

Song Lee

Data Analysis Portfolio

Name: 이송

Phone: 010-6407-5642

Email: slee7@kaist.ac.kr



성장하는 데이터 분석가, 이송입니다

데이터 기반 의사결정을 도와
서비스 개선과 혁신에 기여하고 싶습니다

Summary

- 석사 졸업논문 주제로 고객의 모바일 앱 행동 데이터를 분석했어요
- SQL을 활용하여 데이터 분석 역량을 키웠어요
- 라크로스팀 활동을 통해 커뮤니케이션 역량을 보유하고 있어요
- 금융권 인턴을 하며 데이터 기반의 개선 방안을 도출하고 시각화를 경험했어요

Graduation

- 2024 KAIST 경영공학부 IT경영 트랙 석사
- 2022 이화여대 국제학, 소프트웨어학 전공
 - 통계학, 경영학 부전공
 - 2018-2019 독일 마부르크대 교환학생
- 2016 한영외고 스페인어과

Internship Experience

- 2022 메리츠화재 데이터 사이언스
- 2021 한국전자통신연구원 AI 리서치, 데이터 엔지니어링
- 2019 SAP 독일 본사 IT 컨설팅, test automation
- 2018 망고슬래브 디지털 마케팅

Awards

- 2022 SC제일은행 Women in Fintech 3등
- 2022 딥러닝 기반 논문 장려상 (한국관광학회 대학원생 세션)
- 2021 AI 핀테크 스타트업 지원 선정 (5,300만원 지원금)
- 2018 UNWTO 세계 청년 도시관광총회 대상

Skills

- 상: Python, SQL, MySQL, OracleSQL, R
- 중: Tableau, Tensorflow, Pytorch, Sklearn, AWS
- 하: Spark, MongoDB
- 기타 JS, React.js, Node.js, Flask, Django, Google Analytics, Figma, Huggingface, Worksoft Certify

Additional Info

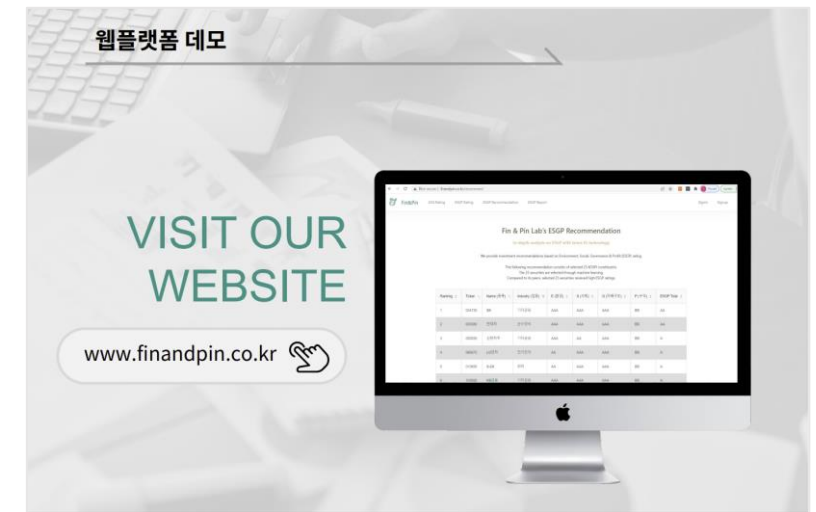
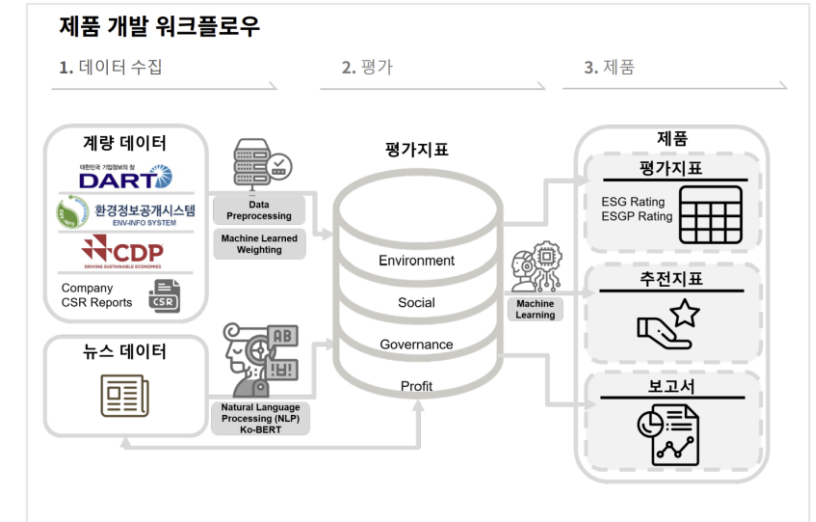
- Notion Portfolio: <https://zrr.kr/00QT>
- Linkedin: <https://www.linkedin.com/in/song-lee->

Table of Contents

- 대표 프로젝트 #1 요약: AI 스타트업 창업 및 ESG 추천지수 개발
- 대표 프로젝트 #2 요약: 메리츠화재 인턴 프로젝트
- 대표 프로젝트 #3 요약: SQL 추천시스템 개발
- 석사학위 논문 요약: 오픈뱅킹 효과 분석

Eco Reco: NLP 딥러닝 모델 기반 ESG 추천지수 서비스 개발

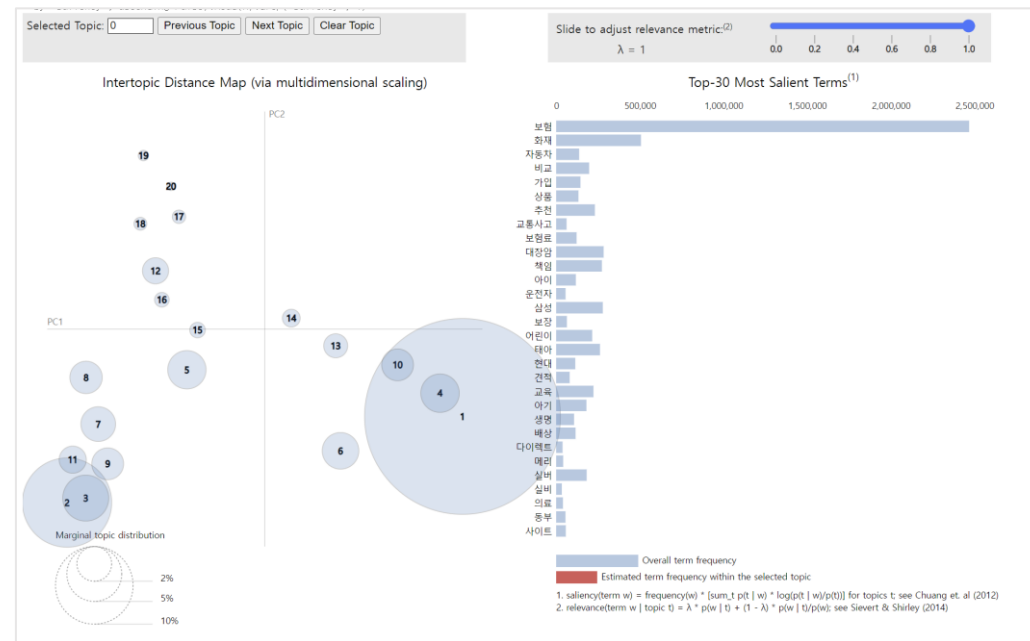
- 요약: Fin and Pin 스타트업을 창업하여 Eco Reco라는 투명하고 자동화된 NLP 모델 기반 ESG 추천지수 서비스를 개발
 - 목적: ESG 평가방식 기준이 공개되지 않고, 지표가 3,000개 존재하여 등급을 쉽게 파악하기 어려움. 투명하고 대중적인 데이터 기반 ESG 지표를 구축하고자 함
 - 수행기관: 정부 창업지원 (예비창업패키지 5,3000만원 지원)
 - 도구: Python, tensorflow, pandas, konlpy, sklearn, selenium, bs4, MongoDB, AWS Open Search, AWS EC2
 - 데이터: 국내 ESG 뉴스 데이터 크롤링
 - 방법: KoBERT 딥러닝 NLP 모델로 fine-tuning 하여 ESG 유형을 분류하고 기업의 ESG 성과를 평가. 상위에 위치하는 종목을 추천
 - 결과: 크롤링과 딥러닝 기반으로 최신 정보를 수시로 반영하는 “ESG 지표의 자동화와 투명성”을 이룰 수 있었음
 - 창업지원 최종 심사에서 우수 등급을 획득
- 비정형 데이터 기반 AI 모델 개발 역량 강화, 데이터 파이프라인 및 분석 (AI, 클라우드, DB) 이해 증진, 팀 관리와 협업 능력 향상



Pretrained LLM과 텍스트마이닝으로 메리츠화재 리뷰 데이터 분석

- 요약: 고객 인식과 청구 예측을 위해 네이버 카페 데이터를 분석하고 기업 전략 제시
- 목적: 보험 상품 개발과 고객 서비스 향상을 위해 여론을 파악하기 함. 온라인 데이터를 분석하여 자사 적용 인사이트 제시
- 수행 기관: 메리츠화재화상보험
- 도구: SQL, PySpark, python, tensorflow, huggingface, konlpy, sklearn, selenium, bs4
- 데이터: 네이버 카페 API, 다이렉트 보험 사이트 리뷰 크롤링
- 방법: Huggingface API로 의도분석과 감성분석을 수행
 - 텍스트마이닝 (n-gram, TF-IDF, LDA)으로 청구 예측과 경쟁사 분석을 수행
- 결과
 - (1) 건강 관련 보험에 긍정적이며, 자동차 보험에 대해 부정적인 경향
 - 보험료보다 보험 정보에 대한 접근성에 대한 관심을 확인
 - (2) 고객이 관심있는 질병과 청구하는 질병은 대체로 일치
 - 어린이보험, 암보험은 온라인 언급빈도가 실제 청구횟수보다 높음
 - (3) 각 보험상품마다 경쟁사가 달라 맞춤형 전략이 필요

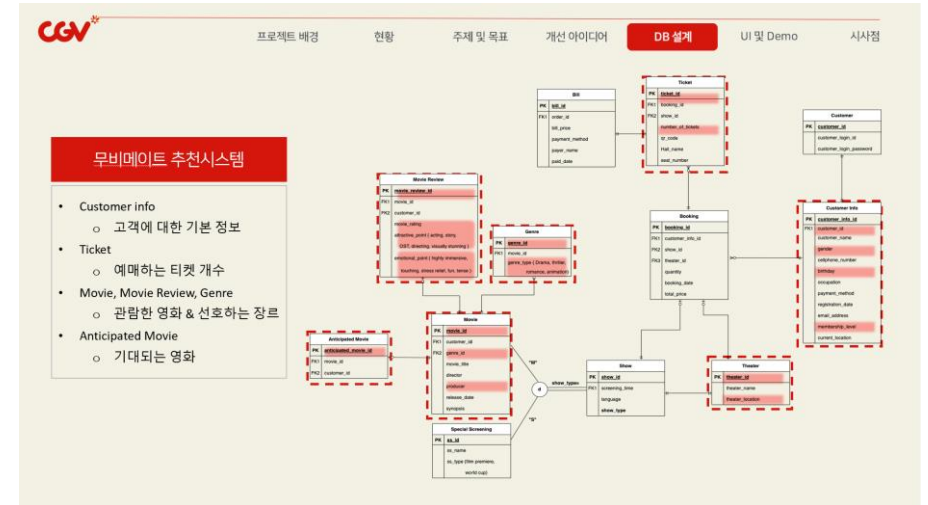
→ End to End로 데이터 분석 프로젝트를 주도적으로 수행,
기업 전략과 목표 설정에 기여



SQL 추천시스템 개발을 통한 영화관 앱 사용성 개선

- 요약: "나만의 무비메이트 by CGV"라는 취향을 기반으로 하는 개인화된 영화 추천 서비스를 제시
- 목적: CGV는 영화관 점유율 1위이나, 모바일 앱 마켓의 평가에서 규모면이나 경쟁사 대비 앱 사용성이 떨어진다는 문제 발견.
개인화 추천 서비스로 앱 사용성 개선과 MAU 증대를 목표
- 수행 기관: KAIST 데이터베이스 관리와 빅데이터 분석
- 도구: SQL, MySQL, Figma
- 방법: 고객이 인식하지 못한 니즈를 중심으로 초개인화된 추천 시스템을 기획
 - SQL을 활용하여 무비메이트 추천시스템의 DB를 설계
 - 개인화된 메인 홈페이지 UI 데모를 제안
- 결과: 취향 기반 추천시스템과 개인화를 통해 앱 사용성을 높였음

→ 현업의 문제점을 파악하여 추천시스템 DB 설계로
사용자 중심의 서비스를 기획하는 능력 증진



오픈뱅킹 시행에 따른 은행 경쟁 변화와 모바일 금융 산업 스피로버 효과

- 요약: 3,500명의 모바일 앱 데이터를 분석하여 오픈뱅킹 도입 후 모바일 은행 산업, 은행 간 경쟁, 타 금융업에 미치는 영향을 파악
- 목적: 오픈뱅킹과 같은 핀테크 혁신이 금융 시장에 미치는 변화를 이해하고, 은행과 금융 업계가 어떻게 대응해야 할지 파악하고자 함
- 도구: Python, pandas, numpy, STATA
- 데이터: 3,500명의 닐슨 모바일 앱 사용 패널 데이터
- 방법: 계량경제 모델 (Difference-in-differences) 활용
 - Feature engineering으로 다중은행 사용, 주거래은행 변화 등 새로운 변수를 도출하여 고객의 행동 포착
- 결과: 모바일 은행과 타 금융 시장의 확대, 은행 간 경쟁 심화, 인터넷은행에게 더 유리함을 발견. 은행, 금융업, 정부 관계자들에게 오픈뱅킹 전략 수립에 있어 인사이트 제공

→ 모바일 앱 사용 패널 데이터를 가공하는 능력과 핀테크 도메인 지식을 습득



감사합니다.

Contact

slee7@kaist.ac.kr

010-6407-5642

<https://www.linkedin.com/in/song-lee->