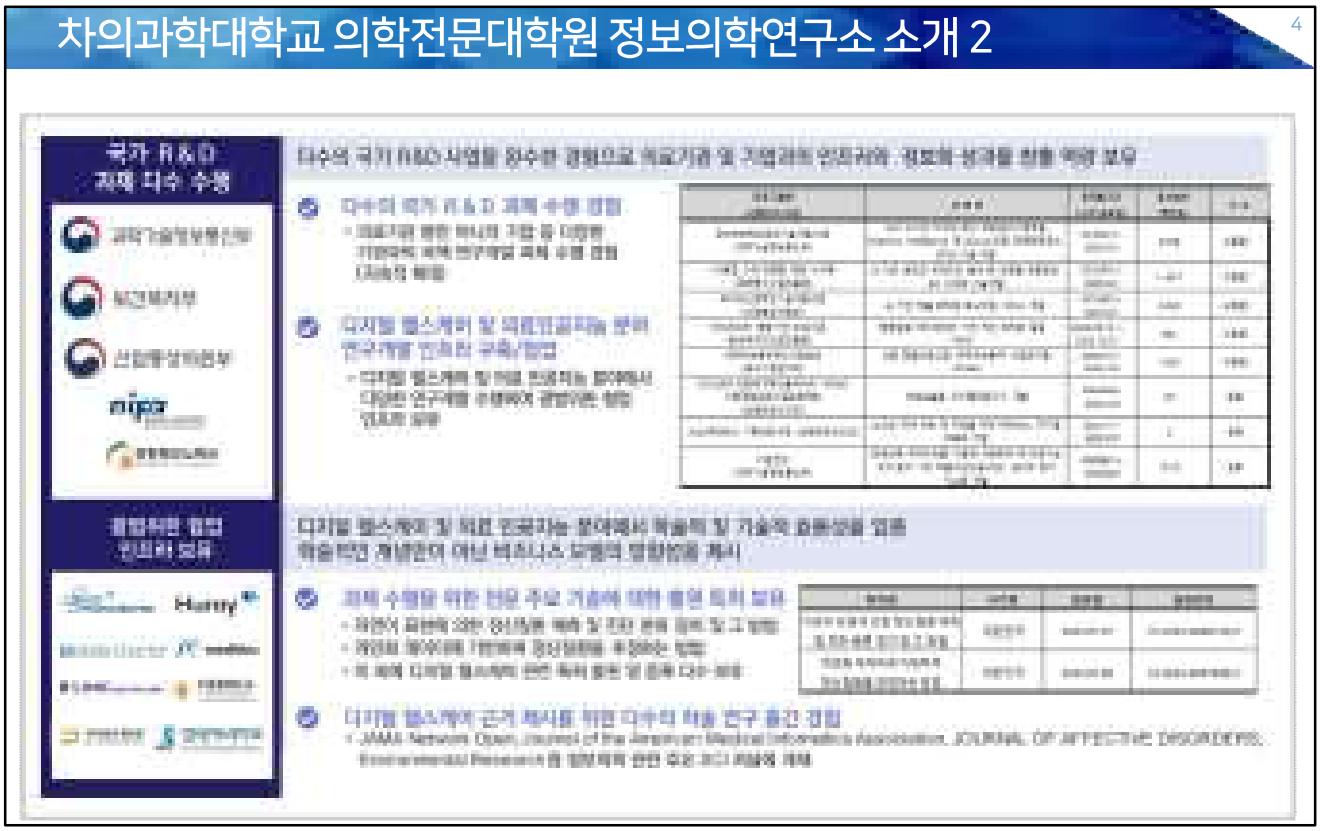
Ⅲ 결과

1. 모델 성능
2. 모델 시연
3. 특허 사항

Ⅳ 결론

1. 한계점
2. 활용방안

차의과학대학교 의학전문대학원 정보의학연구소 소개 1
3
연구소 현황 및 연구조직
차의과학대학교 정보의학연구소의 연구역량
연구소장
연구 총괄
인공지능
신약
빅데이터
디지털 헬스케어
약물부작용
데이터 사이언스
인포매틱스
SCIENTIFIC REPORTS
파트너십
한국IT서비스학회
중소벤처기업부
CHA
Huray
Mobile Doctor


차의과학대학교 의학전문대학원 정보의학연구소 소개 2
4
국가 R&D 과제 다수 수행
Huray

김명관 연구교수 소개

김 명 관 차의과학대학교 정보의학연구소, 연구교수

- curein@naver.com
- 경북대학교 보건학 및 보건의료정보학(박사)
- 보건의료 빅데이터, 임상 의료 데이터 자연어처리, 의료 데이터 유통 및 의료 데이터 가치 평가 전문가

주요 경력

- (현) 대한디지털헬스학회, 상임이사
- (현) 산업통상자원부, 데이터 큐레이션 기반 의료기기 기업 지원용 헬스데이터 유통 플랫폼 개발 및 실증, 연구개발과제 세부 책임자
- (현) 비정형 의료 데이터 가치 평가 전문가 (PCT 특허 보유)
- (현) 디지털 헬스케어 보건의료 빅데이터 인공지능 연구개발 분석 및 컨설팅 전문가
- (현) 과학기술정보통신부, 한국연구재단 의약학단, 평가위원
- (현) 질병관리청, 국립보건연구원 ,평가위원
- (전) 경북대학교 임상오믹스연구소, 산학협력전담 연구원
- (전) 한국생물공학회, 바이오데이터 엔지니어
- (전) 공중보건연구소, 연구소장, 보건의료 및 간호 분야 빅데이터 학술연구 분석 컨설팅
- (전) 경북대학교 예방의학교실, 대구권역 지역사회건강조사, 연구원
- (전) 경북대학교 간호과학연구소, 청소년 흡연예방 및 금연사업단, 연구원

NLP based Mental Diagnosis

Chapter

자연어처리 기반 우울증 중증도 진단 분류 모델 개발

Ⅰ 개요

1. 배경 및 필요성

1. 배경 및 필요성

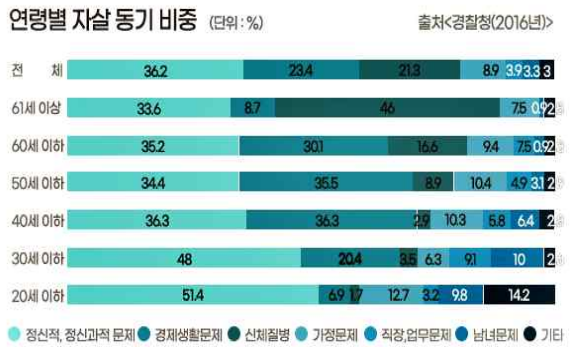
✓ 우울증은 자살 시도의 관문이 될 수 있기에 전문가의 적절한 치료가 필요한 질환

➢ 2주 이상 연속적인 증상으로 곤란한 경우

우울한 감정 / 흥미 상실 / 식욕 저하 / 피로감 / 수면 변화 / 자살 생각

"우울증"으로 진단

➢ 자살시도자의 55.6%가 우울증으로 진단



1. 배경 및 필요성

✓ 대한민국, 최근 5년(2017-2021)간 전 연령층의 우울증 환자수 증가

출처 : 건강보험심사평가원. 최근 5년(2017-2021년) 우울증과 불안장애 진료현황 분석

1. 배경 및 필요성

✓ 현대 의학에서의 정신질환 진단 방식의 한계

정신질환 환자

정신건강 의료기관 방문하여 PHQ-9 , GAD-7, PSS 등 문진표 작성

의사 진료 시 정신적 호소 증상 기록 및 판단

[초진 진단]

"조울증"

기분안정제 처방

[재진 진단]

"우울장애"

항우울제 처방 또는 인지행동치료 처방으로 변동

정신과 진료 특성상 정신질환 환자의 **증상을 호소하는 말에 의해 진단**

주관적 요소가 강하여 **초진 대비 재진 진단 변경에 의해 처방약과 치료법이 변경**되는 경우 존재

정신질환은 정확한 진단이 난해한 특성 때문에 적절한 **초기 대응이 힘든 분야**

숙련된 정신건강의학과 전문의도 초진 대비 재진 **진단 일치도는 평균 40% 수준**

Chapter

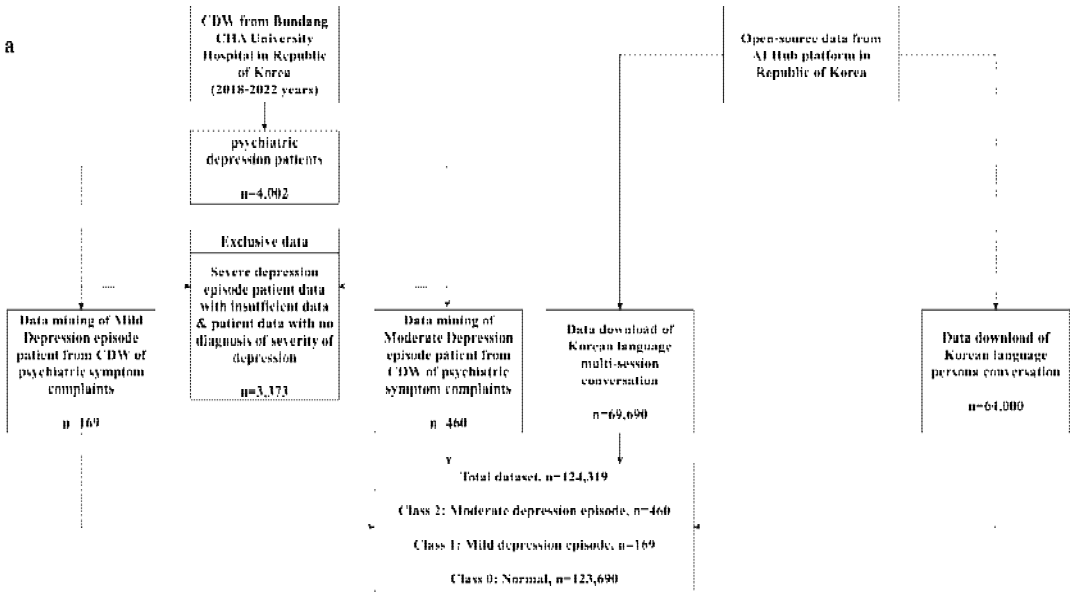
자연어처리 기반 우울증 중증도 진단 분류 모델 개발

Ⅱ 방법

1. 데이터 구축
2. 도구 및 방법

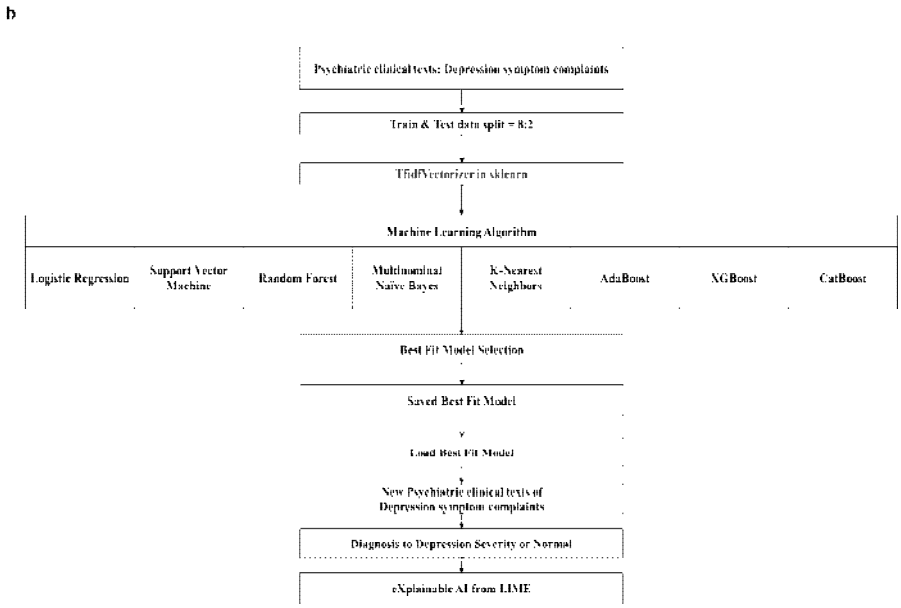
1. 데이터 구축 – 우울증 중증도

✓ IRB : CHAMC 2023-12-018, 분당차병원 CDW & AI HUB 멀티세션 한국어 대화 및 페르소나 데이터 활용



2. 도구 및 방법 – 우울증 중증도

✓ 인공지능 모델 학습 과정



Chapter

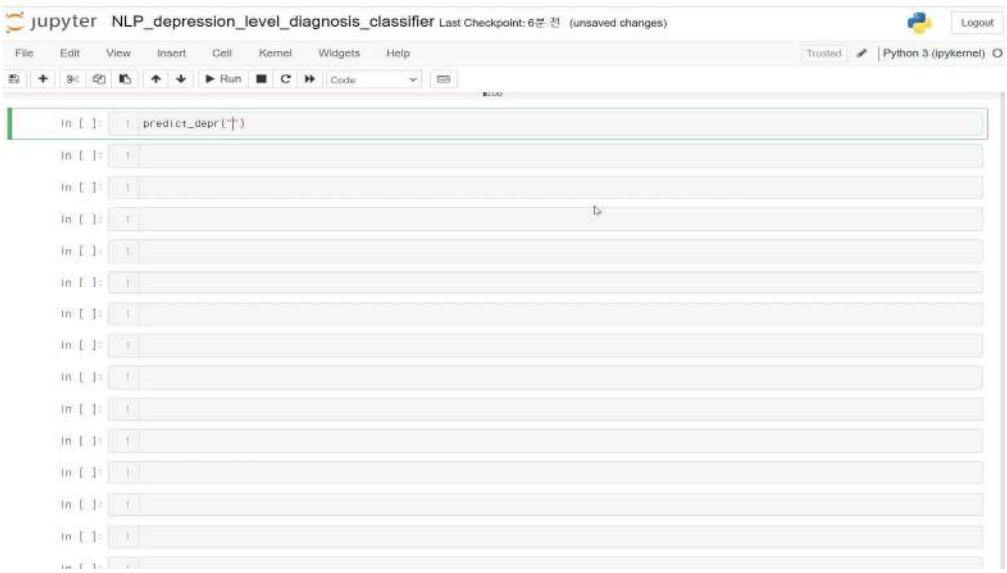
자연어처리 기반 우울증 중증도 진단 분류 모델 개발

III 결과

1. 모델 시연
2. 모델 성능
3. 특허 사항

1. 모델 시연

✓ 우울증 중증도(정상, 경미한 우울증, 중등도 우울증) 진단 분류 모델 시연



2. 모델 성능

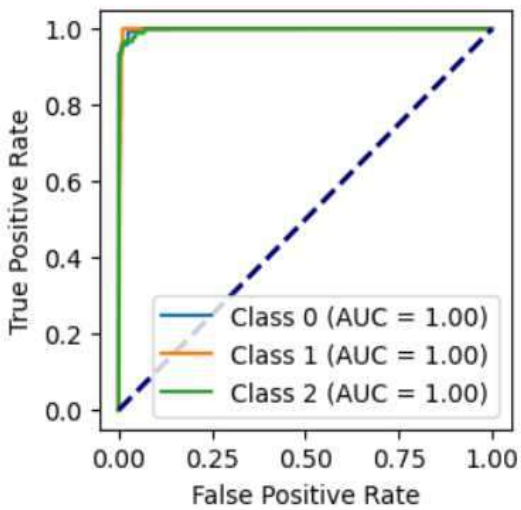
✓ 우울증 중증도 진단 분류 모델 성능 비교

Classifier	Accuracy	Precision	Recall	F1 Score	AUC	Total Performance Average	Rank
XGBoost	99.7	99.6	99.7	99.6	>99.9	99.7	1
Logistic Regression	99.7	99.6	99.7	99.6	>99.9	99.7	1
Support Vector Machine	99.7	99.6	99.7	99.6	>99.9	99.7	1
CatBoost	99.6	99.5	99.6	99.5	99.9	99.6	2
Multinominal Naïve Bayes	99.6	99.1	99.6	99.3	99.9	99.5	3
Random Forest	99.6	99.5	99.6	99.5	97.6	99.2	4
AdaBoost	99.6	99.4	99.6	99.5	95.7	98.8	5
K-Nearest Neighbors	63.7	99.5	63.7	77.4	76.3	76.1	6

2. 모델 성능

✓ 우울증 중증도 진단 분류 모델 성능

XGBoost - Performance	
Accuracy	99.7
Precision	99.6
Recall	99.7
F1 Score	99.6



3. 특허 사항

✓ 우울증 중증도 진단 분류 모델 특허 사항

관 인 생 략

출 원 번 호 통 지 서

출 원 일 자	2024.03.07
특 기 사 항	심사청구(무) 공개신청(무)
출 원 번 호	10-2024-0032298 (접수번호 1-1-2024-0257187-00) (DAS접근코드D7C8)
출원인 명칭	차의과학대학교 산학협력단(2-2005-014743-0)
대리인 성명	특허법인태백(9-2008-100101-3)
발명자 성명	김명관 한현욱 강은석 왕다운 권영훈 박준호
발명의 명칭	자연어처리 및 설명 가능한 인공지능에 의한 우울증 중증도 수준 예측 및 진단 분류 장치 및 그 방법

특 허 청 장

Chapter

자연어처리 기반 우울증 중증도 진단 분류 모델 개발

Ⅳ 결론

1. 한계점
2. 활용방안

1. 한계점

❑ 데이터셋 구축 어려움	무조건적으로 많은 데이터 투입보다 확실하게 우울증을 표현하는 정밀한 텍스트 비정형 데이터 학습 필요.
❑ 토큰화 어려움	한국어 특성상 영어에 비해 토큰화를 통한 형태소 분석을 반영한 인공지능 모델 고도화의 어려움
❑ LLM 활용 난해	정신건강의학적 비정형 텍스트에서 우울증을 표현한 정밀 제어 데이터가 필요하므로 거대 언어 모델 활용이 쉽지 않음

2. 활용방안

➤ 정신건강의학과 CDSS 모듈로 사용

➤ 정신건강의학과 진료 시 더욱 신속하고 정확하게 초진만으로도

- ✓ 우울증 수준을 효율적으로 파악 가능
- ✓ 환자가 호소하는 증상에서 우울증 원인 요소를 의료진이 객관적으로 추출하여 시각적으로 파악할 수 있도록 보조
- ✓ 의료진이 정밀하고 신중하게 근거 기반 정신과 치료를 할 수 있도록 보조

➤ NLP 기반 진단 보조 시스템 확산

➤ 정신건강의학과 특성상 대체로 비정형 텍스트 데이터 활용

- ✓ 불안장애/조울증/조현병도 진단 보조 가능
- ✓ CDSS 모듈로 결합되면 정신건강의학과 인공지능 지능형 SW로서 케이스별 효과적인 치료 정보 제공 보조 가능

NLP based Mental Diagnosis

자연어처리 기반 우울증 중중도 진단 분류 모델 개발

감사합니다

In-context learning for bioelectrical signals

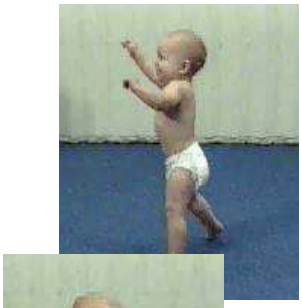
Hosik Choi

Dept. of Urban Big Data Convergence & Dept. of AI
University of Seoul

May/2024

Introduction

Gait data



1Y
Wide based
Stiff knee
Toe-heel

Maturation

3Y6M
Adult pattern:
Heel-toe
Normal knee

Gait data

Development of gait after 3y6m

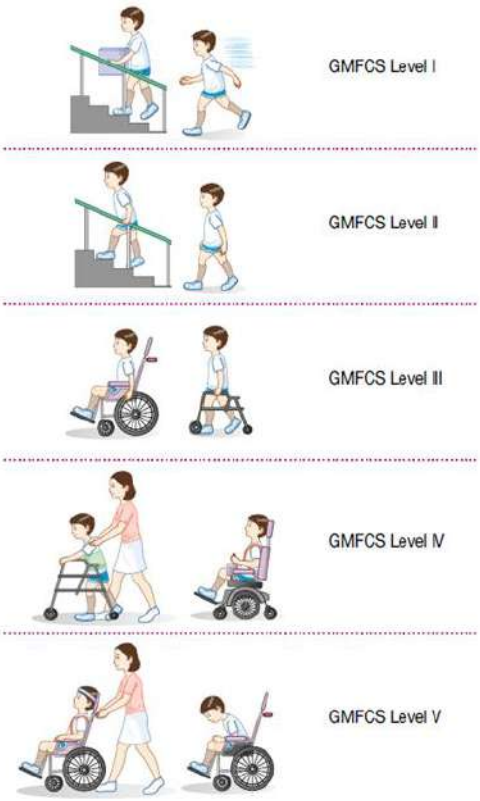
- Cadence : ↓
- Step length : ↑
- Velocity : becames constant

Gait data

❖ 뇌성마비: 뇌의 손상으로 인해 발생하는 운동과 자세의 장애

❖ GMFCS (Gross Motor Function Classification System)

: 뇌성마비 환자의 총 운동 기능을 관찰하여 심각도를 분류하는 시스템



- 1단계: 특별한 제한 없이 보행 가능
- 2단계: 제한이 있는 보행. 점프나 달리기의 어려움
- 3단계: 손으로 잡고 사용하는 보행 보조기를 사용하여 보행
- 4단계: 혼자 앉을 수 있으나 독립적 이동은 제한
- 5단계: 몸통과 목을 가누기 어려움

Gait data

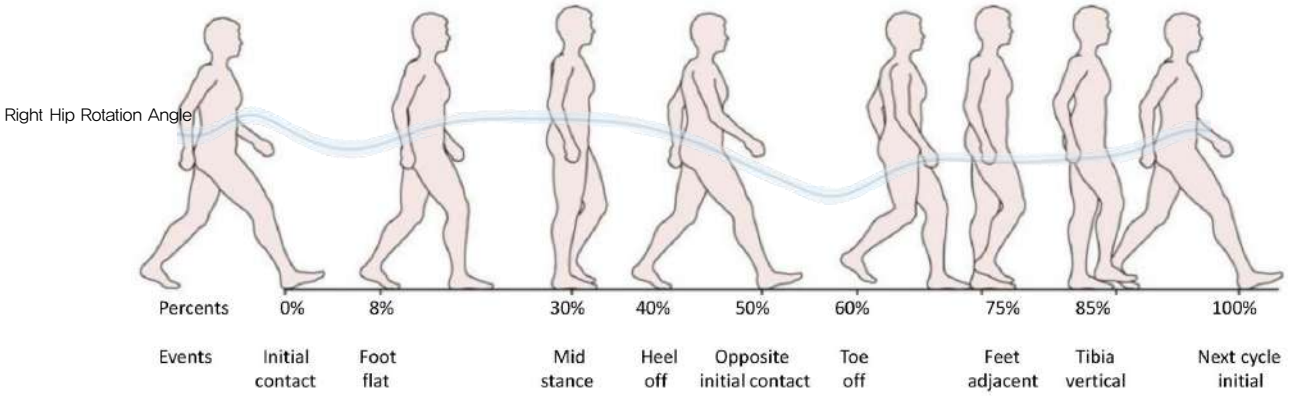
❖ 3D 보행 분석

- Gait analysis is the systematic study of human motion by instrumentation for measuring body movements.
- 3D gait analysis is the gold standard for gait analysis.
- Seoul National University Bundang Hospital

Gait data

❖ 보행 주기 (Gait Cycle) : 한 발이 지면에 닿을 때 부터 시작하여 다시 지면에 닿을 때 까지를 말하며

입각기(stance phase)와 유각기(swing phase)로 나누어짐

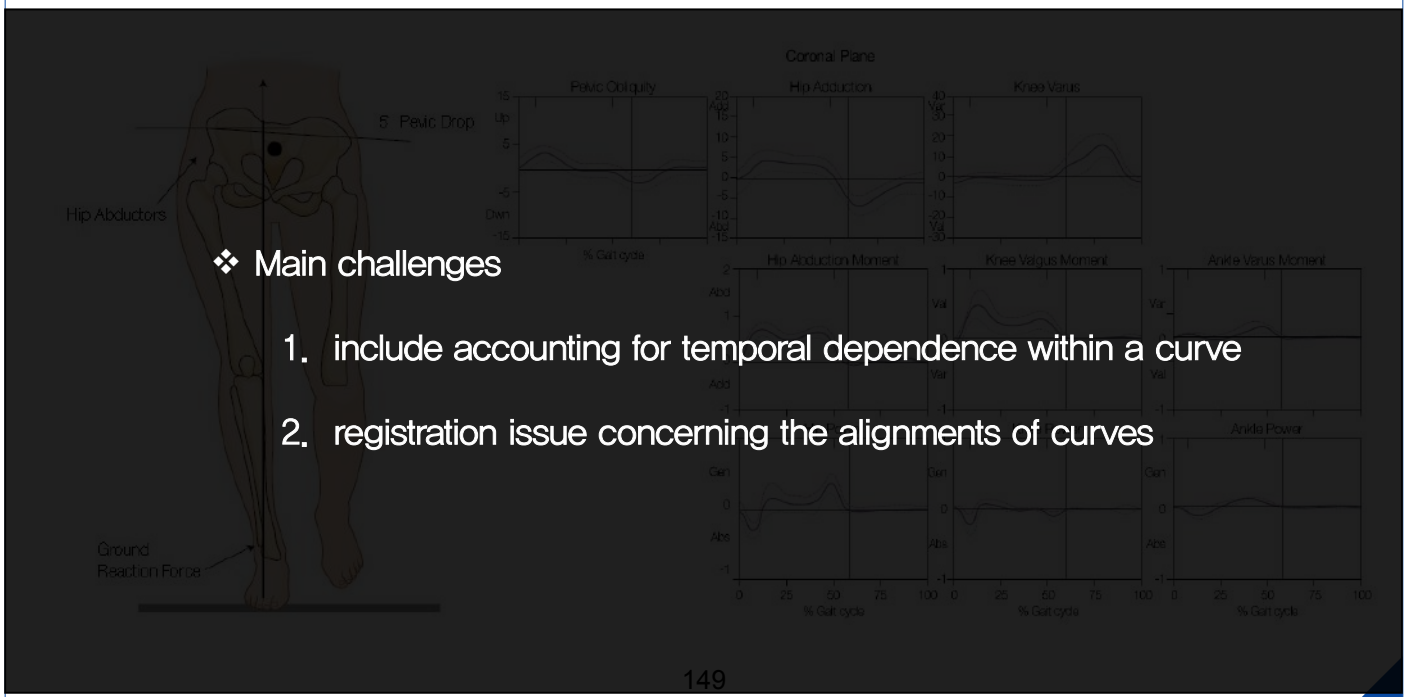


Periods	Loadding response	Mid stance	Terminal stance		Pre swing	Initial swing	Mid swing	Terminal swing
Tasks	Initial double limb support	Initial single limb support			Second double limb support	Second single limb support		
Phase	Stance phase					Swing phase		
Per				Push off				
Cycle	One cycle							

Functional data

❖ 함수형 데이터(functional data)

: 연속체에 걸쳐서 변화하는 곡선으로 구성된 데이터



❖ Main challenges

1. include accounting for temporal dependence within a curve
2. registration issue concerning the alignments of curves

Functional data Analysis

❖ Machine learning methods and Deep learning methods (CNNs)

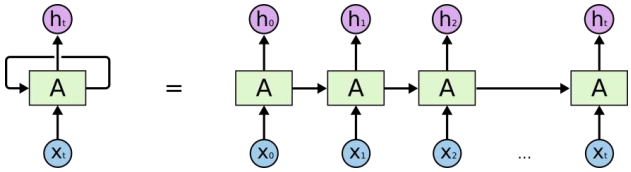
- ex) 로지스틱 FDA(Muller, 2005)
  X(t): t에 따른 functional predictor, Y: 그룹 변수,
  $\beta(t)$: t에 따른 회귀계수 함수

$$\log \frac{\mathrm{pr}(Y=1|X)}{\mathrm{pr}(Y=0|X)} = a_0 + \int_I X(t)\beta(t)\,dt.$$

➪ 관측치가 통계적으로 독립적임을 가정

❖ Deep learning methods (RNNs)

- 반복 학습을 통해 시간적 의존성 명시적 설명 가능



➪ 과거에서 미래로의 단방향 흐름만을 모델링

❖ BERT

- large-scale language model with deep bi-directional learning based on the **encoder** of a transformer

➪ 1) Alignment 과정을 제거하고 2) 특정 변동구간에 대한 이상치 처리 용이

9

Large Scale Model

Changes in nlp **model size** over time

Training Dataset Size for Model

❖ **Data Abundance** + Advance in Computing Hardware

➪ zero-shot 및 few-shot 학습에서 성공을 거두었음

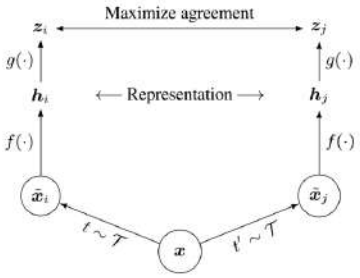
10

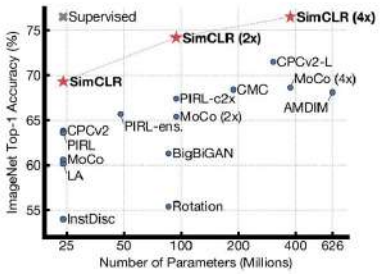
Data Augmentation

❖ 그러나, 보행데이터 수집의 어려움

- 의료데이터는 **개인**의 건강과 관련된 복잡한 요소와 민감한 정보를 포함
- 특정한 장비를 통해 측정되는 데이터로 **수집과 처리 비용이 높음**

❖ 데이터 증강은 BERT에 대한 교육 데이터의 크기와 품질을 효과적으로 개선하는 방안이 됨





- SimCLR과 같은 self-supervised learning 기반의 애플리케이션에서 데이터 증강을 통한 성능 향상 달성
- Computer vision 분야, NLP 분야에서 제시한 증강 기법은 연속적인 scale을 가진 함수형 데이터에 적합하지 않음
  - **복합 다항 벌점 함수** Trend-filtering을 통해 데이터의 패턴을 모델링 하는 새로운 증강 기법 제안

- Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In *International conference on machine learning* (pp. 1597-1607). PMLR.

Self-supervised Learning

❖ 실제 자료 + 증강된 자료로 BERT를 학습시킬 시 이점

1. 토큰화를 통해 데이터를 범주형 자료로 변환
   1. 정렬 프로세스 제거
   2. 이상치 처리 용이
2. Two stage 학습 : 사전학습 〉 미세 조정
   1. 증강된 데이터가 미세조정에서 일반화 성능 향상
   2. 사전학습의 새로운 목적 함수로 Noise Contrastive Estimation 추가하여 학습 효율 향상

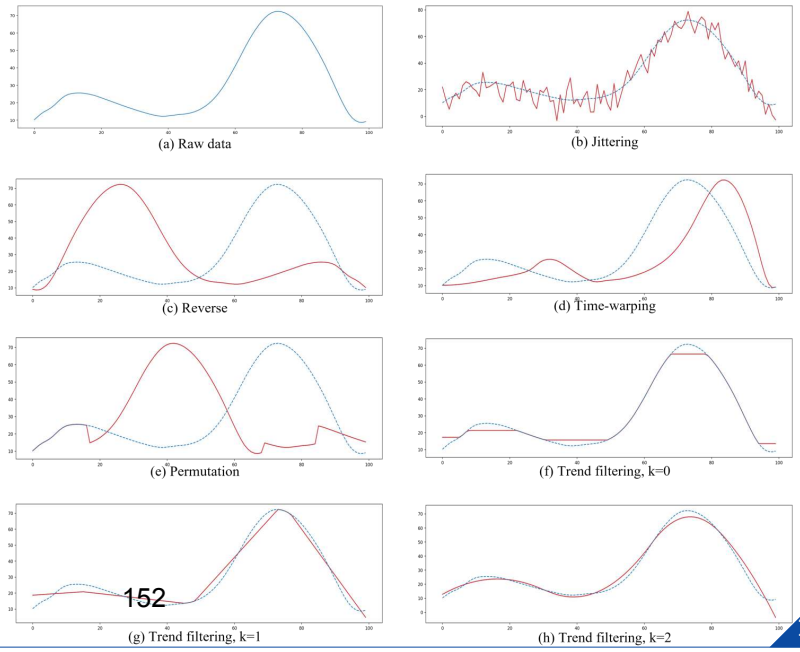
Background

Augmentation Methods

- $m \in \{1, \ldots, M\}$: Index of an observed function of gait data
- $t \in \{1, \ldots, T\}$: T be the number of time points
- $x^{m(t)} \in \mathbb{R}$: the mth observed curve measured at time t

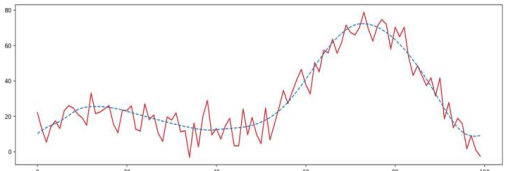
❖ Aumentation methods

- Jittering
- Reverse
- Time–warping
- Permutation
- Trend filtering

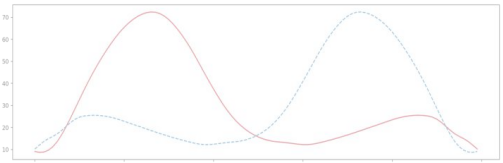


Augmentation Methods

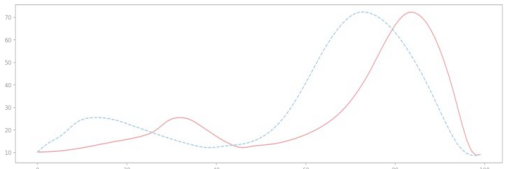
- 평균 μ, 표준편차 σ를 가진 gaussian noise를 원 데이터에 곱하거나 가산하여 random irregular fluctuation을 생성



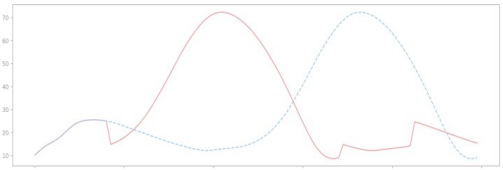
(b) Jittering



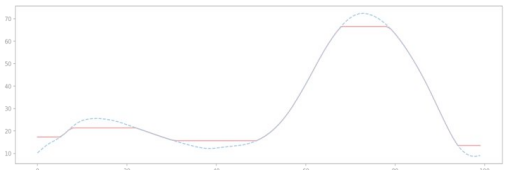
(c) Reverse



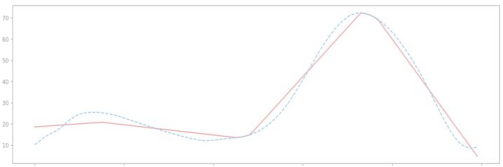
(d) Time-warping



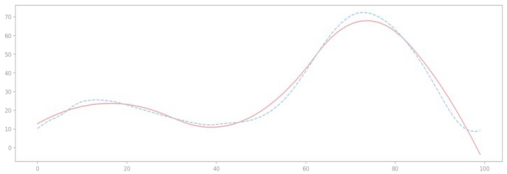
(e) Permutation



(f) Trend filtering, k=0



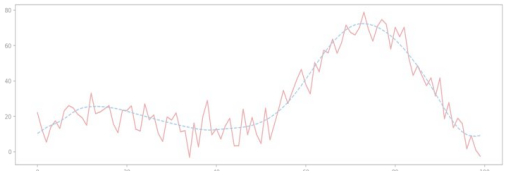
(g) Trend filtering, k=1



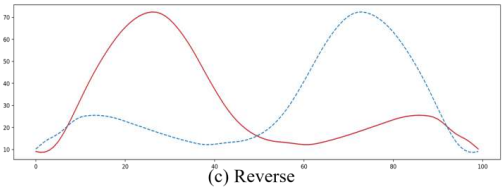
(h) Trend filtering, k=2

Augmentation Methods

(a) Raw data

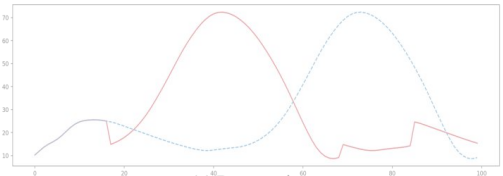


(b) Jittering

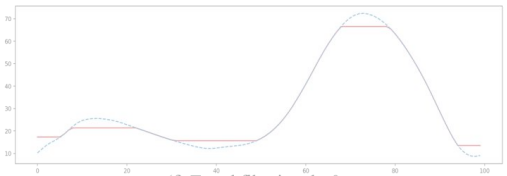


(c) Reverse

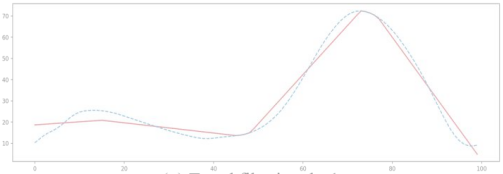
- 시계열을 반전 (reverse)
- Computer vision의 flip 방법에 영감을 받아 $\{x^{m(1)}, \dots, x^{m(T)}\}$를 $\{x^{m(T)}, \dots, x^{m(1)}\}$ 로 변환.



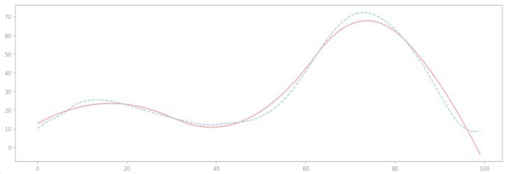
(e) Permutation



(f) Trend filtering, k=0



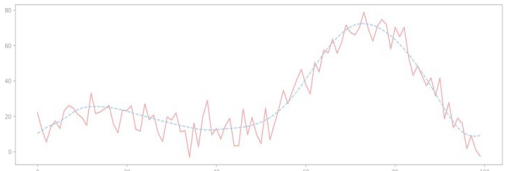
(g) Trend filtering, k=1



(h) Trend filtering, k=2

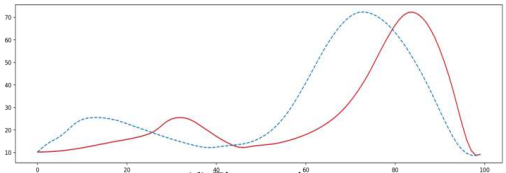
Augmentation Methods

(a) Raw data

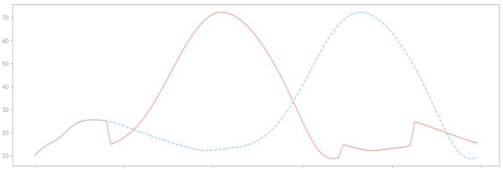


(b) Jittering

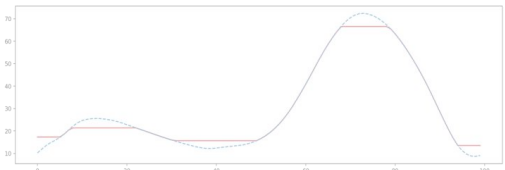
- 보행자의 보폭에 대한 시간적 특성 반영
- 시간 간격을 왜곡함으로써 새로운 데이터를 생성



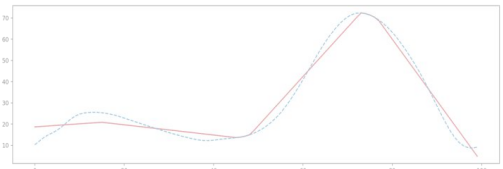
(d) Time-warping



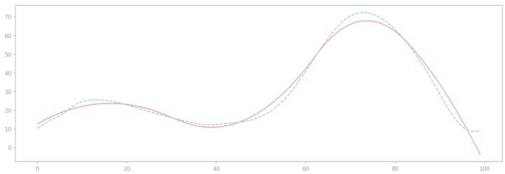
(e) Permutation



(f) Trend filtering, k=0



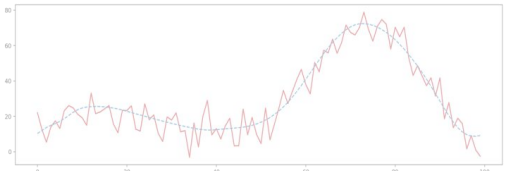
(g) Trend filtering, k=1



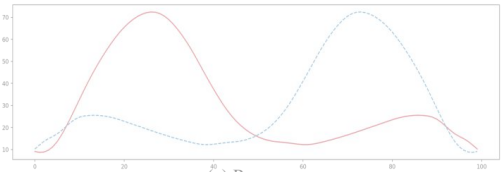
(h) Trend filtering, k=2

Augmentation Methods

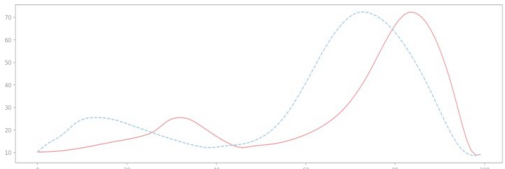
(a) Raw data



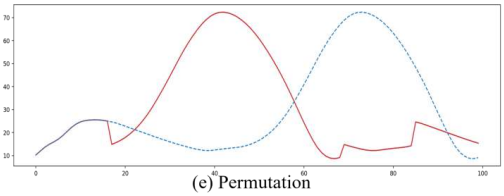
(b) Jittering



(c) Reverse

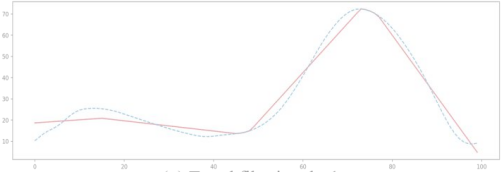


(d) Time-warping

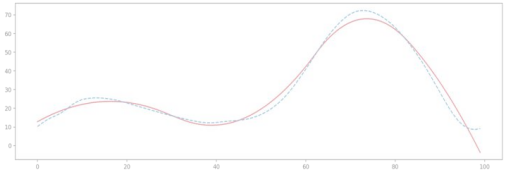


(e) Permutation

- Function 내 시간 위치를 random하게 교란하는 방법
- 관측 구간을 분할하고 이를 무작위 재배치



(g) Trend filtering, k=1



(h) Trend filtering, k=2

Trend-filtering

❖ Trend Filtering (Kim et al. 2009)

- 조각별 선형 함수를 사용한 효율적인 추세 추정 방법을 제안
- l_1 Trend-filtering : nonparametric regression approach

❖ Minimize the Objective Function , $k \geq 0$

$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^T} \frac{1}{2} \sum_{t=1}^{T} (x^m(t) - \beta_t)^2 + \frac{T^k}{k!} \lambda \left\| D^{(k+1)} \beta \right\|_1 ,$$

- λ : regularization parameter
- $D^{(k+1)} \in \mathbb{R}^{(n-k-1) \times n}$: $k+1$의 discrete difference operator

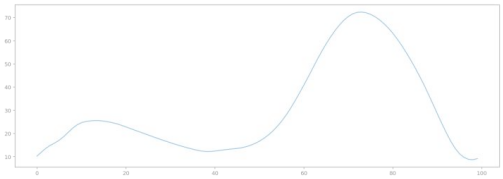
❖ where $k = 0$, penalty term for fused lasso : $\left\| D^{(1)} \beta \right\|_1 = \sum_{t=1}^{T-1} |\beta_t - \beta_{t+1}|$

❖ For $k \geq 1$, $D^{(k+1)} = D^{(1)} D^{(k)}$

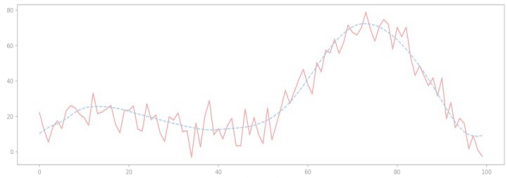
- If $D^{(2)}$ matrix is

$$D = \begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \end{bmatrix}$$

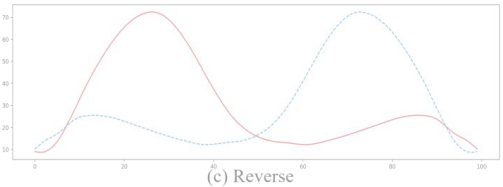
Augmentation Methods



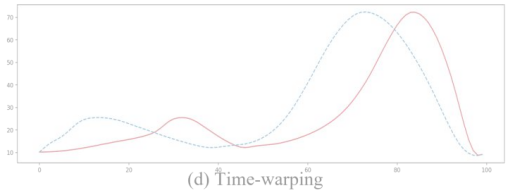
(a) Raw data



(b) Jittering

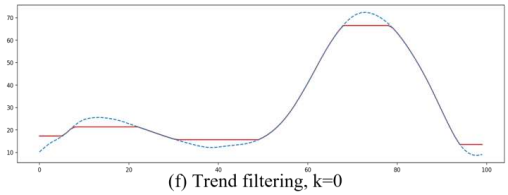


(c) Reverse

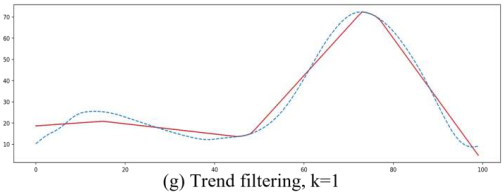


(d) Time-warping

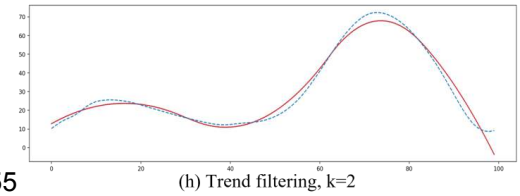
- Trend filtering의 k 차수에 따른 변화
- $k \in \{0, 1, 2\}$



(f) Trend filtering, k=0



(g) Trend filtering, k=1



(h) Trend filtering, k=2

Transformer

- Attention mechanism의 등장
- 인코더의 입력 시퀀스 참조
- RNNs의 병렬처리 문제 해결, self-attention 제안

BERT

- Transformer의 Encoder
- Two objectives in pre-training: MLM, NSP

이전 — 2017 — 2018 — 2019 — 2020

RNNs

- 순차적으로 시퀀스 데이터를 처리
- 긴 입력 시퀀스를 fixed context vector로 압축하며 정보 손실

ELMo, GPT

- NLP에서 pre-train 도입
- shallow bidirectional

GPT3

- Model을 깊게 쌓아 Parameter수를 늘림
- Zero-shot learning

❖ Multi layer bidirectional transformer encoder

- $BERT_{medium}(L=8, H=512, A=8)$, total parameter = 41M

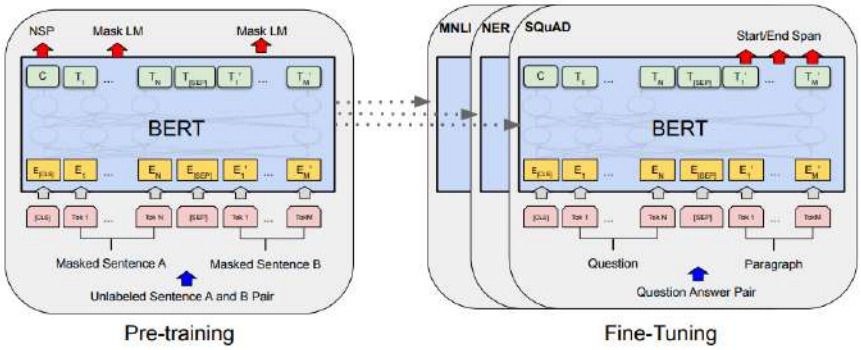
  $(L: number\ of\ transformer\ blocks, H: hidden\ size, A: number\ of\ attention\ heads)$

- Pre-training: unlabeled data training

  Task #1 : Masked Language Modeling (MLM)

  Task #2 : Next Sentence Prediction (NSP)

- Fine-tuning: pre-train BERT로 weight 초기화, downstream task에서 labeled data 사용



$$\min_{\theta,\xi} \alpha_1 \ell_{\text{mlm}}(\mathrm{x};\theta) + \alpha_2 \ell_{\text{nsp}}(\mathrm{x}, \mathrm{y}_{\text{nsp}};\theta) + \alpha_3 \ell_{\text{sup}}(\mathrm{x}, \mathrm{y}_{\text{sup}};\xi)$$

For example, $(\alpha_1, \alpha_2, \alpha_3) = (\frac{1}{2}, \frac{1}{2}, 0)$ in the pre-training step, $(\alpha_1, \alpha_2, \alpha_3) = (0, 0, 1)$ in the fine-tuning step

Proposed Method

Pre-processing

❖ Tokenization

1. 0과 1사이의 각 function ($m \in \{1, \ldots, M\}$)을 표준화 :

$$\tilde{x}^m = \frac{x^m - \min(x^m)}{\max(x^m) - \min(x^m)}$$

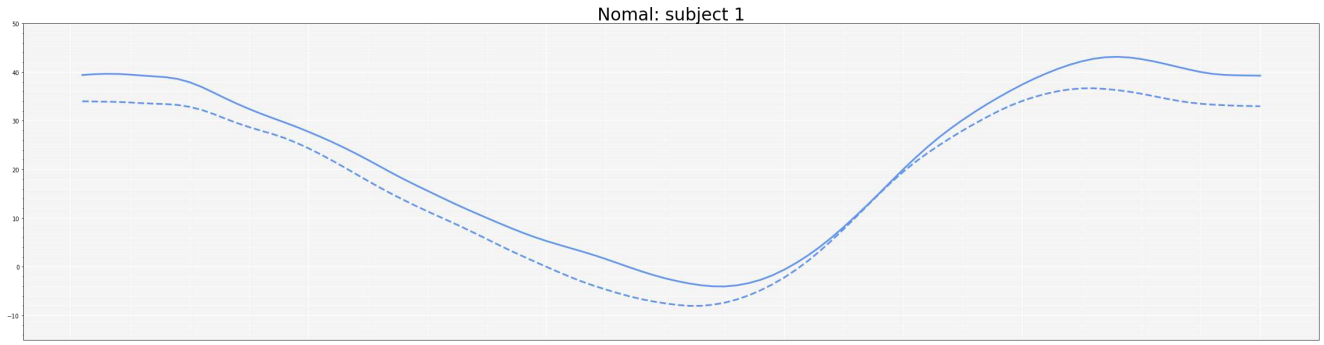
2. 등구간화 하여 다음 사이의 문자로 대체

“ a - z Z - A ”

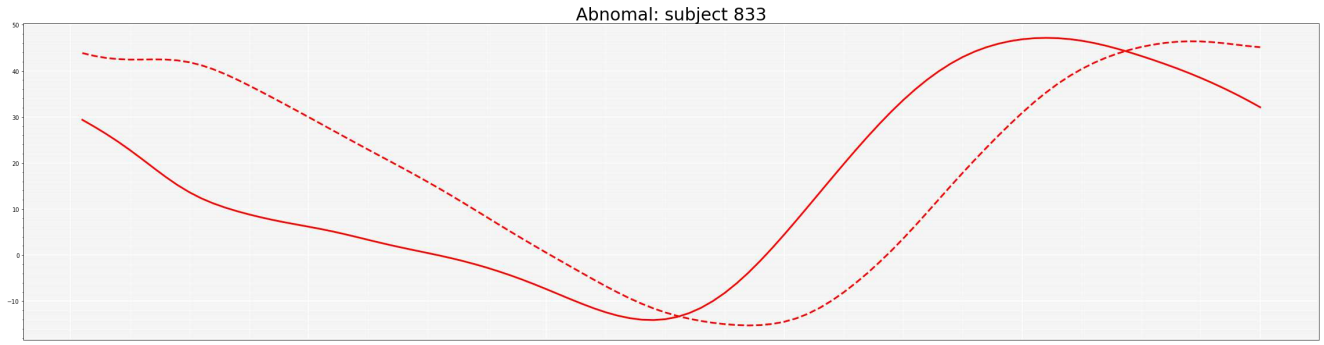
❖ 52개의 토큰과 5개의 특수 토큰을 가진 단어사전 구축

- 특수 토큰: [CLS], [SEP], [UNK], [MASK], [PAD]

Pre-processing



EC EC EC EB CA Bz zy yw wu ut tr rp qo pm ol nk lj lj kj lk nm pp ss vv yy BA DB ED GD GD GC FC EB EB

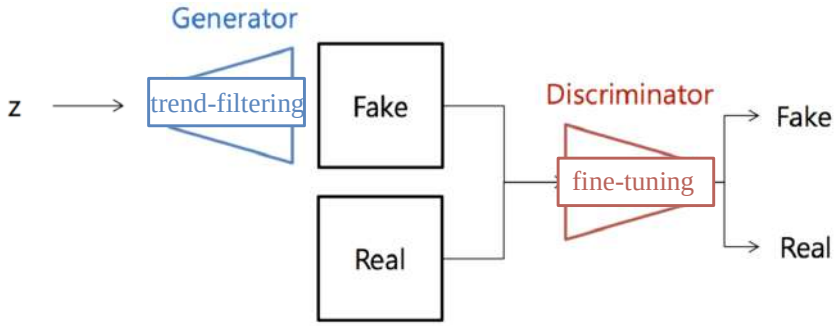


AG yG vG sF rE qC pB pz ox nv mt lr kp jm hk gj gh gg if mf qg ui yk Co Fs Gw lz lC lF HG GH EH DH BH

Pre-processing

❖ Noise contrastive learning

- 비선형 로지스틱 회귀분석을 수행하여 관측된 데이터와 인위적으로 생성된 노이즈를 구별함



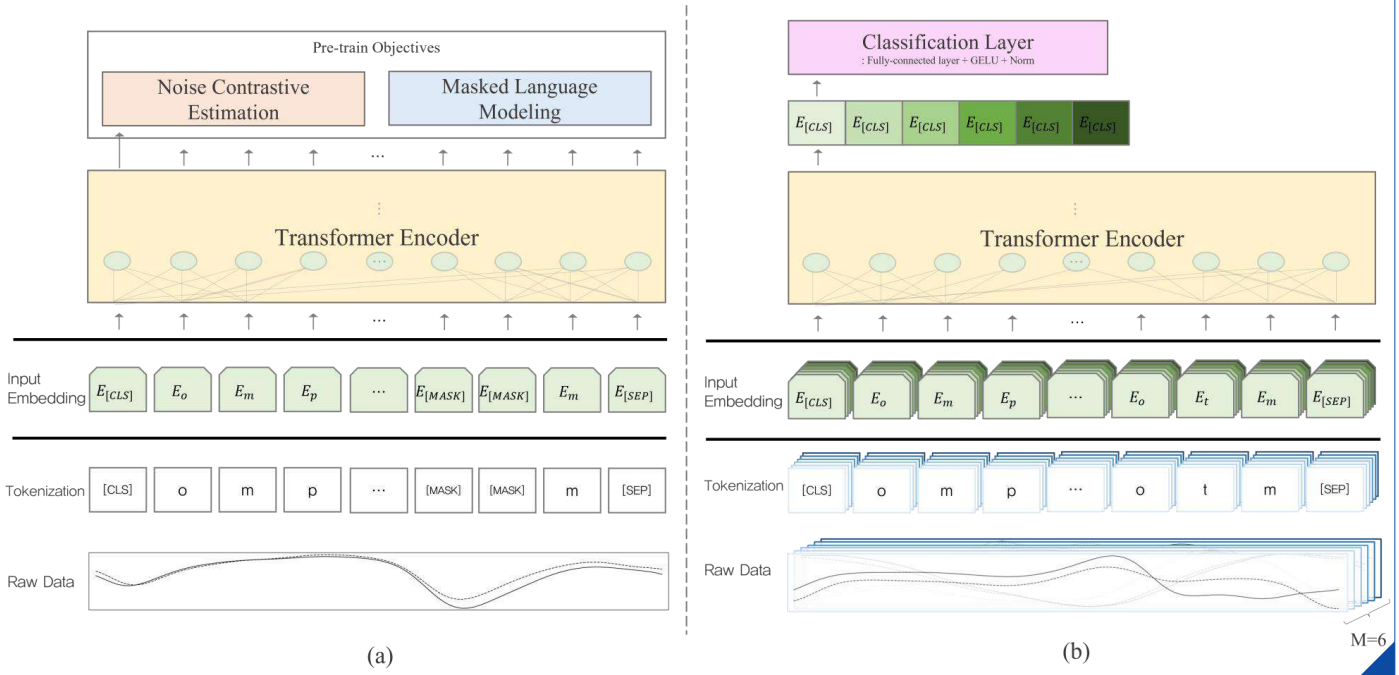
- GAN 관점에서 증강된 데이터를 적대적으로 여기고 분류 하는 문제를 새로운 사전학습 task로 추가
  - 학습의 효율성 향상
    1. MLM task: local contextual learning
    2. NCE task: global contextual learning
- GAN → 생성 네트워크에 대한 unstable 학습 개선

• Gutmann, M., & Hyvärinen, A. (2010, March). Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In Proceedings of the thirteenth international conference on artificial intelligence and statistics (pp. 297-304). JMLR Workshop and Conference Proceedings.

Model architecture

❖ Two stage learning, $\alpha_1, \alpha_2, \alpha_3$ ($\geq 0, \Sigma_{i=1}^{3}\alpha_i = 1$)

$$\min_{\theta,\xi}(\alpha_1 \ell_{\mathrm{mlm}}(\mathrm{x_n}\,;\theta) + \alpha_2 \ell_{\mathrm{disc}}(\mathrm{x_n}, \mathrm{y_{disc}}; \theta) + \alpha_3 \ell_{\mathrm{sup}}(\mathrm{x_n}, \mathrm{y_{sup}}; \theta, \xi))$$



(a) (b)

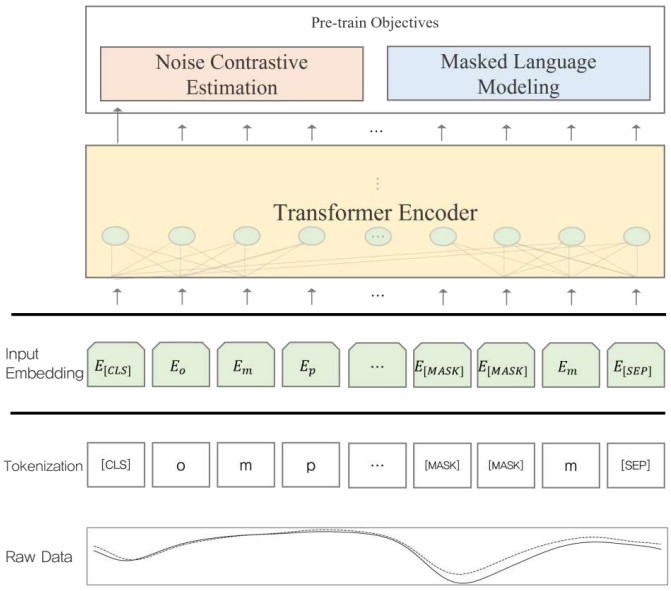
Model architecture

- nth subject has M functions, $\mathbf{x}_n = \{x_n^1, \ldots, x_n^M\}$, ($M = 6$ in our data)
- the training data consist of N inputs, $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, ($N = 833$ in our data)
- the nth individual function can be expressed as $x_n^m = \{x_n^m(1), \ldots, x_n^m(T)\}$

Model architecture

❖ Two stage learning, $\alpha_1, \alpha_2, \alpha_3 \ (\geq 0, \Sigma_{i=1}^{3}\alpha_i = 1)$

$$\min_{\theta,\xi}(\alpha_1 \ell_{\mathrm{mlm}}(\mathrm{x_n}\,;\theta) + \alpha_2 \ell_{\mathrm{disc}}(\mathrm{x_n}, \mathrm{y_{disc}}; \theta) + \alpha_3 \ell_{\mathrm{sup}}(\mathrm{x_n}, \mathrm{y_{sup}}; \theta, \xi))$$



(a)

- 문맥 정보 반영을 위한 reconstruction error 측정
- $\mathrm{x_n^m}$의 토큰 중 일부를 [MASK] 토큰으로 대체
  - 마스킹 된 토큰을 예측

- L : *number of masked tokens*
- $\Pi_n^m = \{\pi_{n1}^m, \ldots, \pi_{nL}^m\}$: *indices of masked tokens of* $\mathrm{x_n^m}$
- $\mathrm{x_n^m}(\Pi_n^m)$: *set of the masked tokens* in $\mathrm{x_n^m}$
- $\mathrm{x_n^m}(-\Pi_n^m)$: *set of the observed tokens*

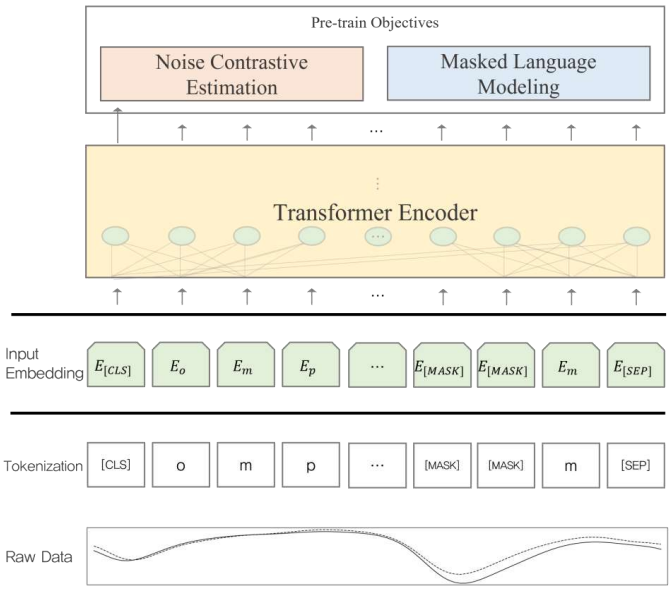
- $\ell_{\mathrm{mlm}}(\mathrm{x_n}\,;\theta)$
  $= -\frac{1}{L}\sum_{l=1}^{L}\sum_{m=1}^{M}\log \mathrm{p}\,(\mathrm{x_n^m}(\Pi_{nl}^m)|\mathrm{x_n^m}(-\Pi_n^m);\theta).$

Model architecture

Proposed Method

❖ Two stage learning, $\alpha_1, \alpha_2, \alpha_3 \ (\geq 0, \Sigma_{i=1}^{3}\alpha_i = 1)$

$$\min_{\theta,\xi}(\alpha_1 \ell_{\mathrm{mlm}}(\mathrm{x_n}\,;\theta) + \alpha_2 \ell_{\mathrm{disc}}(\mathrm{x_n}, \mathrm{y_{disc}}; \theta) + \alpha_3 \ell_{\mathrm{sup}}(\mathrm{x_n}, \mathrm{y_{sup}}; \theta, \xi))$$



(a)

- NCE를 위한 discriminator
- 이진 분류 손실 문제

- $y_n^{\mathrm{disc}} = \{0, 1\}$
- $\mathrm{p_n}$: n번째 관측에 대한 sigmoid 확률

- $\ell_{\mathrm{disc}}(\mathrm{x_n}, \mathrm{y_n^{disc}}; \theta)$
  $= -\left(y_n^{\mathrm{disc}}\log(\mathrm{p_n}) + \left(1 - y_n^{\mathrm{disc}}\right)\log(1 - \mathrm{p_n})\right)$

Model architecture

❖ Two stage learning, $\alpha_1, \alpha_2, \alpha_3 \ (\geq 0, \Sigma_{i=1}^{3}\alpha_i = 1)$

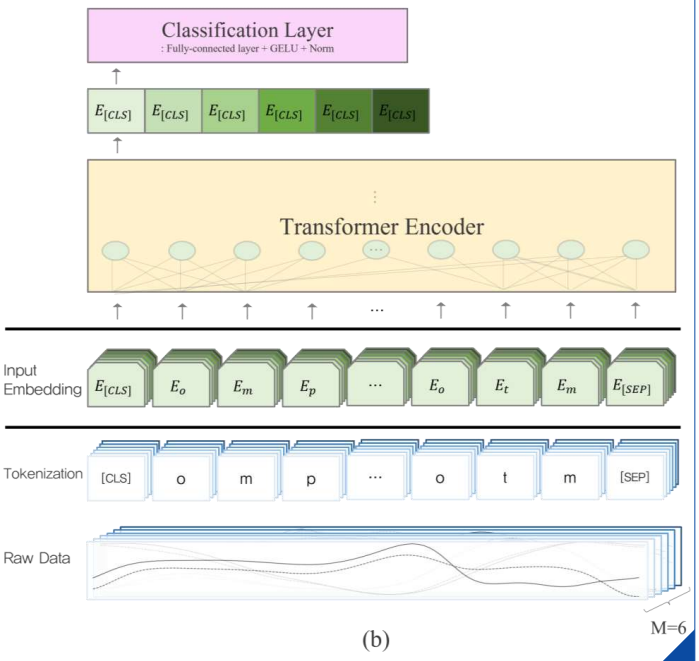
$$\min_{\theta,\xi}(\alpha_1 \ell_{\text{mlm}}(x_n ; \theta) + \alpha_2 \ell_{\text{disc}}(x_n, y_{\text{disc}}; \theta) + \alpha_3 \ell_{\text{sup}}(x_n, y_{\text{sup}}; \theta, \xi))$$

- Fine -tunin을 위한 linear layer
- 이진 분류 손실 문제

- $y_n^{\text{sup}} = \{0, 1\}$
- q_n: n번째 관측에 대한 sigmoid 확률

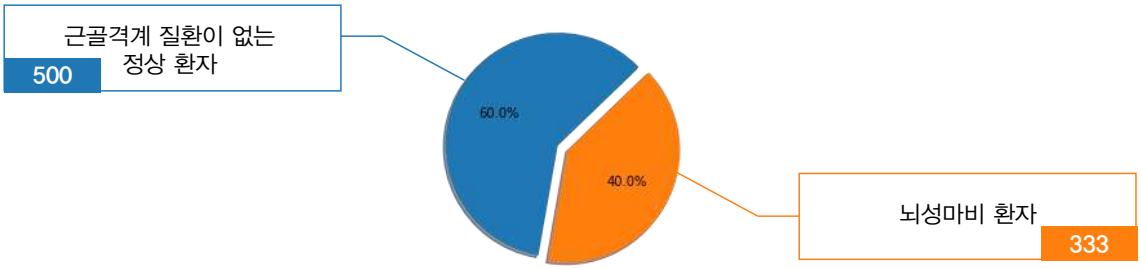
- $\ell_{\text{sup}}(x_n, y_n^{\text{sup}}; \theta, \xi)$

$$= -\left(y_n^{\text{sup}} \log(q_n) + (1 - y_n^{\text{sup}}) \log(1 - q_n)\right)$$



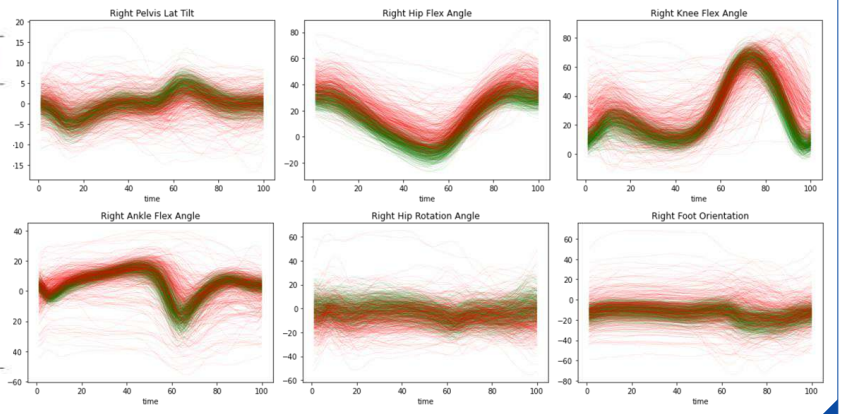
(b)

Data analysis

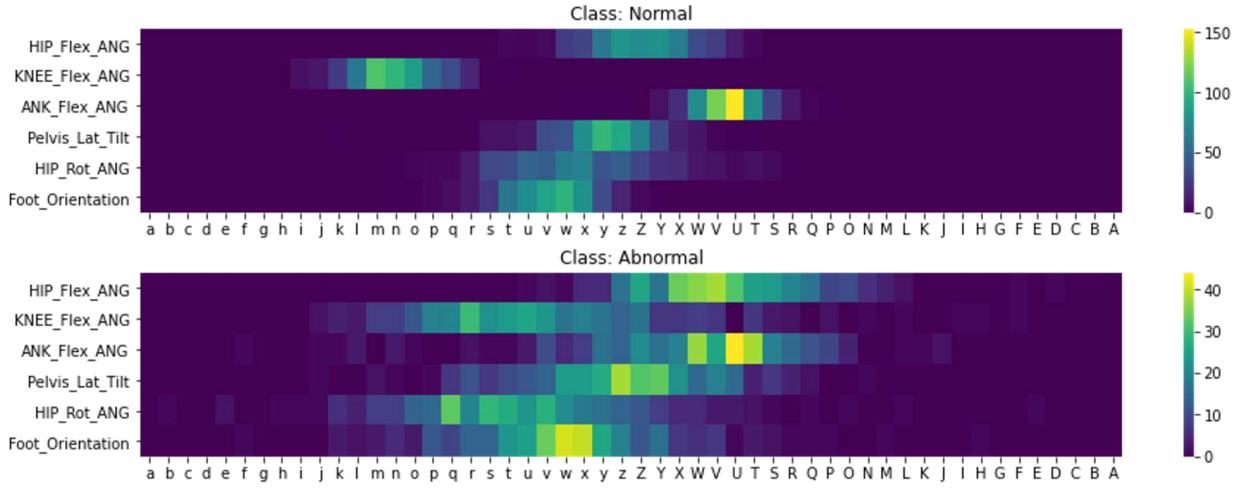


❖ 수집 기관: 분당 서울대학교 병원 (2018.01 ~ 2021.12)

body parts	total	normal	abnormal
Hip flex angle	0.21	0.14	0.25
Knee flex angle	0.18	0.16	0.19
Ankle flex angle	0.22	0.19	0.24
Pelvis lateral tilt	0.20	0.15	0.23
Hip rotation angle	0.17	0.11	0.20
Foot orientation	0.21	0.15	0.24



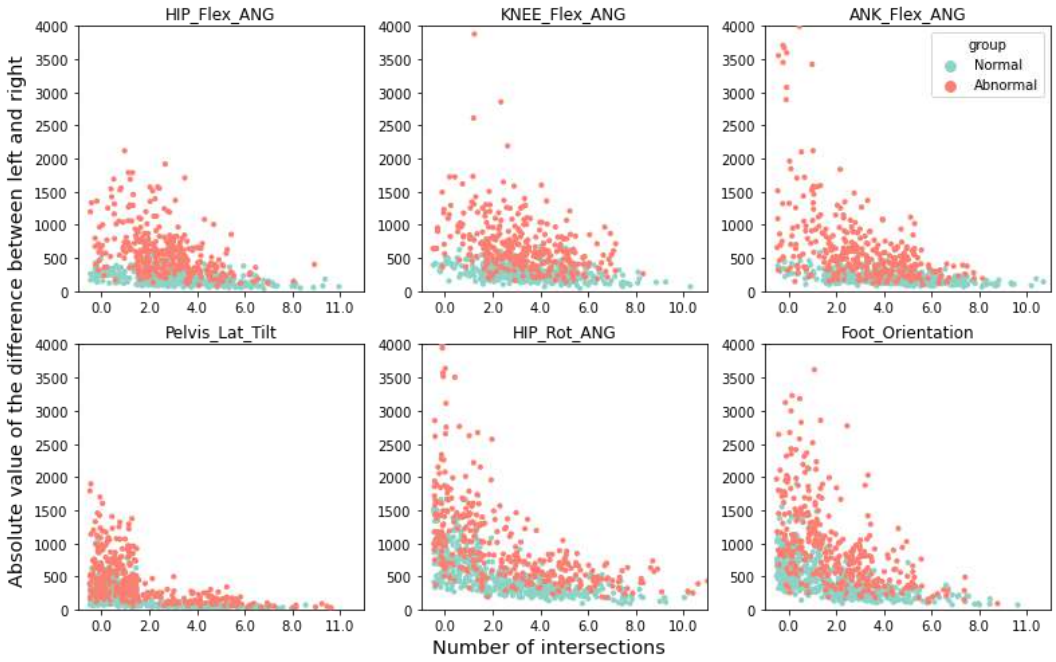
❖ 각 신체 부위의 첫 번째 토큰 수



- 각 신체 부위의 첫 번째 토큰 수에 대한 heatmap
- **비정상** 데이터의 첫 시점의 변동성 〉 **정상** 데이터의 첫 시점의 변동성

  ⇨ FDA 시 alignment 과정이 필수적

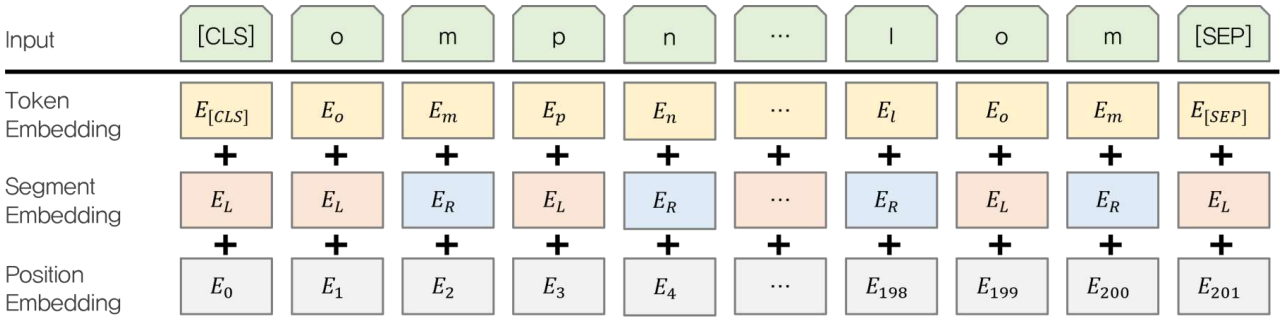
Input representation to BERT



❖ 정상 그룹은 비정상 그룹보다 교차점의 분산이 크고 절댓값이 작다

⇨ 비정상 그룹의 **좌 우 측정치의 차이가 유의미**함을 의미

Input representation to BERT



❖ Segment Embedding

- BERT의 입력에서 segment embedding은 연결되는 두 문장 사이의 구분자 역할을 함
- 걷는 동안 왼쪽과 오른쪽 값이 교차하여 관찰되게 때문에 이러한 순차적 순서를 모델에 입력에 반영

❖ 세 가지 입력값의 합으로 구성된 입력 representation은 모델의 self-supervised learning을 통해 다른 토큰들의 가중치에 영향을 받게 됨.

❖ 따라서, 입력 문장을 구성하는 토큰이 동일하여도 순서에 따라 서로 다른 출력값을 제공함

Computing details

❖ GPU : Nvidia RTX3090 x 4 (memory: 24GB)

❖ Hyper-parameter:

- Learning rate: 0.0001
- Weight decay: 0.01
- epoch: 100

❖ Model setting (BERT medium) – total parameter ($|\theta| + |\xi|$) = 약 41,700,000

- Layer: 8
- Hidden state: 512
- Attention heads: 8

❖ Fine tuning setting

- Train : validation : test = 7 : 1 : 2

Metrics

❖ Accuracy, precision, recall, f1-score

		Predicted	
True			
	TP True Positives	**FN** False Negatives Type II error	
	FP False Positives Type I error	**TN** True Negatives	

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F - score = \frac{2}{1/precision + 1/recall}$$

❖ 초기 값의 무작위성이 존재하는 딥러닝 모델의 신뢰성을 높이기 위해 10번의 반복 실험 결과를 **평균(표준편차)** 로 보고

Analysis results

❖ 언어 모델의 양방향성이 성능에 미치는 영향을 파악하기 위해 5개의 머신 러닝 모델과 1개의 딥러닝 모델과 비교

❖ Machine learning methods:

- 로지스틱 회귀(logistic regression)
- 의사결정나무(decision tree)
- SVM
- 랜덤포레스트(random forest)
- XGBoost

❖ Deep learning method: 4-layer LSTM

❖ Functional BERT : 추가적인 증강 데이터 없이 원시 데이터만으로 학습한 모델

methods	accuracy	f1 score	precision	recall
Logistic regression	0.728 (0.001)	0.633 (0.001)	0.688 (0.001)	0.586 (0.001)
Decision tree	0.791 (0.001)	0.724 (0.001)	0.766 (0.001)	0.687 (0.001)
Random forest	0.857 (0.001)	0.806 (0.002)	0.884 (0.002)	0.741 (0.003)
XGBoost	0.866 (0.001)	0.822 (0.002)	0.875 (0.004)	0.775 (0.003)
SVM	0.883 (0.001)	0.848 (0.001)	0.879 (0.001)	0.819 (0.001)
DNN	0.875 (0.022)	0.868 (0.022)	0.867 (0.021)	0.879 (0.019)
LSTM	0.926 (0.006)	0.905 (0.007)	0.931 (0.021)	0.880 (0.006)
functional BERT	**0.962** (0.006)	**0.961** (0.007)	**0.962** (0.006)	**0.962** (0.006)

Analysis results

❖ 제안된 모델에 대한 증강 방법의 효과를 판단하기 위해 각 방법 당 $r = 10$배 증강하여 사전 학습하고 미세 조정에서 성능을 비교

❖ No augmentation: 추가적인 증강 데이터 없이 원시 데이터만 학습에 사용

augmentation methods	accuracy	f1 score
No augmentation	0.962 (0.006)	0.961 (0.007)
Jittering ($\sigma = 0.1$)	0.974 (0.005)	0.973 (0.005)
Jittering ($\sigma = 0.2$)	**0.975** (0.004)	**0.973** (0.004)
Reverse	0.961 (0.007)	0.959 (0.007)
Time-warping	0.962 (0.004)	0.961 (0.004)
Permutation	0.954 (0.005)	0.952 (0.005)
Trend-filtering ($k = 0$)	0.962 (0.007)	0.960 (0.007)
Trend-filtering ($k = 1$)	0.967 (0.006)	0.965 (0.006)
Trend-filtering ($k = 2$)	0.964 (0.006)	0.963 (0.006)

Analysis results

❖ 제안된 방법에 대한 증강 데이터의 크기가 미치는 영향을 파악하기 위해, 증강 데이터 크기 hyper-parameter r의 탐색 범위를 조정 ($r \in \{10, 20, 100\}$)

❖ Jittering($\mu = 0,\ \sigma = 0.2$)와 trend filtering ($k \in \{0, 1, 2\}$) 두 가지 방법만 고려

❖ No augmentation 의 accuracy : 0.958 (0.008)

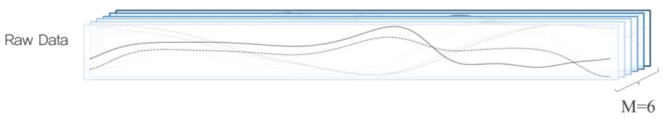
augmentation method	r	jittering	trend-filtering		
			$k = 0$	$k = 1$	$k = 2$
augmented data	×10	0.975 (0.004)	0.960 (0.007)	0.965 (0.006)	0.963 (0.006)
	×20	0.975 (0.007)	0.975 (0.006)	0.964 (0.009)	0.962 (0.003)
	×100	**0.977** (0.004)	0.962 (0.007)	0.965 (0.005)	**0.976** (0.005)

Analysis results

❖ 기존 BERT의 훈련과제인 MLM에 새로운 학습 목표로 NCE를 추가한 경우 성능의 변화를 파악

training objectives	jittering	$k = 1$	$k = 2$
MLM $(\alpha_1, \alpha_2, \alpha_3) = (1, 0, 0)$	0.977 (0.005)	0.965 (0.005)	0.976 (0.005)
MLM + NCE $(\alpha_1, \alpha_2, \alpha_3) = (1, 1, 0)$	**0.979** (0.004)	**0.969** (0.006)	0.974 (0.005)

❖ 모델의 마지막 계층 구성에 따른 성능 평가



model architecture	f1 score
max pooling	0.847 (0.007)
concatenation	**0.961** (0.007)

❖ 모델의 마스킹 비율에 따른 성능 평가

corruption rate	f1 score
10%	0.975 (0.005)
15%	0.975 (0.006)
25%	**0.977** (0.007)
50%	0.972 (0.007)

Conclusion & Discussion

Conclusion & Discussion

❖ Conclusion

1. 변동이 큰 함수형 데이터를 분석하기 위해 언어 모델을 채택한다. 자기 지도 학습을 통해 내재적인 구조를 이해하는 것이 정렬 프로세스를 구현하지 않고도 함수형 데이터 분석을 위한 성능에 긍정적인 영향을 미친다는 것을 보여준다.
2. 함수형 데이터의 증강 방법으로 trend-filtering을 제안한다. 이는 실제 데이터의 패턴을 반영하는 증강 데이터를 효율적으로 생성한다. 제안된 데이터 증강 기법이 딥러닝에 미치는 영향을 기존의 시계열 증강 기법과 비교한다.
3. 노이즈 생성기(generator)로써 trend-filtering을 도입하여 증강한 데이터와 실제 데이터를 판별해 학습 효율을 높인다는 점에서 잡음 대조 추정법의 방식으로 해석 가능하다.

❖ Discussion

- 정상/비정상의 이진 분류 문제를 넘어 비정상 클래스의 GMFCS 수준(1 ~ 5 level)을 구체적으로 예측한다면 뇌성마비 환자의 운동 기능에 맞는 사후 처치 및 재활을 제공 가능

안전한 의료데이터 활용위한 PETs: 동형암호 활용사례

May. 2024

㈜크립토랩 신 성철 이사

HEAAN CRYPTO LAB

Content

I. 의료데이터 양날의 검

II. PETs 소개

III. 동형암호 활용사례

Ⅰ. 의료데이터 양날의 검

❑ 의료데이터, 개방을 넘어 활용의 시대로….

- 2020년 데이터3법 이후 정부의 데이터 공개정책에 힘입어 의료데이터도 개방을 넘어 활용의 시대로 진입

의료데이터 종류 및 동향

>> 의료분야서 가장 중요한 기술분야와 의료 데이터의 종류 및 동향

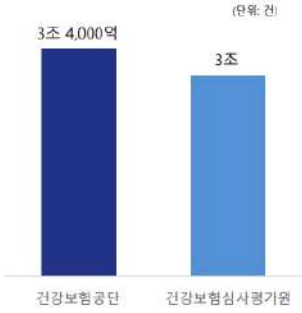
구분	관리자	설명	동향
개인 유전정보	유전체 분석 서비스 업체	• 1 인당 약 30 억 개의 유전자 염기서열 정보 존재 • 개체 간 약 0.1%의 차이 존재	• 유전체 분석비용은 2000년대 초 9,000만 달러였으나, 2017년 100 달러로 하락 • 2016년 기준 79,110건의 유전체 정보 분석 프로젝트 완료
개인건강 정보	개인	• 스마트폰 앱 또는 IoT 디바이스로 수집되는 Life-log*(예: 수면패턴 등)	• 다양한 디바이스와 서비스 증가
전자 의무기록	의료 기관	• 환자의 모든 진료정보를 전산화하여 입력, 저장, 관리하는 형태(예: 진단정보, 처방자료, 처방결과 등)	• 전세계적으로 디지털화 가속
국민 건강정보	공공 기관	• 자격 및 보험료, 진료내역, 건강검진결과, 의료급여 등	• 한국은 단일 건강보험체계를 갖고 있다는 특수성으로 인해 국민의 건강 관련 빅데이터가 공공기관에 집중

Source: 한국정보화진흥원

Source:과학기술정책연구원, KDB 산업기술리서치센터 자료, 삼정KPMG 경제연구원 재구성
Note: Life-log는 개인의 일상생활 활동에 관한 모든 데이터를 의미함

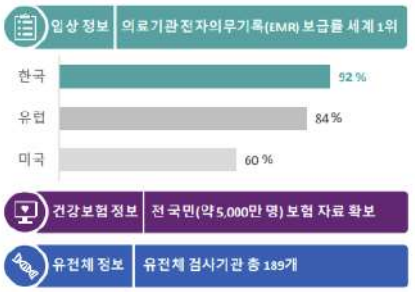
국내 공공 의료데이터 보유현황(2019년)

>> 국내 공공의료 빅데이터 수



Source: 건강보험공단, 건강보험심사평가원
Note: 누적기준

>> 국내 보건의료 데이터 현황



※참조: 제 124호 March 2020, 삼정KPMG

Ⅰ. 의료데이터 양날의 검

❑ 민감정보일수록 활용 니즈 크고, 유출 시 사회적 비용도 막대함.

1. 병원 직원, 제약사 직원이 병원시스템에서 약 처방 및 환자정보 유출

17개 대학병원에서 18만5271명 환자정보 유출…
개인정보위, 16개 병원에 과태료 부과

병원·제약사 직원이 병원 시스템에서 처방 환자정보 외부 반출 `23.07 메디게이트

2. 처방약거래 플랫폼이 광고 목적으로 구글, 페이스북 등에 사용자 정보 제공

150만 달러 벌금… GoodRx 개인 의료정보 유출 - `23.02 HIT news

3. 임상연구소 해커 공격으로 이름, 검사결과 등 250만 정보유출, 집단 소송

NEWS

Medical lab data breach exposes millions of patients' data across states

by Tricia Ennis June 16, 2023

참고 병원별 유출 개요 및 행정처분

법인명	병원명	유출 경위	유출규모(명)	시정조치
(학)가톨릭학원	서울성모병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일, USB)	16,463	• 과태료 360만 원 • 개선권고/결과공표
	여의도성모병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	17,115	• 과태료 360만 원 • 개선권고/결과공표
	은평성모병원	• 제약사 직원이 환자정보를 촬영·다운로드하도록 묵인	3,633	• 과태료 360만 원 • 개선권고/결과공표
	의정부성모병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	20,027	• 과태료 360만 원 • 개선권고/결과공표
	부천성모병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	9,673	• 과태료 360만 원 • 개선권고/결과공표
	성빈센트병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	38	• 과태료 360만 원 • 개선권고/결과공표
(학)건국대학교	건국대충주병원	• 내부직원이 환자정보를 제약사 직원에게 송부(USB)	485	• 과태료 420만 원 • 개선권고/결과공표
(학)고려중앙학원	고려대안암병원	• 제약사 직원이 환자정보를 다운로드하도록 묵인	1,399	• 과태료 360만 원 • 개선권고/결과공표
	고려대구로병원	• 제약사 직원이 로그인 되어있던 시스템의 환자정보를 외부로 반출	14,385	• 과태료 360만 원 • 개선권고/결과공표
	고려대안산병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	14,038	• 과태료 360만 원 • 개선권고/결과공표
(학)동은학원	순천향대서울병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	8,693	• 과태료 420만 원 • 개선권고/결과공표
(학)연세대학교	세브란스병원	• 내부직원이 환자정보를 제약사 직원에게 송부(이메일)	57,912	• 과태료 720만 원 • 개선권고/결과공표
(학)일송학원	성심병원	• 제약사 직원이 환자정보를 다운로드하도록 묵인	8,613	• 과태료 420만 원 • 개선권고/결과공표
	동탄성심병원	• 제약사 직원이 환자정보를 다운로드하도록 묵인	3,095	• 과태료 420만 원 • 개선권고/결과공표
	강남성심병원	• 제약사 직원이 ID·PW를 도용하여 시스템 환자정보를 외부로 반출	5,831	• 과태료 420만 원 • 개선권고/결과공표
	한강성심병원	• 제약사 직원이 ID·PW를 도용하여 시스템 환자정보를 외부로 반출	1,487	• 과태료 420만 원 • 개선권고/결과공표
(의)삼성의료재단	강북삼성병원	• 제약사 직원이 로그인 되어있던 시스템의 환자정보를 외부로 반출	2,366	• 개선권고

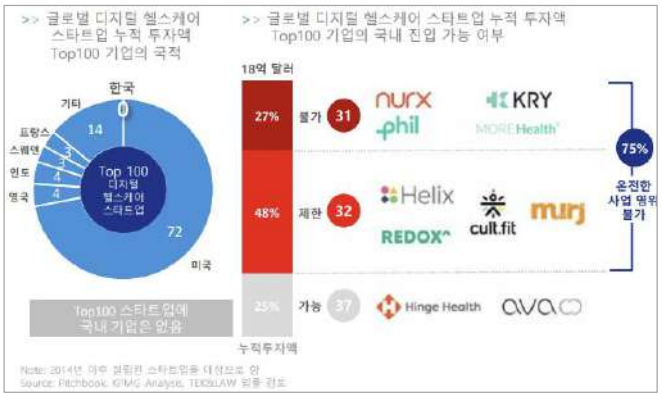
※ 과태료 처분 대상 16개 병원은 모두 안전성 확보조치 의무 위반
(개인정보 보호법 제23조제2항 및 제29조)

Ⅰ. 의료데이터 양날의 검

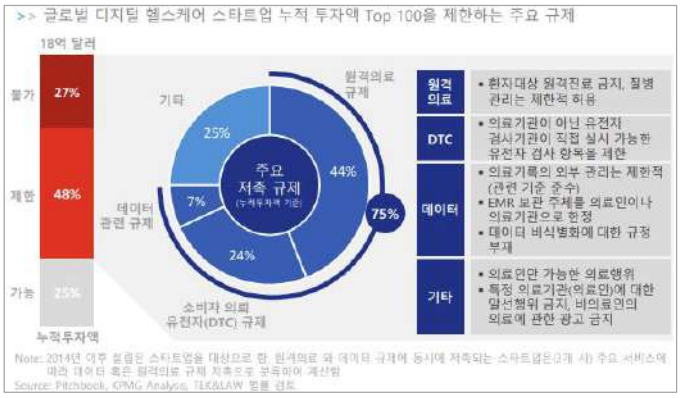
❑ 글로벌 디지털 헬스케어 시장 성장, 국내 시장은 규제들로 발 묶임

- 비식별화 의료정보 개념 법제화, 자율적 활용 규제 명확화, 원격의료 범위 확대, DTC 유전자검사 항목 확대 등 정책 변화 필요

디지털 헬스케어 글로벌 시장 성장



디지털헬스케어 분야 주요 규제

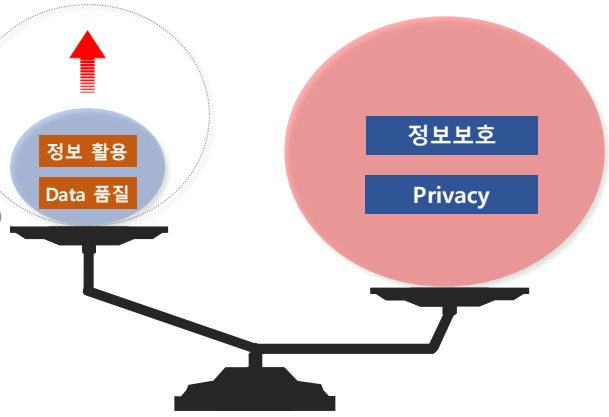


※참조: 제 124호 March 2020, 삼정KPMG

Ⅰ. 의료데이터 양날의 검

❑ 정보 보호와 활용 : Trade-off ? 둘 모두 만족시킬 수 없는가?

정보보호 vs 정보활용



*데이터3법(20.08 개정)
*공공데이터법(20.12)
*데이터 산업진흥 및 이용촉진에 관한 기본법(22.04)
*보건의료데이터 활용가이드라인 개정(24.01)

*신용정보법
*정보통신망법
*개인정보보호법
*의료법
...
*GPPR(유)
*CCPA(미)
*HIPAA(미)

Content

Ⅰ. 의료데이터 양날의 검

Ⅱ. PETs 소개

Ⅲ. 동형암호 활용사례

Ⅱ. PETs 소개

❑ PETs (Privacy Enhancing Technologies, 개인정보 보호 강화기술)

- 개인정보를 필요한 범위내에서만 처리하고 유출 사고 예방에도 도움이 되는 기술 (Privacy + Security + Compliance)
- 대한민국 정부 AI시대 안전한 활용위한 기술로 PETs 중요성 강조

PET 필요성/중요성

개인정보호·활용 위한 정부의 관심

"AI 시대 안전한 활용 위한 'PET' 기술 활용 중요"

- 개인정보보호위원회 개인정보기술포럼 (`23.12)

Ⅱ. PETs 소개

❑ 다양한 영역에서 PETs가 연구되고 있고, 일부 기술들은 현장에서 널리 사용 중임.

- Key Technologies : 동형암호, MPC, DP, ZKP, Synthetic Data, Federate Learning

주요 PETs 기술

No	방법론	설명
1	✓ **동형암호 (Homomorphic Encryption)**	**암호화 상태에서 데이터를 연산하는 기술**
2	✓ **Secure MPC** (Multi-party computation)	다수의 암호화된 데이터를 연산하는 암호기술
3	✓ **Differential Privacy**	데이터에 통계적 노이즈를 추가하여 보호 기술
4	✓ **영지식 증명** (Zero-knowledge Proofs)	데이터를 공개하지 않고 검증하는 암호기술
5	**난독화** (Obfuscation)	민감정보에 노이즈나 변형을 통해 마스킹
6	**가명화** (Pseudonymization)	식별정보를 문자나 가상의 데이터로 대체
7	**데이터 최소화** (Data minimization)	서비스 제공 가능한 최소한의 정보 수집
8	**Communication anonymizers**	온라인 신원(IP,email)을 폐기하고 추적 불가능한 신원으로 대체
9	✓ **합성데이터** (Synthetic data)	원본 데이터의 특성을 유지하는 다양한 알고리즘을 통해 인위적으로 생성한 데이터
10	✓ **연합학습** (Federated Learning)	다수 데이터를 물리적으로 결합하지 않고 수행하는 기계학습 방법

※출처 : AI multiple Research

PETs 기술 구분

*These categories are based on the authors' analysis of the PET space. The authors acknowledge that there are multiple ways to group these technologies, techniques, and processes.

※ 참고 Privacy Enhancing Technologies: Categories, Use Cases, and Considerations, 2021

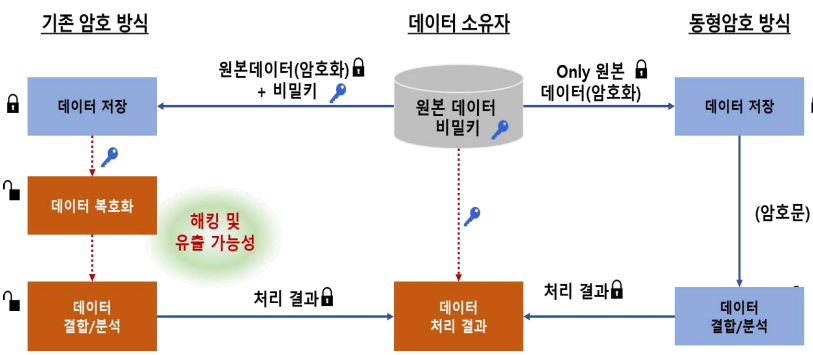
Ⅱ. PETs 소개

❑Homomorphic Encryption(동형암호, 同型暗號)

- 데이터를 암호화한 상태에서 복호화 하지 않고 연산할 수 있는 최신 암호화 기술

동형암호 개념

기존암호 vs 동형암호



Content

Ⅰ. 의료데이터 양날의 검

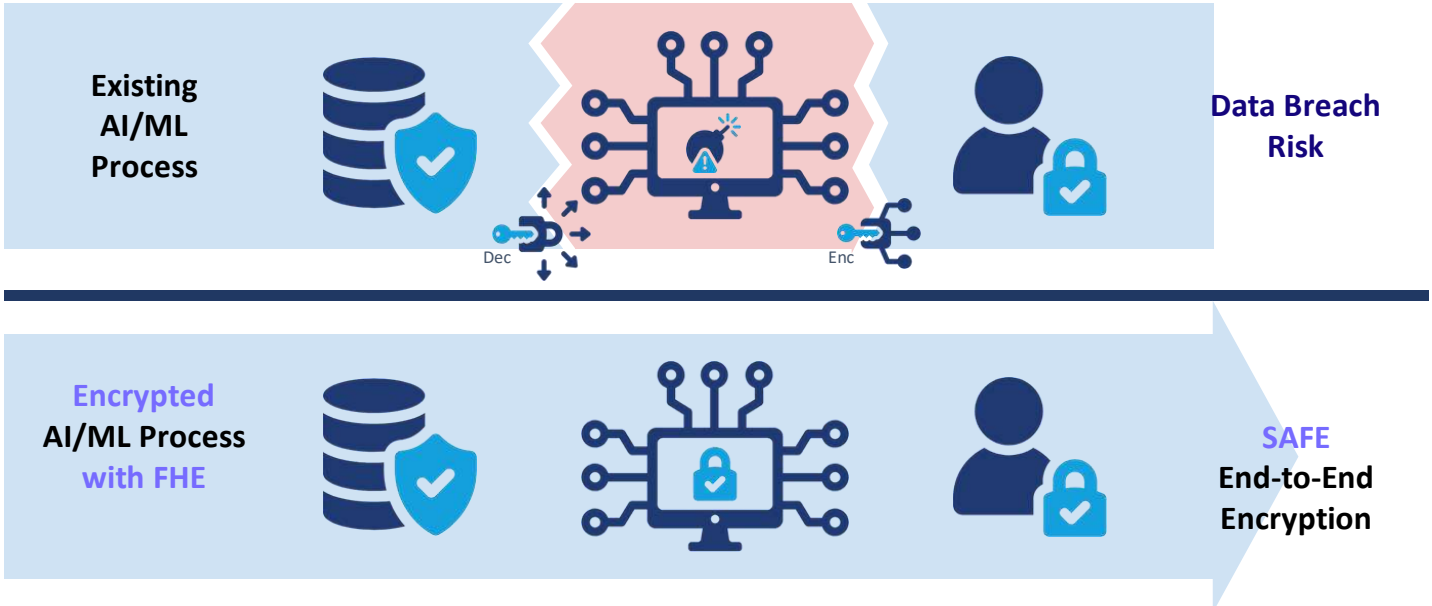
Ⅱ. PETs 소개

Ⅲ. 동형암호 활용사례

Ⅲ. 동형암호 활용사례

❑ End-to-End Encryption with Fully Homomorphic Encryption

- 동형암호기반 AI모델 개발로 안전한 데이터 활용 AI 환경을 만들 수 있음.



Ⅲ. 동형암호 활용사례

❑ 얼굴, 지문 등 생체정보 인식위한 비교연산시 데이터 유출 및 프라이버시 문제 발생

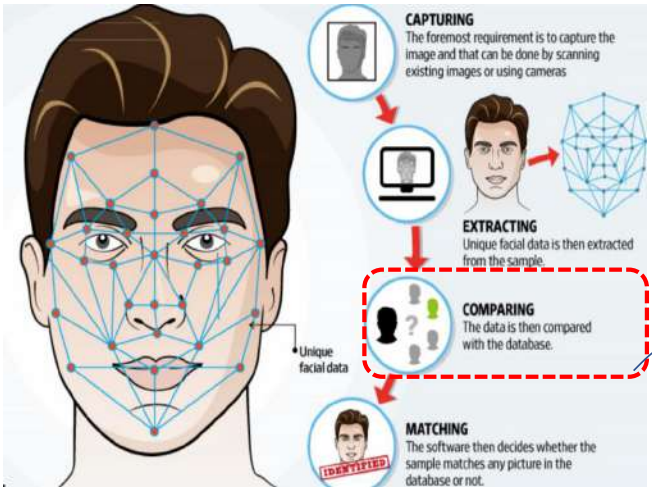
인권위 "실시간 원격 얼굴인식 기술 원칙적 금지해야"

2023.01.25 12:00 입력

윤기은 기자

청부서울청사로 들어가는 공무원과 민원인들이 2021년 11월15일 마스크를 내리고 얼굴을 얼굴면석 출입기에 인식시키고 있다. /

안면 인식 및 생체정보 비교



Privacy 이슈

얼굴 지문

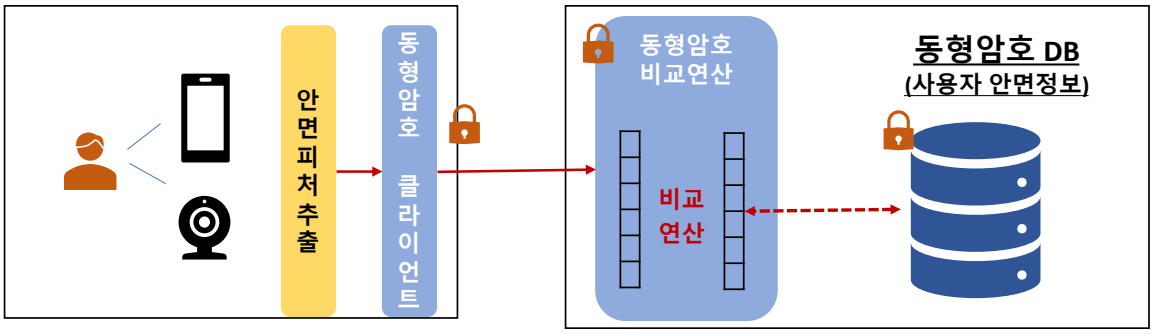
Ⅲ. 동형암호 활용사례

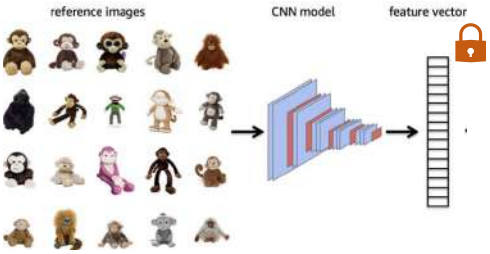
❑ 사용자 생체정보는 동형암호화하여 서버에 안전하게 보관하고, 유사도 연산을 수행함.

- 공항, 병원, 관공서 등에서 사용자 인증에 활용

단말기

동형암호기반 생체인증 서버



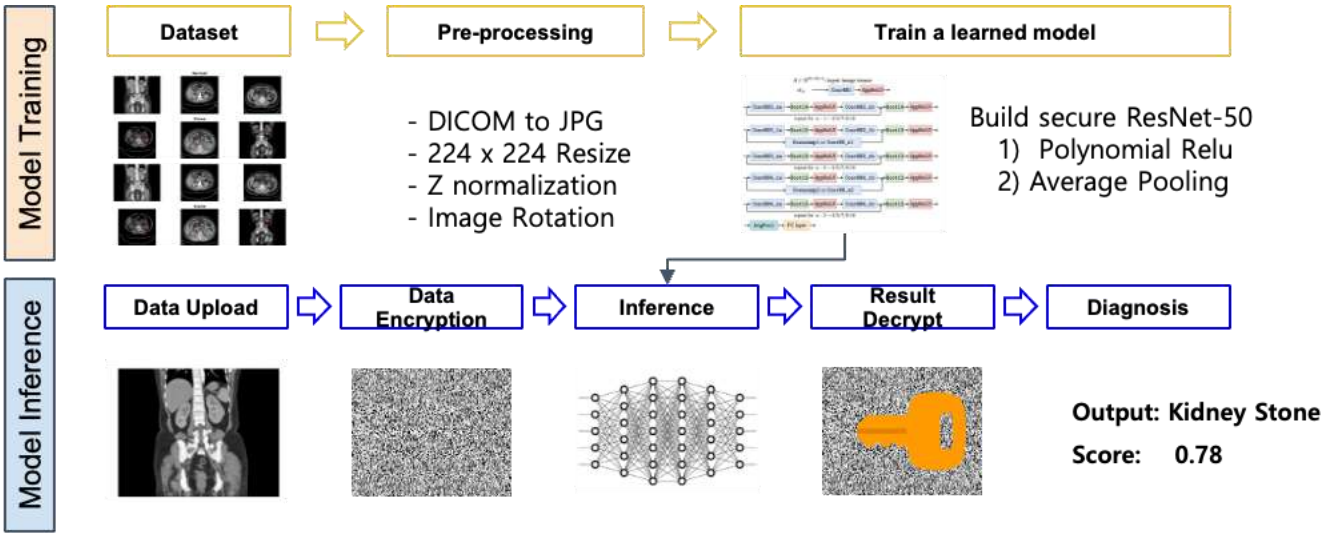


$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

Ⅲ. 동형암호 활용사례

❑ 동형암호 활용 AI Model 및 민감 데이터 보호(아산병원, 2023)

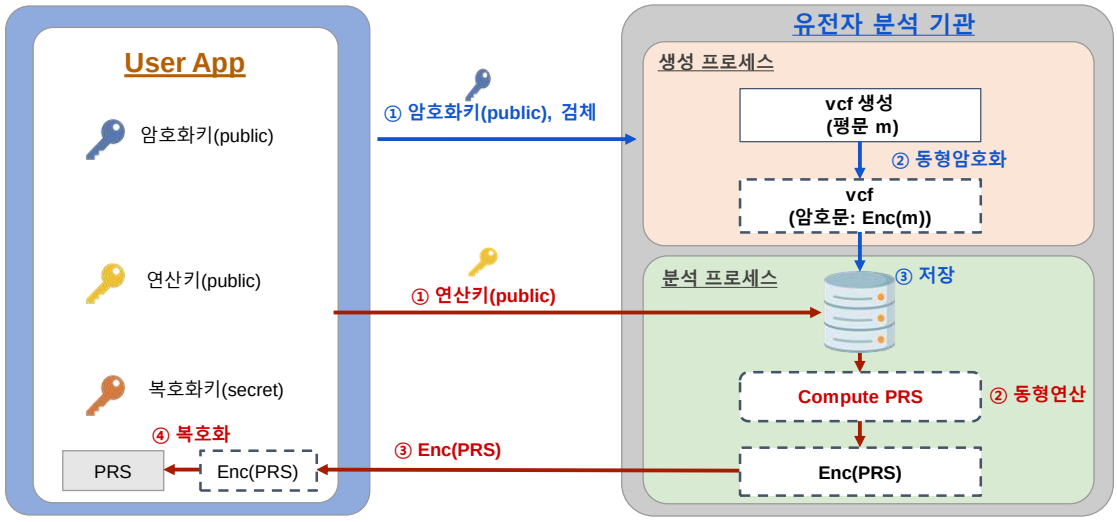
- 환자의 진료정보(이미지)를 보호하면서 질병여부를 진단할 수 있음.



Ⅲ. 동형암호 활용사례

❑ 동형암호 활용 시 민감정보를 보호하는 안전한 B2C 유전자 질병예측(PRS) 서비스가 가능함.

- 유전자 검사 질병 예측 결과는 사용자 앱에서만 확인할 수 있음.(B2C, DTC 서비스)



- 공개키(Public) : 암호화키(Encryption), 연산키(Evaluation)
- 비밀키(Secret) : 복호화키(Decryption)

- VCF(variant call format): 변형호출형식, 참조유전체와 비교 변이와 빈도를 표시하는 파일형식
- PRS(Polygenic Risk Score): 다중유전자위험점수, 질병 위험도 점수
- DTC(Direct to customer): 소비자 직접의뢰 유전자 검사 서비스

Ⅲ. 동형암호 활용사례

❑ ChatGPT와 같은 LLM 활용시 발생할 수 있는 Privacy 이슈 및 정보보호 문제

Why is Facebook shutting down its facial recognition system and deleting 'faceprints'?

https://www.theguardian.com/technology/2021/nov/03/why-is-facebook-shutting-down-its-facial-recognition-system-and-deleting-faceprints

Facebook has decided to limit the use of facial recognition technology due to concerns about user privacy and safety issues.

Samsung Bans Staff's AI Use After Spotting ChatGPT Data Leak

https://www.bloomberg.com/news/articles/2023-05-02/samsung-bans-chatgpt-and-other-generative-ai-use-by-staff-after-leak

Samsung Electronics Co. is banning employee use of popular generative AI tools like ChatGPT after discovering staff uploaded sensitive code to the platform.

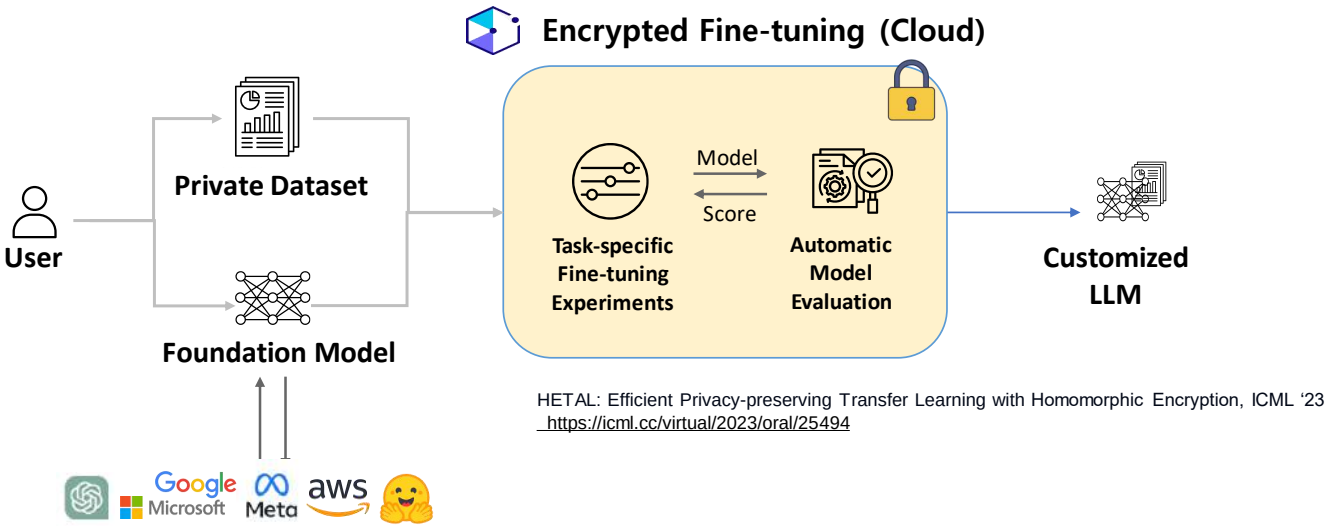
HEAAN CRYPTO LAB

17

Ⅲ. 동형암호 활용사례

❑ 의료, 금융, 국방 등 정보의 기밀성이 높은 Domain에서 안전하게 LLM을 활용하는 방법

- 도메인관련 정보를 충분하게 학습한(fine tuning) LLM



HETAL: Efficient Privacy-preserving Transfer Learning with Homomorphic Encryption, ICML '23 https://icml.cc/virtual/2023/oral/25494

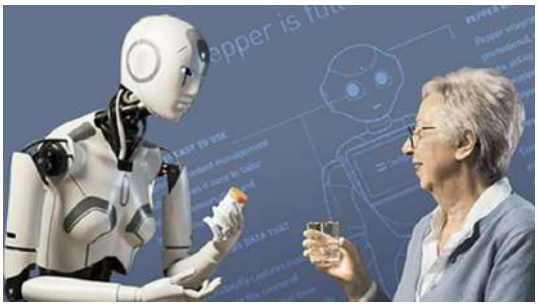
HEAAN CRYPTO LAB

18

❑ 정신건강 디지털 헬스케어에서 민감정보 보호 및 활용을 위한 대안 기술: 동형암호

- AI서비스 = 초개인화 맞춤형 = 개인정보 = Privacy = 정보보호

치매환자 치료 로봇 예시



생성되는 민감정보 예시

- 개인 정보: 이름,성별,나이, 지역 등
- 의료정보: 질병, 투약정보
- Life Log: 생체정보, 바이탈정보, 이동정보, 식습관
- 대화정보: 얼굴, 이미지, 대화 음성, Text정보

치매환자 One Stop 서비스(예시)

EoD

QnA

CryptoLab

Session D

D3. [학술세션]
인공지능과 4차산업혁명 공진화

좌 장: 권 오 병(경희대)

식품 위해 사전과 모델 성능 고도화 방안을 통한 수입식품 위해도 예측모형 성능 제고

이다현
경희대학교
빅데이터응용학과
sq9031@khu.ac.kr

이은서
경희대학교
빅데이터응용학과
les1512@khu.ac.kr

권상지
경희대학교 경영학과
aaaitaaa@khu.ac.kr

신윤종
경희대학교 인공지능학과
yoonjong12@khu.ac.kr

김정연
경희대학교
빅데이터응용학과
donitz8946@khu.ac.kr

권오병
경희대학교 경영학과
obkwon@khu.ac.kr

***Abstract** – 본 연구의 목적은 수입 식품의 위해도 예측 모델을 개발하고 이를 고도화하여 성능을 향상시키는 방안을 제안하는 것이다. 다양한 모델 고도화 실험을 통해 품목명 위해도 변수 생성, 차원축소, 스태킹 앙상블, 이동평균 시계열 가중치 등의 방법이 모델의 성능 향상에 유의미한 영향을 미치는 것을 확인했다. 이러한 연구 결과는 수입 식품 안전성을 보다 효과적으로 관리하기 위한 방향성을 제시하며, 앞으로의 연구에 큰 도움이 될 것으로 기대된다.*

***Keywords** – 위해도 예측 모델, 위해 식품 사전, 스태킹앙상블, 시계열가중치*

본 연구는 2024 년도 식품의약품안전처의 연구개발비(21163MFDS516)로 수행되었습니다.

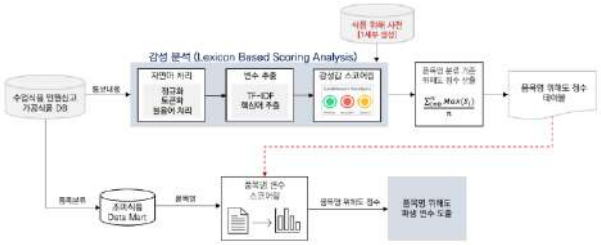
I. 서론

수입 식품의 안전성은 각 국가 및 지역의 공중보건 문제뿐만 아니라 국제적인 문제이다. 그러므로 적시에 수입 식품에 대해서 위해 요인을 식별하여 사고 발생을 예방할 수 있어야 한다. 그러면서도 한정된 인력 및 예산 문제를 고려하면 전수 조사가 어려우므로 조사가 필요한, 즉 식품 안전에 문제가 있을 것으로 보이는 수입 건에 대해 예상 판별할 수 있는 '수입식품 안전 위해도 예측모형'이 필요하다. 식품의약품안전처에서는 매년 증가하는 수입 검사에 대한 효율적인 업무처리를 위해 사전적 안전 관리를 위한 자동화 시스템을 구현하려는 연구가 진행되고 있다. 기존 연구들은 과거 수입이력, 부적합정보, 국내외 위해정보 등을 종합 분석한 수입식품 사전예측 검사시스템(OPERA, 2014)과 해외 사전관리 및 통관검사에 걸친 전 단계 통합 정보관리 체계 및 위해도 분석시스템인 지능형 수입식품 통합시스템(2018)이 존재하지만, 정밀한 검사를 위한 고도화 작업을 필요로 한다. 이에 본 연구는 위해도 예측 모델의 개발과 고도화를 통해 성능을 향상시키는 방안을 탐색하였다. 위 두 모델 모두 정밀한 검사를 위한 고도화 작업을 필요로 하며 위해도 예측 모형 고도화 방안을 다양한 관점에서 탐구하는 연구는 미비한 실정이다. 따라서 본 연구에서는 위해도 예측 모델 개발 뿐 아니라 이 모델을 고도화하여 성능을 향상시키기 위한 다양한 방안은 탐색하였다.

II. 방법론

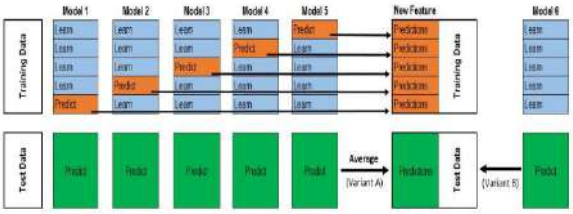
위해도 사전 제작은 <그림 1>과 같은 과정을 거쳐 진행하였다. 먼저, 정확한 값들이 감성점수로 매겨질 수 있도록 사용자 정의 사전을 이용하여 형태소 분석을 실시하였다. 위해점수는 식품 위해적 관점에서 하나의 품목에서 부적합 사유로 검출된 위해물질들의 개수와 상관없이 가장 높은 위험을 가진 물질 하나로 인해 해당 품목의 부적합이 결정되므로 품목명 위해도 점수 산출 식을 이용하여 최종점수로 반영하였다. 이후, 품목명 변수에 해당

위해도 점수를 맵핑하여 최종 점수를 환산하였다. 마지막으로, 산출한 품목명 위해도 점수를 해당 품목에 맞게 매핑하여 '품목명 위해도' 변수를 생성하였다.



<그림 1> 품목명 위해도 변수 생성 Framework

모델 고도화 작업은 차원축소, 스태킹앙상블, 이동평균 시계열 가중치 등을 실행하였다. 차원축소는 연속형 변수에는 PCA 기법, 범주형 변수에는 MCA 기법을 적용하여 기존에 가장 높았던 기본 Ensemble 모델의 성능 결과와 비교 검증을 실시하였다. 스태킹 앙상블 실험은 <그림 2>와 같은 과정으로 진행되었는데, 기존 SoftVoting Ensemble 과 Stacking Ensemble 의 성능 비교 검증방식으로 진행되었으며, 소요시간을 줄이기 위해 차원축소 기법을 먼저 적용한 후에 Stacking Ensemble 을 실시하였다.



<그림 2> 스태킹 앙상블 실험 Framework

시계열 가중치는 가중이동평균(WMA)과 지수이동평균(EMA) 기법을 사용하여 '부적합률_품목명'에 가중치를 부여한 후 변수 중요도를 확인하였다.

III. 결론

다양한 모델 고도화 실험을 통해, 본 연구에서 활용하는 데이터 셋에 어떤 고도화 방식이 가장 적합한지를 검증한 결과, 식품 위해 사전 제작, 차원축소, 스태킹 앙상블, 이동평균 시계열 가중치 방법으로 모델의 성능이 유의하게 개선된 것을 확인하였다. 식품 위해 사전을 제작하여 생성한 품목명 위해도 변수의 경우 전반적으로 모든 모델의 Recall 이 개선되거나 비슷한 것을 확인하였으며, 가장 성능이 우수한 앙상블 모델의 성능이 앙상블 기술을 사용하지 않은 기존 모델의 성능과 비교하였을 때 0.502 에서 0.555 로 향상된 것을 확인할 수 있었다. 차원축소에 있어서는 연속형 변수에 PCA, 범주형 변수에 MCA 를 적용하는 것이 가장 효과적이었으며, 스태킹 앙상블에 있어서는 Final_estimator 가 AdaBoost 일 때 가장 효과적이었다. 이동평균 시계열 가중치를 부여할 때는 지수평활계수를 0.8 로 부여한 지수이동평균이 가장 효과적이었다. 이러한 고도화의 우수한 결과는 앞으로의 추가적인 성능 고도화 방법에 대한 방향성을 제시해준다. 추후 새로운 고도화 작업 수행 시, 기존 고도화에서 효과가 좋았던 방법론을 적절히 혼합해 연구를 진행한다면, 수많은 시행착오를 줄이고 더욱 좋은 성능을 제고할 수 있다. 추후 모델 고도화를 실시할 때, 변수 중요도가 낮은 변수들을 제거하고 차원축소를 진행해 볼 필요가 있다. 해당 방식을 통해 데이터 노이즈를 효과적으로 제거할 수 있으며, 학습 시간 단축 등 성능 개선의 여지가 충분하다.

IV. 참고문헌

Choubey, D. K., Kumar, M., Shukla, V., Tripathi, S., & Dhandhania, V. K. (2020). Comparative analysis of classification methods with PCA and LDA for diabetes. Current diabetes reviews, 16(8), 833-850.

Goldberg, D. M., Khan, S., Zaman, N., Gruss, R. J., & Abrahams, A. S. (2022). Text mining approaches for postmarket food safety surveillance using online media. Risk Analysis, 42(8), 1749-1768.

IZSTO, Ru, G., Crescio, M. I., Ingravalle, F., Maurella, C., UBESP, ... & Rosat, L. (2017). Machine Learning Techniques applied in risk assessment related to food safety. EFSA Supporting Publications, 14(7), 1254E.

Kaviani, M., Dara, R., & Farber, J. M. (2022). Machine learning application in food safety, production, and quality. Harnessing Big Data in Food Safety, 1-21.

Zhou, T., & Jiao, H. (2023). Exploration of the stacking ensemble machine learning algorithm for cheating detection in large-scale assessment. Educational and Psychological Measurement, 83(4), 831-854.

Engaging Audiences with Varicose Veins Information: Insights from YouTube Content Analysis

이홍주
가톨릭대학교 경영학과
hongjoo@catholic.ac.kr

김연정
가톨릭대학교 경영학과
yeonjeongg12@gmail.com

이준우
가톨릭대학교 경영학과
junwoo3610@gmail.com

***Abstract** – This study explores the relationship between medical information and user response in varicose vein-related video content on YouTube. Through data collection, medical term extraction and classification, evaluation term analysis, topic modeling, and linear regression analysis, the following findings were obtained.*

First, videos with low view counts lacked information applicable to daily life. Videos that only included specialized information about diseases and surgeries, without information on exercise or diet that can be applied in daily life, seemed to fail to attract the attention of general viewers.

Second, positive evaluation terms were mainly used for terms related to cause, treatment, and examination, while negative evaluation terms were mainly used for terms related to surgery. For the terms related to the disease itself, both positive and negative evaluation terms were used. This confirms that the sentiment of evaluation terms can be distinguished for each topic.

Third, the analysis of the correlation between view count and topics of varicose vein-related video content revealed that Topic - Physical Education and Muscles and Topic - Diet and Health were associated with a decrease in view count. This suggests that these topics are not directly related to varicose veins and therefore fail to attract viewers' interest.

Our findings suggest that when providing information on major topics such as blood circulation, symptoms, and treatment methods, it is necessary to also provide information that can be applied in daily life in order to attract viewers' attention. Additionally, it is important to be careful about the use of evaluation terms, as they can affect viewers' perception of the disease. Finally, it is necessary to avoid composing videos solely with topics that have low relevance, such as physical education and muscles, and diet and health.

This study serves as a preliminary investigation of the relationship between medical information and user response in varicose vein-related video content. Future research with more data can yield more specific and practical results.

Keywords** – **Social Media, Medical Information, Information Dissemination, Text Mining

사사표기

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020S1A3A2A 02093277).

참고문헌

Fortuna, G., Schiavo, J. H., Aria, M., Mignogna, M. D. & Klasser, G. D. (2019). The usefulness of YouTube™ videos as a source of information on burning mouth syndrome. *Journal of Oral Rehabilitation* 46, 657–665.

Liu, X., Zhang, B., Susarla, A., & Padman, R. (2020). Go to youtube and call me in the morning: Use of social media for chronic conditions. *MIS Quarterly 44*(1), 257-283.

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S. & Gramopadhye, A. K. (2015). Healthcare information on YouTube: A systematic review. *Healthcare Informatics Journal* 21, 173–194.

Onder, M. E. & Zengin, O. (2023). Quality of healthcare information on YouTube: psoriatic arthritis. *Zeitschrift für Rheumatologie* 82, 30–37.

ReFaey, K. *et al.* (2018). The reliability of YouTube videos in patients education for Glioblastoma Treatment. *Journal of Clinical Neuroscience* 55, 1–4.

Saffi, H., Do, T. P., Hansen, J. M., Dodick, D. W. & Ashina, M. (2020). The migraine landscape on YouTube:

A review of YouTube as a source of information on migraine. *Cephalalgia* 40, 1363–1369.

Yuksel, B. & Cakmak, K. (2020). Healthcare information on YouTube: Pregnancy and COVID-19. *International Journal of Gynecology Obstetrics* 150, 189–193.

생성형 AI에 의한 이미지 생성이 미술관 방문의도에 미치는 영향

– 작가와 미술관 조절효과를 중심으로 –

최륜희*
*경희대학교 경영학과, 박사과정
(ryunh@khu.ac.kr)

이은서*
*경희대학교 빅데이터응용학과, 석사과정
(les1512@khu.ac.kr)

권오병**
**경희대학교 경영학과
(obkwon@khu.ac.kr)

CAITech
research center for advanced information technology

목 차

Ⅰ 서론

Ⅱ 이론적 배경

Ⅲ 연구모형

Ⅳ 연구방법

Ⅴ 연구 진행 사항

CAITech
research center for advanced information technology

1. 서론

Ⅰ Ⅱ Ⅲ Ⅳ Ⅴ

1.1 연구배경 및 목적

- 생성형 AI는 새로운 툴로 미술, 음악, 영상, 소설 등 다양한 분야에서 활약 중
  - 고품질의 결과를 제시하고, 사람이 한 것과 매우 유사함
  - 유명 화가의 그림기법, 시인의 작문법, 음악 창작 등 예술/문화 분야에서 많은 이슈가 되고 있음
  - 생성형AI는 교육/엔터테인먼트/ 창작 활동 등 다양한 분야에서 활용 중(Chiu, (2023); Schuckart, (2023))

분야	생성형AI
그림	**DALL-E, Midjourney, Stable Diffusion, …**
글쓰기	ChatGPT, NovelAI, Bard, Bing, wrtn, clova, …
음악	Mix.audio, viodio, …
가상인간&음성	Lensa, OnAir Studio, ploonet, typecast, clovadubbing, …

- 생성형 AI가 창작한 작품 감상 후, 실제 작품을 보기 위해 미술관 방문의도에 관한 연구는 미비한 실정
  - 미술관과 박물관의 체험, 평소 인식, 방문 인식, 온라인 미술관의 방문의도 등에 대한 연구는 진행되었으나, 스타일 전이(style transfer)를 활용한 생성형 AI 이미지 작품이 미술관 방문의도에 영향을 미치는 연구는 미비함
- 스타일 전이를 활용한 생성형AI 예술 작품이 심미적 체험 중심의 체험경제이론에 따라, 미술관 방문 의도에 미치는 영향을 연구하고자 함

2. 이론적 배경 (1/2)

Ⅰ Ⅱ Ⅲ Ⅳ Ⅴ

2.1 GAN의 스타일 전이(Style Transfer)를 적용한 생성형AI 이미지 생성

- GNN 기반의 스타일 전이를 적용한 생성형AI를 사용하고자 함
  - CNN기반 Neural Style Transfer를 사용하여, 원본 이미지와 적용할 스타일의 이미지 쌍을 학습시켜 새로운 이미지를 생성함
  - GAN기반 스타일 전이는 원본 컨텐츠 구조를 유지하면서, 변환하고자 하는 스타일을 적용해 새로운 이미지를 생성함

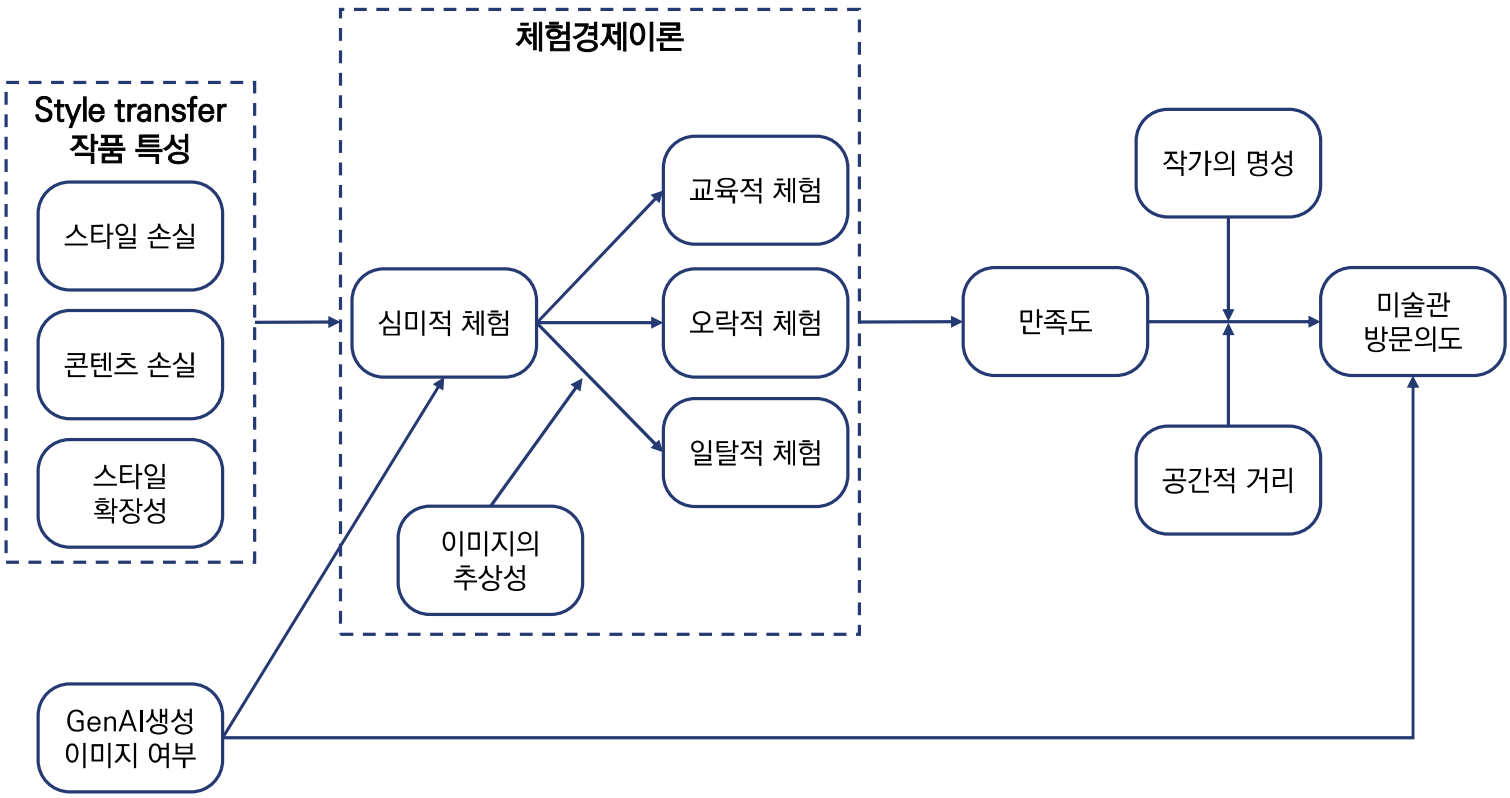
	CNN 기반 Style transfer	GAN 기반 Style transfer
비교 예시		
주요 기능	이미지의 컨텐츠와 스타일을 분리 및 재조합	생성자와 판별자 두 모듈로 새로운 이미지 생성
핵심 기술	컨볼루셔널 신경망의 중간층 특성 활용(스타일을 적용할 이미지와 목표 스타일 이미지 간의 쌍이 필요함)	생성적 적대 신경망 구조 활용(스타일 다양성을 학습하기 위한 레이블 없는 데이터로부터 학습하여 이미지 쌍이 필요 없음)
목표 및 적용 범위	이미지의 컨텐츠 유지 및 스타일 변경으로, 스타일 복제 및 적용에 초점	다양하고 혁신적인 이미지 생성으로, 스타일 전이 및 다양한 이미지를 생성함
특징	• 정확한 스타일 복제 가능 • 높은 컨텐츠 일관성을 유지한 상세 스타일 적용 가능 • 생성 가능한 스타일의 다양성 한계 • 사전에 정의된 스타일 적용에 적합	• 창의적이고 새로운 결과물 생성 가능 • 고도의 학습과 정교한 모델 설계 필요 • 광범위한 스타일 변환 및 창의적 이미지 생성에 유용

2. 이론적 배경 (2/2)

2.2 체험경제이론

- Pine & Gilmore(1998,1999)의 체험경제이론은 심미적/오락적/교육적/일탈적 체험 요소로 구성되어, 체험 마케팅, 관광 등의 분야에서 자주 사용됨
  - 전통적으로 교육, 관광, 상거래 등 소비자의 체험을 강조하는 등 전 비즈니스 영역에서 소비자들의 만족감 혹은 구매 의도를 제고하는 현상을 유용하게 설명해 활용되는 이론임
  - 생성형AI 이미지를 통한 심미적 체험에 따라, 오락적/교육적/일탈적 체험 요소에 미치는 영향에 대해 연구하고자 함
- 체험경제이론의 체험요소와 미술관 방문의도에 미치는 연구를 진행하고자 함
  - 미적 체험 영역에서 활동이나 환경에 거의 영향을 미치지 못하지만, 미학적 관점에서 참여자가 있는 공간이 실제로 그곳에 있는 것과 같은 미적 체험을 경험함(Oh et al.,2007)
  - 미술 교육/홍보 등을 위해 다양한 프로그램을 기획하며, 소장 작품 또는 컬렉션 홍보를 위해 VR, 홈페이지 등을 이용, 이에 생성형AI를 활용하여 미술관 방문의도를 연구하고자 함

3. 연구모형



4. 연구방법 (1/2)

4.1 조작적 정의

변수	아이템	조작적 정의	참고 문헌
스타일 손실	STYLOSS	적용한 스타일 이미지의 질감이 새 이미지에 반영된 정도(이미지 A와 B의 차이 값↓, 유사도↑)	Kotovenko, et al.(2019); Sanakoyeu, et al.(2018) ; Nicholas, et al.(2022)
컨텐츠 손실	CONLOSS	원본 이미지와 생성된 이미지 간의 내용 차이가 최소화되어야 함	
스타일 확장성	STYCAL	다양한 스타일을 효과적으로 적용 가능하게 하고, 컨텐츠의 왜곡없이 스타일을 생성함	
심미적 체험	EST	체험을 통해, 보게 되는 생성형AI의 이미지 작품을 감상하는 행위	Pine&Gilmore(1998); Oh,&Jeoung, (2007); 김지희 외,(2010) 박수경 외,(2007)
교육적 체험	EDU	체험을 통해, 관련 지식/정보를 알려주고, 향상시켜주는 요소	
오락적 체험	ENT	체험을 통해, 재미있고, 즐겁고, 관심을 사로 잡고, 기분이 좋아지는 상태를 만드는 요소	
일탈적 체험	ESP	체험을 통해, 현실에서 벗어나는 몰입감과 시간과 노력을 투자할 만한 가치가 있는 요소	
이미지의 추상성	ABS	생성된 이미지에 적용된 컨텐츠 이미지가 추상화의 모습을 얼마나 갖고 있는지를 뜻함	Created
작가의 명성	REP	세상에 널리 퍼져 평판 높은 이름의 사전적 정의로, 타인이 인정하는 정도로 일반적으로 이름이 알려진 정도를 뜻함	최영신&현은정 (2018)
공간적 거리	PRX	개인이 주관적으로 인식하는 거리감, 공간적 거리에 대한 인지수준을 뜻함 (예: 얼마나 먼 거리에 위치해 있는가)	Trope&Liberman, 2010; Trope et al., 200; 권미혜 외,(2016)

4. 연구방법 (2/2)

4.2 생성형AI 이미지 제작 과정

- 원본 이미지에 적용할 스타일(겸재정선)을 입혀, 결과물을 생성함
  - Ostagram 활용
- 원본 이미지와 결과물을 같이 보여주면서, 설문조사 실시 예정

가짜뉴스가 주식시장에 미치는 영향: 과거 사례의 조절효과

박준휘
한양대학교 일반대학원
wnsgnl8229@hanyang.ac.kr

김종우
한양대학교 경영대학
kjw@hanyang.ac.kr

Abstract *– 최근 정보 기술의 발전으로 주식시장 참여자들은 투자의사결정 과정에서 뉴스를 참고해 의사결정을 한다. 그러나, 최근 가짜뉴스의 증가는 주식시장의 불확실성을 증가시키며, 시장 참여자들의 권익을 침해하고 있다. 본 연구는 가짜뉴스가 주식시장의 불확실성에 미치는 효과를 분석하고, 투자자들이 과거 정보를 활용하여 효과적으로 가짜뉴스에 대응하는지에 대한 여부를 과거 기업관련 가짜뉴스 발생 이력과, 기사의 발화자의 과거 거짓 진술 이력을 조절변수로 사용하여 검증을 수행한다.*

Keywords *– 가짜뉴스, 자연어처리, 조절효과, 주식 이상 거래, Natural Language Inference.*

이 논문은 2020 년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2020S1A3A2A02093277)

I. 서론

정보 기술의 발전은 정보의 확산 및 공유 속도를 증가시켜, 정보 격차를 해소하는데 도움을 준다는 긍정적 측면을 갖는다. (Kaplan et al., 2019) 이러한 정보의 공유는 시장 참여자들의 능동적 정보 탐색활동을 지원하며, 투자 의사결정에 중요한 역할을 한다. 하지만, 인터넷에서 최근 기관과 개인이 의도적으로 거짓 정보를 퍼뜨리는 현상이 증가하고 있으며 (Siering et al., 2017; Subramanian, 2017), 소셜 미디어, 블로그 등과 같은 감시가 어려운 정보 공유 플랫폼이 폭발적으로 증가함에 따라 가짜 및 편향된 뉴스의 발생과, 이로 인한 잠재적 영향이 커지고 있다. 일례로, 2024 년 1 월 9 일 미국 금융당국의 소셜미디어(SNS) 공식 계정에 비트 코인의 현물 상장지수펀드(ETF)가 승인됐다는 가짜뉴스가 게시되었고, 당국이 “계정이 해킹됐다”며 곧바로 승인 사실을 부인하고 이를 삭제하는 소동이 벌어졌으며, 이 사건으로 비트 코인 가격은 1 개당 4 만 8000 달러 부근까지 치솟았다가 당국의 부인으로 급락하는 큰 변동성을 보였다. 가짜뉴스는 진짜 뉴스에 비해 더 빠른 확산 속도와 넓은 확산 범위를 갖기 때문에, 정보에 따른 의사결정이 활발한 금융시장에서, 거짓 정보의 확산은 심각한 경제적 피해를 야기할 수 있다. 또, 가짜뉴스의 발생 이력은 공개정보에 대한 대중의 신뢰도를 저해하기 때문에, 새로운 정보가 주식시장에 미치는 효과를 교란한다는 연구 결과가 존재한다. 따라서, 본 연구에서는 과거 기업관련 가짜뉴스의 발생이력과 가짜뉴스 발화자의 과거 거짓 진술 이력을 조절변수로 하여, 가짜뉴스가 주식관련 진짜뉴스 및 가짜뉴스가 주식시장에 미치는 효과를 교란하는가에 대한 여부를 검증하는 연구를 수행한다.

II. 문헌검토

현재, 가짜뉴스와 연관된 선행연구들은 주로 뉴스의 내용 및 형식에 대한 논의가 지배적인 추세이다. 내용의 감성적 특성에 따른 주식시장의 변화를 살펴보거나 (Arcuri et al., 2023), 기사 내용의 감정적 특성 (Alzahrani et al,2023)에 따른 특징을 살펴보는 등 기사의 내용적 특성에 기반한 논의가 이루어지고 있으며, 이러한 연구들은 투자자의 가짜뉴스 식별 가능 여부에 대한 논의를 다루지 않고 있다. 주식 투자자들은 과거 발생한 가짜뉴스 사례를 기반으로 뉴스의 내용을 차별적으로 투자의사결정에 반영한다. 선행연구에 따르면, 미국 증권거래위원회가 2011 년 발표한 ‘홍보성 주식 뉴스’단속 사례를 기점으로, 가짜뉴스가 주가에 미치는 효과가 감소했을 뿐 아니라, 진짜뉴스가 주식 가격에 미치는 효과 또한 감소했다는 연구 결과가 존재한다 (Kogan et al., 2019). 따라서, 본 연구는 투자자들이 과거 정보를 활용하여 효과적으로 가짜뉴스에 대응하는가에 대한 분석을 진행한다. 구체적으로 해당 연구에서는 기업과 관련된 과거 가짜뉴스 발생이력과, 가짜뉴스 발화자의 과거 거짓말 이력을 조절변수로 사용하여, 해당 정보가 투자의사결정에 영향을 미치는가에 대한 분석을 진행한다. 앞선 논문에서는 SEC 공표 사례를 기점으로 가짜뉴스의 효과의 변화를 탐색했으나, 이는 시장 전체의 수준에서의 변화를 추적했기 때문에, 해당 효과를 뉴스의 진위여부, 기업 단위로 확장해, 특정 기업에 대한 가짜뉴스가 미치는 효과를 추가적으로 분석하고자 한다. 또, 이 논문에서는 발화자의 특성을 과거 해당 발화자의 거짓말 이력을 사용하여 사용한다.

거짓 정보의 최초 유포자의 과거 거짓말 이력은 뉴스의 신뢰성에 영향을 미치는 주요 요인 중 하나일 것이다. 과거 선행연구에서는 인물의 직업적 특성 (Hong, Ziyang, et al., 2023)이나 인물의 영향력을 반영하는 방식의 연구(Vosoughi et al., 2018;Arcuri et al., 2023) 등, 인물의 영향력에 집중한 연구는 다수 존재했으나, 인물의 신뢰 가능성을 중심으로 한 연 구는 존재하지 않았다. 따라서, 본 연구에서는 과거 해당 인물의 거짓말 이력을 인물에 대한 신뢰 가능성의 지표로 두고, 이를 조절변수로 활용하는 연구를 진행한다.

Ⅲ. 데이터

가짜뉴스의 경우, 발생 이후 수일 내에 원문에 대한 정보가 삭제되는 경향이 강해 데이터에 다소 어려움이 존재한다. 선행연구들의 경우, 공공기관에서 발표하는 가짜뉴스 단속 사례를 기반으로 분석을 진행하거나 (Arcuri et al., 2023; Hong, Ziyang, et al., 2023), 뉴스의 진위여부를 평가하는 공공 웹사이트의 데이터를 기반으로 연구를 진행하였다. (Vosoughi et al., 2018; Ha Louisa et al., 2023; Alzahrani et al,2023). 공공기관 발표 데이터의 경우, 그 규모가 다소 작고, 원본의 내용에 대한 정보를 얻기에 제한이 있기 때문에, 해당 연구는 대표적인 '팩트 체크' 웹사이트인 politifact.com 의 데이터를 수집해 사용한다. Politifact.com 은 미국의 비영리단체 'Poynter Institute'에서 운영하는 프로젝트로, 단체에 소속된 저널리스트들이 기사의 핵심 내용을 대상으로 진실 여부에 대한 평가를 진행해, 평가의 근거와 함께 기사를 6 개의 진실 등급으로 분류한다. 또, 최초 기사가 발행된 매체에 대한 정보와, 발화자의 이름 정보를 갖기 때문에, 분석 목적을 달성하기 위한 가장 합리적인 데이터를 포함하고 있는 웹사이트이다. 연구는 해당 사이트를 대상으로 S&P500 지수에 편입되어 있는 기업명과 Nasdaq 상장 기업을 대상으로 검색을 진행하여, 데이터를 수집해 사용하였다. 발화자의 과거 거짓 이력은 수집된 데이터를 대상으로 최초 발화자의 정보를 추출하였으며, politifact.com 내의 해당 인물관련 페이지에서, 과거 거짓말 이력을 수집하였으며, 과거 기업 관련 가짜뉴스의 발생은 데이터 내의 대상 기업명을 기준으로 데이터를 집계해 사용하였다.

한편, 가짜뉴스의 경우, 원문의 정보와 대부분 삭제되어, 기사에 대한 대중의 관심을 측정할 수 있는 '조회수', '댓글의 수','추천 수'등 1 차적 관심도 지표를 수집하기에 어려움이 존재한다. 이전의 가짜 뉴스 관련 선행연구들의 경우, 개별 가짜뉴스에 대한 대중의 관심도 차이를 통제하지 못했다는 한계를 지닌다. 따라서, 본연구에서는 기사에 대한 관심도 지표를 온라인 커뮤니티(Reddit)에서의 각 기사에 대한 게시글 및 댓글의 수를 개별 기사에 대한 관심도 통제변수로 사용한다. 구체적으로, 연구에서는 기업관련 가짜뉴스가 발생한 이후 시점의 기업관련 reddit 게시물과 댓글 정보를 수집하고, NLI(Natural Language Inference) 방법론을 적용해, 가짜뉴스의 내용을 담고 있는 게시글과 댓글 정보를 추출해, 전체 게시글 대비 해당 뉴스에 대한 언급 량을 변수화해 통제변수로 사용한다.

또한 연구에서는 기업의 특성에 따른 영향력을 통제하기 위한 변수를 사용한다. 가짜뉴스의 효과는 기업 및 기업 주식의 특성에 따라 그 영향력의 크기가 달라지는 경향이 있다. 특히, 가짜뉴스가 주식시장에 미치는 효과는 기업의 규모가 클수록 작아지는 경향이 있으며(Kogan et al., 2019), 기업이 속한 산업에 따라 그 영향력이 달라지는 경향(Arcuri et al., 2023)이 있기 때문에, 선행연구에 따라 기업 특성과 산업 특성에 따른 통제변수를 추가해 연구를 진행하였다.

IV. 참고문헌

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business horizons, 62(1), 15-25.

Hong, Z., Liu, Q., Tse, Y., & Wang, Z. (2023). Black mouth, investor attention, and stock return. International Review of Financial Analysis, 90, 102921.

Arcuri, M. C., Gandolfi, G., & Russo, I. (2023). Does fake news impact stock returns? Evidence from US and EU stock markets. Journal of Economics and Business, 125, 106130.

Alzahrani, A. I., Sarsam, S. M., Al-Samarraie, H., & Alblehai, F. (2023). Exploring the sentimental features of rumor messages and investors' intentions to invest. *International Review of Economics & Finance*, *87*, 433-444.

Kogan, S., Moskowitz, T. J., & Niessner, M. (2019). Fake news: Evidence from financial markets. Available at SSRN, 3237763.

Ha, L., Rahut, D., Ofori, M., Sharma, S., Harmon, M., Tolofari, A., ... & Khan, A. (2023). Implications of source, content, and style cues in curbing health misinformation and fake news. Internet Research, (ahead-of-print).

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. science, 359(6380), 1146-1151.

메타데이터 기반 3D 애셋 검색 기능 강화를 위한 사용자 친화적 추천 시스템 설계 방안 연구

권상지
경희대학교 경영학과
aaaitaaa@khu.ac.kr

허지혜
경희대학교 빅데이터응용학과
floramiss@khu.ac.kr

이은서
경희대학교 빅데이터응용학과
les1512@khu.ac.kr

최상호
경희대학교 경영학과
reo.choi@khu.ac.kr

최륜희
경희대학교 경영학과
ryunh@khu.ac.kr

권오병
경희대학교 경영학과
obkwon@khu.ac.kr

Abstract – ***본 연구의 목적은 3D 애셋의 메타데이터를 기반으로 사람의 감성 데이터를 추가하여 3D 애셋 검색기능을 강화하고, 누적된 검색결과를 통해 감성 기반 추천 시스템을 개발하여 사용자 경험을 향상시키는 것이다. 이를 위해, 누적된 감성데이터 분석을 통해 사용자의 선호도와 반응을 파악하고, 이를 추천 알고리즘에 통합함으로써 더욱 정교하고 개인화된 추천을 제공하고자 한다. 또한, 협업 필터링 기법을 결합하여 사용자 간의 상호작용과 유사성을 분석함으로써 컨텐츠의 발견성을 높이고, 사용자에게 더욱 적합한 3D 애셋을 추천할 수 있는 시스템을 설계한다. 본 연구는 K-콘텐츠를 활용한 3D 애셋 플랫폼의 검색 및 추천 시스템 효율성과 정확도를 증진시키는 새로운 접근 방식을 제시하며, 사용자 중심의 서비스 경험을 제공함으로써 3D 콘텐츠 시장의 확장에 기여할 것으로 기대된다.***

Key Terms – ***감성분석, 추천 시스템, 협업 필터링***

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2024 년도 문화기술 연구개발 사업으로 수행되었음 (과제명: 디지털 3D 애셋의 지능형 축적 및 공유 플랫폼 개발, 과제번호: RS-2024-00225159)

I. 서론

현재 디지털 콘텐츠 시장의 성장으로 3D 애셋의 중요성이 더욱 부각되고 있다. 특히 K-콘텐츠 산업 활성화를 위해서는 한국 고유의 창작 가치와 기술 혁신을 융합한 새로운 비즈니스 모델과 사용자 경험 창출이 필요하다. 그러나 아직 3D 콘텐츠에 대한 성공적인 비즈니스 모델이 미흡한 실정이다. 이에 사용자에게 고품질의 맞춤형 3D 콘텐츠를 제공하는 서비스가 필요하다. 그러나 기존의 3D 애셋 플랫폼들은 기본적인 메타데이터 기반의 검색 기능에 의존하고 있어 사용자의 구체적이고 정교한 검색 요구와 감성을 반영하기에는 한계가 있다.

따라서 본 연구의 목적은 K-콘텐츠를 활용한 3D 애셋 플랫폼에서 기존 메타데이터에 사용자 감성 데이터 정보를 추가하여 로그에 누적된 감성 데이터를 기반으로 추천 시스템 강화, 협업 필터링을 활용한 사용자 간의 상호작용과 선호도를 파악하는 방법을 제안하는 것이다. 이를 통해 개인화된 추천을 제공할 수 있는 알고리즘을 개발하여 사용자 경험을 향상시키고자 한다. 본 연구에서는 3D 애셋에 내장된 감성데이터를 활용하여 사용자의 선호도와 반응을 보다 정밀하게 파악하고, 사용자에게 더욱 적합한 3D 애셋 콘텐츠를 추천함으로써, 최종적으로 사용자 중심의 풍부한 서비스 경험을 제공하고자 한다.

II. 선행 연구

2.1 감성기반 추천 시스템

감성기반 추천은 사용자의 감정을 인식하고 분석하여 그에 맞는 개인화된 추천을 제공하는 기술이다. 최근에는 감성기반 추천 시스템이 더욱 세분화되고 정교해지고 있다. 특히 자연어 처리(NLP), 기계학습, 딥러닝 같은 기술의 발전으로, 복잡한 사용자 데이터에서 세밀한 감정 상태를 파악하고 이를 기반으로 사용자에게 맞춤형 추천을 제공하는 연구가 활발히 이루어지고 있다.

기존의 추천 시스템에서는 사용자와 3D 애셋 아이템 간의 상호작용이 충분하지 않아 효과적인 추천을 하는데 어려움을 겪고 있다. 그러나 추천 시스템에 정서 요인을 추가함으로써 사용자의 선호와 반응을 더 깊이 이해할 수 있고, 이로써 사용자 만족도를 높이는 개인화된 추천을 가능하게 한다(Dang, et al., 2021).

2.2 협업 필터링 추천 시스템

협업 필터링은 특정 아이템에 대한 선호도가 유사한 사용자들은 다른 아이템에 대해서도 비슷한 선호도를 나타낼 것이라는 기본 가정을 바탕으로 사용자 혹은 아이템 간 유사도를 기반으로 선호도를 예측하는 방법이다(현지연 외, 2019). 특히, 로그 데이터를 활용한 협업 필터링은 사용자의 암묵적 행동 데이터를 깊이 분석하여 이를 기반으로 사용자의 선호도를 정밀하게 예측한다. 또한, 사용자의 감정 분석 이력에서 얻은 데이터는 콘텐츠에 대한 감정적 반응을 보여주어, 추천 시스템이 사용자의 정서적 선호를 더 잘 이해하는 데 도움을 줄 수 있다.

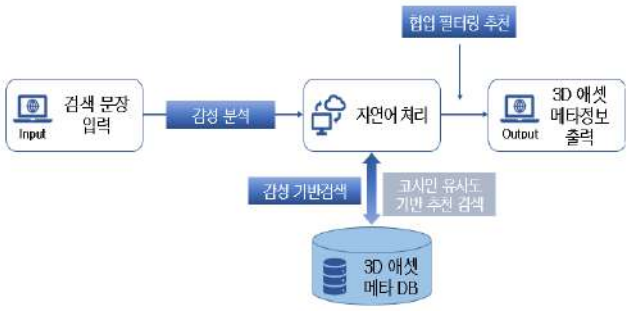
이처럼 다양한 형태의 로그 데이터를 분석함으로써, 협업 필터링 시스템은 사용자 개개인의 독특한 행동 패턴과 선호도를 모델링 할 수 있다. 이 모델을 바탕으로 시스템은 사용자가 발견하지 못한 새로운 콘텐츠나 상품을 추천함으로써 개인화된 사용자 경험을 제공할 수 있다.

III. 방법론

3.1 3D 애셋의 감성기반 추천 시스템

3D 애셋 데이터베이스에 통합된 감성 데이터를 추가하여 이를 활용한 감성기반 추천 시스템을 구축한다. 이는 사용자의 검색 쿼리에 대한 감성 분석으로부터 시작되며, 고급 자연어 처리 모델인 Sentence BERT(SBERT) 및 Korean Online That-gul Emotions(KOTE)로 사전학습 된 KcElectra 모델을 활용하여 검색어에서 감성을 정밀하게 분석하도록 하였다(Jeon, et al., 2022).

한편, 감성 분석의 정확성을 제고하기 위해, 다차원 감성 요인 중 감성 점수가 특정 임계치(예: 0.5) 이상인 것 중 최대 상위 3 개의 감성을 추출한다. 이 과정에서 KOTE 모델은 텍스트에 나타난 감정의 뉘앙스를 심층적으로 이해하고, 문맥적 요소를 고려하여 더 정확한 감성 분석 결과를 도출한다. 추출된 감성 데이터는 데이터베이스 내에서 검색어와 코사인 유사도가 높은 3D 애셋을 식별하고, 해당 감성에 부합하는 3D 애셋을 추천 목록에 포함시킨다. 사용자가 제공한 검색어를 기반으로 누적된 감성 데이터는 감성 기반의 추천 시스템을 위한 핵심 요소로 작용한다. 사용자가 n 회 이상 검색할 경우, 누적된 감성 데이터로 감성 기반 추천 시스템의 정확성을 지속적으로 개선할 수 있다.



<그림 1> 감성기반 추천 시스템 프로세스

3.2 감성 기반 협업 필터링 추천 시스템

다음으로 사용자의 로그 데이터를 활용하여 협업 필터링 메커니즘을 강화한다. 사용자 간의 로그 데이터를 분석하여, 유사한 검색 및 감성분석 패턴을 보이는 사용자 그룹을 식별하고, 이 그룹 내에서 상호 추천을 통해 사용자 경험을 풍부하게 한다. 먼저 사용자의 로그 데이터를 수집하고 분석하여, 패턴 인식을 통해 유사한 행동(감성)을 보이는 사용자 그룹을 식별한다. 이 과정에서 클러스터링 알고리즘을 사용하여, 사용자들을 그들의 행동(감성) 유사성에 기반한 세분화된 그룹으로 분류한다. 분류된 사용자 그룹은 상호 추천 메커니즘의 핵심 요소로 작용한다. 시스템은 그룹 내 각 개인의 로그

데이터를 실시간으로 종합하여, 그룹에 속한 다른 사용자에게도 호응을 얻을 가능성이 높은 3D 애셋을 추천한다.

IV. 결론

현재 3D 애셋 데이터베이스에 감성데이터를 구축하는 중이며, 샘플 데이터를 활용하여 감성데이터와 협업 필터링을 활용한 사용자 맞춤형 3D 애셋 추천 시스템을 개발 중이다. 개발이 완료되면 실제 데이터셋을 가지고 추천 시스템의 정확성과 사용자 대상 만족도 실험을 통해 검증할 예정이다.

V. 참고문헌

Jeon, D. & Lee, J. & Kim, C. (2022). User Guide for KOTE: Korean Online Comments Emotions Dataset.

Dang, C.N. & Moreno-Garcia, M.N. & De la Prieta, F.. (2021). An Approach to Integrating Sentiment Analysis into Recommender Systems

현지연&유상이&이상용. (2019). 평점과 리뷰 텍스트 감성분석을 결합한 추천시스템 향상 방안 연구

Session D

D4. [학술세션]
산업별 데이터 기반 수요 예측 모델

좌 장: 류 민 호(동아대)

Daily tourism demand forecasting with iTransformer model

황가혜
국민대학교 경영대학
huangjiahui611@163.com

장청롱
국민대학교 경영대학
JohnZhang@kookmin.ac.kr

Abstract – *This study is based on an inverted transformer (iTransformer) approach with a self-attention mechanism, which forecasts the daily tourism volume of a city by combining multivariate time series data, including historical tourism volume, search engine queries, weather, holiday, seasonal and tourism activities data. Compared to baseline models, the effectiveness of the iTransformer in predicting daily urban tourism demand has been confirmed. The findings of this study can be applied to forecast daily visitor arrivals in tourist cities, assisting tourism-related decision-makers in making more accurate predictions and formulating more reliable plans.*

Key Terms – Tourism demand, Forecasting, Inverted transformer, iTransformer, Multivariate time series

I. Introduction and Literature Review

The tourism industry has gradually become a pillar industry in the economic development of many countries around the world. Due to internal and external factors, tourism demand exhibits significant fluctuations（Song et al., 2019). Tourism management departments and enterprises require accurate tourism demand forecast, as accurate tourist volume forecast is an important part of tourism destination management and efficient business operation (Palmer et al., 2006). In existing tourism demand forecasting research, in terms of forecasting granularity, the focus is mainly on long-term international tourism forecasts (Havranek et al., 2021), as well as short-term forecasts for tourist attractions (Xu et al., 2023) and partially on hourly-level forecasts for tourist volumes at attractions (Xue et al., 2023). Due to international tourism being easier to track and measure compared to domestic tourism, existing research mainly focuses on international tourist volumes. However, since the primary market for the tourism industry consists of domestic tourists, short-term forecasts of domestic tourism demand for a city are more critical for destination governments and tourism enterprises.

In this study, we take the famous Chinese tourist city of Hangzhou as an example and employ a deep learning model to establish a daily urban tourism demand forecasting model. The forecast results usually depend on the features of the data and the characteristics of the research. To achieve more accurate forecasts, two main factors are primarily considered: the predictive variables used and the methods or techniques employed (Bi et al., 2020). Regarding the first factor, historical tourism volume, search engine queries and weather data are commonly used variables in tourism demand forecasting. As for the second factor, the most commonly used methods for tourism demand forecasting include time series models, econometric models, and artificial intelligence models (Wu et al., 2024). The Transformer introduced the self-attention mechanism, which does not rely on recursion and convolution to generate outputs, and it has shown significant advantages in the field of time series forecasting (Bi et al., 2023). In recent years, some improved models based on Transformer have gradually been used for tourism demand prediction.

This study aims to adopt a novel Transformer model called the inverted transformer (iTransformer), which introduces a new inverted perspective proposed by Liu et al. (2023). Compared to other Transformer models, the iTransformer achieves comprehensive leadership in time series forecasting tasks without the need to modify any modules. This model addresses the performance degradation and computational explosion issues present in traditional Transformer models for time series prediction.

II. Data and methodology

This study validates the effectiveness of the iTransformer model in tourism demand forecasting using the daily volume of domestic tourist arrivals to Hangzhou as an example. The research process includes five stages: data collection, data pre-processing, data division, model training and forecasting, and performance measures.

Data

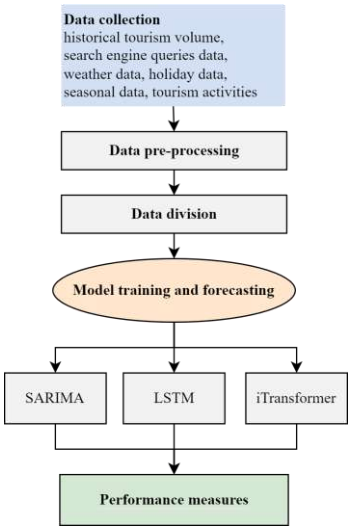
This study obtained the daily volume of domestic tourist arrivals to Hangzhou from January 1, 2023, to March 31, 2024, through the Chinese Tourism Management Department. Drawing on previous research, we included historical tourism volume, search engine queries, weather, holiday and seasonal data as feature variables for

forecasting daily tourism volume. Additionally, this study considered tourism activities data. Our search engine queries data are obtained from the Baidu website due to its largest market share in China, with keywords selected related to tourism. The weather data is collected from the China Weather website, including weather conditions and temperature. The date pattern is represented by four values, namely workdays, weekends, short holidays, and long holidays. Tourist travel shows clear seasonality, thus we categorize the seasonal variable into spring, summer, autumn, and winter. Tourism activities are reflected by whether the government organizes tourism consumption events.

Models

The iTransformer was proposed by Tsinghua University and the Ant Group team (Liu et al., 2023), that simply applies the attention and feed-forward network on the inverted dimensions. Specifically, the time points of individual series are embedded into variate tokens which are utilized by the attention mechanism to capture multivariate correlations; meanwhile, the feed-forward network is applied for each variate token to learn nonlinear representations. The article conducts experiments comparing the variants of the Transformer model before and after inversion. It demonstrates that inversion is a structural framework more consistent with the characteristics of time-series data and proves that iTransformer has achieved comprehensive leadership in complex time-series forecasting tasks.

For comparison, we use SARIMA and LSTM models as baseline models to evaluate the effectiveness of the iTransformer model in predicting daily urban tourism demand.



<Figure 1> The research framework

III. Research expectations

Firstly, to our knowledge, this study represents the first attempt to forecast tourism demand using daily urban domestic tourist arrivals. Compared to forecasts of inbound tourist volumes and tourist volumes at attractions, forecasting domestic tourism volume in a city provides decision support for destination government macro-management and the efficient operation of tourism enterprises.

Secondly, this study incorporates historical tourism volume, search engine queries, weather, holiday, seasonal and tourism activities data as feature variables into tourism demand forecasting. Multiple data sources can provide a stronger theoretical foundation and a more comprehensive overview for tourism volume prediction.

Thirdly, to our knowledge, this study is the first to apply the iTransformer model to the field of tourism demand forecasting. Through comparative analysis of predictive performance with benchmark models, the superiority of this model in the field of time series forecasting is confirmed. The interpretability analysis of tourism demand forecasting can also provide more persuasive insights for tourism-related decision-makers, assisting them make more accurate predictions and formulate more reliable plans.

IV. References

Bi J W, Liu Y, Li H. Daily tourism volume forecasting for tourist attractions[J]. Annals of Tourism Research, 2020, 83: 102923.

Bi J W, Han T Y, Yao Y, et al. Tourism demand forecasting under conceptual drift during COVID-19: an ensemble deep learning model[J]. Current Issues in Tourism, 2023: 1-20.

Havranek T, Zeynalov A. Forecasting tourist arrivals: Google Trends meets mixed-frequency data[J]. Tourism Economics, 2021, 27(1): 129-148.

Liu Y, Hu T, Zhang H, et al. itransformer: Inverted transformers are effective for time series forecasting[J]. arXiv preprint arXiv:2310.06625, 2023.

Palmer A, Montano J J, Sesé A. Designing an artificial neural network for forecasting tourism time series[J]. Tourism management, 2006, 27(5): 781-790.

Song H, Qiu R T R, Park J. A review of research on tourism demand forecasting: Launching the Annals of Tourism Research Curated Collection on tourism demand forecasting[J]. Annals of Tourism Research, 2019, 75: 338-362.

Wu X X, Shi J, Xiong H. Tourism forecasting research: a bibliometric visualization review (1999–2022)[J]. Tourism Review, 2024, 79(2): 465-486.

Xu S, Liu Y, Jin C. Forecasting daily tourism demand with multiple factors[J]. Annals of Tourism Research, 2023, 103: 103675.

Xue G, Liu S, Ren L, et al. Forecasting hourly attraction tourist volume with search engine and social media data for decision support[J]. Information Processing & Management, 2023, 60(4): 103399.

패션 수요예측을 위한 TimeGAN 기반 데이터 생성 방법론 연구

나민수
연세대학교 디지털애널리틱스학과
minsu97@yonsei.ac.kr

박상욱
연세대학교 산업공학과
lostmywatch@yonsei.ac.kr

이주필
연세대학교 디지털애널리틱스학과
wnvlf1883@yonsei.ac.kr

이서현
연세대학교 인공지능학과
lsh2684360@yonsei.ac.kr

백상덕
연세대학교 산업공학과
eagleeagle@yonsei.ac.kr

김우주
연세대학교 산업공학과
wkim@yonsei.ac.kr

Abstract *– 수요 예측은 거의 모든 산업에서 비즈니스 계획을 수립하기 위한 기본이 되어왔다. 최근 인공지능 및 머신러닝 기술을 활용해 수요 예측을 시도하는 연구도 늘고 있는 추세이다. 그러나 패션 산업 분야에서의 수요예측은 다른 산업에 비해 더욱 어려운 분야로 인식되어 왔다. 패션 산업의 특성상 고객의 선호도에 변화가 많고, 신제품이 나오는 주기가 짧기에 데이터의 양이 적고 신뢰성 있는 데이터를 대규모로 확보하기 힘들다.*

기존의 패션 수요 예측은 통계적 모델링 방식을 활용하거나, 선형 회귀, 의사결정나무 등의 머신러닝 모델을 활용하는 방식으로 이루어졌다. 그러나 실제 현업의 패션 데이터는 대부분의 경우에 길이가 짧은 데이터만이 활용 가능한 경우가 많기에 수요 예측에 대한 성능이 보장되지 않는다.

따라서 본 연구는 원본 판매량 데이터를 기반으로 시계열 데이터를 다량으로 생성해 예측에 활용하여, 미래에 대한 불확실성을 반영하며 예측의 성능을 높이는 방식을 제안하다. 본 연구는 (주)NFL 마레몬떼와 협업하여 13 가지의 패션 카테고리에 대한 판매량 수요 예측을 시도한다.

방법론을 요약하면, 원본 데이터의 길이보다 짧은 데이터를 복수로 생성한 후, 이를 적절히 Shifting 한 후에 Padding 을 통해 남은 부분을 원본 데이터로 대체하여 데이터셋을 증강하였다. 이후 이를 수요 예측 모델에 원본 데이터와 증강된 데이터를 함께 예측 모델에 인풋하여 수요 예측을 시도하였다. 그 결과 13 개의 패션 카테고리 중에서 대부분의 패션 카테고리에서 원본 데이터만을 활용한 경우에 비해 SMAPE 를 포함한 다양한 지표를 기준으로 10% 이상 성능이 향상된 것을 확인하였다.

Keywords – Time-series forecasting, Generative Adversarial Network, Demand Forecasting, Deep Learning, Data Augmentation

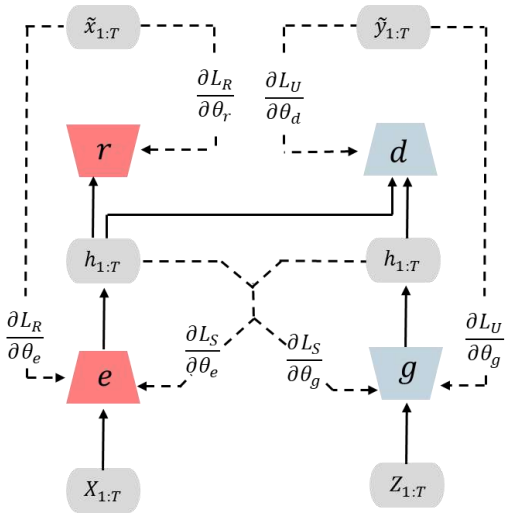
I. 서론

수요예측은 패션 산업에서 매우 중요한 과제로 간주되어왔다. 기업은 정확한 수요 예측을 통해 패션 제품에 대한 재고를 최소화해야 한다. 또한 패션 기업의 마케팅 및 프로모션과 관련된 의사결정에도 수요예측이 반영되어 비용 절감 및 이익 극대화에 기여할 수 있다. 그러나 패션 산업에서 수요예측은 다른 산업에 비해 어렵다고 인식되어 왔다. 패션 산업의 특성상 트렌드가 빠르게 변화하고 고객의 선호도에 영향을 미치는 요인이 다양하다. 또한 신제품이 출시되는 빈도가 잦고, 제품의 주기가 짧기에 예측에 활용할 수 있는 장기 시계열 데이터를 확보하기 매우 힘들다. 따라서 본 연구는 비교적으로 길이가 짧은 판매량 데이터를 데이터 생성을 통해 증강한 후, 예측에 활용하여 미래 판매량에 대한 예측의 성능을 높이는 것을 목적으로 한다.

II. 방법론 및 그림

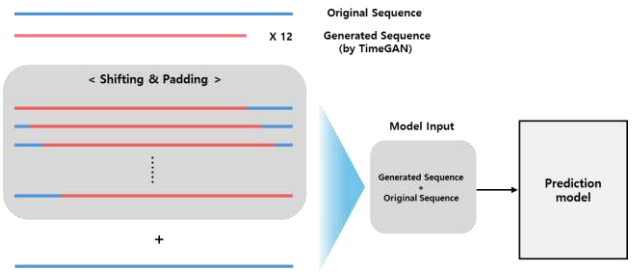
판매 수요 예측을 정확하게 하기 위해서는 과거 판매 수요를 바탕으로 미래의 수요를 예측하여야 한다. 이때 원본 판매 수량 데이터와 더불어 이를 증강한 데이터를 함께 예측에 활용하면 더 정확한 예측이 가능할 수 있다. 본 연구는 TimeGAN(Yoon, 2019) 모델을 통해 새로운 시계열을 12 개 생성하여 이를 Shifting & Padding 이라는 과정을 거친 후, 원본 데이터와 함께 활용해 예측을 수행한다.

<그림 1>은 연구에 활용한 TimeGAN 모델의 아키텍처이다.



<그림 1> TIMEGAN 아키텍처

<그림 2>는 Shifting & Padding 방법론을 포함한 판매량 예측 과정을 나타내는 그림이다.



<그림 2> SHIFTING & PADDING 방법론

III. 결론

본 연구의 실험 결과는 패션 카테고리에 대한 판매량 데이터를 생성 기법을 통해 증강하여 같이 예측에 활용할 시에 기존의 원본 판매량 데이터만을 활용하여 예측한 결과보다 더 성능이 향상되는지를 분석하고자 하였다. 이를 검증하기 위해 TimeGAN 모델과 Shifting & Padding 방법론을 활용하여 모델에 대한 인풋(input)을 구성하였다. 이를 Long Short-Term Memory(Hochreiter, 1997)과 같이 전통적인 시계열 예측 모델과 Adversarial Sparse Transformer(Wu, 2020), SCINet(Liu, 2022)과 같이 비교적으로 최근에 개발된 시계열 예측 모델에 적용하여 성능을 확인한 결과 기존 대비 SMAPE, MAPE, MAE 등의 지표에서 평균 10% 이상의 성능이 향상되었다. 또한 대부분의 패션 카테고리에서 기존 대비 성능이 향상되어 본 연구가 제시한 방법론이 수요예측의 성능 향상에 기여함을 증명하였다.

IV. 참고문헌

Yoon, J., Jarrett, D., & Van der Schaar, M. (2019). Time-series generative adversarial networks. *Advances in neural information processing systems*, *32*.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Wu, S., Xiao, X., Ding, Q., Zhao, P., Wei, Y., & Huang, J. (2020). Adversarial sparse transformer for time series forecasting. *Advances in neural information processing systems*, *33*, 17105-17115.

Liu, M., Zeng, A., Chen, M., Xu, Z., Lai, Q., Ma, L., & Xu, Q. (2022). Scinet: Time series modeling and forecasting with sample convolution and interaction. *Advances in Neural Information Processing Systems*, *35*, 5816-5828.

주파수 분해 및 멀티모달 기반의 의류 아이템 판매수량 예측에 관한 연구

백상덕
연세대학교 산업공학과
eagleeagle@yonsei.ac.kr

박상욱
연세대학교 산업공학과
lostmywatch@yonsei.ac.kr

김우영
연세대학교 산업공학과
timothy@yonsei.ac.kr

박병윤
연세대학교 산업공학과
bonbak@yonsei.ac.kr

나민수
연세대학교 산업공학과
minsu97@yonsei.ac.kr

이서현
연세대학교 산업공학과
lsh2684360@yonsei.ac.kr

이주필
연세대학교 산업공학과
wnvlf1883@yonsei.ac.kr

이원준
연세대학교 산업공학과
streo222@yonsei.ac.kr

김우주
연세대학교 산업공학과
wkim@yonsei.ac.kr

Abstract *– 시계열 예측은 시점에 따라 변화하는 데이터에 대해 과거 데이터의 분포를 조건으로 미래 데이터의 분포를 예측하는 분야이다. 또한, 특정 제품에 대한 고객의 수요 또는 판매 수량을 예측하는 수요 예측은 시계열 예측의 대표적인 과업(Task)이다.*

최근 딥러닝 기반의 시계열 예측 연구는 하나의 장기 시계열 데이터에 대해 이를 여러 구간으로 나누고 각 구간에 대하여 과거 시점의 데이터를 기반으로 미래 시점의 데이터를 예측한다. 그리고 이 과정에서 주파수 분해(frequency decomposition) 기법을 활용해 시간 영역(time domain)에서 포착할 수 없는 시계열의 계절성(seasonal component) 및 경향성(trend component)을 포착하여 미래 시점의 데이터 예측에 반영하는 기법으로 발전해왔다. 하지만 국내 패션 산업의 의류 판매수량 데이터는 일반적으로 길이가 매우 짧은 시계열에 해당하기 때문에 장기 시계열 예측을 위한 딥러닝 기반의 방법론을 적용하기에 제한적이다.

본 연구는 국내 패션 기업 M사의 제품 별 판매수량 데이터에 대해 주파수 분해 및 멀티 모달 기반의 방법을 활용하여 단일 딥러닝 모델을 기반으로 다수 제품의 미래 판매수량을 예측할 수 있는 방법론을 제안한다. 이 과정에서 출시일, 판매기간이 상이한 1,200 여개의 제품 별 판매수량 데이터, 각 제품에 대한 특성 데이터(제품 이미지, 제품 설명 텍스트)를 활용하며, 푸리에 변환(Fourier transform) 및 웨이블릿 변환(Wavelets transform)을 기반으로 시계열의 계절성과 추세성을 포착하는 FEDformer(Frequency Enhanced Decomposed Transformer) 모델에 대해 구조적 변경을 적용한다. 이러한 구조적 변경은 단일 모델이 다수의 시계열을 구분해 학습 및 예측하도록 한다. 또한 제품의 자체 특성을 미래 시계열 예측에 반영하기 위해 각 제품의 특성 데이터와 제품 판매수량 간의 연관성을 크로스 어텐션(Cross Attention) 기법을 이용하여 포착한다.

본 연구는 발생 시점 및 길이가 서로 상이한 다수의 단기 시계열 데이터를 활용하며, 시계열 외의

아이템 자체 특성을 활용한다는 점에서 기여점이 있을 것으로 판단된다.

Keywords – Time Series Forecasting, Demand Forecasting, Multimodal Deep Learning, FEDformer

"이 논문은 국토교통부의 스마트시티 혁신인재육성사업으로 지원되었습니다."

".III. 참고문헌

Zhou, T., et al. (2022). FEDformer: Frequency Enhanced Decomposed Transformer for Long-term Series Forecasting (arXiv:2201.12740v3; Version V3) [Preprint]. arXiv. https://doi.org/10.485

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: a survey. Philosophical Transactions of the Royal Society A, 379(2194), 20200209. https://doi.org/10.1098/rsta.2020.0209

소셜 미디어 기반 국내 소형 SUV 자동차 구매요인 분석

최혁준
아주대학교 경영대학
hyukjun3874@ajou.ac.kr

강주영
아주대학교 경영대학
jykang@ajou.ac.kr

Abstract –
본 연구는 소셜 미디어 분석을 통해 국내 자동차의 구매요인을 분석하고자 한다. 자동차 구매요인에는 많은 요인이 존재하는데, 그 중에서 어떤 요인이 구매요인에 중요한 부분을 차지하고 있는지 분석했다. 국내 자동차에 대한 소셜 미디어에서의 대중의 반응과 관심을 조사하고, 이를 통해 구매요인들을 분석했다. 특히, 네이버-오너평가 항목을 분석해 실제 자동차 오너들의 평가에 대해 분석하고, 유튜브의 자동차 리뷰 댓글을 통해 소비자들이 해당 자동차의 어떤 점에 집중하는지 분석했다. 차량의 범위를 소형 SUV로 한정시켜, 어떤 요인이 가장 큰 구매요인으로 다가오는지 분석했다.

소셜 미디어에서 자동차의 주요 특징과 기능에 대한 사용자들의 의견과 반응을 분석했다. 자동차 구매결정에 영향을 주는 요인에는 비용, 크기, 디자인, 안전성 등이 있다. 소비자들은 자동차 리뷰를 볼 때 디자인을 1순위로 본다는 결과가 나왔고, 공간성과 주행성능은 그 뒤를 이었다. 특히 2023년 판매량 1위인 기아의 셀토스는 가격 대비 공간성이 좋다는 평가와 디자인적인 요소로 인해 독점적인 판매량을 보였다.

일부 소형 SUV 차량들 중에는 준중형 SUV 가격과도 구간이 겹치는 차들이 존재했다. 이런 차들이 셀토스에게 밀린 이유는 같은 가격으로 차급을 높일 수 있다는 의견이 주요했고, 소비자들이 차 크기에 맞는 적정가격을 어떻게 생각하는지에 대해 분석해 자동차 크기에 따른 적정가격에 대해 분석했다. 또, 판매량이 가장 적었던 현대자동차의 베뉴의 경우에는 가격대비 낮은 공간성과 출력으로 인해 소비자들의 선택을 받지 못했다는 결과를 도출했다.

마지막으로, 소셜 미디어 분석 결과를 토대로 소비자들이 원하는 국내 소형 SUV에 대한 이상적인 모델을 제시했다. 소비자들이 원하는 공간성, 디자인, 가격에 대해 분석하여 향후 자동차 기업들이 새로운 자동차를 출시할 때 방향성을 제시해줄 것이다.

소셜 미디어 분석을 활용하여 국내 소형 SUV 자동차의 구매요인을 종합적으로 분석하고, 자동차 시장에 대한 방향성을 제공할 것으로 기대된다.

Keywords – 자동차, 구매요인 분석, 소셜미디어 분석, 소형 SUV

A Study on Predicting Employee Turnover Intention using Machine Learning: Using SMOTE technique and Stacking technique

강석훈
한성대학교 경영학과
sh941023@gmail.com

박영식
한성대학교 경영학과
youngsik.park@hansung.ac.kr

박정민
한국발명진흥원
paspark1@gmail.com

이형용
한성대학교 경영학부
leemit@hansung.ac.kr

Abstract – *Since employee turnover entails significant costs for companies, one recent and rapidly advancing approach to prevent turnover involves utilizing machine learning methods to proactively identify employees' intentions to leave, thereby preventing turnover. However, in the case of turnover data, an imbalance between instances of turnover and non-turnover can restrict the predictive performance of simple machine learning models. In this study, we first utilized various single machine learning models such as K-Nearest Neighbors (KNN), Random Forest, Decision Trees, Support Vector Machines, and Adaptive Boost (AdaBoost) to predict turnover intentions and calculate accuracy. In the second stage, we balanced the dataset using the Synthetic Minority Over-sampling Technique (SMOTE), an oversampling method, and then applied single machine learning models to predict turnover intentions and measure accuracy. Finally, we introduced the Stacking technique to predict turnover intentions using data before and after applying SMOTE. The predictive accuracy was higher when using SMOTE compared to not using it, and the predictive accuracy was even higher when applying the Stacking technique compared to using single algorithms. Overall, this study demonstrated that employing the SMOTE technique to handle imbalanced data and utilizing the Stacking technique can enhance the model's performance when predicting turnover intentions.*

Key Terms – *Turnover intention, SMOTE technique, Stacking technique*

Session E

E1. [학술세션]
데이터 주도 에너지 및 환경 정책 연구

좌 장: 방 영 석(연세대)

LDA 와 BERTopic 을 통한 전문가와 대중의 원자력 에너지 시선 비교 : 균형잡힌 정책 수립을 향한 고찰

홍채영
아주대학교
비즈니스애널리틱스학과
hongcy@ajou.ac.kr

이한솔
아주대학교
경영인텔리전스학과
lsh15 @ajou.ac.kr

강주영
아주대학교
경영인텔리전스학과
jykang @ajou.ac.kr

Abstract *– 최근 기후 변화와 에너지 안보 문제가 심각해지고 있어 에너지 정책이 중요한 역할을 한다. 원자력 에너지 정책은 전문가와 다양한 이해관계자들의 의견 조율을 통해 수립되지만, 전문가들의 의견이 더 큰 비중을 차지할 수 있다. 전문가 의견 비중이 높은 에너지 정책은 기술적 측면이 강조되어 안전, 환경 등의 사회적 측면이 부족할 수 있다. 또한 5 년 주기로 수립하고 시행되기 때문에 대중의 의견을 파악하고 반영하는 것이 중요하다. 따라서 본 연구는 에너지 관련 신문 데이터를 활용하여 전문가 의견을 수집하고, 원자력 관련 영상 댓글을 활용하여 대중의 의견을 수집한다. 전문가는 LDA, 대중은 BERTopic 으로 토픽 모델링하고 토픽별 감성분석을 통해 각 그룹의 인식을 파악한다. 전문가와 대중의 인식 비교 결과, 두 집단 모두 환경 문제와 에너지 정책에 관심이 있지만 전문가는 기술과 정책에 집중하고 대중은 안전과 정치적 측면에 초점을 맞추고 있다. 또한 두 집단 모두 에너지 발전 방식 선택 시 경제성을 중요하게 생각하지만 전문가는 탄소 중립과 신뢰성에 좀 더 주목한다. 사용후핵연료 폐기물 처리 및 관리의 중요성도 공통적으로 인식하지만 이 문제를 해결하는 방안에 대한 관점이 다르다. 결과적으로 전문가와 대중 간의 인식 차이를 파악하여 에너지 정책 수립에 반영할 수 있는 지침을 도출하고 안전과 환경, 경제성 등 다양한 측면을 고려한 포괄적인 에너지 정책의 방향성을 제시한다. 또한, 전문가와 대중 간의 의견 조율을 통해 에너지 정책을 수립하는 방향 설정에 있어 균형 잡힌 의사결정으로 지속 가능하고 효과적인 정책이 수립되기를 기대한다.*

Key Terms *– 원자력 에너지, 인식, 텍스트 마이닝, BERTopic, LDA*

Time-LLM 을 활용한 해양 기상 데이터 분석

윤지원
부산대학교
ziwonyun@gmail.com

김호원
부산대학교
howonkim@gmail.com

***Abstract** – 해양 기상 데이터의 정확한 예측은 해양 산업의 효율적인 운영과 안전한 항해에 중요한 역할을 한다. 본 연구에서는 **Large Language Model (LLM)**을 기반으로 한 시계열 데이터 예측 프레임워크인 **Time-LLM** 을 활용하여 해양 기상 데이터를 예측하는 실험을 진행한다. **Time-LLM** 은 시계열 데이터와 자연어 간의 연관성을 강화하고, **Patch Embedding** 과 **Prompt-as-Prefix** 를 활용하여 시계열 데이터의 패턴을 더욱 정확하게 이해한다. 기상청 기상자료개방 포털에서 수집한 해양 기상 데이터를 전처리한 후, 모델을 훈련시켜 해양 기상 조건을 예측한다. 결과적으로, **Time-LLM** 은 해양 기상 데이터를 효과적으로 예측하는 것을 확인했다. 이러한 연구 결과는 해양 산업 및 항만 운영에서의 해양 기상 데이터 예측에 적극적으로 활용될 수 있을 것으로 기대된다.*

***Keywords** – 다변량 분석, 시계열 분석, 항만 물류, LLM, Smart logistics*

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 융합보안핵심인재양성사업의 연구 결과로 수행되었음" (IITP-2023-2022-0-01201)

I. 서론

항만 물류 산업은 세계적으로 수출입 물품의 움직임과 교역에서 중추적인 역할을 수행한다. 특히 우리나라는 3 면이 바다로 둘러싸여 있어 수출입 물동량의 98%가 해상을 통해 이동하는 특성을 가지고 있어 이 산업의 중요성이 더욱 부각되고 있다. 해당 산업에서 기상 조건의 변화는 선박의 출입, 화물의 하역 및 적재 작업, 물류의 이동 및 보관 등에 큰 영향을 미치며, 때로는 물류의 지연, 손상, 비용 증가 등의 부정적인 결과를 초래할 수 있다. 그러므로 정확한 기상 예측과 분석은 항만 물류 산업의 효율성과 안정성을 유지하는 데 매우 중요하다.

인공지능 기술의 발전은 다양한 산업 분야에 혁신적인 변화를 가져오고 있다. 특히 대규모 언어 모델(Large Language Models, LLM)의 발전으로 인해 다양한 문제를 새로운 방식으로 접근할 수 있게 되었다. 이러한 LLM 의 발전으로 인해 시계열 데이터를 LLM 을 활용하여 분석하는 방법론들이 등장하고 있다.

본 논문에서는 해양 기상 데이터를 Time-LLM(Jin et al., 2023) 프레임워크를 활용하여 분석한다. 기상청 기상자료개방포털에서 수집한 데이터를 전처리하고 모델을 학습하여 예측 성능을 평가한다. 이를 통해 해당 프레임워크를 활용한 해양 기상 데이터 예측의 유용성과 적용 가능성을 평가한다.

II. 관련 연구

LLM 의 문맥 이해, 일반화 능력, 전이 학습 능력과 같은 많은 장점들이 강조되고 있다. 최근에는 이러한 장점들을 활용하여 시계열 분석 모델에 LLM 을 적용하는 방법에 대한 다음과 같은 연구들이 이루어지고 있다.

One Fits All(Zhou et al., 2023): 시계열 데이터를 LLM 으로 다루고자 한 초창기 연구로, LLM Transformer(Vaswani et al., 2021) 블록의 self-attention 과 feedforward layer 는 고정하고 embedding layer, normalization layer, output layer 만을 학습하는 방법을 사용하였다.

LLM4TS(Chang et al., 2024): supervised pre-training 을 통해 시계열 데이터와 자연어의 modalities 를 align 하고, 이후 task-specific fine-tuning 을 진행하는 two-stage fine-tuning process 방식으로 설계되었다. 이는 LLM 을 시계열 예측에 사용하기 위해 시계열 데이터와 자연어 데이터의 modality align 을 최초로 제안했다.

Time-LLM: 시계열의 Patch Embedding 을 자연어 표현 공간으로 reprogramming 하여 시계열과 자연어를 align 하고, 특정 template 기반의 prompt 를 prefix 로 사용하여 모델이 시계열 patch 를 더 잘 이해할 수 있도록 하며 downstream task 수행 능력을 향상시켰다.

LLMTime(Gruver et al., 2024): 시계열 데이터를 숫자로 구성된 문자열로 표현하고 텍스트의 다음

토큰을 예측하도록 하는 방식으로 zero-shot time series forecasting 을 수행하였다.

본 논문에서는 대부분의 실험에서 높은 성능을 보인 Time-LLM 을 활용하여 실험을 진행하고자 한다.

III. 본론

3.1 데이터 설명

본 논문이 제안하는 실험을 진행하기 위해 기상청 기상자료개방포털에서 제공하는 해양기상부이 데이터를 사용하였다.

<표 1> 데이터 정보

Feature	설명
일시	1 시간 간격(YYYY-MM-DD HH:MM)
풍속(m/s)	바람의 세기
풍향(deg)	바람의 방향
GUST 풍속(m/s)	순간적으로 강한 바람이 불었을 때의 풍속
현지기압(hPa)	해당 지역(거제도)의 대기의 압력
습도(%)	대기 중의 수증기의 양
기온(° C)	해당 지역(거제도)의 기온
수온(° C)	해당 지역(거제도)의 수온
최대파고(m)	특정 시간 동안 관측된 파도 중 가장 높은 높이
유의파고(m)	특정 시간 동안 관측된 파도 중 가장 높은 3 분의 1 에 해당하는 파고의 평균
평균파고(m)	특정 시간 동안 관측된 파도의 평균 높이
파주기(sec)	연속된 파도가 발생하는 주기
파향(deg)	파도가 발생하는 방향

2020 년부터 2022 년 사이 측정된 데이터를 훈련 데이터로, 2023 년에 측정된 데이터를 테스트 데이터로 실험을 진행하였다.

3.2 실험 설정

```
prompt_ = (
    f"<|start_prompt|>Dataset description: {self.description}"
    f"Task description: forecast the next {str(self.pred_len)} steps given the previous {str(self.seq_len)} steps information; "
    "Input statistics: "
    f"min value {min_values_str}, "
    f"max value {max_values_str}, "
    f"median value {median_values_str}, "
    f"the trend of input is {'upward' if trends[b] > 0 else 'downward'}, "
    f"top 5 lags are : {lags_values_str}<|<end_prompt>|>"
)
```

<그림 1> Time-LLM prompt template

<그림 1>과 같은 Time-LLM prompt 의 dataset description 부분에 <그림 2>와 같이 데이터셋 정보, feature 정보, 시간에 대한 정보를 입력하였다.

Measurements of marine weather conditions such as wind speed, wind direction, barometric pressure, humidity, air temperature, water temperature, wave height, wave period, and wave direction at one-hour intervals at sea level from 2020 to 2022.

<그림 2> 해양 기상 정보 dataset description

실험에서 백본 LLM 모델으로 Llama-7B(Touvron et al., 2023)를 사용하였다.

3.3 실험 결과

입력 길이 512, 예측 길이 92 로 테스트 데이터에 대한 성능 지표를 Transformer 기반 SOTA 모델인 PatchTST(Nie et al., 2022) 모델과 비교하였을 때, 표 2 와 같은 결과를 확인했다.

<표 2> 실험 결과

	Time-LLM	**PatchTST**
MSE	0.58	0.71
MAE	0.57	0.64

IV. 결론

본 논문에서는 Time-LLM 프레임워크를 활용하여 해양 기상 데이터를 예측하기 위한 실험을 수행하였다. 기상청 기상자료개방포털에서 데이터를 수집하고 전처리한 후 Time-LLM 을 사용하여 모델을 훈련시켰다. 또한, Transformer 기반 SOTA 모델인 PatchTST 와의 예측 성능 비교를 통해 해당 프레임워크를 사용하여 해양 기상 데이터를 예측하는 것이 효과적임을 확인하였다.

해당 프레임워크에서는 LLM 의 정교한 추론과 패턴 인식 능력을 활용하여 시계열의 패턴을 인식하고 향후 시점에 대해 효과적으로 예측하는 것을 확인하였다.

이러한 해양 기상 데이터의 효과적인 예측을 활용해 스마트 물류 및 항만 운영에서 해양 기상 조건에 따라 운송 노선을 최적화하고, 선박의 운항 일정을 효율적으로 관리할 수 있다. 또한, 예측된 기상 정보를 활용하여 로지스틱 네트워크를 최적화하고 운송 비용을 절감할 수 있다. 뿐만 아니라, 정확한 기상 예보는 안전한 항해를 보장하고 해양 사고의 위험을 감소시키는 데 도움이 될 수 있다.

V. 참고문헌

Jin, M., Wang, S., Ma, L., Chu, Z., Zhang, J. Y., Shi, X., ... & Wen, Q. (2023). Time-llm: Time series forecasting by reprogramming large language models. arXiv preprint arXiv:2310.01728.

Zhou, T., Niu, P., Sun, L., & Jin, R. (2024). One fits all: Power general time series analysis by pretrained lm. Advances in neural information processing systems, 36.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Chang, C., Wang, W. Y., Peng, W. C., & Chen, T. F. (2024). Llm4ts: Aligning pre-trained llms as data-efficient time-series forecasters.

Gruver, N., Finzi, M., Qiu, S., & Wilson, A. G. (2024). Large language models are zero-shot time series forecasters. Advances in Neural Information Processing Systems, 36.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.

Nie, Y., Nguyen, N. H., Sinthong, P., & Kalagnanam, J. (2022). A time series is worth 64 words: Long-term forecasting with transformers. arXiv preprint arXiv:2211.14730.

에너지·환경 분야의 국가 R&D 연구 동향 분석

김은경
한국과학기술정보연구원
eungyeong@kisti.re.kr

Abstract *– 에너지·환경 관련 산업은 신산업 창출, 일자리 증대, 국민 생활 향상 등을 목표로 산업의 육성과 활성화를 위해 부처 차원에서 진행하고 있다. 국내 산업은 글로벌 시장과 비교 시 기술력, 산업 현장 인력 부족, 기존 산업과의 접목 등 여러 요소가 필요하다고 평가하고 있다. 본 논문에서는 에너지·환경 관련 분야의 산업 활성화를 위해 최근 10 년 동안의 국가 R&D 과제와 참여인력 정보 등을 분석하였다. 분석 결과, 참여 인력은 년도별로 증가하지만, 특정 지역에 집중되어 있으며, 석·박사 수준의 고급 인력이 부족한 것으로 나타났다. 그리고 매년 정부의 연구비 투자가 증가하였고, 기업과 대학에서 공동 연구가 활발하게 이루어지고 있었다. 이러한 분석 정보는 정부가 추진하는 에너지·환경 산업 활성화를 위해 앞으로의 정부 재정적 지원, 산학연 협력관계 구축, 실무에 적합한 전문 인력 양성 등 정책결정에 활용될 수 있을 것이다.*

주거 공간 사용자 관리 서비스 활성화 가설 도출을 위해 귀납적 학습으로 분석한 탐색적 다중 사례 연구

김인수
경희대학교 경영학과
k2p2@khu.ac.kr

정백
경희대학교 산학협력단
ylbaek@khu.ac.kr

이선미
대구한의대학교 K-뷰티비즈니스학과
leesunmi@dhu.ac.kr

Abstract *– 효율적인 주거 공간 사용자의 관리를 위해서 다양한 플랫폼 서비스들이 사용되고 있는데, 이러한 서비스의 활성화 요인과 가설을 탐색하기 위해, 다중 사례 연구 방법을 사용하였다. 먼저, 사례 분석을 위해 2012 년부터 2024 년 4 월까지 존재하고 있는 주거 공간 사용자 관리 플랫폼 서비스 20 개를 조사 후 해당 서비스들의 주요 기능들을 나열하였다. 조사된 기능들은 '관리비 확인', '관리 납부 기능', '중고장터 기능', '시설물 현황 확인', '공지사항 및 소식지 기능', '이웃 소통 기능', '에너지 사용량', '멤버십 기능', '건물 시세' 등 총 11 가지로 확인되었으며, 중요 핵심 변수로 작용하지 않을 것으로 예상되는 3 가지 '에너지 사용량', '멤버십 기능', '건물 시세'를 제외하고 총 8 가지를 변수로 채택하였다.*

활성화의 여부는 다운로드 수 및 플레이스토어 별점을 기준으로 하였으며, 총 20 개의 사례 중 10 개 사례는 Success(S), 10 개의 사례는 Room to be Developed(D)로 분류하였다. 인공지능 분야의 대표적 기계 학습 방법론 중의 하나인 ID3 알고리듬에 기반한 귀납적 학습 접근 방식을 사용하여 분석하였다. 특히, 본 연구에서는 OpenAI GPT-4 의 Advanced Data Analytics 기능을 사용하여, 사례분석 엑셀파일을 그대로 업로드하고, GPT-4 와 대화형 인터페이스를 통하여 의사결정나무를 도출하였다. 분석 결과, '이웃 소통 기능'이 여덟 가지 변수 중 가장 영향력 있는 변수로, 이후 '소식지 기능', '시설물 현황 확인' 기능이 각각 중요 변수로 나타났다.본 연구에서는 이러한 결과를 토대로 다음과 같이 주거 공간 사용자 관리 서비스 활성화에 관한 4 가지 가설이 생성되었다.

1. 이웃소통 기능을 가진 주거 공간 사용자 관리 플랫폼은 서비스가 활성화된다.

2. 이웃소통 기능과 공지사항 및 소식지 기능 이 모두 없는 경우 활성화가 되지 않는다.

3. 이웃소통 기능이 부재하나, 공지사항 및 소식지 기능, 시설물 현황 확인 기능이 있는 주거 공간 사용자 관리 플랫폼은 서비스가 활성화된다.

4. 이웃소통 기능이 없는 경우, 공지사항 및 소식지 기능이 있더라도, 시설물 현황 확인 기능이 부재한 경우, 활성화가 되지 않는다.

도출된 가설은 주거 공간 사용자 관리 플랫폼과 같은 주거공간 사용자 관리 서비스의 상업적 성공을 위한 전략 수립에 통찰력을 제공할 수 있을 것으로 예상된다.

또한, 세개의 주요 변수 중 이웃소통 기능은 IT 서비스에서의 Interaction 서비스에 해당하고, 공지사항 및 소식지 기능은 Information 서비스에 해당하며, 시설물 현황 확인 기능은 Monitoring 서비스에 해당한다. 결국 사용자 관리 서비스에서 Interaction 서비스가 필수적이라는 이론적 함의를 도출할 수 있으며, 주거 공간 사용자 관리 플랫폼과 같은 주거공간 사용자 관리 서비스의 경우는 사용자간 Interaction 서비스가 없더라도, 주거공간이라는 특성을 대표하는 Monitoring 서비스와 Information 서비스가 모두 제공될 경우는 성공할 수 있다는 이론적 함의를 제공한다.

Key Terms *– 건물 관리 플랫폼, ID3 알고리즘, 귀납적 학습, 다중 사례 연구*

Session E

E2. [특별세션]
데이터로 보는 글로벌 R&D 인사이트

좌　장: 전 승 표(KISTI)

과학기술문헌(WOS) 데이터를 활용한 국내 기초자치단체간 과학협업 분석 : 인용영향력 증분(Incremental Citation Impact, ICI) 지표를 중심으로

권오진
한국과학기술정보연구원
글로벌 RnD 분석센터
dbajin@kisti.re.kr

이정우
한국과학기술정보연구원
글로벌 RnD 분석센터
jungwoolee@kisti.re.kr

안세정
한국과학기술정보연구원
글로벌 RnD 분석센터
sjahn@ kisti.re.kr

***Abstract** – 국제공동 연구의 중요성이 커짐에 따라 관련 연구들이 증가하고 있다. 그러나 과학계량학 기반의 지역간 공동연구는 다소 부족한 실정이다. 특히 논문(**Web of Science**)의 경우 분석의 기본 단위가 표준화되어 있지 않고 광역과 시군구가 혼재되어 데이터 획득에 어려움에 직면하고 있다. 본 연구는 **KISTI** 글로벌 **R&D** 분석센터에서 수행중인 연구자 이동성 연구에 대한 중간 결과물인 인공지능을 활용한 국내 지역명 식별체계 분류시스템을 활용하여 국내 시군구간 공동연구에 대해 **ICI** 지표를 이용하여 협력연구에 대한 득과 실을 살펴보고자 한다.*

***Keywords** – 인용 영향력증분 지표(**Incremental Citation Impact**), 기초자치단체간 공동 연구, **WOS Data**, 시군구 지역명 식별*

이 논문은 2024 년
한국과학기술정보연구원(KISTI)의 기본사업으로
수행된 연구입니다.(과제번호: K24L3M1C1)

과학기술 분석을 위한 KISTI 분석용 데이터베이스 소개

권태훈
한국과학기술정보연구원
kth78@kisti.re.kr

이 철
한국과학기술정보연구원
clee@kisti.re.kr

Abstract *– 최근 과학의 발전을 정량적으로 이해하려는 Science of Science 에 대한 관심이 증가하고 있다. 정략적인 이해를 위해서는 당연히 데이터가 중요하므로 분석을 위한 데이터베이스에 대한 중요성이 강조되고 있다. 하지만, 기존 데이터베이스는 문헌검색에 특화된 검색용 데이터 베이스로 통계분석 등의 분석에 최적화되어 있지 않다. 이에 KISTI 에서는 논문, 특허 등 다양한 문헌정보에 대한 분석을 지원하기 위해 분석에 특화된 데이터베이스를 구축하고 지속적으로 업데이트하고 있다. 효율적이고 지속적인 구축 및 업데이터 작업을 위해 분석 데이터 구축, 식별 데이터 구축, 속성 데이터 구축으로 크게 3 가지 구축 프로세스를 설계하였으며, 해당 프로세스간 연계를 통해 데이터의 정합성, 무결성 및 정규화를 추구하였다. 구축된 KISTI 분석용 데이터베이스는 논문분야 Web of Science, Scopus, PubMed, Crossref, KCI, OpenAlex, 특허분야 PATSTAT, 뉴스미디어분야 NEWSCATCHER, 기업정보분야 NICE, 연구자정보분야 ORCID 등의 다양한 데이터를 포함하며 데이터 연계성 및 확장성을 위해 KISTI 에서 정제 및 표준화한 연구기관, 지역 등의 분석용 속성 데이터도 포함하고 있다. 본 논문에서는 KISTI 분석용 데이터베이스의 활용도 제고를 목표로 해당 데이터베이스의 구축 프로세스를 포함하여 구축 현황에 대해 소개하고 활용 방안을 제시하고자 한다.*

Keywords *– 계량서지학, 과학기술지표 분석, 관계형 데이터베이스, 데이터베이스 시스템, 이종데이터*

이 논문은 2024 년
한국과학기술정보연구원(KISTI)의 기본사업으로
수행된 연구입니다.(과제번호: K24L3M1C1)

논문 데이터를 활용한 글로벌 국가연구기관의 과학기술 프로파일 분석

안세정
한국과학기술정보연구원
글로벌 R&D 분석센터
sjahn@kisti.re.kr

서주환
한국과학기술정보연구원
글로벌 R&D 분석센터
jungwoolee@kisti.re.kr

이준영
한국과학기술정보연구원
글로벌 R&D 분석센터
road2you@ kisti.re.kr

박진서
한국과학기술정보연구원
글로벌 R&D 분석센터
jayoujin@ kisti.re.kr

Abstract – 본 연구에서는 **Web of Science** 논문 데이터를 활용하여 전세계 주요 국가연구기관들의 연구성과를 다양한 관점에서 분석한 결과를 다룬다. **KISTI** 와 네덜란드 라이덴대학 **CWTS** 에서 산출한 기관별 라이덴랭킹 지표와 **KISTI** 에서 구축한 글로벌 **R&D** 인사이트 시스템을 활용하여 주요 국가연구기관들의 영향력, 협력, 오픈 액세스, 젠더에 따른 과학기술지표 데이터를 종합적으로 분석하였다. 특히, 주요 과학기술 5 대 분야의 과학기술지표를 비교하여 각 기관들의 특성을 파악하고, 해당 국가 내에서의 국가연구기관으로서의 역할과 위치를 살펴보고자 한다. 또한, 한국의 국가과학기술연구회 소속 기관들과 해외 벤치마킹 대상 기관들의 현황을 비교 분석하여, 국내외 연구기관 간의 차이점과 시사점을 고찰해본다.

Keywords – 논문데이터, 국가연구기관, 과학기술 프로파일, 과학기술지표

이 논문은 2024 년 한국과학기술정보연구원 (KISTI)의 기본사업으로 수행된 연구입니다. (과제번호: K24L3M1C1)

학술논문 데이터를 활용한 글로벌 ESG 연구 지형 분석 - BERTopic 과 ChatGPT4.0 을 결합하여

이정우
한국과학기술정보연구원
글로벌 R&D 분석센터
jungwoolee@kisti.re.kr

안세정
한국과학기술정보연구원
글로벌 R&D 분석센터
sjahn@kisti.re.kr

이인혜
NH 농협금융지주
NH 금융연구소
ihlee22@nonghyup.com

Abstract – 최근 세계적으로 지속가능한 경제를 만들어야 한다는 문제 의식이 국제사회에 대두되면서 ESG(Environmental, Social, Governance)가 주목받고 있다. 이 연구는 UN Global Compact 가 ESG 용어를 처음 언급한 2004 년부터 2023 년까지 20 년간의 글로벌 ESG 연구 지형을 분석한다. Clarivate 의 Web of Science 데이터베이스에서 수집된 3,464 건의 학술논문 데이터를 기반으로, BERTopic 과 ChatGPT 를 사용하여 연구 주제를 도출하고 주제명을 명명하였다. 특히 ChatGPT 는 최근 ESG 분야에서 부상한 신조어와 전문 용어를 효과적으로 식별하기 위해 임베딩과 주제 프로파일링 단계에 적용하였다. 또한, 기존 BERTopic 의 HDBSCAN 이 유발하는 이상치 발생에 따른 문제점을 제어하고자 Mini Batch K-Means 를 사용한 클러스터링으로 데이터의 복잡성을 관리하고 효율적으로 주제를 분류하였다. 본 연구 결과는 국가별 ESG 연구의 양적 및 질적 차이를 비교하고, 글로벌 협력 패턴을 분석하여 ESG 연구의 글로벌 지형을 가시화하는 데 기여를 한다.

Keywords – ESG, 학술논문 데이터, 과학계량 분석, BERTopic, ChatGPT

이 논문은 2024 년도
한국과학기술정보연구원(KISTI)의 기본사업으로
수행된 연구입니다.(과제번호: K24L3M1C1)

과학기술 문헌 데이터베이스와 Sentence Transformer 모델을 활용한 기술 문헌 집단 구성 프로세스 연구

이철
한국과학기술정보연구원
글로벌 RnD 분석센터
clee@kisti.re.kr

이정우
한국과학기술정보연구원
글로벌 RnD 분석센터
jungwoolee@kisti.re.kr

김영진
한국과학기술정보연구원
글로벌 RnD 분석센터
kimyoungjin06@kisti.re.kr

이준영
한국과학기술정보연구원
글로벌 RnD 분석센터
road2you@kisti.re.kr

안세정
한국과학기술정보연구원
글로벌 RnD 분석센터
sjahn@kisti.re.kr

서주환
한국과학기술정보연구원
글로벌 RnD 분석센터
nano@kisti.re.kr

권태훈
한국과학기술정보연구원
글로벌 RnD 분석센터
kth78@kisti.re.kr

박진서
한국과학기술정보연구원
글로벌 RnD 분석센터
jayoujin@kisti.re.kr

Abstract *– 최근 국가전략기술 등 첨단기술을 중심으로 기술개발 주기가 짧아지면서 R&D 전략을 수립하는 과정에서 국내 및 국외 기술 현황 분석의 중요성이 증가하고 있다. 기존 기술 분석은 주로 전문가가 제시하는 키워드로 검색식을 작성하여 기술 문헌을 검색하는 방식에 기반하지만 이는 정성적 판단에 따른 한계가 존재한다. 이에 본 연구에서는 과학기술 문헌 데이터베이스와 딥러닝 기반 언어모델을 활용하여 기술과 관련된 문헌 집단을 계량적인 기법을 통해 구성하는 프로세스를 제안한다. 이를 위해 전세계 논문의 서지정보를 제공하는 Web of Science 데이터베이스 내 약 9,000 만건 논문 정보를 활용하여 핵심 문헌 집단을 파악하였다. 또한 딥러닝 기반 Sentence Transformer 모델을 통해 핵심 문헌과 유사도가 높은 기술 문헌을 도출하여 최종적으로 기술 분석을 위한 문헌 집단을 구성하였다. 본 연구를 통해 기술 현황 분석에 필요한 문헌 구성에 있어 인지적 한계를 극복하고, 보다 정량적인 기준을 통해 기술 분석이 가능해질 것으로 기대한다. 특히 글로벌 기술경쟁 심화에 따라 첨단기술에 대한 환경 분석에 활용함을 통해 데이터 기반 과학기술 정책 수립을 위한 근거로 활용될 수 있는 의의를 지닌다.*

Keywords *– 기술분석, 문헌검색, 계량분석, 언어모델, Sentence Transformer*

이 논문은 2024 년
한국과학기술정보연구원(KISTI)의 기본사업으로
수행된 연구입니다. (과제번호: K24L3M1C1)

데이터로 바라본 일본 수출 규제의 영향: 사회적 변화와 불매운동을 중심으로

전승표
한국과학기술정보연구원
글로벌 R&D 분석센터
spjun@kisti.re.kr

서주환
한국과학기술정보연구원
글로벌 R&D 분석센터
nano@kisti.re.kr

이철
한국과학기술정보연구원
글로벌 R&D 분석센터
clee@kisti.re.kr

Abstract – 지정학적 긴장과 더불어 새롭게 대두된 무역규제가 기술혁신은 물론 사회적으로 어떤 영향을 주는지 시사점을 찾기 위해서 본 연구는 우리나라와 일본 사이에서 4 년간 진행된 혁신 제품의 국가간 무역 분쟁 사례를 다양한 데이터를 기반으로 분석했다. 구체적으로 일본의 우리나라에 대한 3 종 품목의 새로운 수출 규제에서 보인 양국가의 대응과 규제대상 3 종 제품의 교역량 변화를 살펴보았고, 이 과정에서 나타난 사회적 영향을 살펴보았다.

분석결과에 따르면 일본의 수출 규제는 우리나라 기술혁신을 촉진한 것으로 나타나서 관련 특허출원이 증가했다. 사회적 변화를 살펴보기 위해서 뉴스와 검색 데이터를 분석했는데, 불매운동이 집중된 의류, 자동차, 주류의 특정 브랜드에서 뉴스가 증가했고, 연령별 검색 트랜드가 크게 변화한 것으로 나타났다. 검색에 나타난 관심의 변화는 관련 브랜드의 매출에 영향을 준 것으로 나타났으며, 불매운동의 영향은 2 년이상 동안 지속되었는데, 연령별로 지속 기간에 차이가 나타났다.

본 연구는 기술혁신과 무역을 중심으로 분석되었던 우리나라와 일본의 무역분쟁에 대한 기존 연구를 사회적 변화와 불매운동이라는 관점으로 확대했고, 즉시성이 높은 다양한 이종 데이터를 중심으로 분석했다는 측면에서 사회적 영향 모니터링 시스템 개발과 같은 다양한 방면에서 기여가 기대된다.

Keywords *– 검색데이터, 무역규제, 불매운동, 사회적 변화*

이 논문은 2024 년
한국과학기술정보연구원(KISTI)의 기본사업으로
수행된 연구입니다.(과제번호: K24L3M1C1)

Session E

E3. [특별세션]
과학기술사업화 AI플랫폼 개발과 활용전략

좌 장: 최 윤 정(KISTI)

인공지능 기반 공공기술사업화 플랫폼 구축에 관한 연구

최윤정 *한국과학기술정보연구원* yjchoi@kisti.re.kr

배국진 *한국과학기술정보연구원* baekj@kisti.re.kr

변정은 *한국과학기술정보연구원* jebyun@kisti.re.kr

김은선 *한국과학기술정보연구원* yjchoi@kisti.re.kr

***Abstract** – 우리나라 정부의 R&D 예산 투자규모는 확대되고 있으나, 공공기술 활용(거래) 실적은 예산증가율에 미치지 못하고 있어 연구개발 성과를 활용한 경제적 가치창출에 한계를 보이고 있다. 이는 전략적인 기술사업화 방안 미흡과 기술사업화 혁신주체간 연계 부족으로 기술 수요자와 공급자간 다양한 갭(gap)이 여전히 존재하기 때문이다. 따라서 본 연구에서는 데이터와 인공지능 기술 적용을 통해 사업화 유망성이 높은 공공 R&D 를 조기에 탐색·선별하고 기술공급-기술수요 정보의 정확성과 연계 효율성을 높이며 시장수요에 기반한 R&D 기획이 가능하도록 수요자별 맞춤형 지원을 위한 공공 R&D 유망 사업화 AI 플랫폼 구축 방안을 제시한다.*

***Keywords** – 공공기술, 기술사업화, 인공지능, 유망 사업화 플랫폼*

이 논문은 2024 년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다. (과제번호: (KISTI) K24L3M2C3)

I. 서론

우리나라 정부의 공공 R&D 예산은 2019 년 처음으로 20 조원을 넘어 2023 년 약 31 조원으로 지속적인 증가 추세를 보이고 있다(KISTEP, 2022). 또한 최근에는 12 대 국가전략기술 육성을 위한 R&D 투자 역시 증가하고 있다. 그러나 국가 연구개발비 대비 공공기술의 기술이전 및 사업화 성과창출은 크게 증가되지 않고 있다.

공공연구기관(대학, 연구소) 기술이전 사업화 실태조사 보고서(2023 년)에 따르면 2022 년 기준 공공기술 이전율 (조사대상 년도 기술이전건수/조사대상 년도 신규개발(확보) 기술건수) 은 10 년전(2014 년)과 같은 수치(31.7%)를 보이면서 국내 공공기술사업화의 성과는 정체되어 있는 실정이다. 공공연구소의 기술이전 건수는 전년도 대비 24.9% 감소하였고, 대학은 17.8% 기술이전율의 하락을 보이고 있다.

<그림 1> 연도별 공공기술이전율 (산업통상자원부, 2023)

또한 최근 10 년간 공공기술의 연구생산성(기술이전 수입/연구개발비)를 보면 연평균 약 1.5%로 해외 선진국(예. 미국 공공연구소, 10.8%)의 절반에도 못 미치는 매우 저조한 상황이다(산업통상자원부 2018)

이중 기술도입자 유형별 기술이전 수입 비중을 살펴보면 중소기업으로부터 수취가 전체 수입의 65.6%로 가장 많고, 다음으로 대기업(19.2%), 해외기관(9.9%), 중견기업(3.9%) 순이다. 중소기업 중 일반 중소기업은 전체의 57.9%이며, 공공연구기관 자체기술 기반 창업 기업은 전체 수입의 7.7%이다.

<그림 2> 기술도입자 유형별 기술이전 수입 비중 (산업통상자원부, 2023)

공공기술의 사업화 과정에 장애가 되는 주요 원인은 전문인력, 예산 부족과 같은 하드웨어 측면을 제외하면 기술도입 및 사업화를 진행할 수요, 창업기업의 발굴과 기술사업화의 대상이 될 수 있는 사업적 가치가 높은 공공기술의 부족으로 나타났다.(손수정, 2015; 산업통상자원부, 2023). 이러한 문제를 해결하기 위해 본 연구는 인공지능 기술을 활용하여 공공영역에서

플랫폼 구축을 통해 사업화 가능성이 높은 유망 공공기술을 적시 선별하며, 기술공급-기술수요 정보의 정확성과 연계 효율성을 높이고, 수요자별 맞춤형 지원을 위한 공공기술사업화 플랫폼 구축 방안을 제시한다.

II. 인공지능 기반 공공기술사업화 플랫폼 구축

인공지능 기술을 접목하여 활용성이 높은 공공기술 사업화 플랫폼을 구축하기 위해서는 다음과 같은 이슈를 고려할 필요가 있다. ① 실효성 있는 지원 서비스 발굴, ② 기술사업화 데이터 통합 관리체계 수립, ③ 안정적인 인프라 환경 구축, ④ 체계적 운영 관리방안 수립 및 파트너쉽 구축으로 구분된다.

먼저, 실효성 있는 지원 서비스를 발굴하기 위해서는 기존 시스템 현황을 비교분석하여 공백영역을 도출하고, 플랫폼 이용자를 포함한 이해관계자들의 의견을 수집하여 추진 방향성을 수립한다. 본 플랫폼의 경우, 기술사업화 초기 단계(R&D 기획, 아이디어 개발 및 기술시장 측면의 구체화)에서 사업 유망성이 높은 공공 R&D 를 조기에 선별하고 기업 유망성 및 기술제품의 시장현황을 분석하며, 글로벌 유망아이템을 탐색이 주요 기능으로 도출되었다.

공공기술사업화 데이터 구축 이슈로는 첫째, 기술사업화에 필요한 기관별 분산 데이터를 연계 통합하고 경제적 비용을 고려하여 부족한 데이터를 추가 확보하는 것이다. 둘째, 형태가 각기 다른 기술사업화 데이터에 대한 정제 가공을 통해 정합성을 확보하고 데이터 최신성 문제를 해결하는 것이다.

안정적인 서비스를 위한 인프라(HW/SW) 환경을 구축하기 위해 플랫폼 운영을 위한 조직 내부 유관 시스템 자원 연계 활용성을 고려하고, 신규 도입 인프라를 설계하며, 대용량의 기술사업화 데이터를 신속히 분석하고 가시화 할 수 있는 인프라 환경 구축이 요구된다. 마지막으로 체계적인 운영과 지속가능한 활용 방안을 모색하기 위해 원활한 플랫폼 구축·운영을 위한 운영관리 방안(인력.예산 등)을 수립하고 실 수요 기반의 플랫폼 활용과 성과 확대를 위해 이해관계자 협의체 구성이 필요하다.

하기 <그림 3>은 한국과학기술정보연구원에서 구축한 공공 R&D 사업화 유망성 탐색 AI 플랫폼이다. 본 플랫폼은 기술사업화의 초기 단계에서 기술적 성과를 시장기회와 접목시킬 아이디어를 구체화 하거나, 사업성이 높은 기술을 조기에 발굴하고자 하는 공공연구기관, 기업, 기술중개기관, 예비창업자에게 공공 R&D 사업화 분석 정보를 제공한다.

<그림 3> KISTI 공공 R&D 사업화 유망성 탐색 AI 플랫폼(Apollo)

주요 기능으로는 유망 사업화 공공 R&D 예측 서비스, 유망 사업화 수요기업 예측 서비스, 기업 유망성 분석 서비스, 글로벌 유망 아이템 탐색 서비스, 제품 시장 분석 서비스로 구분된다. 플랫폼 이용자에 따라 기술공급자(공공연구기관 개발자)는 공공 R&D 사전기획, 공공 R&D 개발 후, 기개발된 공공 R&D 사업화 구상 시 본 플랫폼을 활용할 수 있고, 기술수요자(기업)은 유망기술 발굴, 공공기술 탐색, 보유 기술 스케일업에 적용 가능하다. 또한 기술중개기관(공공 및 민간 전문기관)은 기술공급자-기술수요자 중개지원, 중소기업 기술사업화 지원에 적용 가능하며, 국민/개인(예비창업자)는 사업 아이템 구상 시 활용할 수 있을 것이다.

III. 결론

본 연구는 인공지능 기반의 공공기술사업화 지원 플랫폼 구축을 위한 기반 연구로, 사업화 유망성이 높은 공공 R&D 탐색을 위한 플랫폼 개발의 프레임과 주요 서비스 기능을 제시하였다. 이러한 인공지능 플랫폼 구축은 공공기술 거래와 사업화 제고를 위한 전략적 정책적 도구로 활용가능 할 것이다. 그러나 본 연구의 한계점으로서 연구결과의 신뢰성과 활용성을 높이기 위해서는 사용자별 플랫폼에 대한 결과 실증과 이를 위한 전문 커뮤니티 및 생태계 구축이 수반되어야 할 것이다.

IV. 참고문헌

한국과학기술기획평가원. (2022). 2022 년도 글로벌 R&D 투자동향 분석.

산업통상자원부. (2023). 2023 년도 공공연구기관 기술이전 · 사업화 실태조사보고서

손수정 (2015). 기술사업화: 갭(Gap)의 인식과 브릿지(Bridge)의 설계, 178. STEPI Insight. pp.1-28.

공공기술 매칭 플랫폼의 진단 및 사용자 경험 개선 전략

김민주 변정은 김은선 최윤정

한국과학기술정보연구원 *한국과학기술정보연구원* *한국과학기술정보연구원* *한국과학기술정보연구원*

mjkim@kisti.re.kr jebyun@kisti.re.kr yjchoi@kisti.re.kr yjchoi@kisti.re.kr

***Abstract** – 본 연구는 국내 공공기술이 기술이전 및 사업화 단계에서 겪고 있는 저조한 성과의 문제점을 해결하기위해 국내 공공기술 매칭 플랫폼의 현황과 개선 방안을 탐구하였다. 공공기술 매칭을 지원하는 세가지 플랫폼을 분석하고, 해당 플랫폼의 주요 사용자인 기술수요자(기업)와 기술중개 전문가를 대상으로 표적 집단 인터뷰(FGI)를 실시하였다. 분석 결과, 주요 사용자들은 기술사업화 과정에서 정보 접근성 및 매칭의 어려움을 겪고 있음을 확인하였다. 이를 바탕으로, 본 연구는 기술수요자 중심의 검색 엔진 개발, 공공기술 및 관련 시장 정보의 연계 제공 등을 포함한 공공기술 매칭 플랫폼 개선안을 제시한다.*

***Keywords** – 공공기술, 공공기술 매칭 플랫폼*

이 논문은 2024 년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다. (과제번호: (KISTI) K24L3M2C3)

I. 서론

최근 수년간 한국 정부의 공공 R&D 예산은 지속적인 증가 추세를 보이며, 2021 년 약 27 조원에서 2023 년 약 31 조원으로 성장하였다. 이러한 재정적 투자에도 불구하고, 공공기술의 기술이전 및 사업화 실적은 주요 선진국들과 비교했을 때 낮은 수준에 머물러 있는 실정이다. 특히, 2017 년 기준 한국의 기술이전 효율성은 1.55%로 미국의 4.66%에 크게 미치지 못하며, 2020 년 기준 유효 기술이전 비율은 24.5%에 불과하다.

선행연구에 의하면, 공공기술의 사업화 과정에 장애가 되는 주요 원인은 적합한 공공기술의 발굴 어려움과 사업적 가치가 높은 공공기술 발굴의 어려움으로 밝혀졌다 (손수정, 2015; Vohora et al., 2004; Yun et al., 2019). 이러한 문제에 대응하기 위해 여러 공공기술 매칭 플랫폼이 개발되었고, 최근에는 정부가 사업화 가능성이 높은 공공기술을 조기에 선별하고 이를 사업화 역량이 있는 기업과 신속히 공유하는 전략을 수립함에 따라 이러한 플랫폼들의 중요성이 더욱 커지고 있다. 하지만, 이 플랫폼들이 공공기술 매칭과 기술이전, 사업화 촉진에 있어서 실제로 활용되는 데는 여전히 한계가 있는 것으로 나타났다.

본 연구는 공공기술 매칭을 지원하는 국내 주요 플랫폼 현황을 분석하고, 이를 활용하는 사용자들이 직면한 문제점을 식별하여 플랫폼 개선 방안을 모색하는 것을 목표로 한다. 이를 위해 대량의 국가 공공기술 정보를 제공하며 인공지능 알고리즘을 통한 맞춤형 공공기술 정보를 제공하는 NTB 기술은행, Smart Tech-Bridge, 국가과학기술지식정보서비스(NTIS) 플랫폼을 선정하여 분석하였다. 또한, 실제 플랫폼 사용자들과의 표적 집단 인터뷰(FG)를 통해 얻은 인사이트를 바탕으로 플랫폼 개선 방안을 제시한다.

II. 연구방법

본 연구는 인공지능 알고리즘을 적용하여 국내 공공기술 데이터를 심층적으로 분석하고, 이를 통해 맞춤형 공공기술 발굴을 지원하는 플랫폼을 분석 대상으로 하였다. 또한, 이들 플랫폼이 실제 기술사업화의 과정에서 어떻게 활용되고 있는지에 대한 실증적 분석을 수행하였다.

1) AI 기반 공공기술 매칭 플랫폼 분석

- 기술은행 (National Tech Bank): 국가의 기술 자산을 효과적으로 관리하며 약 21 만건 이상의 공공기술 정보를 산업계에 제공하는 플랫폼이다. 분석 알고리즘을 통해 기술명과 관련 키워드, 그리고 키워드 간 연관성을 파악하여, 사용자가 특정 기술명을 명확히 알지 못할 때에도 관련 기술을 신속하게 검색할 수 있는 지능형 검색 기능을 제공한다. 또한, 이 플랫폼은 기술개발 상태와 산업별 개발 현황을 시각화하여 제공하며, 타임라인 기반의 기술 변화 시뮬레이션 기능을 제공하는 특징이 있다.
- 스마트 테크브릿지(Smart Tech-Bridge): 공공기술의 이전 및 기술금융을 지원하는 플랫폼으로, 약 17 만건의 공공기술, 20 만 건의 특허, 58 만 개의 논문 정보를 바탕으로 기술 추천 기능을 제공한다. 사용자 활동 로그와 사용자가 입력한 키워드를 기반으로 키워드 패턴 매칭 및 심층신경망을 활용한 공공기술 추천

알고리즘을 통해, 사용자가 자신의 관심사에 적합한 공공기술을 탐색하도록 지원한다. 또한, 성장단계별 공공기술 추천 기능을 제공하고 탐색된 공공기술과 관련된 시장 동향 리포트를 제공한다.

- 국가과학기술지식정보서비스(National Science & Technology Information Service): 국가 R&D 관련 사업, 과제, 연구자, 성과 등에 대한 지식정보 전수를 종합하여 제공하는 플랫폼이다. 이 플랫폼은 사용자가 공공기술을 효율적으로 탐색할 수 있도록 단계별, 디텍토리, 연관 키워드 추천 기반 검색 등 기능을 제공한다. 또한, 검색된 공공기술에 대한 자동화된 요약정보 및 해당 기술과 연계된 특허, 논문 등의 성과 정보를 제공하는 것이 특징이다.

2) 표적 집단 인터뷰(FGI)

- 표적집단 및 조사방법: 기술수요자 4 명(C1-C4) 및 기술중개 전문가 4 명(T1-T4)으로 구성된 총 2 개의 그룹을 대상으로 공공기술 탐색 시 애로사항 및 요구사항을 수렴하기 위해 심층 인터뷰를 총 2 회 진행하였다. 모든 대화는 전사하였고 인터뷰 과정에서 파악된 공통되는 주제는 어피니티 다이어그램을 통해 분류하였다.

- 분석결과: 첫째, 참가자들은 일관되게 기술수요자의 사업 컨셉과 일치하는 공공기술을 찾는 과정의 어려움을 지적했다. 현재 사업 영역과 매칭되는 공공기술을 단순한 키워드 매칭이 아닌, 기업의 구체적인 상황과 요구 사항에 부합하는 맞춤형 기술 매칭이 요구된다고 강조했다. 특히, 참가자 C1, C3, T2 는 현 플랫폼에서 제공하는 방대한 공공기술 데이터베이스에서 키워드를 통해 검색한 결과를 산업 분야, 기술 개발 연도 등 추가적인 정보로 필터링하는 수동적 접근 방식에서 벗어나, 플랫폼이 사용자의 상황과 검색 의도를 정확히 이해하고, 그에 따라 최적의 공공기술을 매칭해주는 시스템이 필수적이라고 제안했다.

또한, 다수의 참가자들은 적합한 공공기술을 발견했을 때 현재와 같이 단편적인 정보 제공을 넘어, 해당 기술의 발전 트렌드, 응용 가능성 및 관련 산업과 시장의 종합적인 정보를 연동해 제공하는 서비스의 필요성을 강조했다. 특히, C4, T1, T4 는 현재 상황에서 참가자들은 추가 정보를 위해 시장 보고서 등을 개별적으로 탐색하며 정보를 수집하고 있음을 지적했다.

그 외에도 참가자들은 개발 중인 공공기술에 대한 연구개발 진척도 정보를 제공받는 것의 중요성을 강조했다. 이를 통해 기술 수요자가 초기 개발 단계에서부터 협력을 시작할 수 있는 기회를 얻을 수 있을 것이라고 언급했다. 또한, 공공기술 담당자나 소유자와의 직접적인 연결 채널 및 정보의 정기적인 업데이트가 필수적이라고 언급했으며, 기술 용어에 대한 표준화의 필요성도 지적했다.

III. AI 기반 공공기술 매칭 플랫폼 개선 제안

본 연구에서는 현재 공공기술 매칭 플랫폼들의 정보 제공 수준과 사용자들의 서비스 기대치 간의 간극을 파악하여, 보다 효과적인 공공기술 매칭을 위한 개선 방안을 제안하고자 한다.

먼저, 공공기술 검색 엔진은 단순 키워드 검색을 넘어 사용자의 상황을 고려한 맞춤형 추천 시스템을 포함해야 한다. 예를 들어, 기술 수요자의 사업 연차, 기술 분야에 대한 관심, 연구개발 수준, 사업장 위치, 주요 사업 목적과 생산 제품 및 서비스 정보를 통합해 특화된 기술 매칭이 가능해야 한다. 이를 실현하기 위해서는, 공공기술 매칭에 필요한 주요 속성을 식별하고 이를 바탕으로 실제 기술 이전의 효과를 예측할 수 있는 알고리즘 개발이 필수적이다. 더 나아가, 거대언어모델(LLM)과 같은 최신 기술을 활용하여, 기업이 제출한 사업 계획서나 IR 자료 분석을 통해 자동으로 중요 특징값을 추출하고 이에 맞는 공공기술을 매칭하는 기능의 개발도 고려해볼 수 있다.

둘째, 기술사업화 과정에 필요한 공공기술 정보, 관련 기술 문서, 시장 정보 등을 연계하는 기능이 필요하다. 현재는 사용자가 각 정보를 별도로 검색하고 조사해야 하는 번거로움이 있는데 이를 해소하기 위해 각 제공 정보들 간의 연계 설계가 필요하다. 이를 위해서는 기술과 시장 간의 키워드 네트워크 구축, 기술-제품 간의 사전 구축, 국문과 영문 정보의 자동 연계 등 데이터 통합 기능을 추가할 필요가 있다.

IV. 결론

본 연구는 공공기술 매칭 플랫폼의 개선을 위한 기반 연구로, 국내 공공기술 매칭 플랫폼의 현 상황을 분석하고 전문가 그룹에 대한 표적 집단 인터뷰를 실시하였다. 연구 결과로 도출된 개선 방안이 공공기술의 효과적인 이전 및 사업화 과정에 기여할 것으로 기대된다.

V. 참고문헌

손수정 (2015). 기술사업화: 갭(Gap)의 인식과 브릿지(Bridge)의 설계, 178. STEPI Insight. pp.1-28.

Vohora, A., Wright, M., & Lockett, A. (2004). Critical junctures in the development of university high-tech spinout companies. Research policy, 33(1), 147-175.

Yun, J. K., Kwon, J. C., & Choi, S. H. (2019). Government R&D Technology Commercialization Policy Case Study. Journal of Distribution Science, 17(2), 53-69.

공공 R&D-수요기업 추천모델의 벡터 유사도 기반 경량화 연구

김지민
한국과학기술정보연구원(KISTI)
기술사업화연구센터 RnBD 분석연구팀
kjm@kisti.re.kr

변정은
한국과학기술정보연구원(KISTI)
기술사업화연구센터 RnBD 분석연구팀
jebyun@kisti.re.kr

최윤정
한국과학기술정보연구원(KISTI)
기술사업화연구센터 RnBD 분석연구팀
yjchoi@kisti.re.kr

***Abstract** – 공공 R&D 기술의 활용 및 사업화를 활성화하기 위해서는 R&D 기술 개발자 및 수요기업에 대한 체계적인 정보제공이 매우 중요하다. 또한 공공 R&D 와 수요기업에 대한 데이터 기반의 과학적인 매칭 및 추천을 통해, 보다 신속하고 정확한 기술사업화 의사결정을 지원할 수 있다. 따라서 본 연구는 인공지능(AI) 기반 공공 R&D-수요기업 추천모델을 개발함으로써, 기술 개발자 및 기술 수요기업에 대해 최적의 사업화 정보를 제공하고자 한다. 또한 추천 모델을 서비스하기 위한 벡터 유사도 기반 알고리즘 및 벡터 데이터베이스(DB)를 설계하여, 모델의 배포 및 추론 속도를 개선하는 경량화 방안은 연구하였다. 향후 본 연구를 확장하여 보다 자원 효율적이고 사용자 친화적인 AI 기반 공공 R&D-수요기업 추천 모델의 개발 및 서비스가 가능할 것으로 보인다.*

***Key Terms** – 공공 R&D, 기술사업화, 경량화, 벡터화, 유사도*

이 논문은 2024 년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다.(과제번호: K-24-L03-C03)

I. 서론

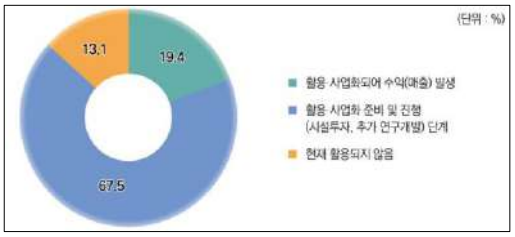
대한민국 정부 공공 R&D 예산은 2022 년 약 29 조원, 2023 년 약 31 조원을 넘어섰고, 국가 총 연구개발비(GRED)를 보면 약 8.7%로 높은 연평균 성장률을 보이고 있다 (KISTEP, 2022). 또한 대한민국 GRED 는 주요국 중 중국 다음으로 높은 성장세를 보이고 있으며, GDP 대비 총 연구개발비 비중은 세계 최고 수준을 기록하고 있다 (KISTEP, 2022).

반면 정부의 공공 R&D 예산으로 이루어진 공공 R&D 기술의 활용(기술거래) 실적은 예산 증가율에 미치지 못하고 있어, 공공 R&D 성과를 활용한 경제적 가치창출에는 한계를 보이고 있다.

국내 공공 R&D 기술 사업화 현황 중 가중 중요한 수치인 기술이전 현황을 살펴보았을 때, 2022 년도에 새롭게 체결된 기술이전 계약건수는 총 9,387 건으로 2021 년 10,263 건 대비 감소한 것으로 나타났다 (산업통상자원부, 2023). 또한 공공연구소의 기술이전 계약은 2022 년 4,572 건으로 2021 년 4,653 건에 비해 다소 감소하였고, 대학의 경우 2022 년 4,815 건으로 전년도 대비 약 14% 감소하여 저조한 사업화 성과를 보이고 있다 (산업통상자원부, 2023).

<그림 1> R&D 투자 총예산 추이(산업통상자원부, 2023)

또한 유효한 기술이전 계약 중, 이전 기술이 성공적으로 제품 생산 및 공정 개선에 활용되어 매출이 발생하는 경우는 약 19.4%로 낮아, 사업화가 이루어지더라도 실제적인 경제적 가치창출에 기여하는 정도는 낮은 것으로 보인다.



<그림 2> 이전기술 사업화 현황(산업통상자원부, 2023)

공공 R&D 기술의 활용 및 사업화를 활성화하기 위해서는 기술 개발자 및 수요기업에 대한 체계적인 정보제공이 매우 중요하다. 또한 공공 R&D 와 수요기업에 대한 데이터 기반의 과학적인 매칭 및 추천을 통해, 보다 신속하고 정확한 기술사업화 의사결정을 지원할 수 있다.

따라서 본 연구는 인공지능(AI) 기반 공공 R&D-수요기업 추천모델을 개발함으로써, 기술 개발자 및 기술 수요기업에 대해 최적의 사업화 정보를 제공하고자 한다. 또한 추천 모델을 서비스하기 위한 벡터 유사도 기반 알고리즘 및 벡터 DB 구조를 설계하여, 모델의 배포 및 추론 속도를 개선하는 경량화 방안은 제언하고자 한다.

II. 본론

본 연구는 한국과학기술정보연구원(KISTI)에서 보유한 공공 R&D 사업화 통합 데이터베이스(DB)로부터 공공기술 및 기술공급자(개발자)의 특징을 추출하고, 대한민국 기업 DB(기업개요, 재무 등)로부터 기술 수요기업의 특징을 추출하였다.

공공R&D-수요기업 추천모델은 크게 두가지 연구모델로 구성되어 정확도 및 신뢰성을 확보하고자 하였다. 첫번째 특징추출 모델은 기업정보와 기술정보로 부터 기업특징 및 기술특징을 추출하는 모델이다. 두번째 추천 모델은 추출된 기업특징 및 기술특징을 기반으로 기술거래를 학습하고, 기술거래 가능성을 판단하는 모델이다.

공공 R&D 사업화 특징추출 모델은 특징 추출(Feature extraction)이 가능한 Lasso(L1 Regulation) logit model, Ridge(L2 Regulation) logit model, ElasticNet Regulation Model 로 구현하였다 (Zou, H., & Hastie, T., 2005).

(**1**) L1 Regulation Penalty : $R(w) := \frac{1}{2}\sum_{i=1}^{n} |w_i|$

(**2**) L2 Regulation Penalty : $R(w) := \frac{1}{2}\sum_{i=1}^{n} {w_i}^2$

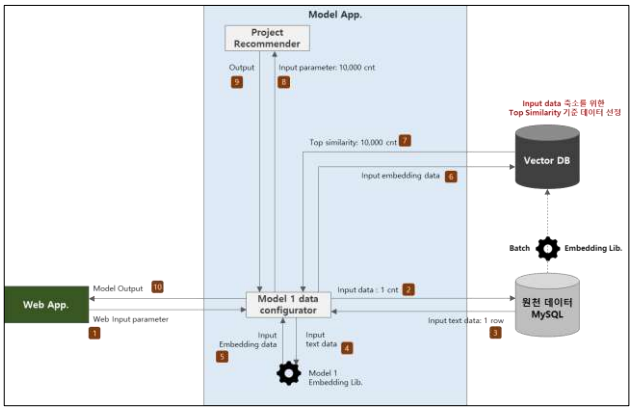
(**3**) ElasticNet Regulation Penalty
: A convex combination of L1 and L2 regulation penalty

$$R(w) := \frac{\varphi}{2}\sum_{i=1}^{n} {w_i}^2 + (1-\varphi)\frac{1}{2}\sum_{i=1}^{n} |w_i|$$

추천 모델은 기존의 기술거래 성공사례 데이터를 기초로 하여 학습 데이터셋을 구축하고, 학습 데이터를 입력 데이터로 하여 지도 학습 방법으로 공공기술 및 기술 수요자 특징을 추출하였다. 추출된 공공 R&D 특징 및 기술수요자 특징을 기초로 하여 Cross-modal Autoencoder 기반으로 공공기술 및 기술수요자 기술거래 가능성 판단결과를 도출하였다.

추천 모델은 웹 플랫폼을 통해 사용자가 활용할 수 있도록 개발되었으며, 이과정에서 모델의 배포와 원활한 서비스를 위해 추론 속도의 개선이 요구되었다. 따라서 본 연구에서는 추천 모델에 대한 경량화 구조를 설계하였는데, 첫번째는 원천데이터로부터 벡터(Vector) 임베딩(Embedding) DB 를 생성하는 것이고, 두번째는 유사도 점수를 계산하여 벡터 간의 유사도를 정량화하고 주어진 벡터의 가까운 값을 효율적으로 검색하는 구조이다.

공공 R&D 데이터에서는 과제 요약문의 '한글 키워드', '한글 기술명', '한글 키워드'를 활용하여 벡터 DB 를 생성하였으며, 수요기업 데이터에서는 '한글 주요제품', '사업목적'을 활용하여 벡터 DB 를 생성하였다. 이를 바탕으로 내적 유사도 계산을 통해 총 10,000 개의 유사 공공 R&D 및 수요기업을 추출해낼 수 있다. 위 방식으로 추천 대상이 된 공공 R&D 및 수요기업 데이터를 추천 모델을 통해 추론하여 최종적인 기술거래 가능성을 판단하고 결과를 사용자에게 제시하였다.



<그림 3> 공공 R&D-수요기업 추천모델 경량화 구조

III. 결론 및 시사점

벡터 유사도 계산 알고리즘 및 벡터 DB 를 통하여 공공 R&D 및 수요기업 추천 서비스를 개발한 결과, 아래와 같이 추론속도를 개선할 수 있었다. <표 1 참조>

<표 1> 기존 및 개선 추천 모델 추론 속도 비교

* 기준: 10,000 count	소요 시간	
기존 추천 모델 추론	45~47s	**약 65% 개선**
개선 모델 추론	14~18s	
Vector DB 적용 소요시간	0.05~0.35s	

본 연구를 통해 향후 자원 효율적이고 사용자 친화적인 AI 기반 공공 R&D-수요기업 추천 모델의 개발 및 서비스가 가능할 것으로 기대되며, 본 연구에서 제시된 기법을 확장하여 다양한 AI 기반 모델의 서비스 응용이 가능할 것으로 보인다.

향후 추가적인 벡터 유사도 계산 알고리즘 정교화, 벡터 DB 확대 등을 통하여 추천결과의 정확도 및 신뢰도를 높이고 모델의 효과성을 제고할 수 있을 것으로 기대된다.

IV. 참고문헌

산업통상자원부. (2023). 2023 년도 공공연구기관 기술이전·사업화 실태조사보고서.

한국과학기술기획평가원. (2022). 2022 년도 글로벌 R&D 투자동향 분석.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2), 301-320.

데이터 기반 제품 시장경쟁력 평가모델 개선에 관한 연구: 기업 경쟁전략관점에서의 가격결정 중심으로

조민제
한국과학기술 정보연구원 기술사업화연구센터
mjcho@kisti.re.kr

배국진
한국과학기술 정보연구원 기술사업화연구센터
baekj@kisti.re.kr

변정은
한국과학기술 정보연구원 기술사업화연구센터
jebyun@kisti.re.kr

김은선
한국과학기술 정보연구원
데이터분석본부
kimes@kisti.re.kr

Abstract *– – 기술 기반의 기술 사업화 질적 성장을 위해서는 기업의 기술 수준과 규모를 고려한 스케일업 지원 체계를 구축하는 것이 중요하다. 이러한 지원 체계에서 기존의 지원 기업 선별 방식은 주로 재무 성과에 의존하고 있어 혁신 기업 창출에 한계가 있었다. 이에 KISTI 는 국내 최초로 기술/제품/기업/시장 데이터를 통합하고 분석할 수 있는 데이터 기반의 글로벌 시장 경쟁력 평가 모델을 개발하여 다양한 공공 부문에 기술 이전을 통해 기술 및 기업의 스케일업을 지원하고 있다. 본 연구에서는 기업의 경쟁 전략 관점에서의 제품가격을 중심으로 제품별 시장 세그먼트를 분할하고, 각 시장 세그먼트별 품질계수 및 가격계수의 변화에 따른 가치계수의 결과를 파악하였다. 향후 공공 연구개발 결과물의 사업화 촉진을 목표로 하는 다양한 지원 사업 등에 평가 모형의 활용이 가능할 것이다.*

Key Terms *– 제품 시장경쟁력 평가모델, 가격계수, 제품가격, 민감도 분석*

이 논문은 2024 년도
한국과학기술정보연구원(KISTI)의 기본사업으로
수행된 연구입니다. (과제번호: (KISTI) K-24-L03-M02-C03).

I. 서론

국내 기술 기반 사업화의 증가 추세에도 불구하고 유니콘 기업의 부재와 고성장 기업의 쇠퇴는 질적 성장의 부족을 의미한다. 이에 정부는 기술 수준과 기업 규모를 고려한 기업 지원 체계를 구축하여 기술 사업화의 질적 성장을 육성하고 있다. 기존의 지원 기업 선정을 위한 평가 방식은 주로 재무적 성과에 의존한다. 이러한 방식은 혁신적인 기업 선정에 한계가 있다(이재석 외, 2021).

KISTI 기술사업화연구센터에서는 1990 년대 이후부터 국내·외 기술·시장정보 분석을 통해 기술사업화 과정의 주요 의사결정을 지원할 수 있는 체제를 개발하였으며, 2000 년대 들어 정부의 기술사업화 지원이 본격화됨에 따라 KISTI 가 보유한 방대한 데이터와 축적해 온 정보분석 기술을 토대로 기술사업화 분석모델 및 시스템을 개발하여 기술사업화 지원의 효율성 및 효과성을 높이고자 노력해왔다.

본 연구에서는 기업 경쟁전략관점에서의 가격결정 결과인 제품가격을 중심으로 제품별 시장 세그먼트별 분할하고, 각 시장 세그먼트별 품질계수와 가격계수의 변화에 따른 가치계수의 결과를 파악하여 평가모형의 개선을 통해 공공 R&D 사업화 촉진을 활성화하는데 그 목적이 있다.

II. 본론

가치공학을 기반으로 한 글로벌 시장 경쟁력 평가모델은 제품의 원가 절감과 품질 향상을 동시에 고려하여, 기술이나 제품이 세계 시장에 진입했을 때 어느 정도의 경쟁력이 있는지를 체계적이고 과학적으로 분석하고 예측하는 모델이다. 이 평가 모델에서는 평가대상 제품과 벤치마킹 글로벌 최고수준 제품 간의 품질과 가격을 비교 분석하여 정량화한다. 평가에서 적절한 경쟁제품 선정은 중요하다. 경쟁제품은 평가대상 제품과 동일한 시장 가격 세그먼트에서 선정되어야 한다. 각 시장 가격 세그먼트에서 요구하는 가격과 품질에 대한 요구수준이 다르기 때문이다. 따라서 가치계수 산정 단계에서는 가격 세그먼트의 특징을 반영할 수 있는 방법에 대한 검토가 필요하다.

본 평가모델에서 제품의 경쟁우위는 가치계수로 정의한다. 가치계수의 핵심변수는 품질 관점의 스펙 비교 값, 기술 확신도, 기술 중요도 등이 있으며, 가격 관점의 가격 계수가 있다. 이전 연구에서 가치계수에 가장 큰 영향을 주는 요인으로 가격계수가 제시되었다. 가격계수는 평가대상 제품가격과 벤치마킹 제품가격 간의 상대적 비율(ratio)이다.

제품가격은 여전히 기업과 소비자 모두에게 가장 중요한 결정요소 중 하나이다. 기업은 제품가격을 제품 생산비용에 근거해 생산원가에 기업의 적정이윤을 더하거나, 소비자가 느끼는 효용 및 가치를 기반으로 산정한다. 제품가격은 기업의 경영전략에 따라 접근방식이 다르다. 저원가 전략에서는 경쟁사보다 낮은 가격으로 제품을 제공하여 시장에서 경쟁 우위를 확보하므로, 제품가격이 고객들이 구매 결정을 내릴 때 주요한 요인이다. 반면 차별화 전략에서는 경쟁사 대비 차별화된 특징을 통해 고객들에게 고유한 가치를 제공하여 경쟁 우위를 확보하며, 고객들은 제품가격과 품질을 종합적으로 고려하여 제품을 선택하게 된다.

제품 시장 세그먼트는 제품 가격을 기준으로 구분된다. 저가 시장에서는 경쟁 우위를 위해 가격이 중요하다. 따라서 경쟁사보다 낮은 가격을 유지해야 하며, 가격을 낮추면서도 품질을 일정 수준 유지해야 한다. 중가 시장에서는 가격과 품질이 모두 중요하다. 저가 시장에 비해 가격이 높은 만큼 상대적으로 높은 품질 수준을 기대한다. 고가 시장에서는 소비자들이 고가격 제품을 구매할 때 높은 가격을 지불하는 대신 뛰어난 품질과 성능을 기대한다. 따라서 최고 수준의 품질이 필수적이다. 이는 제품 가격(시장 세그먼트)에 따라 기업의 접근 방식이 달라진다는 것을 의미한다.

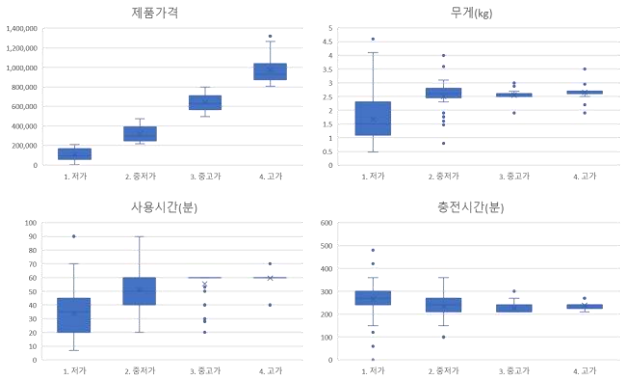
본 연구에서는 무선 청소기 제품을 대상으로 각 시장 세그먼트별 가격과 품질에 대한 기초 통계량을 분석하고, 각 세그먼트별 가격과 품질에 따른 가치계수의 감도분석 및 상관분석을 수행하였다. 온라인 쇼핑몰 사이트를 대상으로 300 개의 무선 청소기 제품에 대한 제품가격과 대표적인 품질특성을 수집했다. 제품 가격을 기준으로 K-means 클러스터링을 수행하여 총 4 개의 시장 세그먼트로 분할했다.

각 시장 클러스터별로 제품 가격 및 품질특성 데이터를 대상으로 수치적 특성을 파악하기 위해 박스플롯 차트 분석을 진행했다. 이를 통해 각 데이터별 제 1 사분위(Q1), 중앙값(median), 평균값(average), 제 3 사분위(Q3) 및 이상치(Outlier)를 파악했다.

그리고 각 시장 세그먼트별 가격계수 및 품질계수의 변화에 따른 가치계수의 민감도 분석을 실시했다. 제품 가격 및 품질특성별 평균값을 평가대상의 데이터로, Q1/중앙값/Q3 데이터를 각각 경쟁제품의 데이터로 가정하여 가격계수와 품질계수, 그리고 이에 따른 가치계수를 산출했다.

III. 결론 및 시사점

그림 1 과 같이 제품가격은 저가 세그먼트에서보다 상위 세그먼트로 갈수록 그 편차가 증가했다. 반면 품질특성에서는 그와 반대로 상위 세그먼트에서보다 하위세그먼트에서 그 편차의 폭이 컸다(무게, 사용시간). 충전시간의 경우 그 편차의 폭은 다른 품질특성에 비해 작으나 상위 세그먼트일수록 그 폭은 감소했다.



<그림 1> 제품가격 및 품질특성 데이터 기초통계량

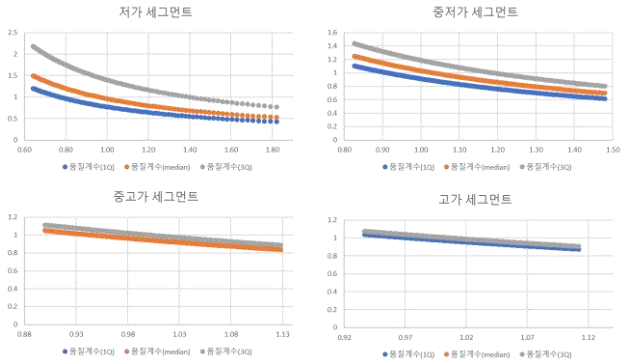
그림 2 와 같이 가격계수 및 품질계수의 변화에 따른 가치계수의 민감도 분석 시뮬레이션에서는 가격 세그먼트에 따라 결과가 달리 나타났다. 저가 세그먼트에서는 가격계수 및 품질계수의 변화에 따라 가치계수는 달라지게 되는데 품질계수가 높을수록 그 변동폭은 커졌다. 특히, 가격계수가 1.2 미만일 경우, 가격계수에 따른 가치계수의 변화가 크며, 그 변동폭은 품질계수가 높을수록 커졌다.

중저가 세그먼트에서는 품질에 따른 가치계수의 폭이 이전보다 크진 않으나 구별가능할 정도로 차이가 발생했다. 마찬가지로 가격계수가 1.2 미만이었을 때, 가격계수에 따른 가치계수의 변화가 1.2 이상일 때보다 변화가 커졌다.

중고가 세그먼트에서는 품질에 따른 가치계수의 폭이 중저가에 비해 줄어들었으며, 가치계수는 주로 가격계수에 의해 변동되었다.

마지막으로 고가 세그먼트에서는 품질에 따른 가치계수의 차이는 중고가에 비해 거의 없어졌으며,

가치계수도 가격계수에 의해서만 거의 변동되었음을 알 수 있었다.



<그림 2> 제품가격 및 품질특성에 따른 가치계수 감도 분석결과

본 연구는 가치계수에 따른 가격계수 및 품질계수의 영향력을 파악하기 위한 방안으로 이를 기반으로 제품가격에 따른 평가모형의 개선 필요성을 도출했다는 측면에서 의의가 있다. 향후에는 가치계수에 영향을 주는 품질계수와 가격계수의 가중치를 가격 세그먼트에 따라 검증하는 후속연구 및 개발이 요구된다.

Ⅳ. 참고문헌

이재석, 이기호, & 이상명. (2021). 창업기업의 성장단계별 지원체계에 관한 연구: 국내외 유니콘 기업의 사례 비교. 중소기업연구, 43(1), 165-186.

기술제품의 글로벌 경쟁력 제고를 위한 데이터기반 기술사업화 지원 플랫폼 활용전략 : 치료 및 진단기기 분야를 중심으로

신주연
한국과학기술정보연구원
기술사업화연구센터
jyshin0417@kisti.re.kr

김지수
한국과학기술정보연구원
기술사업화연구센터
zisu@ kisti.re.kr

변정은
한국과학기술정보연구원
기술사업화연구센터
jebyun@ kisti.re.kr

최윤정
한국과학기술정보연구원
기술사업화연구센터
yjchoi @ kisti.re.kr

***Abstract** – 최근 원자재 가격 폭등 및 환율 변동성 확대 등 글로벌 교역환경은 악화되고 있으며, 이로인해 기술기반 중소기업이 수출 현장에서 겪는 애로는 커지고 있다. 이러한 환경에 대응하기 위해 중소기업에서는 자체적으로 제품 경쟁력 제고를 시도하고 있지만 글로벌 시장진출을 위한 제품/가격 경쟁력 확보 및 기술제품 혁신이 미진한 상황이다. 이에 빅데이터, AI 등 데이터기반의 기술사업화 지원 플랫폼을 활용하여, 시장 중심의 기술제품 경쟁력을 효과적, 정량적으로 평가하고, 글로벌 시장 진출 전략을 지원할 필요성이 커지고 있다. 본 연구에서는 치료 및 진단기기 분야를 중심으로 기술제품의 글로벌 경쟁력을 제고하기 위해 데이터기반 기술 사업화 지원 플랫폼을 활용한 사례를 제시하고, 향후 다양한 분야 기술제품의 글로벌 경쟁력 제고를 위한 플랫폼 활용 전략을 제시하고자 한다.*

***Keywords** – 기술사업화, 데이터기반, 시장경쟁력, 시장진출전략, 플랫폼활용*

본 연구는 한국과학기술정보연구원 주요사업(K-24-L03-C03-S01)의 지원을 받아 수행되었음

I. 서론

한국 무역협회에서 진행한 '오픈 이노베이션 트렌드 설문 조사' 에 의하면 포춘 글로벌 500 기업은 한국 스타트업에 대해 기술력(55%)은 좋지만 글로벌 진출 준비도(4%)에 대해서는 상대적으로 낮게 평가하고 있다. 또한 연구소기업을 비롯한 중소기업, 기술벤처 스타트업의 경우, 제품 경쟁력 제고와 해외시장 개척에 대해 필요성을 느끼고 있으나, 해외시장 제품 정보에 대한 애로사항을 겪고 있으며, 초보수출기업의 경우, 시장중심의 제품 개발 및 글로벌 진출 전략이 부족하여 해외시장 진출에 어려움을 겪고 있다.

이에 본 연구에서는, 바이오 · 의료 분야 중 치료 및 진단기기 분야 기술제품의 글로벌 시장진출을 위한 효과적인 전략 수립을 위해, 한국과학기술정보연구원 데이터분석본부 기술사업화연구센터에서 개발한 데이터기반 기술사업화 지원 플랫폼을 활용한 사례를 제시하고, 향후 다양한 분야 기술제품의 글로벌 경쟁력 제고를 위한 플랫폼 활용전략을 제안해 보고자 한다.

II. 본론

본 연구에서는 데이터기반 기술사업화 지원 플랫폼의 여러 분석 모듈 중 '글로벌 시장경쟁력 평가' 모듈을 활용하였고 치료 및 진단기기 분야에 속하는 2 개의 제품(AI 기반 두피케어 디바이스, 호기가스 분석기)에 대한 글로벌 시장경쟁력 분석을 수행하였다.

<표 1> 치료 및 진단기기 분야 평가대상 제품

제품명	특징
AI 기반 두피케어 디바이스	1. AI 를 활용한 두피 진단 솔루션을 제공 2. 일렉트로포레이션 기법 적용으로 두피 탈모 케어 성능 개선
호기가스 분석기	1. 장 질환관련 3 종 (수소, 메테인, 황화수소) 호기가스 분석 가능 2. 시료 전처리 과정 없이 실시간 분석 가능

상기 2 개의 평가대상 제품에 대해 주요 품질특성 및 소비자가 요구하는 품질특성을 도출하고, 각 품질특성에 대한 중요도 분석 및 글로벌 선도제품과의 품질특성 및 가격 비교분석을 통해 분석 대상 제품의 글로벌 시장에 대한 경쟁력을 정량적인 등급으로 나타내었다. 또한 경쟁제품 대비

분석대상 제품의 품질특성이 우수한 부분과 부족한 부분을 정량적으로 제시하였다.

III. 결론 및 시사점

본 연구에서는 데이터기반 기술사업화 플랫폼을 활용하여 치료 및 진단기기 분야 기술제품에 대한 글로벌 시장경쟁력을 과학적, 체계적으로 분석한 사례를 제시하였다. 해당 플랫폼 활용을 통해, 기업에서는 분석 대상 제품의 품질특성별 경쟁력을 파악하여, 우수한 부분에 대해서는 제품 마케팅을 위한 레퍼런스 자료로 활용이 가능하며, 부족한 부분은 개선할 수 전략 수립이 가능하다. 또한 해당 플랫폼은 다양한 기술제품에 적용이 가능하므로 기술사업화 지원 기관에서는 경쟁력이 있는 기업을 선별할 수 있을 뿐만 아니라, 중소기업의 성공적인 기술사업화를 위한 전략 수립의 근거로도 활용이 가능할 것으로 보인다.

IV. 참고문헌

김지민, 변정은, 최윤정. (2020). 글로벌 기술사업화 지원 모델의 적용사례-에너지 분야 기업을 중심으로. 2020 *한국지능정보시스템학회 추계학술대회 논문집*, 124-125.

최윤정, 변정은, 김은선. (2023). 데이터를 이용한 공공기술 기반 제품의 시장경쟁력 평가 모형에 관한 연구. 2023 *한국지능정보시스템학회 춘계학술대회 논문집*, 177-178

Session E

E4. [학술세션]
AI를 활용한 의료 및 건강 데이터의 혁신적 응용

좌 장: 김 상 진(동아대)

퇴행성 뇌질환 조기진단을 위한 Mel-Spectrogram 기반 Ensemble 삼진 분류 모델

성상하
동아대학교 경영대학
sangha@donga.ac.kr

김상진
동아대학교 경영대학
skim10@dau.ac.kr

Abstract *– 고령화가 가속화됨에 따라 이와 관련된 퇴행성 뇌질환자가 증가하고 있습니다. 특히, 알츠하이머와 파킨슨병은 고령층에서 흔히 발생하는 퇴행성 뇌질환으로, 인지 저하 및 운동 장애와 같은 증상을 동반합니다. 이와 같은 퇴행성 뇌질환의 경우 조기 진단을 통한 치료가 필수적으로 요구됩니다. 퇴행성 뇌질환의 초기 진단은 질병 진행의 지연 뿐만 아니라 환자의 삶의 질을 개선하기 위해 중요합니다. 본 연구에서는 퇴행성 뇌질환의 조기 진단을 위해 700 명의 음성 데이터를 활용한 앙상블 기반 삼진 분류 알고리즘을 제안합니다. 정상인과 환자의 음성 데이터를 대상으로 주요 변수를 추출하고 질병 분류를 위한 합성곱 신경망(Convolution Neural Network , CNN) 을 사용합니다. 또한, 앙상블 기법을 기반한 모델 고도화를 통해 향상된 분류 정확도를 제시합니다. 본 연구는 음성 데이터 기반의 질병 삼진 분류 알고리즘을 제시함으로써, 향후 퇴행성 뇌질환 조기 진단에 기여할 수 있을 것으로 기대됩니다.*

Keywords *– 삼진분류, 음성 데이터, 퇴행성 뇌질환, 합성곱 신경망(Convolution Neural Network), Ensemble.*

I. 서론

최근, 고령 인구의 급격한 증가로 인해 퇴행성 뇌질환 유병률이 크게 증가하고 있다(Al Shoukry et al., 2020). 이와 관련된 대표적인 퇴행성 뇌질환으로는 알츠하이머병(Alzheimer's disease, AD)과 파킨슨병(Parkinson's disease, PD)이 있다. 이러한 고령층의 퇴행성 뇌질환은 고령 인구의 자립성과 삶의 질에 상당한 영향을 미치며, 인지 저하, 운동 장애 및 일상 기능 저하를 유발한다(Pulido et al., 2020; Rowe et al., 2021). 아직 퇴행성 뇌질환을 치료하기 위한 치료제가 개발되지 않았기 때문에 퇴행성 뇌질환 조기 진단의 중요성은 계속해서 증가하고 있다. 이에 따라, 인공지능(artificial intelligence, AI)을 기반으로 한 퇴행성 뇌질환 조기 진단 연구가 주목받고 있다(Chakravarthy et al., 2023). 따라서 본 연구는 AD 와 PD 의 조기 진단을 위해 AI 기술을 활용한 알고리즘을 제안하는 것을 목표로 한다.

퇴행성 뇌질환의 조기 진단에 대한 주요 연구는 뇌 이미징과 생리학적 검사인 자기 공명 영상 촬영(Magnetic Resonance Imaging, MRI) 및 뇌파 검사(Electro Encephalo Graphy, EEG)를 대상으로 연구되었다. 하지만, MRI 는 검사에 비용 및 시간이 많이 소요되며, 진단을 위해 의학적인 전문 지식이 동반된다. EEG 의 경우 MRI 에 비해 비용 효율적이지만 검사 중 눈 깜빡임과 움직임 등으로 인한 간섭이 결과에 영향을 미칠 수 있다는 단점이 있다(Cassani et al., 2018). 이와 같은 연구에서 활용된 검사 도구는 비용, 시간 및 지리적 제약으로 인해 퇴행성 뇌질환 조기 진단 접근성에 대한 제한 사항이 있다(Billeci et al., 2020). 따라서, 본 연구에서는 조기 진단의 접근성을 개선하기 위해 음성 데이터를 기반한 퇴행성 뇌질환 분류 모델을 제시한다. 음성 데이터는 환자의 음성 특성, 발음 및 억양을 기반으로 뇌 손상을 진단한다. 또한 비침습적이므로, 신속하게 결과를 확인할 수 있다는 장점이 있다(López-de-Ipiña et al., 2015).

수집된 음성 데이터의 양과 품질이 크게 향상되면서, 기계 학습과 딥러닝을 활용한 연구가 증가하고 있다. 기계 학습 기술을 활용하여 뇌 질환을 진단하는 연구는 나이브 베이즈 방법을 사용하여 퇴행성 뇌질환을 진단하는 연구가 수행된 바 있으며, 분류 정확도가 74.11%로 나타났다(Rana et al., 2022). 또한, 또 다른 음성 데이터를 기반으로 한 SVM 기반의 연구는 85%의 분류 정확도를 나타낸다(Rusz et al. 2011). 최근 SVM 과 딥러닝을 결합한 조기 진단 모델도 발표되고 있으며, CNN 과 결합한 모델의 경우 파킨슨 환자 진단 시, 86.9%의 분류 정확도를 제시한다(Vigo et al. 2022). 하지만 이와 같은 기존의 선행연구들은 상대적으로 적은 환자 수를 대상으로 실험을 진행하여 표본의 수가

적다는 한계점이 있다. 또한, 일부 연구의 경우 다중 클래스에 대한 분류를 수행했지만, 대부분의 연구는 AD 와 PD, 단일 클래스를 대상으로 연구를 수행합니다. 따라서 본 연구에서는 선행 연구의 한계점을 개선하기 위해 총 700 명의 참가자로 구성된 대규모 음성 데이터셋을 활용하여 실험의 신뢰성을 확보한다. 또한, AD 와 PD 를 동시에 진단하는 삼진 분류 모델을 제안하고, 앙상블 기반의 결합 알고리즘을 제시하여 고도화된 퇴행성 뇌질환 삼진 분류 알고리즘을 제시한다.

II. 연구방법론

본 연구에서는 퇴행성 뇌질환 삼진 분류 모델을 제시하기 위해 Mel-Spectrogram 과 CNN 백본 네트워크 중 하나인 Resnet 을 활용한다.

Mel-Spectrogram 은 음성 데이터의 주파수를 시간에 따른 변화로 시각화한 그래프이다. 이는 음성 신호의 주파수를 Mel 스케일로 변환하여 나타낸다. Mel 스케일은 사람의 음을 인지하는 기준을 반영한 변환 함수이다. 이는 사람의 음성 인식의 반응을 모델링하기 위해 설계된 주파수 척도로, 사람의 주파수를 인식하는 방식에 가깝게 설계되어 있다.Mel-Spectrogram 은 이러한 특징을 활용하여 음성 처리 및 음성 인식과 같은 음성 관련 작업에서 활용되고 있는 전처리 기법이다. 본 기법을 통해 시간 흐름에 따른 주파수 특성을 시각화하여 음성 데이터의 주요 특징을 강조하고, CNN 모델의 입력 데이터로 활용할 수 있다.

ResNet 는 대표적인 CNN 백본 네트워크이다. ResNet 은 네트워크의 깊이가 증가함에 따라 발생하는 그래디언트 소실 문제를 해결하기 위해 제안된 모델로, 잔차 블록(Residual Block)을 활용해 깊은 네트워크를 구성한다. 이러한 구조는 네트워크의 성능을 크게 향상시키면서도 모델의 파라미터 수를 크게 증가시키지 않아 더 깊은 네트워크를 효과적으로 학습할 수 있다. ResNet 은 이미지 분류, 객체 탐지 등 다양한 컴퓨터 비전 작업에서 활용되고 있으며, 본 연구에서는 이를 기반으로 전처리 된 음성 데이터를 분류한다.

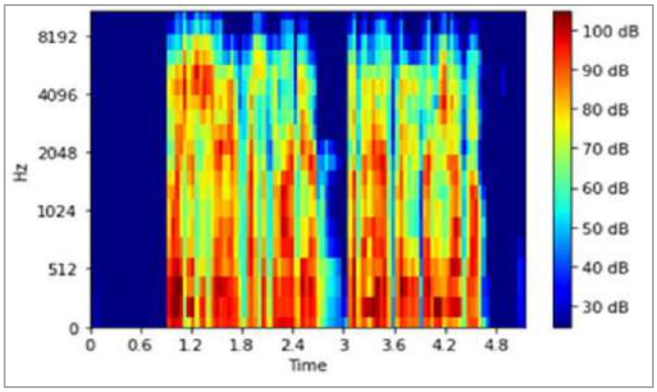
III. 실험

3-1. Data

본 연구에서는 고령자 700 명을 대상으로 수집된 음성 데이터를 대상으로 실험을 진행한다. 700 명의 고령자 중 300 명은 정상인, 400 명은 퇴행성 뇌질환 환자였으며, AD 와 PD 의 수집 비율은 1:1 로 각 200 명의 데이터로 구성된다. 데이터는 병원에 등록된 환자들이나 요양원에 거주하는 고령자를 대상으로 수집되었다.

각 실험 대상자의 음성 데이터는 단음절 읽기, 단어 읽기, 문장 읽기 및 단락 읽기를 포함하고 있으며, 총 100 개의 음성 데이터로 구성된다.

수집된 음성 데이터의 특징 추출을 위해 Mel-Spectrogram 이 사용되었으며, 이를 통해 음성 데이터의 특징 맵을 2D 데이터로 추출한다. Mel-Spectrogram 을 통해 전처리 된 데이터는 그림 1 에 나타난다. 전처리 된 데이터는 히트맵 형식으로 제시되어 데이터의 특징을 시각적으로 확인할 수 있다. 생성된 2D 데이터는 인간의 음성 영역에 반응하는 40 개의 밴드로 구성된다. Mel-Spectrogram 이 적용된 데이터는 훈련 데이터 80%, 테스트 데이터 20%로 구분하여 실험을 수행한다.



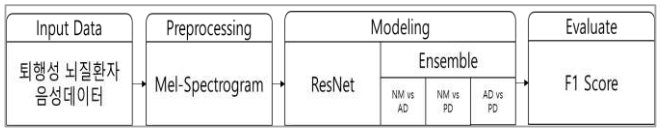
<그림 1> Mel-Spectrogram 적용 예시

3-2. Modeling

본 연구에서는 녹음된 개별 음성 데이터를 활용해 퇴행성 뇌질환자를 각 클래스(NM,AD, PD)로 분류하기 위해 ResNet 을 구축한다. 이 과정에서 총 100 개의 문장 중 충분한 음성 길이가 확보된 40 개의 문장에 대한 모델링을 수행하였다. 총 40 개의 학습 모델은 입력된 음성 데이터를 기반으로 각 피험자에 대한 삼진 분류를 수행한다.

최종 결과 제시에 앞서, 본 연구는 구축된 40 개 분류 모델에 대한 각 모델의 성능을 평가하고, 병변 분류에 사용될 핵심 발화 모델을 선택했다. 선택된 핵심 발화를 사용하여 최종적으로 정상 - AD - PD 에 대한 삼진 분류를 수행했다.

또한, 본 모델을 앙상블 기반의 고도화를 추가적으로 수행하였다. 기존 삼진분류 모델과 달리 이진 분류 모델을 구성한 후, 결과를 앙상블하여 최종 삼진 분류를 수행한다. 각 이진 분류 모델은 정상-AD, AD- PD, PD-정상 을 구분하며, 이진 분류 모델의 결과를 빈도수 기반으로 집계한다. 본 연구에서 제시된 모델은 최종적으로 F1-Score 를 통해 최종적으로 평가한다. 본 연구에서 활용된 기존 삼진 분류 및 앙상블 기반 삼진 분류 모델의 프로세스는 그림 2 와 같다.



<그림 2> 분석 프로세스

Ⅳ. 실험결과

본 연구는 퇴행성 뇌질환 환자를 분류하는데 사용된 ResNet 과 앙상블 기법을 통해 고도화 된 ResNet 간의 성능을 비교한다. 본 연구의 실험 결과는 표 1 과 같다. 각 행은 각각의 클래스(정상, AD, PD)에 대한 F1 Score 를 나타낸다. 실험 결과, 모든 클래스 군에 대해 Ensemble 기법 적용 이후 성능이 향상되었으며, 특히 환자군(AD, PD)의 경우 약 10%의 성능 향상이 나타난다. 또한 각 클래스별 샘플 수에 따라 가중치가 부여된 Weighted F1 Score 를 고려한 결과, 전체 분류 정확도가 약 8.65% 향상됨을 확인할 수 있다. 따라서, Ensemble 기법을 적용한 ResNet 이 단일 ResNet 에 비해 더 우수한 성능을 나타낸다.

<표 1> 음성 데이터 기반 퇴행성 뇌질환자 삼진 분류 결과

구분	ResNet	Ensemble ResNet
NM F1 SCORE	88.89%	95.93%
AD F1 SCORE	78.48%	88.37%
PD F1 SCORE	74.67%	84.50%
Weighted F1 SCORE	81.85%	90.50%

Ⅴ. 결론

본 연구에서는 정상, AD, PD 로 구분된 총 700 명의 고령자의 음성 데이터를 수집하여 퇴행성 뇌질환 분류 모델을 제시하였다. 녹음된 음성 데이터의 주요 특징 추출 및 입력 데이터 변환을 위해 Mel-Spectrogram 이 활용되었으며, Mel 스케일은 사람의 음성 주파수 범위를 포함하는 40 개로 선정하여 실험을 수행하였다. 이후 전처리 된 데이터에 대해 ResNet 기반의 삼진 분류 알고리즘을 활용하여 퇴행성 뇌질환자를 분류하였다. 실험을 수행한 결과, 이진 분류 모델의 결과 값을 Ensemble 한 모델의 결과가 약 8.65% 우수하게 나타났다. 특히, 환자를 대상으로 한 분류 정확도가 약 10% 가량 우수하게 나타남으로써 본 연구에서 제시한 고도화 된 Ensemble 기반 분류 모델이 퇴행성 뇌질환자를 조기진단하는데 효과적으로 활용될 수 있음을 시사한다.

Ⅵ. 참고문헌

Al-Shoukry, S., Rassem, T. H., & Makbol, N. M. Alzheimer's diseases detection by using deep learning algorithms: a mini-review. IEEE Access, 2020, 8, 77131-77141.

Pulido, M. L. B., Hernández, J. B. A., Ballester, M. Á. F., González, C. M. T., Mekyska, J., & Smékal, Z. Alzheimer's disease and automatic speech analysis: a review. Expert systems with applications, 2020, 150, 113213.

Rowe, T. W., Katzourou, I. K., Stevenson-Hoare, J. O., Bracher-Smith, M. R., Ivanov, D. K.,Escott-Price, V. Machine learning for the life-time risk prediction of Alzheimer's disease: a systematic review. Brain communications, 3(4), 2021, fcab246.

Chakravarthy, A., Panda, B. S., & Nayak, S. K. Review and Comparison for Alzheimer's Disease Detection with Machine Learning Techniques. International Neurourology Journal, 2023, 27(4), 403-409.

Cassani, R., Estarellas, M., San-Martin, R., Fraga, F. J., & Falk, T. H. Systematic review on resting-state EEG for Alzheimer's disease diagnosis and progression assessment. Disease markers, 2018.

Billeci, L., Badolato, A., Bachi, L., & Tonacci, A. Machine learning for the classification of alzheimer's disease and its pro-dromal stage using brain diffusion tensor imaging data: A systematic review. Processes, 2020, 8(9), 1071.

López-de-Ipiña, K., Solé-Casals, J., Eguiraun, H., Alonso, J. B., Travieso, C. M., Ezeiza, A., ... & Beitia, B. Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach. Computer Speech & Language, 2015, 30(1), 43-60.

Rana, A., Dumka, A., Singh, R., Rashid, M., Ahmad, N., & Panda, M. K. An Efficient Machine Learning Approach for Di-agnosing Parkinson's Disease by Utilizing Voice Features. Electronics, 2022, 11(22), 3782.

Rusz, J., Cmejla, R., Ruzickova, H., Klempír, J., Majerova, V., Picmausova, J., ... & Ruzicka, E. Acoustic analysis of voice and speech characteristics in early untreated Parkinson's disease. In MAVEBA. 2011. pp. 181-184.

Vigo, I., Coelho, L., & Reis, S. Speech-and language-based classification of Alzheimer's disease: a systematic review. Bioen-gineering, 2022, 9(1), 27.

3D-to-2D 지식 증류를 통한 의료 영상 분류 성능 개선 방안

윤혜민
동아대학교 경영정보학과
yoonlea1205@donga.ac.kr

김상진
동아대학교 경영정보학과
skim10@dau.ac.kr

Abstract – 최근 의료영상 분석에 딥러닝, 특히 3D Convolution Neural Network(CNN) 기술의 적용이 확대되고 있으나, 의료 영상의 비싼 획득 비용과 진단 레이블의 어노테이션에 대한 인력의 획득 등으로 인해 데이터 양에 제한이 따른다. 이에 대한 대안으로 많은 연구들이3D 영상에서 추출한 대표 슬라이스를 사용한 2D 영상분석을 시도했으나, 이 방법은 전체 볼륨 정보의 일부만을 활용한다는 한계가 있다. 본 연구에서는 3D volumetric 데이터의 지식을 2D CNN 학습에 활용하는 새로운 3D-to-2D 지식 증류 프레임워크를 제안한다. 3D 데이터로부터 전체 볼륨 지식을 인코딩하는 3D teacher network, 2D 데이터로부터 부분적 볼륨 정보를 인코딩하는 2D student network, 그리고 두 네트워크 간 특징 임베딩 공간에서의 데이터 연관성 간극을 줄이기 위한 증류 손실함수 모듈을 제안한다. 제안한 방법은 PPMI 데이터셋에서 2D projection 방법에 대해 전반적인 분류 성능을 개선을 보였으며, FuseMe 접근방식을 적용한 결과, 3D teacher network(97.66%)보다 더 높은 98.30 %의 F1 score 를 달성하였다.

Keywords – 3D to 2D, Deep learning, Functional neuroimaging, Knowledge distillation, Parkinson's disease

This research was supported by Management Information systems, Dong-A University in Republic of Korea.

I. 서론

의료 영상 데이터는 퇴행성 질환으로 인한 뇌 변화를 탐지하고 모니터링 하는데 있어 필수적이며, CT, SPECT, PET, MRI 와 같은 비침습적 기술과 3D 프린팅 팬텀, AI 기반 진단과 같은 기술을 통해 의료 영상 분석 분야는 지속적으로 발전하고 있다(Filippou and Tsoumpas, 2018). 이러한 기술의 발달은 방사선 노출 위험을 줄이고, 이미지 해상도를 향상시키며 다중 모달리티 데이터를 활용하고 데이터가 제한된 상황에서 진단의 신뢰성과 안전성을 높이는데 기여하였다(Zhou et al., 2021, Huang et al., 2020). 이러한 기술의 진보는 분석의 초점을 수작업 특징추출에서 직접 데이터 해석으로, 단일 모달리티에서 다중 모달리티로 전환시키며, 데이터의 양과 복잡성이 증가함에 따라 보다 정확하고 효율적인 분석 방법에 대한 요구가 증가하고 있다.

의료 영상 분석 분야에서 딥 러닝 기술 채택으로 인한 성능 향상이 보고되고 있다. ImageNet 에서 2D Convolution Neural Network(CNN)의 효과가 입증된 이후, 다양한 CNN 아키텍처가 도입되었고(Russakovsky et al., 2015), 특히 3D CNN 의 사용은 파킨슨병(PD)이나 알츠하이머(AD) 탐지와 같은 신경 영상 분석 분야에서 다양한 연구로 이어졌다(Choi et al., 2017). 3D CNN 은 풍부한 volumetric 특징 정보를 관찰할 수 있는 많은 매개변수를 갖추지만, 대부분의 딥 러닝 모델과 마찬가지로 과도하게 많은 매개변수 모델은 훈련 데이터의 분포에 과적합이 될 가능성이 있으며 실시간 상황에서 마주치는 유형에 취약 할 수 있다. 이를 위해 많은 데이터 셋이 필요하지만 데이터 셋 구축에 시간과 비용이 많이 든다는 문제가 있다.

딥 러닝 분야에서 작은 데이터 셋의 한계를 극복하기 위해 대규모 데이터 셋에서 학습된 지식 시스템을 다른 작업에 적용하는 전이 학습 개념이 제안되었다. ImageNet 에서 사전 훈련된 모델의 매개변수를 고정시키고 분류헤드를 추가하여 미세 조정하는 방식으로 전이 학습이 주로 적용된다. 또한, 3D 볼륨 영상을 2D 형식으로 추출 및 변환하는 연구가 제안되었다. 이 접근법은 한 슬라이스 또는 몇 개의 슬라이스 영상만으로 볼륨 정보를 충분히 반영하지 못해 성능이 저하 될 수 있으며, 파킨슨 병과 같이 특정 뇌 영역의 상세한 관찰이 필수적인 경우에는 충분한 정보에 기반한 판단이 이루어지지 않을 수 있다. 한편, 2D projection 방법은 제한된 깊이 정보를 사용하여 부분적인 볼륨 특징을 표현할 수 있지만, 이러한 특징들이 원본 3 차원 공간에서 어떻게 구조적으로 연결되어 있는지 모델에 설명하지 못한다.

본 연구에서는 소규모 데이터 셋에서 훈련된 2D CNN 네트워크가 원본 3D 볼륨 정보의 전체 범위를

충분히 발견하지 못하는 해결하기 위해 3D 볼륨 영상과 이 영상의 2D 표현간의 차이를 연결하는 새로운 접근방식인 3D 에서 2D 로의 지식증류(3D-to-2D Knowledge Distillation; 3D-to-2D KD)를 제안한다. 연구에서 3D 볼륨 영상과 projection 된 2D 표현간의 모달리티 간 차이를 줄이는 것과 관련된 과학적 문헌이 부족함을 강조하며, 제안된 방법은 2D CNN 네트워크와 사전 훈련된 3D CNN 네트워크 간의 그래프 표현 차이를 최소화 함으로써 개별 샘플에 포함된 부분 볼륨 특징 관계의 사용을 극대화 하고자 한다. 이 접근 방법은 PPMI 데이터 셋에서 얻은 다중 센터 SPECT 영상을 사용한 PD 분류 작업을 통해 검증되며, 2D projection 방법에 따라 모달리티 간 차이 감소 정도가 다르다는 것에 주목한다. 본 연구는 의료 영상 데이터 분석의 복잡성과 중대한 성격을 강조하며 다양한 인공지능 기반 3D 볼륨 영상 분석 작업에 직교적으로 적용 가능한 간단하지만 효과적인 해결책으로 3D-to-2D KD 를 제안한다. 본 연구의 목적은 3D 정보를 활용하여 2D CNN 을 훈련시키는 프레임워크, 다양한 2D projection 방법에 걸쳐 모달리티간 차이를 분석, 데이터 관계 정보를 활용하여 모든 2D 방법에서 PD 분류 성능을 개선하는 지식 증류의 사용이 포함된다.

II. Method

우리는 Teacher network 인 3D CNN 및 Student network 에서 생성된 인코딩된 특징으로부터 유사성 행렬 간의 차이를 직접 줄임으로써 추출한다. 입체적 특징을 기반으로 하는 3D-to-2D KD loss 는 다음과 같이 정의하고 적용하였다.

$$L_{total} = L_{CLS}s(\hat{y}_S, y_S) + L_{fg}(S_T, S_S)$$

여기서, 유사성 행렬 간의 차이는 다음과 같이 구한다.

$$L_{fg}(S_T, S_S) = \frac{1}{B^2} \sum_{(i,j)\in I} ||S_T - S_S||_2^2$$

추출된 표현에 대해서 우리는 L2norm 을 적용하고 2D student network 의 최종 이전 레이어에서 특성 표현을 위한 유사성 행렬도 유사하게 계산하여 최종 손실함수에 적용하였다

$$\tilde{f}_S = \frac{f_S}{||f_S||_2};\ S_S = \tilde{f}_S \cdot \tilde{f}_S^T$$

III. Results

우리의 제안 방법의 효과를 검증하기 위해 PPMI database(www.ppmi-info.org/data)로부터 2023 년 3 월 24 일에 획득하였다. 해당 연구 코호트는 Dopamine transporter imaging 를 촬영한 121 Normal Control(NC)와 816 Parkinson's disease(PD) 진단을 가진 총 1028 subjects 를 포함한다. 이 SPECT 영상은 HMRMES system 을 통한 iterative reconstruction 과 PMOD 를 통해 전처리 과정을 거쳤다. 보다 상세한 데이터 또는 전처리에 대한 정보가 필요한 독자는 PPMI documentation 을 참고할 수 있다.

전체 데이터는 6:2:2 로 나뉘었고 각각 훈련, 검증, 테스트 데이터셋으로 사용되었다. Backbone network 는 ResNet18 을 사용하였다.

본 연구의 제안에 평가를 위해 선행연구에서 취했던 대표적인 2D projection method 를 재현하고, 각 방법에 3D-to-2D KD 를 적용했을 때 성능을 비교한다. Single slice as input(Khan et al., 2019, Yagis et al., 2021)은 original 3D 영상에서 Dopamine transfer 가 표현되는 슬라이스 index 에서 중간 슬라이스와 전후 슬라이스 1 장씩, 총 3 장의 슬라이스를 non-i.i.d 입력으로써 사용하고 Adjacent slices with EF(Aderghal et al., 2016)는 앞서 추출한 3 장을 RGB-band 에 할당하여 채널차원으로 결합을 수행한 입력을 모델에 사용한다. 위 영상추출방법을 2D+e approach 로 참조한다. Adjacent slices with JF(Aderghal et al., 2017)는 original 3D 영상에서 axial, coronal, sagittal 각 축마다 2D+e approach 를 수행하고 각각의 2D CNN 을 통해 얻어진 신경망의 intermediate layer 에서 결합하여 전달한다. 이 방법은 선행연구에서 FuseMe 라는 말로 참조된다. Rank-pooling 은 3D 영상에서 임의의 축에따라 슬라이스를 순회하며 cofficients 를 할당하여 가중합을 취함으로써 단일 슬라이스에 3D 정보를 담는다. Rank-pooling with EF 는 각 축에대한 가중합의 단일 슬라이스를 RGB-band 에 결합하여 모델에 적용하는 것이고 Rank-pooling with JF 는 입력은 각 축에대한 가중합의 단일슬라이스이지만 모델에 적용하는 방법은 FuseMe 와 동일하다.

<표 1> 3D-TO-2D KD 적용한 입력 별 실험 결과

Configuration	Weighted F1 (%)	
3D Teacher Network		97.66
2D Student Network with 2D Projection method	without 3D-to-2D KD	with 3D-to-2D KD (Ours)
(A) Single slice as input		
Axial	97.07	97.24
Coronal	94.83	97.07
Sagittal	80.72	80.30
(B) Adjacent slices with EF (2D+e)		
Axial	90.47	96.92
Coronal	91.68	97.14
Sagittal	90.47	96.92
(C) Adjacent slices with JF (FuseMe)	96.98	**98.30**
(D) Rank-pooling		
Axial	91.10	92.99
Coronal	92.15	94.96
Sagittal	94.13	94.81
(E) Rank-pooling with EF	93.25	94.69
(D) Rank-pooling with JF	94.59	**96.45**

<표 1>에서 본 연구의 방법은 (A)의 sagittal 에서의 단일 슬라이스를 제외한 거의 모든 방법에서 지식이 강화되고 성능이 향상됨을 관찰할 수 있다. 특히 FuseMe 설정의 (C)에서의 성능은 3D Teacher model 의 성능과 비교할 만한 결과를 보였다. (A)에서 파킨슨병에서 도파민 변성을 관찰할 수 있는 선조체 영역에서의 변화를 대표할 수 있는 적절한 평면을 선택함으로써, 충분한 데이터 증강 및 다양한 모델을 만들 수 있다는 것을 발견했으며, 약간의 성능개선도 관찰 할 수 있었다.

IV. 참고문헌

ADERGHAL, K., BENOIS-PINEAU, J., AFDEL, K. & GWENAíLLE, C. FuseMe: Classification of sMRI images by fusion of Deep CNNs in 2D+ ε projections. Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing, 2017. 1-7.

ADERGHAL, K., BOISSENIN, M., BENOIS-PINEAU, J., CATHELINE, G. & AFDEL, K. Classification of sMRI for AD diagnosis with convolutional neuronal networks: A pilot 2-D+ study on ADNI. International Conference on Multimedia Modeling, 2016. Springer, 690-701.

CHOI, E., SCHUETZ, A., STEWART, W. F. & SUN, J. 2017. Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association,* 24.

FILIPPOU, V. & TSOUMPAS, C. J. M. P. 2018. Recent advances on the development of phantoms using 3D printing for imaging with CT, MRI, PET, SPECT, and ultrasound. 45**,** e740-e760.

HUANG, S.-C., PAREEK, A., SEYYEDI, S., BANERJEE, I. & LUNGREN, M. P. J. N. D. M. 2020. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. 3**,** 136.

KHAN, N., HON, M. & ABRAHAM, N. 2019. Transfer Learning with intelligent training data selection for prediction of Alzheimer's Disease. arXiv 2019. *arXiv preprint arXiv:1906.01160*.

RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATHY, A., KHOSLA, A. & BERNSTEIN, M. 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision,* 115**,** 211-252.

YAGIS, E., ATNAFU, S. W., DE HERRERA, A. G. S., MARZI, C., GIANNELLI, M., TESSA, C., CITI, L. & DICIOTTI, S. 2021. Deep learning in brain MRI: Effect of data leakage due to slice-level split using 2D convolutional neural networks.

ZHOU, S. K., GREENSPAN, H., DAVATZIKOS, C., DUNCAN, J. S., VAN GINNEKEN, B., MADABHUSHI, A., PRINCE, J. L., RUECKERT, D. & SUMMERS, R. M. J. P. O. T. I. 2021. A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. 109**,** 820-838.

A Study of Improving Medical Response Performance Using RAFT

Su-hyeon Bae
Ewha womans University Computer Science
susu19@ewha.ac.kr

Ji-won Park
Ewha womans University Computer Science
mary000605@ewha.ac.kr

Ho-yeon Park
Ewha womans University Computer Science
hoyeonpark@ewha.ac.kr

Abstract – We propose the introduction of RAFT technology into the medical QA system to address the issues of information accessibility and usability faced by healthcare professionals in the medical industry. Through experiments conducted on the MedQuAD and PubMedQA datasets, the paper demonstrates the effectiveness of integrating RAFT in improving the response performance of medical QA systems. This suggests ways to improve the retrieval and use of medical information, enabling healthcare professionals to access more accurate and reliable information rapidly, thus improving the quality of medical services.

Keywords – Finetuning, LLM, Medical QA, QA, RAG

I. Introduction

In today's medical industry, the accessibility and utility of information are critical factors that directly affect the quality of healthcare services. Healthcare professionals need to search and analyze vast amounts of medical information for various purposes such as disease diagnosis, treatment planning, and medical research. However, according to the Expert Search Strategies study (Russell-Rose et al., 2017), healthcare professionals spend an average of 4 hours on search tasks, which is a significant time burden that can diminish their work efficiency. Moreover, the complexity of information retrieval, as well as the heterogeneity and specificity of medical information, necessitate the development of more specialized retrieval solutions. Based on this background, this paper proposes to introduce Retrieval-Augmented Fine-Tuning (RAFT) technologies into medical QA systems to demonstrate the effect of enhancing response generation performance and to suggest approaches for searching and utilizing medical information.

II. Theoretical Background

Information retrieval in the medical field is essential but time-consuming. To address this issue, there have been studies applying QA to large datasets, such as PRNA (Datla et al., 2017) and CMU-OAQA (Wang & Nyberg, 2017), where CMU-OAQA achieved top performance in medical task domains for re-identification and ranking responses. This research demonstrates the ability of QA systems to effectively process medical data. However, in QA tasks requiring accurate response generation, Large Language Models (LLMs) have limitations in producing incomplete (Singhal et al., 2022; Bian et al., 2023; Xu et al., 2023) and non-logical (Chen et al., 2023c; Zhong et al., 2023; Kang et al., 2023) responses.

RAG (Retrieval-Augmented Generation) can be used to improve the accuracy and reliability of generated responses by retrieving relevant documents from a knowledge base as model input (Lewis, P. et al., 2020). Self-BioRAG (Jeong et al., 2024) demonstrated the potential accuracy improvement of RAG in medical QA across multiple biomedical benchmark datasets. Applying RAG to medical QA, demonstrated the potential improvement on generated response performance for medical queries by using RAG. In this study, we propose the application of RAFT (Retrieval-Augmented Fine-Tuning) (Zhang et al., 2024) in the medical domain by leveraging the successful elements of RAG combined with fine-tuning.

III. Methodology

Question answering (QA) systems are information retrieval technologies that automatically provide answers to questions posed in natural language (Lopez et al., 2007). These systems derive a variety of question types, including fact, definition, and list, based on structured databases of information or knowledge (Barskar et al., 2012; Mutabazi et al., 2021).

The RAG model is used in knowledge-intensive natural language processing (NLP) tasks by combining pre-trained seq2seq models with dense vector indices from Wikipedia. In open-domain QA tasks, the RAG model shows superior performance compared to pre-trained seq2seq-based models. RAFT (Retrieval Augmented Fine Tuning) is a novel approach that combines RAG and SFT (Supervised Finetuning) to enhance model performance for open-book questions within a specific domain.

IV. Experiment

This study aimed to validate the effectiveness of RAFT technology in improving the response performance in medical QA systems. Experiments were conducted using datasets based on MedQuAD and PubMedQA, a total of 16,729 medical question-answer pairs. The MedQuAD dataset, generated by crawling 12 NIH websites, covers 37 question types including diseases, drugs, and screenings. PubMedQA is a biomedical QA dataset collected from PubMed abstracts, divided into training, validation, and test datasets in a 7:2:1 ratio.

The experiments proceeded in three steps: evaluating the performance of the BioMedGPT-LM-7B (vanilla model), applying PEFT (parameter-efficient fine-tuning) to fine-tune only certain parameters of the vanilla model, and finally evaluating the performance of the RAFT model, which combines RAG with the fine-tuned model. In the RAFT process, a validation dataset was stored in Chroma DB to search for the three most similar datasets to the input question. The retrieved data, along with the input question, were then entered into the prompt to generate responses using the fine-tuned model.

The performance of the three models was evaluated using four metrics: ROUGE, BLEU, F1 score, and Perplexity. The experimental results showed that the RAFT model outperformed the vanilla model and the fine-tuned model in all performance metrics. This demonstrates the contribution of RAFT technology to increasing the accuracy and reliability of response generation in medical QA systems.

V. Conclusion

This experiment demonstrated the effectiveness of RAFT technology in improving response performance in medical QA systems. The implementation of the RAFT model allows healthcare professionals to quickly access more accurate and reliable information, which has the potential to improve the quality of medical services. However, one of the major limitations of this study is the exclusive reliance on automated evaluation metrics to assess the accuracy of the generated response. Given the specialized nature of medical responses, the inclusion of human evaluation by medical experts is essential for a more accurate and in-depth evaluation of generated response performance. With the addition of human evaluation, the medical QA systems presented in this paper are expected to provide a solid foundation for practical use in real-world medical settings.

References

Barskar, R., Ahmed, G.F., Barskar, N. (2012). An approach for extracting exact answers to Question Answering (QA) system for english sentences. Procedia Eng. Volume 30, 1187–1194. https://doi.org/10.1016/j.proeng.2012.01.979

Datla, VV., Arora, TR., Liu, J., Adduru, V., Hasan, SA., Lee, K., et al. (2017). Open domain real-time question answering based on asynchronous multiperspective context-driven retrieval and neural paraphrasing. Proceedings of The Twenty-Sixth Text Retrieval Conference, TREC 2017, Gaithersburg, Maryland, USA, https://trec.nist.gov/pubs/trec26/papers/prna-QA.pdf.

Jeong, M., Sohn, J., Sung, M., Kang, J. (2024). Improving Medical Reasoning through Retrieval and Self-Reflection with Retrieval-Augmented Large Language Models. https://www.semanticscholar.org/reader/848772a50cee68e88988ded7522e280d1c490598

Lewis, P. et al. (2020). Retrieval-augmented generation for knowledge intensive nlp tasks. Advances in Neural Information Processing Systems 33 (NeurIPS 2020), Vancouver, Canada. https://dl.acm.org/doi/abs/10.5555/3495724.3496517

Lopez, V.; Uren, V.; Motta, E.; Pasin, M. AquaLog. (2007). An ontology-driven question answering system for organizational semantic intranets. *J. Web Semant. 5*, 72–105. https://doi.org/10.1016/j.websem.2007.03.003

Mutabazi, E., Ni, J., Tang, G., Cao, W. (2021). A Review on Medical Textual Question Answering Systems Based on Deep Learning Approaches. *Applied Sciences*. 11(12):5456. https://doi.org/10.3390/app11125456

Russell-Rose, T., & Chamberlain, J. (2019). Expert Search Strategies: The Information Retrieval Practices of Healthcare Information Professionals. JMIR Med Inform. 2017;5(4):e33., http://medinform.jmir.org/2017/4/e33/.

Wang, D., Nyberg, E. (2017). CMU OAQA at TREC 2017 LiveQA: A Neural Dual Entailment Approach for Question Paraphrase Identification. Proceedings of The Twenty-Sixth Text Retrieval Conference, Gaithersburg, Maryland, USA. https://www.semanticscholar.org/paper/CMU-OAQA-at-TREC-2017-LiveQA%3A-A-Neural-Dual-for-Wang-Nyberg/ab6276e4d5bb832e8073bdfccfb53809df7d8381

Zhang, T., Patil, S. G., Jain, N., Shen, S., Zaharia, M., Stoica, I., & Gonzalez, J. E. (2024). RAFT: Adapting Language Model to Domain Specific RAG. https://doi.org/10.48550/arXiv.2403.10131

데이터 마이닝을 활용한 바이오의약 분야의 기술 융합 도출 연구

김남현
아주대학교 경영대학
nhk9503@ajou.ac.kr

김세형
현대엔지비
sehyoung@hyundai-ngv.com

강주영
아주대학교 경영대학
jykang@ajou.ac.kr

***Abstract** – 최근 고령화 현상이 심화되면서 바이오 산업이 전세계적으로 차세대 주력산업으로 급부상하고 있다. 글로벌 바이오의약 시장은 AI 를 비롯한 4 차 산업혁명 기술혁신과 함께 지속적으로 성장하고 있는데, 국내 바이오의약 산업도 글로벌 시장의 성장세에 맞추어서 커지고 있다. 또한 코로나 19 팬데믹이 감염병 예방을 위한 백신과 치료제 개발의 중요성을 일깨워주면서 바이오의약 산업 규모도 점차 커지고 있는 실정이다. 바이오의약 분야에서 전략적 우위를 차지하기 위해서는 빠르게 변화하는 바이오의약 분야의 기술 융합 양상을 파악해야 한다. 본 연구에서는 바이오의약 분야의 기술 융합 양상을 분석하기 위해 바이오의약 분야와 함께 융합되는 기술을 파악하도록 한다.*

***Keywords** – 바이오의약, 모듈성 분석, 특허 분석*

I. 서론

고령화 현상과 코로나 19 팬데믹 현상으로 인해 바이오 산업은 최근 들어 전세계적으로 차세대 주력산업으로 주목받고 있다(정지은 등, 2022). 국가기술표준원에 따르면 대한민국의 바이오산업은 크게 여덟 가지로 분류된다. 바이오의약 산업은 하나의 대분류로서 지속적으로 글로벌 시장 규모가 성장하고 있는 추세이다. 또한 코로나 19 팬데믹 현상이 가져온 위기 상황은 국가 안보 측면에 있어서 바이오의약 산업의 중요성을 부각시켰다(한국제약바이오협회, 2023). 이러한 국제 바이오산업 흐름과 함께 국내 바이오의약 산업도 지속적으로 커지고 있다. 특히, 바이오시밀러와 위탁개발생산 (Contract Development and Manufacturing Organization, CDMO)과 같은 생산 부문에 대한 투자가 활발하다(경희권, 2023; 정지은 등, 2022).

본 연구는 바이오의약 분야의 CPC 코드의 네트워크 분석을 기반으로 연구를 진행한다. 바이오의약 분야의 CPC 코드와 함께 활용되는 다른 분야의 CPC 코드를 분석함으로써 기술의 융합 발전 양상을 파악할 수 있다. 바이오의약 분야 특허를 중심으로 어떠한 분야의 CPC 코드가 함께 융합되는 것인지 파악해서 융합되는 분야를 도출하는 것이다. CPC 코드의 네트워크에서 모듈성 분석을 통해 기술 클러스터들을 도출하고, 해당 클러스터에서 바이오의약 CPC 코드와 함께 나타나는 주요한 기술 내용을 해석함으로 융합 분석을 진행할 수 있다. 본 연구는 국가 전략 사업인 바이오의약 분야의 융합 발전 양상을 데이터 마이닝 기법을 활용해 파악한다는 점에서 향후 바이오의약 산업 전략 수립에 효과적으로 활용할 수 있을 것이라 판단한다

II. 연구개요

바이오의약 분야에 속하는 주요 CPC 코드는 의약품 제조 및 개발, 유전자 치료 및 기술, 바이오의약품 및 생물학적 제제, 임상 시험 및 연구로 크게 네 가지로 분류할 수 있다. 본 연구에서는 아래의 CPC 코드와 함께 융합되는 다양한 분야를 도출함으로써 융합 분야를 도출하는 연구를 진행하려고 한다.

<표 1> 바이오의약 분야의 주요 CPC 코드

분류	CPC 코드	설명
의약품 제조 및 개발	A61K 31/00	의약품의 화학적 조성
	A61K 38/00	의약품의 생산 방법, 조제 과정 등
	A61K 45/00	의약품, 약제 또는 화학적 조성물 관련 정보 처리 및 저장

유전자 치료 및 기술	C12N 15/00	동물세포 또는 식물세포의 유전자 조작
	C12N 15/63	동물세포 또는 식물세포의 유전자 조작에 관한 것, 특허화 된 생산 및 사용을 특정한 유전자에 적용하는 것
	C12N 15/90	병원체 또는 바이러스의 유전자 조작
바이오 의약품 및 생물학적 제제	C07K 1/00	단백질 또는 펩타이드의 일반적인 방법 또는 기술
	C07K 16/00	면역학 또는 생화학적 방법 또는 기술을 이용한 단백질 또는 펩타이드 생산
	C07K 21/04	합성 RNA 또는 DNA; 캐리어의 구조, 화학적 조성 또는 생물학적 특성; 이들의 제조, 정제 또는 일반적인 사용
임상 시험 및 연구	G16H 40/00	의료, 보건 또는 임상 시험을 위한 데이터 처리

III. 결론

국가 간 경쟁이 치열해지면서 유망한 기술 분야에서 경쟁우위를 선점하는 게 중요하다. 글로벌 바이오의약 산업의 규모가 커지면서 국내 바이오산업도 국가 전략 산업으로서 핵심 역량을 재고할 필요가 있다. 본 연구는 바이오의약 분야의 CPC 코드를 바탕으로 네트워크 분석을 실시해서 기술 융합 분석을 진행한다. 이를 통해 바이오의약 산업 전략 수립 및 연구 기획에 기여할 수 있기를 기대한다.

IV. 참고문헌

경희권. (2023). 경제 안보 시대, 전략산업의 미래와 우리의 대응 방안. 산업연구원.

정지은, 최윤희, 허선경. (2022). 바이오의약품산업의 가치사슬별 경쟁력 진단과 정책 방향. 산업연구원.

한국제약바이오협회. (2023). KPBMA Briefn 25호.

실제 피트니스 환경에서의 운동 동작 분석을 위한 AI 기반 합성 데이터 활용

박성진
국민대학교 일반대학원
데이터사이언스학과
tjdfudfhs@kookmin.ac.kr

박종혁
국민대학교 경영대학
AI 빅데이터융합경영학과
jonghyuk@kookmin.ac.kr

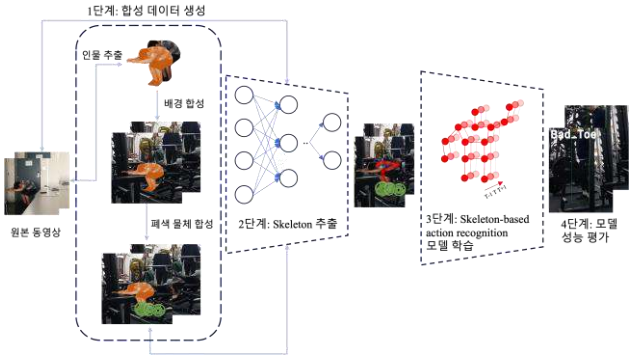
***Abstract** – 인공지능의 발달과 함께 피트니스 분야에서도 운동 동작을 자동으로 인식하고 피드백을 제공하기 위한 여러 연구가 진행되어 왔다.*
*그러나 기존 연구들은 학습 데이터의 부족, 동작 오류 라벨의 부재와 같은 이유로 현실에서의 활용이 어려웠다. 이에 본 연구는 실제 피트니스 환경과 유사한 합성데이터를 활용하여 폐색으로 인한 성능 저하를 줄이고 모델의 강건성을 향상시켜 현실에서의 활용성을 높였다. 합성데이터는 **Waseda squat dataset** 을 기반으로 다양한 피트니스 환경을 반영한 데이터를 생성하였다. 이를 통해 **Skeleton-based action recognition** 모델을 학습 시킨 결과 합성데이터의 사용이 운동 동작 오류 구분에 효과적임을 확인할 수 있었다. 이러한 연구 결과는 정확한 운동 피드백을 제공하여 운동 효과를 극대화하고 부상 위험을 최소화하는데 기여할 것으로 기대된다.*

***Keywords** – 동영상, 동작 구분 , 인공지능, 피트니스 , 합성 데이터*

I. 서론

꾸준한 운동 및 피트니스 활동은 근력 증진, 기초 대사량 향상, 스트레스 완화와 같은 다양한 신체적, 정신적 이점을 제공한다고 알려져있다. (Garber et al., 2011; Niermann et al., 2013; Abou et al., 2016). 그러나 부정확한 운동 동작 수행은 근육 발달 저하, 부상 위험 증가 등의 부정적인 결과를 초래할 수 있다.(Lorenzetti et al., 2018) 이에, 올바른 자세로 근력운동을 수행하는 것의 중요성이 강조되지만, 많은 초보자들이 정확한 자세를 파악하기 어렵고, 전문가의 조언을 얻는 데에는 상당한 비용이 발생한다. 기존 여러 연구들이 이런 문제를 해결하기 위해, 운동 동작을 자동으로 인식하고 이로부터 피드백을 제공하는 방법을 제시했다. (Rungsawasdisap et al., 2018; Ogata et al., 2019; Parmar et al., 2022). 그러나 이런 연구들은 연구실이나 가정집 같은 운동장소가 아닌 곳에서 촬영된 데이터를 사용하여 실제 피트니스 환경에서 발생할 수 있는 운동 장비에 의한 가려짐, 주변 사람과의 겹침과 같은 폐색이 존재하는 데이터에서 성능 저하가 발생한다.(Cheng et al., 2019) 또한 dataset 의 구체적인 동작 라벨이 존재하지 않아 동작 오류분석에 한계를 보였다. 본 연구는 이러한 한계를 극복하기 위해 Waseda squat dataset(Ogata et al., 2019)의 스쿼트 오류 동작 라벨을 기반으로 합성 데이터를 생성했다. 이렇게 생성된 합성 데이터로 여러 폐색 상황을 가정하여 폐색에 의한 성능 저하 문제에 대응 하였다.

본 연구는 그림 1 와 같이 총 4 단계의 구조를 갖는데, 먼저 1 단계는 합성데이터를 생성하는 단계이다. 이미지에서 인물을 추출하기 위해 COCO dataset(Chen et al., 2015)으로 사전 학습된 최신의 이미지 객체 분할 모델인 YOLOv8 을 사용하여 인물을 추출하였다. 그리고 배경과 폐색 물체를 합성하여 합성데이터를 생성한다.



<그림 1> 제안 방법론 구조

이후 2 단계 에서는 Skeleton 추출을 위해 2D Skeleton 추론 분야에서 우수한 성능을 보이는 HRNet(Sun, et al. 2019)을 COCO dataset(Chen et al., 2015)으로 사전 학습하여 2D Skeleton 을 추출하였다. 3 단계에서는

2 단계에서 얻어진 Skeleton 을 Skeleton-based action recognition 모델의 입력 값으로 사용한다. 이러한 접근 방식은 비디오에서 동작 정보만을 활용하여 폐색에 의한 영향을 최소화하며, 모델의 강건성을 높인다고 알려져 있다. (Varol et al., 2021; Duan et al., 2022) 생성된 합성 데이터는 그림 2 에서 확인할 수 있다.

<그림 2> 원본 및 합성 데이터

마지막 4 단계에선 실제 피트니스 시설에서 촬영한 데이터를 평가 데이터로 사용하여 ST-GCN(Yan et al., 2018), AA-GCN(Shi et al., 2019), PoseC3D(Duan et al., 2022)등의 다양한 Skeleton-based action recognition 모델의 예측 성능을 Top 1 정확도로 비교한다. 그 결과 표 1 에서 보이는 바와 같이 본 논문이 제시하는 방법을 적용하여 생성한 합성데이터와 원본데이터를 함께 사용했을 때 기존 데이터만을 사용했을 때보다 더 높은 정확도를 보였다. 이 결과는 본 연구가 제안하는 합성 데이터 활용 방법론이 실제 피트니스 시설에서 촬영된 데이터에 대한 운동 동작 오류 구분에 효과적임을 입증한다. 따라서, 본 연구는 더 정확한 운동 피드백을 제공하여 운동 효과를 극대화하고 부상 위험을 최소화하는데 기여할 것으로 기대된다.

<표 1> 성능 표

모델	데이터	Top1 정확도
ST-GCN	원본	0.285
	원본+합성	0.325
AA-GCN	원본	0.285
	원본+합성	0.32
PoseC3D	원본	0.36
	원본+합성	0.37

III. 참고문헌

GARBER, Carol Ewing, et al. (2011). Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise.

REINER, Miriam, et al. (2013). Long-term health benefits of physical activity–a systematic review of longitudinal studies. BMC public health, 13: 1-9.

ABOU ELMAGD, Mohammed. (2016). Benefits, need and importance of daily exercise. Int. J. Phys. Educ. Sports Health, 3.5: 22-27.

LORENZETTI, Silvio, et al. (2018). How to squat? Effects of various stance widths, foot placement angles and level of experience on knee, hip and trunk motion and loading. BMC Sports Science, Medicine and Rehabilitation, 10: 1-11.

RUNGSAWASDISAP, Nantana, et al. (2018). Squat movement recognition using hidden Markov models. In: 2018 International Workshop on Advanced Image Technology (IWAIT). IEEE. p. 1-4.

OGATA, Ryoji, et al. (2019). Temporal distance matrices for squat classification. Proceedings of the ieee/cvf conference on computer vision and pattern recognition workshops.

PARMAR, Paritosh, et al. (2022). Domain knowledge-informed self-supervised representations for workout form assessment. In: European Conference on Computer Vision. Cham: Springer Nature Switzerland. p. 105-123.

CHENG, Yu, et al. (2019) Occlusion-aware networks for 3d human pose estimation in video. Proceedings of the IEEE/CVF international conference on computer vision. p. 723-732.

CHEN, Xinlei, et al. (2015). Microsoft coco captions: Data collection and evaluation server. arXiv preprint arXiv:1504.00325.

SUN, Ke, et al. (2019). Deep high-resolution representation learning for human pose estimation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. p. 5693-5703.

VAROL, Gül, et al. (2021). Synthetic humans for action recognition from unseen viewpoints. International Journal of Computer Vision, 129.7: 2264-2287.

DUAN, Haodong, et al. (2022). Revisiting skeleton-based action recognition. Proceedings of the IEEE/CVF

conference on computer vision and pattern recognition. p. 2969-2978.

YAN, Sijie et al. (2018). Spatial temporal graph convolutional networks for skeleton-based action recognition. Proceedings of the AAAI conference on artificial intelligence.

SHI, Lei, et al. (2019). Two-stream adaptive graph convolutional networks for skeleton-based action recognition. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. p. 12026-12035.

Session F

F1. 이사회

여 백

Session F

F2. [특별세션]
Comparative Genomics 2.0 and AI (I)

좌 장: 유 혁(고려대)

여 백

Session F

F3. [특별세션]
Comparative Genomics 2.0
and AI (I)

좌 장: 정 대 균(경희대)

여 백

Session F

F4. [학술세션]
AI 기반 첨단 이미지 및 패턴 분석 기술

좌 장: 김 은 미(부산대)

넙치 이미지의 딥러닝 기반 질병 탐지 및 XAI 에 관한 연구

정다희
연세대학교 공과대학
kate1205@yonsei.ac.kr

손상호
연세대학교 공과대학
nikekobe9202@gmail.com

김우주
연세대학교 공과대학
wkim@yonsei.ac.kr

Abstract *– 환경 문제와 자원 감소 등의 변화로 인해 양식산업의 중요성이 증대되며, 국내 주요 식품자원인 수산물의 안정적인 공급과 기존 양식산업의 스마트화에 대한 필요성이 강조되고 있다. 국내에서는 스마트양식 기술과 모델 개발에 대한 연구가 활발히 이루어지고 있으며, 어류양식 기술의 첨단화를 위해 정부에서도 스마트양식 관련 투자를 진행 중이다.(이동길, 2023)*

저어류인 넙치와 같은 어종은 전 세계적으로 높은 양식 및 소비량을 자랑함에도, 이들의 낮은 활동성과 양식 환경 내에서의 밀착 경향 때문에 이상징후 탐지에 있어 기술적 도전이 크다.

본 논문은 스마트 양식 기술의 활성화를 위해 넙치 등 저어류 어종에 대한 이상 탐지 시스템을 제안하고자 한다. 객체의 질병 부위 탐지를 위해 DeSTSeg 모델을 적용하였다.(Zhang et al., 2023, p.3914-3923) 이는 정상 이미지와 자체적으로 Describable Textures Dataset 을 기반으로 한 합성 이상 이미지를 구축하여 학습함으로써 기존의 PatchCore 모델에 비해 성능이 우수하였다. (Roth et al., 2022, p.14318-14328) 더불어 오탐지 개체들애 대해 XAI 를 적용하여, 모델이 정상 혹은 이상 판단을 내릴 때 어떤 영역에 주목했는지 분석했다.(Zhang, 2021) 모델 학습을 위해 사용된 데이터는 지미션(AI 솔루션 기업)의 넙치 양식장의 수족관 촬영 이미지를 사용하였다. 본 논문에서 제안한 방법으로 기존에 제안된 방법론에 비해 양식 어류의 질병을 더 정확하게 탐지함으로써, 어류의 폐사 위험을 감소시키고 스마트 양식 기술을 한 단계 발전시킬 것으로 기대된다.

Keywords – Anomaly Detection, Deep Learning, Olive Flounder image, XAI

"이 논문은 국토교통부의 스마트시티 혁신인재육성사업으로 지원되었습니다."

III. 참고문헌

이동길, 정해승, 서준혁, 김형수, & 박정환. (2023). 기술· 경제· 환경적 측면에서의 스마트양식 산업화 모델과 기술개발 방향. *한국수산과학회지*, *56*(6), 759-765.

Zhang, Xuan, et al. (2023). DeSTSeg: Segmentation Guided Denoising Student-Teacher for Anomaly Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR) (pp. 3914-3923).

Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14318-14328).

Zhang, Q., Rao, L., & Yang, Y. (2021). Group-cam: Group score-weighted visual explanations for deep convolutional networks.*arXiv preprint arXiv:2103.13859.*

Li, X., Shang, M., Qin, H., & Chen, L. (2015, October). Fast accurate fish detection and recognition of underwater images with fast r-cnn. In *OCEANS 2015-MTS/IEEE Washington* (pp. 1-5). IEEE.

Kang, J. Y., Son, H. S., & Choi, H. S. (2022). A Comparison of Pre-Processing Techniques for Enhanced Identification of Paralichthys olivaceus Disease based on Deep Learning. *The Journal of the Korea Contents Association*,*22*(3), 71-80.

Iqbal, U., Li, D., & Akhter, M. (2022). Intelligent diagnosis of fish behavior using deep learning method. *Fishes*, 7(4), 201.

Local Sparse PCA 기반 공간정보 탐색을 통한 공간계량경제모델의 공간 패턴 인사이트 구축에 관한 연구

박건희
이화여자대학교 빅데이터분석학협동과정
califppy24@ ewhain.net

강윤철
이화여자대학교 경영대학 경영학과
yckang@ewha.ac.kr

Abstract*– *최근 공간데이터를 활용한 공간계량분석기법과 방법론에 대한 종합적이고 실용적인 연구는 계속해서 진행되고 있지만, 사상의 공간 분포와 편향된 대규모 공간적 배치에 내포된 진정한 의미를 이해하기 위한 공간분석방법론에 대한 연구는 아직 초기단계에 머무르고 있다. 본 연구에서는 AirBnB 데이터를 활용하여 매출 예측모델링 시 공간효과에 대한 탐색적 정보에 도움이 되는 해석적인 공간 패턴기법에 대해 연구한다. 이를 위해 Local Sparse PCA 기법을 활용하여 공간 특성 요인에 대해 기존 공간계량경제 모델만으로는 발견할 수 없었던 해석적인 사이트를 제시하고자 한다. 본 연구의 결과는 다양한 공간적 활용형태 및 분석목적에 따라 적절한 공간계량경제모형과 공간적 탐색 기법을 선택하는 데에 도움이 될 수 있다.

Key Terms – Spatial modeling, PCA, Sparse PCA, Local Sparse PCA, Airbnb

This work has supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No. NRF-2020R1A2C1101065).

I. 서론

주택가격이나 업종별 매출 등과 같이 어떠한 시상(phenomenon)의 공간분포에 따른 공간효과가 나타나는 공간데이터로 측정된 변수의 예측모형 구축 시, 일반적으로 사용되는 공간계량분석기법 및 GIS 기반 공간분석방법론은 꾸준히 연구되고 있다((김영표, 2004), (홍상기, 2005)). 또한 수요형태와 활용분야에 따라 데이터의 적절한 공간적 탐색 및 검정 방법에 대한 연구도 활발한 실정이다.

공간데이터에서 나타나는 공간효과는 크게 공간적 자기상관과 공간적 이질성으로 나뉘며, 이들에 대한 가설검정은 다양한 방식으로 이루어진다(변필성, 2007). 공간적 자기상관 가설검정은 Moran's I 와 같은 기술적 통계분석기법을 활용한 통계적 검정 기법과 OLS 회귀모형에서의 라그랑지 승수 검정(Lagrange Multiplier test)이 대표적이다(BALTAGI, Badi H., et al.,2003). 공간적 이질성을 검토하는 방법으로는 종속변수와 설명변수 간 비고정성(non-stationary)을 나타내는 Koenker 통계량의 유의성을 확인하는 방법이 있다(이창효, 2021).

본 연구에서 사용될 공간효과에 대처하기 위한 공간계량경제모델을 기술하면 다음과 같다. 첫째, 공간적 자기상관을 고려한 공간자기회귀모델에 속하는 1)공간시차모형(SLM: Spatial Lag Model), 2)공간오차모형(SEM: Spatail Error Model)이다. 둘째, 공간적 이질성을 고려한 모델로는 3)지리가중회귀모델(GWR: Geographically Weighted Regression model)을 활용하였다(구형모, 2019).

한편 위와 같은 공간계량경제모델을 활용한 회귀분석 진행 시, 유의미한 영향요인 선택 및 Spatial neighbor 을 정의하는 과정에서 발생하는 주관적 요인으로 인해 예측모델의 성능 측정을 넘어 공간데이터에 내포된 진정한 의미를 이해함으로써 공간적 분포를 명확히 평가하는 것은 항상 쉽지 않다(HONG, Seong-Yun, et al., 2022). 따라서 종속변수에 유의미한 요인 결정 및 다차원 공간데이터의 탐색을 통해 의미있는 공간적 패턴을 찾기 위한 주성분분석(PCA: Principal Component Analysis)이 공간분석 프로세스에서 활용되고 있다.

하지만, 기존의 주성분분석은 주성분(PC)간 직교성을 유지하면서 가능한 한 많은 분산을 설명하도록 수학적으로 결정되어(ZOU, Hui; HASTIE, et al., 2006), 각 PC 를 이루는 선형 조합은 일반적으로 0 이 아닌 loading 을 가진 모든 original 변수를 포함하므로 PC 의 정확한 의미를 해석하기

어렵다(ERICHSON, N. Benjamin, et al., 2020). 따라서 이를 보완하기 위해 희소 주성분분석(Sparse PCA)를 사용하는 것을 고려하는데, 본 연구에서는 희소성(sparsity)을 부과하기 위해 기존 PCA 의 PC loadings(β)에 대한 Optimization equation 에 lasso 와 ridge 가 결합된 elastic-net penalty 항을 추가한 <equation 1>이 사용된다.

$$\beta_j = \underset{\beta}{\operatorname{argmin}}(\alpha_j - \beta)^T X^T X(\alpha_j - \beta) + \lambda\|\beta\|^2 + \lambda_{1,j}\|\beta\|_1$$

<equation 1>

하지만 이러한 Sparse PCA 를 전역적으로 적용하면 차원 축소 과정 중에 유사한 값의 공간 클러스터링(LI, Pei Heng, et al., 2020)을 고려하지는 못하므로 데이터의 서로 다른 지리적 위치에 대해 지역적으로 Sparse PCA 를 적용하는 지역적 주성분분석(Local Sparse PCA) 또한 진행하여 대규모 공간데이터에 대한 공간 이질성 평가를 용이하게 하는 것을 제안하는 연구가 존재한다(HONG, Seong-Yun, et al., 2022).

본 연구에서는 플로리다 주에 위치한 에어비앤비를 대상으로 한 연매출 예측모형을 구축하고자 한다. 에어비앤비 연매출 역시 공간영향력에서 자유롭지 못하는 종속변수로서 위치의 고정성과 이질성 등과 같은 개별적 특성에 의해서 차이가 발생할 뿐만 아니라, 거시적 지역·공간적 특성요인에 의해서도 달라진다. 따라서 에어비앤비 연매출에 영향을 주는 요인을 파악하고 공간적 의미를 해석하기 위해서는 개별요인과 지역·공간적 요인의 복합적 분석을 통해 정확하고 체계적인 분석이 필요하다. 따라서 Local Sparse PCA 와 같은 차원축소 및 공간적 탐색과 함께 공간계량경제모델링의 비교분석을 진행한다.

II. 연구방법

1. Dataset

본 연구에서 사용하는 Florida Airbnb Dataset 은 2008 년~2019 년 플로리다에 존재하는 에어비앤비의 여러 특성요인에 대한 데이터로, 데이터 개수는 총 645,020 개이며 70 개의 특성요인으로 이루어져 있다. 이 중 차원축소에 활용되는 요인은 37 개의 수치형 변수이다. 같은 시간대에 존재하면서 공간적 효과를 주고받는 것에 대한 공간분석을 진행하기 위해 시간적 범위는 2019 년으로 제한하였으며, 이 때 데이터 개수는 240,829 개로 분석하기 충분한 양임을 확인하였다.

2. Framework

본 연구에서 비교분석에 쓰이는 모델은 OLS, SLM, SEM, GWR 총 4 가지이며, 분석 단계는 크게 주성분분석 단계, 시각화 단계, 회귀분석 단계 총 3 단계로 나뉜다.

첫 번째 주성분분석 단계에서는 일반 PCA, Sparse PCA, Local Sparse PCA 를 각각 진행한다. Local Sparse PCA 를 진행하기 위해서는 Local 단위와 공간가중행렬을 정의하는 것이 선행되어야 한다. Spatial neighbor 의 기준은 i=(u,v) 지점에서 100km 이내에 에어비앤비가 있으면 1, 없으면 0 으로 구성하기로 하였다. 본 연구에서 정의하는 Local 은 미국에서 사용하는 우편번호인 'Zipcode'를 사용한다. 따라서 데이터셋에 존재하는 Zipcode 고유 개수인 946 회만큼의 Sparse PCA 가 수행되며 이 때 Optimization eqaution 의 가중행렬 W 은 946*946 의 정방행렬이 된다. 본 연구에서 사용되는 Local Sparse PCA 에서의 PC loading 도출 시 적용되는 Optimization equation 식은 <equation 1>의 X tranpose 와 X 사이에 W 을 끼워넣은 식과 같다.

Local Sparse PCA 를 수행한 결과로 도출된 Local 각각에서의 PC1(첫번째 주성분)의 Sparse loading 들 중 non-zero loading 에 해당하는 특성요인의 집합을 조합 그룹으로 라벨링(group1, group2, …)하여 집계한다. 그 중 가장 많이 출현하는 조합 그룹 상위 10 개가 에어비앤비 간의 공간효과가 반영된 group1, group2, … , group10 이라는 새로운 Local Sparse PCA 기반의 특성요인으로 도출된다.

새로운 Local Sparse PCA 기반 특성요인들을 원데이터에 병합하여 OLS, SLM, SEM, GWR 을 수행하고 병합 전의 원데이터로만 수행하였던 OLS, SLM, SEM, GWR 결과와 비교분석한다. 성능 비교 시에 사용되는 측도는 예측 오차와 RMSE 이며, PCA, Sparse PCA, Local Sparse PCA 의 Zipcode 별 시각화를 통해 공간계량경제모델링 회귀분석 결과 해석 시 공간적 패턴에 대한 유의미한 인사이트를 확인한다.

III. 논의

본 연구를 통해 플로리다 에어비앤비 매출 예측 모델링 시 공간효과에 대한 탐색적 정보에 도움이 되는 의미 있는 공간 패턴을 나타내는 Local Sparse PCA 기반 새로운 특성요인을 도출함으로써 기존 공간계량경제모델링에서의 한계였던 해석력 측면에서의 인사이트를 발견할 수 있다.

이를 통해 다양한 공간적 활용형태 및 분석목적에 따라 적절한 공간계량경제모형과 공간적 탐색 기법을 선택하는 데 도움이 될 수 있다.

IV. 참고문헌

BALTAGI, Badi H.; SONG, Seuck Heun; KOH, Won. Testing panel data regression models with spatial error correlation. *Journal of econometrics*, 2003, 117.1: 123-150.

ERICHSON, N. Benjamin, et al. Sparse principal component analysis via variable projection. *SIAM Journal on Applied Mathematics*, 2020, 80.2: 977-1002.

HONG, Seong-Yun, et al. Local Sparse Principal Component Analysis for Exploring the Spatial Distribution of Social Infrastructure. *Land*, 2022, 11.11: 2034.

LI, Pei Heng; LEE, Taeho; YOUN, Hee Yong. Dimensionality reduction with sparse locality for principal component analysis. *Mathematical Problems in Engineering*, 2020, 2020: 1-12.

ZOU, Hui; HASTIE, Trevor; TIBSHIRANI, Robert. Sparse principal component analysis. *Journal of computational and graphical statistics*, 2006, 15.2: 265-286.

구형모. 모런 고유벡터 공간 필터링 기반 공간 가변 계수를 이용한 서울시 아파트 가격 결정요인의 공간적 이질성 탐색. *한국지리학회지*, 2019, 8.2: 321-335.

김영표; 임은선. GIS 기반 공간분석방법론 적용 연구. *안양: 국토연구원. Kim Yeongpyo and Lim Eunseon*, 2004

변필성. 공간계량경제모델링: 지리학의 제 1 법칙과 공간회귀모델. *국토*, 2007, 111-119.

이창효; 김가중. 공간효과를 고려한 지역 간 빈집 증감의 영향요인 차이. *Journal of Korea Planning Association Vol*, 2021, 56.3.

홍상기. GIS 기반 공간분석을 실제 사례에 적용하여 유용성 검증,[GIS 기반 공간분석방법론 적용 연구], Application of GIS-based Spatial Analysis Methodology, 김영표·임은선. *국토*, 2005, 148-149.

Character-Specific Image Generation through LoRA Adapters on Stable Diffusion Models

Dongyoung Kim
Finance AI Center, KB Kookmin Bank
dongyoung.kim@me.com

Doohyung Kim
Finance AI Center, KB Kookmin Bank
enguddkdlel@gmail.com

Changhyuk Jang
Finance AI Center, KB Kookmin Bank
chjang@kbfg.com

Sungjun Kwon
Finance AI Center, KB Kookmin Bank
sjkwon@kbfg.com

Hosung Park
Finance AI Center, KB Kookmin Bank
hosungpark@kbfg.com

Hyunsoo Son
Finance AI Center, KB Kookmin Bank
sonhyunsoo@kbfg.com

Sungwoo Oh
Finance AI Center, KB Kookmin Bank
swoh@kbfg.com

Dongwon Choi
Finance AI Center, KB Kookmin Bank
dwchoi@kbfg.com

Abstract – In the realm of digital art and character design, achieving consistent and high-fidelity generation of specific characters using AI-driven models presents a significant challenge. This paper introduces a novel approach that leverages Low-Rank Adaption (LoRA) in conjunction with stable diffusion models to generate images of specific characters with remarkable consistency and detail. Our methodology begins with the collection of approximately 20 images per target character, followed by the generation and refinement of character descriptions using a multimodal Large Language Model (LLM). This process is enhanced through in-context learning, ensuring descriptions are both accurate and consistent across iterations. Subsequently, LoRA adapters are trained on stable diffusion models, incorporating these descriptions to guide the generation process. Through this innovative approach, we successfully create ten-character models, demonstrating our method's effectiveness across different artistic dimensions. Our results reveal that the use of LoRA adapters significantly improves the consistency and fidelity of character image generation, outperforming existing baseline methods. This research not only advances the capabilities of AI in artistic endeavors but also opens new avenues for applications in gaming, virtual reality, and brand representation, where character consistency is paramount. By bridging the gap between AI and creative design, our work paves the way for future explorations in automated character generation and the broader implications of AI in art and design.

Keywords –Generative AI, Stable diffusion, Multimodal large language model, PEFT, LoRA

This work is supported by KB Kookmin Bank.

I. Introduction

Recent advancements in AI and deep learning for computer vision have revolutionized various fields, such as optical character recognition (Kim et al., 2023) and image generation (Rombach et al., 2021). Stable diffusion models have been at the forefront of these innovations, enabling the creation of high-quality digital artwork through a diffusion process that incrementally refines images to achieve unprecedented levels of detail (Saharia et al., 2022; Brooks et al., 2022; Luo, 2022). However, despite their capabilities, these models often struggle with maintaining consistency and specificity when generating character images, which is essential for applications involving recurring characters or branding.

Low-Rank Adaption (LoRA) introduces an innovative solution to this problem by enabling precise and efficient fine-tuning of pre-trained models (Hu et al., 2021). These adapters adjust model parameters in a low-rank space, allowing for significant modifications to the generated images with minimal computational resources. When applied to stable diffusion models, LoRA adapters can enhance the model's ability to consistently generate

specific character traits across different instances, thus overcoming one of the significant limitations of traditional generative models (Gandikota et al., 2023).

This paper explores the synergy between LoRA adapters and stable diffusion models, demonstrating through experimental results that this combination markedly improves the accuracy and consistency of character-specific image generation. By leveraging the strengths of both technologies, we present a methodological advancement that not only enhances the capabilities of AI in artistic endeavors but also sets a new standard for the fidelity and consistency of digital character design.

II. Methodology

In the data preparation phase of our study on character image generation, we meticulously curated a dataset to train LoRA adapters on stable diffusion models. This process involved several carefully structured steps to ensure the generation of high-fidelity and consistent character images. For each target character, identified by a set within our dataset $C = c_1, c_2, \ldots, c_n$, we gathered an initial set of approximately $N \approx 20$ images. These images, denoted as $I_{c_i} = i_1, i_2, \ldots, i_N$ for each character c_i, were selected to span a broad spectrum of poses, expressions, and contexts, aiming to encompass the wide range of attributes and scenarios characteristic of each target, ensuring a comprehensive representation.

Upon gathering the images, the next step involved compiling key features and attributes from the collected image sets into structured documents. For each target character c_i, this compilation yielded a preliminary document $d_{c_i}^0$. These documents served as inputs to a multimodal Large Language Model (LLM, Yang et al., 2023; Alayrac et al., 2022), which processed them to generate preliminary textual descriptions $d_{c_i}^1$. These descriptions aimed to encapsulate the essence of the characters as portrayed in the images, translating visual information into textual format. Recognizing the potential for initial descriptions to miss nuances or contain inaccuracies, we employed an iterative refinement process through in-context learning. Starting with the initial description $d_{c_i}^1$, we iteratively fed each description back into the LLM. At each iteration k, the model was provided with additional context or corrections to refine the description, producing an enhanced version $d_{c_i}^{k+1}$.

This process continued until the refined description achieved a high level of consistency, detail, and completeness, at which point it was deemed final ($d_{c_i}^{\text{final}}$) and added to the dataset of refined descriptions D_{refined}. The culmination of this meticulous data preparation process was the compilation and formatting of the refined descriptions, $D_{\text{refined}} = d_1, d_2, \ldots, d_n$. These descriptions were specifically tailored for subsequent use in training LoRA adapters, facilitating the generation of character images with enhanced fidelity and consistency through stable diffusion models. This systematic approach to data preparation not only ensured the quality and consistency of the input data but also laid a robust foundation for the successful application of advanced machine learning techniques in character image generation.

III. Result

Our innovative approach to character-specific image generation using LoRA adapters on stable diffusion models has demonstrated remarkable success, as illustrated in Figure 1. We created detailed and consistent images for ten distinct characters in both 2D and 3D shapes. These images, generated by the stable diffusion models enhanced with LoRA adapters, exhibited exceptional fidelity and consistency. The use of LoRA adapters enabled precise control over the image generation process, allowing for fine-tuning without the need for extensive retraining. This tailored adaptation was critical in maintaining character consistency across various poses and expressions, a notable improvement over traditional methods which often struggle with such variations. Quantitative assessments showed that our images surpassed baseline models in both visual quality and consistency, confirming the efficacy of our approach.

<FIGURE 1. CHARACTER IMAGES GENERATED BY STABLE DIFFUSION MODEL WITH LORA ADAPTERS TRAINED USING THE METHODOLOGY DESCRIBED IN THIS STUDY>

IV. Discussions

The integration of LoRA adapters with stable diffusion models represents a significant advancement in the field of AI-driven digital art and character design, as evidenced by the generated images shown in Figure 1. This combination enhances the quality and consistency of character images, crucial for applications in gaming, virtual reality, and brand representation. Our methodology, leveraging a multimodal Large Language Model for generating and refining character descriptions, ensures that the generated images faithfully represent the intended attributes of each character.

Despite the promising results, there are areas for future research, such as exploring the integration of additional modalities in the training process to enrich the model's understanding of complex character traits and emotions. Another potential area is the real-time animation of these characters for interactive media (Liu et al., 2024). Moreover, continuous improvements in model training techniques and dataset expansion could further enhance the accuracy and efficiency of character-specific image generation.

In conclusion, our study sets a new standard for digital character design, demonstrating the potential of this technology to revolutionize fields that rely on detailed and consistent character representations. This opens up new avenues for both practical applications and creative exploration in the digital arts.

V. References

Kim, D., Kim, D., Kwak, M., Son, H., Sohn, D., Lim, M., … Choi, D. (2023). Deep Learning OCR based document processing platform and its application in financial domain. Journal of Intelligence and Information Systems, 29(1), 143–174. https://doi.org/10.13088/JIIS.2023.29.1.143

Gandikota, R., Materzynska, J., Zhou, T., Torralba, A., & Bau, D. (2023, November). Concept Sliders: LoRA Adaptors for Precise Control in Diffusion Models. arXiv. https://arxiv.org/abs/2311.12092

Brooks, T., Holynski, A., & Efros, A. A. (2022, November). InstructPix2Pix: Learning to Follow Image Editing Instructions. arXiv. https://arxiv.org/abs/2211.09800

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., ... Norouzi, M. (2022, May). Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. arXiv. https://arxiv.org/abs/2205.11487

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., ... Chen, W. (2021, June). LoRA: Low-Rank Adaptation of Large Language Models. arXiv. https://arxiv.org/abs/2106.09685

Luo, C. (2022, August). Understanding Diffusion Models: A Unified Perspective. arXiv. https://arxiv.org/abs/2208.11970\

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021, December). High-Resolution Image Synthesis with Latent Diffusion Models. arXiv. https://arxiv.org/abs/2112.10752

Alayrac, J.-B., Donahue, J., Luc, P., Miech, A., Barr, I., ... Simonyan, K. (2022, April). Flamingo: a Visual Language Model for Few-Shot Learning. arXiv. https://arxiv.org/abs/2204.14198

Yang, Z., Li, L., Lin, K., Wang, J., Lin, C.-C., ... Wang, L. (2023, September). The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision). arXiv. https://arxiv.org/abs/2309.17421

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., ... Amodei, D. (2020, May). Language Models are Few-Shot Learners. arXiv. https://arxiv.org/abs/2005.14165

Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., Yuan, Z., Huang, Y., Sun, H., Gao, J., He, L., & Sun, L. (2024). Sora: A Review on Background, Technology, Limitations, and Opportunities of Large Vision Models. arXiv. https://arxiv.org/abs/2402.17177

멀티스케일 특성을 활용한 비전 트랜스포머 기반 딥페이크 탐지에 관한 연구

정해민
한국교통대학교 산업경영공학과
hmjung@ut.ac.kr

조혁주
연세대학교 산업공학과
huckjoo1@gmail.com

김우주
연세대학교 산업공학과
wkim@yonsei.ac.kr

이광연
숭실대학교 전자정보공학부
kylee@ssu.ac.kr

Abstract *– 딥페이크는 이미지나 영상에서 특정 사람의 얼굴을 다른 사람으로 대체하는 딥 러닝 기술, 또는 이 기술을 이용해 생성한 가짜 이미지나 영상을 지칭한다. 딥 러닝 기술이 널리 보급되면서 딥페이크 기술에 대한 접근성이 높아졌고, 결과적으로 이를 악용한 범죄도 증가하고 있다. 이에 따라 효과적인 딥페이크 탐지 기술의 필요성이 점점 더 커지고 있다. 딥페이크 생성은 주로 신원 교체와 표정 재연이라는 두 가지 방식으로 이루어지는데, 기존의 탐지 기술은 딥페이크가 어떤 방식으로 생성되었는지에 따라 탐지 성능의 편차를 보인다. 본 연구에서는 딥페이크 탐지 모델의 성능 편차를 줄임으로써 기존 방법론들의 한계를 보완할 수 있는 연구를 제안하고자 하였다. 제안하는 모델은 먼저 영상을 프레임 단위의 이미지들로 자른 다음, 딥페이크의 주된 대상 영역인 얼굴 부분과, 일종의 지역 정보라고 할 수 있는 입 부분을 각각 추출하여 멀티스케일 특성으로 활용한다. 각 특성을 서로 다른 비전 트랜스포머 구조에 입력한 다음, 출력되는 예측 결과들을 종합하여 동영상이 딥페이크인지 아닌지를 효과적으로 판단하게 된다. 특히, 얼굴 부분은 신원 교체 방식으로 생성된 딥페이크를 대응하는 데 도움이 되고, 입 부분은 표정 재연 방식의 딥페이크를 대응하는 데 도움이 되기 때문에 모델은 서로 다른 딥페이크 생성 방식에 대한 강건성을 갖게 된다. 제안하는 방법론을 두 개의 데이터셋에 대해 실험한 결과, 상대적으로 높은 탐지 성능과 함께 다양한 딥페이크 생성 방식에 대해 보다 범용적으로 대응할 수 있는 가능성을 확인하였다.*

Key Terms *– 딥페이크 탐지, 멀티스케일 특성 분석, 비전 트랜스포머, 합성곱 신경망*

Acknowledgement (사사)
2024 년 국립한국교통대학교 지원을 받아 수행하였음

근육병 환자를 위한 단일 카메라 기반 시선 추적 연구

김시원
아주대학교
소프트웨어학과
kimsiw42@ajou.ac.kr

이대호
아주대학교
인공지능융합학과
daeho5000@ajou.ac.kr

곽세현
아주대학교
인공지능융합학과
boring1230@ajou.ac.kr

박정훈
아주대학교
인공지능융합학과
stevejobs@ajou.ac.kr

***Abstract** – 기술의 발전은 인류가 직면한 난제들을 해결하는 데 중요한 역할을 하고 있다. 특히, 의료 분야에서의 기술 혁신은 우리의 삶을 크게 개선할 수 있는 가능성을 제시한다. 이 연구는 루게릭병, 파킨슨병 등과 같은 근육병 환자들의 의사소통에 대한 효과적인 방법을 모색하고자 한다. 기존의 시선 추적 기술은 고가의 웨어러블 장비나 높은 성능의 컴퓨팅 자원이 필요했다. 그러나 본 연구에서는 단일 카메라만을 이용하여 홍채를 촬영하고 홍채 타원의 모양 변화를 분석하여 시선의 변화를 감지했다. 이를 통해 단일 카메라를 제외한 별도의 추가 장비 없이 낮은 성능의 컴퓨팅 자원에서도 비교적 정확한 시선 추적 알고리즘을 제안한다. 이러한 기술은 환자들이 편리하게 이용할 수 있도록 도와줄 뿐만 아니라, 의사소통과 일상생활에 대한 독립성을 향상할 수 있다. 또한, 머리의 위치나 방향 변화를 보정하는 추가 알고리즘을 통해 머리가 계속 흔들리는 환경에서도 적용할 수 있었다. 이러한 방법론은 근육병 환자만이 아니라, 일반인에게도 적용 가능한 기술적 해결책을 제공함으로써 의료 분야에 도움이 될 것으로 기대한다.*

***Keywords** – 루게릭병, 시선추적, 실시간처리, 영상처리, Calibration*

I. 서론

근육병은 신체의 근육을 쇠약 시키는 질병으로, 일상생활에서의 기능과 의사소통에 큰 어려움을 겪는 환자들이 많다. 독일에서 진행된 연구에 따르면 루게릭병 환자의 46%가 의사소통을 위한 기기를 필요로 하는 것으로 나타났다(A. Funke et al., 2018). 따라서, 노인 또는 몸이 불편한 환자를 대상으로 원활한 의사소통을 돕는 연구가 많이 진행되고 있다. 그 중, 루게릭병 환자를 대상으로는 눈을 통해 의사소통을 돕는 연구가 활발히 이루어지고 있다(Spataro et al., 2018). 그들이 마지막까지 사용할 수 있는 유일한 근육이 안구 근육이기 때문이다. 다만, 지금까지 연구되어 온 방법은 별도의 웨어러블 장비 또는 장치를 필요로 하거나, 높은 가격의 시스템을 요구한다는 한계점이 있다.

이에 본 연구에서는 단일 웹 카메라만을 사용하여 저 비용 시스템에서도 효율적으로 실시간 시선 추적이 가능한 Gaze estimation 방법론을 제시한다. 우리의 주요 기여를 다음과 같이 요약한다:

Contributions: (1) 기존의 동공 추적 기반의 영상처리 알고리즘을 사용한 Gaze estimation 연구를(Wan et al., 2022) 흑색 홍채를 보유한 사람에게 쉽게 적용할 수 없었던 한계점을 개선한 홍채 타원 기반의 시선 추적 방법론을 제시한다. **(2)** 실시간 처리가 어려운 한계점을 지닌 딥 러닝 기반의 시선 추적 방식(Jalilian et al., 2021)과 별도의 웨어러블 장비 및 추가 장치를 요구하는 Corneal Reflection 방식(Park et al., 2015)의 한계점을 상호보완하는 성능의 시선 추적 방식을 제안한다. **(3)** 더욱 효율적인 실시간 처리를 위해 Head position 을 보정하는 알고리즘을 adaptive 하게 적용하는 방법론을 제시한다.

II. 관련 연구

Electrooculography: 고전적인 시선 추적 연구는 Electrooculography(EOG)라는 기계적 기록 방식으로부터 시작되었다. 각막과 망막은 양전하와 음전하의 서로 다른 전하를 가지고 있다. 이로 인해 전위차가 발생하는데, 안구가 움직일 때 전위의 차이가 바뀌게 된다. 이에 따라 수직, 수평 움직임을 기록하는 연구가 진행되었다(Ford et al., 1959). 하지만, 높은 정확성을 가져오기 어려웠다.

Requiring Additional Equipment based Gaze Estimation: 고전적인 방법으로는 정확도 문제를 해결할 수 없었기에 머리에 안경과 같은 장비를

<그림 1> 데이터 수집을 진행한 실험 환경

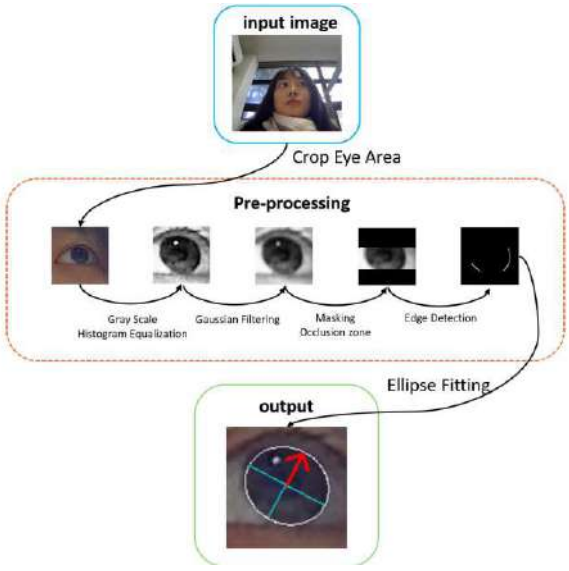
장착하는 Head Mount System 기반의 시선 추적 연구가 진행되거나(Cheng et al., 2021), 외부에 IR LED 와 적외선 카메라를 두어 시선추적을 하는 방식이 많이 연구되었다(Li et al., 2021). 그러나 이러한 방법론은 추가적인 장비 구매를 위한 금전적 부담이 따르며, 추가 장비를 부착한다는 점이 육체적 피로로 작용한다는 한계점이 있다.

Deep Learning based Gaze Estimation: 추가적인 장비를 요구한다는 한계점을 보완하기 위해 회귀 모델과 CNN 모델을 사용해 정확도를 높이는 연구가 많이 진행되었다(Jalilian et al., 2021). 그러나 이러한 방법론은 딥 러닝 모델의 실행을 위해 높은 시스템 자원이 필요하며 저비용 시스템에서는 실시간 추적을 어렵게 한다는 한계점이 있다.

III. 연구방법

우리는 앞선 관련 연구의 한계점을 개선하기 위해 딥 러닝보다 빠른 영상처리 알고리즘을 사용해 높은 정확도와 빠른 추론 시간의 목표를 동시에 달성하고자 했다. 이러한 목표를 달성하기 위해 동공을 기반으로 한 연구가 진행되었지만, 일반적인 웹 카메라에서 홍채와 동공의 구분이 어려운 흑색 홍채를 가진 사람에게 이를 적용할 수 없었다. 따라서 우리는 홍채 타원 모양을 기반으로 시선을 추적하는 연구를 진행하였다.

먼저 입력된 영상에서 눈 영역을 찾아야 한다. 머신러닝 방법론으로는 OpenCV 의 dlib 이나 Google 의 Mediapipe 를 많이 사용한다. 그러나 우리의 환경에서 더욱 정확한 검출을 위해 정확도와 실시간 처리 면에서 모두 우수한 YOLO 모델을 사용하여 우리의 데이터에 맞춰 학습했다. 눈 검출 이후에는 Gray Scale 로 변환하여 Histogram Equalization 을 적용한다. 이를 통해 어두운 영역과



<그림 2> 제안하는 시선 추적 방법의 순서도

밝은 영역의 대비를 더욱 강조할 수 있으며, 홍채 경계선 부분을 추출하는 것에 도움을 준다. 이후 Gaussian Filtering 을 적용해 노이즈를 제거하여 노이즈가 경계선으로 추출되는 것을 방지한다. 이후 눈꺼풀 영역에서 생기는 오인식 영역을 Masking 하는 작업을 거친 후 최종적으로 Canny Edge Detector 를 사용하여 홍채의 경계를 추출한다. 추출된 경계를 통해 타원 모양을 예측하고 해당 홍채 타원의 장축과 단축을 기준으로 시선 벡터의 방향과 크기를 결정한다.

머리의 방향을 고려하지 않은 시선 벡터가 일차적으로 계산되었지만, 머리의 흔들림에도 강인한 시선 추정을 위해 Head Pose Estimation 을 적용해 이를 보완하였다. 이를 위해 얼굴의 움직임을 계산할 때 요(yaw), 피치(pitch), 롤(roll) 각도를 통해 현재 얼굴의 각도와 방향을 고려하여 Calibration 과정을 진행하였다. 영상의 매 프레임마다 이를 적용하면 더욱 높은 정확도를 얻을 수 있겠지만, 이는 많은 계산을 요구하므로 처리 속도가 느려지게 된다. 이러한 방법은 저비용 시스템에 적용하기에는 부담스러운 작업이다. 이를 보완하기 위해 Adaptive 하게 Head Position 을 갱신하는 방법론을 추가로 적용했다:

$$MSE_{hedapose} = \left(yaw_{after} - yaw_{before}\right)^2 + \left(roll_{after} - roll_{before}\right)^2 + \left(pitch_{after} - pitch_{before}\right)^2$$

$$timestamp_{update} = (1 + a) * timestampe_{before} + (-a) * MSE_{headpose}$$

과거 시점에 계산되었던 yaw, pitch, roll 값을 현재 시점의 값과 MSE(Mean Squared Error)값으로 계산한다. 그렇게 구해진 MSE 값과 timestamp 를 계산하는 주기를 업데이트하기 위한 파라미터 α를 추가하여 MSE 값이 커질수록 더 자주 head position 을 업데이트하며, MSE 값이 작을수록 time stamp 를 점점 늘려서 head position 을 업데이트하기 위한 계산 량을 더욱 줄일 수 있다.

이렇게 계산된 얼굴 방향의 벡터를 사용하여 얼굴이 바라보는 방향의 Plane 을 정의할 수 있다. 이를 사용하는 기기의 화면과 평행한 위치가 되도록 시선 벡터를 보정하여 얼굴 방향에 의한 시선 벡터의 왜곡을 보정하였다.

이를 통해 각 Calibration Point 에 대한 시선 벡터를 계산하였다. 다만, 안구는 구의 형태로 이루어져 있어 움직이는 거리가 선형적이지 않고, 경로가 사람마다 다르다는 문제점이 있다. 이에 본 연구에서는 개인의 안구 움직임이 어느 정도의 가중치를 가지고 이동하는지 계산하는 파라미터를 추가한 interpolation 알고리즘을 추가로 적용하였다.

IV. 결론

본 연구에서는 신체가 불편한 환자를 위해 적용할 수 있는 단일 카메라 기반의 시선 추적 방법을 제시하였다. 기존의 시스템들과 대비하여 웨어러블 장비를 사용하지 않으면서도 준수한 정확도를 가지며 기존 방법들의 한계점을 상호보완 함을 확인하였다. 또한, 머리의 흔들림에도 강건한 시선 추적을 위한 알고리즘을 적용하면서도 빠른 추론을 위해 adaptive 하게 적용하는 식을 제안하였으며, 홍채 타원 모양을 기반으로 하는 시선추적 방법 또한 제안하였다. 우리가 제안하는 시선추적 방법론을 통해 루게릭병, 근육병 등의 환자들이 시선 추적 기술에 가볍게 접근할 수 있으며, 더욱 손쉽게 의사소통할 수 있을 것으로 기대한다.

V. 참고문헌

Cheng, Y., Wang, H., Bao, Y., & Lu, F. (2021). Appearance-based Gaze Estimation With Deep Learning: A Review and Benchmark. doi.org/10.48550/arXiv.2104.12668

Ford, A., White, C. T., & Lichtenstein, M. (1959). Analysis of Eye Movements during Free Search. Journal of the Optical Society of America, 49(3), 287. doi.org/10.1364/JOSA.49.000287

Funke, A. et al. (2018). Provision of assistive technology devices among people with ALS in Germany: a platform-case management approach. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 19(5-6), 342-350. doi.org/10.1080/21678421.2018.1431786

Jalilian, E., Karakaya, M., & Uhl, A. (2021). CNN based off angle iris segmentation and recognition. IET Biometrics, 10(5), 518-535. doi.org/10.1049/bme2.12052

Li, Z., Guo, P., & Song, C. (2021). A Review of Main Eye Movement Tracking Methods. Journal of Physics: Conference Series. 1802(4), 42-66. doi.org/10.1088/1742-6596/1802/4/042066

Park, J., Jung, T., & Yim, K. (2015). Implementation of an Eye Gaze Tracking System for the Disabled People. 2015 IEEE 29th International Conference on Advanced Information Networking and Applications. doi.org/10.1109/AINA.2015.286

Spataro, R., Ciriacono, M., Manno, C., & La Bella, C. (2013). The eye-tracking computer device for communication in amyotrophic lateral sclerosis. Acta Neurologica Scandinavica. 130(1), 40-45. doi.org/10.1111/ane.12214

Wan, Z., Xiong, C., Chen, W., Zhang, H. & Wu, S. (2022). Pupil-Contour-Based Gaze Estimation With Real Pupil Axes for Head-Mounted Eye Tracking. 18(6), 3640-3650. doi.org/10.1109/TII.2021.3118022.