

통계데이터사이언스학과전공! 전공+경험+실무의 스페셜리스트!

손지우 1996년 (27세/만 26세) | 남

✉ bungaedm@naver.com ☎ 010-9120-6675 📞 -

🏠 (03699) 서울 서대문구 연희동

학력사항	경력사항	희망연봉	희망근무지/근무형태	포트폴리오
대학원(석사) 재학중	신입	회사내규에 따름	서울전체 인턴직	 Portfolio_손지우.pptx

학력 최종학력 | 대학원 석사 재학중

재학기간	구분	학교명(소재지)	전공	학점
2021.09 ~ 현재	재학중	연세대학교(서울) 대학원(석사) (서울)	통계데이터사이언스학과	4.3 / 4.3
2015.03 ~ 2021.08	졸업	연세대학교(서울) (서울)	심리학과 응용통계학과 (복수전공)	3.91 / 4.3
2012.03 ~ 2015.02	졸업	하나고등학교		-

경력 신입

대외활동

기간	구분	기관/장소	내용
2015.03 ~ 2018.02	사회활동	KSCY 한국청소년학술대회	KSCY 한국청소년학술대회 컨퍼런스 총괄
2015.07 ~ 2017.07	자원봉사	고등학교 심리학 멘토링	고등학교 심리학 멘토링 (장훈고, 하나고, 불곡고 등)
2015.07 ~ 2015.08	사회활동	일본문화원	쿨재팬 리포터
2015.09 ~ 2015.12	교내활동	연세대학교	Psy-world 교내 심리학 학술모임 창립 및 운영
2016.01 ~ 2017.01	자원봉사	위메이저	전공소개 강연봉사
2016.03 ~ 2016.09	동아리	멋쟁이사자처럼	멋쟁이사자처럼 4기 (Ruby on Rails)
2016.09 ~ 2016.12	교내활동	연세대학교 성격사회심리학 연구실	서은국 교수님 성격사회심리학 연구실 RA
2019.03 ~ 2019.12	사회활동	통계청	통계청 대학생 기자단
2019.03 ~ 2019.05	자원봉사	연세대학교 YDMC 전공소개 봉사단체	전공소개 봉사
2019.03 ~ 2019.12	자원봉사	양천구 대학생 멘토단	전공소개 봉사

2019.09 ~ 2020.08	교내활동	연세대학교 통계학회 ESC	학술부 및 총무(임원)
2022.03 ~ 2022.06	수행과제	연세대학교 데이터사이언스연구소	통계상담

자격증/어학/수상내역

취득일/수상일	구분	자격/어학/수상명	발행처/기관/언어	합격/점수
2018.06	자격증/면허증	한자능력검정시험 3급	한국어문화	최종합격
2020.09	자격증/면허증	데이터분석전문가(ADsP)	한국데이터베이스진흥원	최종합격
2020.10	자격증/면허증	SQL개발자(SQLD자격)	한국데이터베이스진흥센터	최종합격
2015.12	수상내역/공모전	Freshman Seminar 창업 101 최우수상	연세대학교	-
2021.01	수상내역/공모전	Y&Z 투자자 프로파일링 시각화 경진대회 입선상	NH투자증권	-
2021.12	수상내역/공모전	빅데이터분석 경진대회	연세대학교 통계데이터사이언스 학과	-

보유기술

보유기술명/수준/상세내용
Python
R Studio
SQL
SPSS
Tableau

포트폴리오/기타문서

파일 구분	파일명
포트폴리오	 Portfolio_손지우.pptx 30.9MB

자기소개서

[1. 빅데이터에 의미 주기]

'무의미에 의미 주기'는 저의 좌우명입니다. 빅데이터에는 많은 정보들이 들어있지만, 너무 방대하다보니 놓치는 의미들도 있습니다. 데이터의 가치를 발견하는 일은 보물찾기처럼 설레는 일입니다.

지난 학기, 정제되지 않은 실제 기업데이터를 분석할 기회가 있었습니다. 저희 팀이 발견한 데이터의 의미를 기업관계자에게 공유하고, 인사이트에 대해 고마움의 표시를 받았습니다. 그때 느낀 뿌듯함은 아직까지도 저한테 강렬하게 남아있습니다. 이번에도 실제 데이터를 다뤄 의미를 찾아보고자 지원하게 되었습니다.

전문지식을 단단하게 쌓기 위해 저는 '어렵게' 공부할 것입니다. 예를 들어, 배포된 알고리즘을 사용하지 않고 때로는 직접 데이터에 알맞은 방법을 찾을 것입니다. 그런 성장 과정에서 사내 스튜디오와의 교학상장이 뒷받침된다면, 크게 두 가지 목표가 이뤄질 것입니다. 첫째는 독자적인 분석 알고리즘 개발, 둘째는 논문작성입니다. 계속 성장에 가속도가 붙는 사람이 되고자 합니다.

[2. 협업이란, 함께 할 때는 함께 하고 따로 할 때는 따로 하기]

NH투자증권 주최 Y&Z세대 투자자 프로파일링 시각화 경진대회에 참가하였습니다. 저의 역할은 주식데이터 크롤링으로 시작하였습니다. 주식종목들을 WICS 기준으로 분류하고, PER과 같은 추가적인 정보를 수집하는 일은 팀원 중 유일하게 주식을 하는 제가 맡았습니다. 협업 초반에는 각자 데이터를 탐색하고, 매주 회의를 통해 새롭게 알아낸 정보들을 공유하며 중요한 변수들을 추렸습니다.

이후에는 머리를 맞대고 분석모델을 만들었습니다. 의견을 자유롭게 쏟아내고, 통계적 근거와 현실적 합리성을 서로 즉각 점검해주었습니다. 예를 들어, 저는 '휴면'이라는 변수를 만들때, 장기투자자를 고려하여 마지막 거래일 외에 거래 주기를 추가로 고려해야 한다고 제안했습니다.

한 달 간 각자 잘할 수 있는 부분은 서로 믿고 맡겼고, 나머지는 함께 고민을 하며 채워나갔습니다. 그 결과, 공동 5위 입선상을 수상하였습니다.

[3. 데이터 전처리로 감동을 주다]

연세대학교 빅데이터 분석 경진대회에서 우수상을 수상하였습니다. '효돌'이라는 어르신 생활관리 AI 로봇인형에 대해 인사이트를 도출해내는 대회였습니다.

저희 팀은 유사한 특징을 가진 사용자들을 파악하여 그룹별 차별화 전략을 제안하고자 했습니다. 하지만, 제공받은 DB는 문제가 많았습니다. 예를 들어, 비현실적으로 하루 사용기록이 천 개 이상 있거나, 센서의 고장으로 사용횟수가 음수로 기록되어 있기도 했습니다.

이상치를 탐지하기 위해 iForest 알고리즘을 사용하였고, 이동평균을 통해 데이터를 제거하거나 변환했습니다. 이러한 접근법은 효돌 관계자분들의 평소 고민거리를 시원하게 해결해주었고, 감동적이라는 평가를 들었습니다.

정제된 데이터로 분석한 결과, 적극적인 사용자들 '기능 위주' 사용자와 '친밀감 위주' 사용자로 세부분류할 수 있었습니다. 전처리의 중요성을 체감하고, 이후 다양한 이상치 탐지 알고리즘에 대해 공부를 하는 계기가 되었습니다.

[4. Python과 머신러닝, 그리고 R과 통계분석]

Python과 R을 주언어로 사용하고 있습니다. Python으로는 머신러닝을 연구하고 있습니다. XGBoost와 같은 방법론들을 사용하여 데이터분석을 하기도 하고, 데이터수집을 위해 크롤링도 합니다. 현재는 SMOTE 알고리즘을 특징 공간으로 확장시켜 적용하는 방법에 대해서 연구 중입니다. 이외에 CV와 NLP 딥러닝도 GAN이나 Attention과 같은 대표적인 모델을 직접 구현해보고 있습니다.

R은 통계분석이 필요할 경우 활용하고 있습니다. 예를 들어, GLMM 같은 복잡한 통계모형을 상황에 맞게 적합할 때 사용합니다. 또한, 블로그는 Rstudio에서 깃헙과 연동하여 관리하고 있습니다. (주소: jwooblog.netlifyapp)

이외 다른 언어도 언제든 배우고자 합니다. C++언어로 EM 알고리즘 등을 구현해보기도 했고, SQLD 자격증을 취득하였으며, 분석결과를 효과적으로 전달하는 태블로도 공부하였습니다.

[5. 다양한 데이터분석 프로젝트 경험]

빅콘테스트 주최 챔피언스리그에 참가하여 NS Shop+ 홈쇼핑 데이터를 분석하였습니다. 2020년 6월 실적을 예측하고, 최적수익을 고려한 편성안을 제시하는 대회였습니다. 상품군별로 데이터 수가 달라서, 성능 좋은 분석모델을 만드는 데에 어려움이 있었습니다. 그래서 편성시간이 가장 많은 제품 대비 절반도 안 되는 제품에 한해서 RUS를 통해 불균형도를 완화한 결과, 성능이 향상되었습니다.

학회 프로젝트로, 아파트 경매가격 예측 경진대회 데이터를 분석하였습니다. 우선 제가 베이지안 회귀분석에서 조건부 확률부터 g-prior까지 직접 코딩하고, BMA를 통해 최종변수를 선택하였습니다. 그 과정에서 플래트 스케일에서 영감을 받아, 명목형 변수를 원핫인코딩하지 않고 확률로 변환하는 새로운 시도를 해보았습니다. 이후에는 명목형 변수의 불균형도를 완화하고 XGBoost를 선택해 1등에 가까운 성능을 구현했습니다.