

# 이 력 서

## 기본정보

- 이름 : 김영훈
- 생년월일 : 1985 년 5 월 1 일(만 38 세)
- 성별 : 남
- 주소 : 서울시 강서구 마곡중앙 5 로 6
- 이메일 : genholy77@gmail.com
- 연락처 : 010-9004-6355



## 핵심역량

- AI 관련 데이터 수집, 분석
- Python
- Linux
- Excel
- English : TOEIC 775 점(취득일자: 2023.09.24)

## 학력사항

|                                    |                   |
|------------------------------------|-------------------|
| 숭실대학교 정보과학대학원 인공지능학과 졸업 (본교:야간) 석사 | 2021.03 ~ 2023.08 |
| 한국외국어대학교 경제학과 졸업 (본교:주간) 학사        | 2005.03 ~ 2013.02 |
| 부산중앙고등학교 인문계 졸업                    | 2001.03 ~ 2004.02 |

## 경력사항

|                            |                             |
|----------------------------|-----------------------------|
| 디플렉스씨앤씨 (KT DS 협력사)        | 2023.07~2023.12 (5 개월)      |
| 에스맨파워 (KT DS AICC 서비스 1 팀) | 2021.08~2023.07 (1 년 11 개월) |
| CJ 바이오사이언스 (재경팀)           | 2017.11~2021.03 (3 년 5 개월)  |

## 병역사항

- 만기제대(2006.08~2008.07, 의무경찰)

## 교육과정 및 기타사항

| 교육과정명       | 교육기관      | 교육기간                     | 교육시간   | 수료여부 |
|-------------|-----------|--------------------------|--------|------|
| 파이썬 빅데이터 분석 | 코드랩아카데미학원 | 2022.03 ~ 2022.06 (4 개월) | 112 시간 | 정상수료 |

**디플렉스씨앤씨(KT DS 협력사)**

2023.07~2023.12 (5 개월)

프리랜서, 퇴사사유 : 계약만료

**데이터 수집, 분석**

**NH 투자증권 음성자산화 사업 - 지식구축**

- 상담원문 분석 : 의도분류 및 개체명 태깅
- 검증 및 테스트 결과 분석 : precision, recall, f1-score 수치 및 에러문구
- 개체명 사전, 사용자 사전, 유의어 사전 추가 : 주식종목명, 금융상품명, 금융 관련 업무명 등
- 형태소 분석 : 오추출된 단어를 형태소 분절 조정을 통해 개선
- 상담원문 전사 : 주식 주문 관련 고객발화

**에스맨파워**

2021.08~2023.07 (1 년 11 개월)

파견직 사원, KT DS AICC 서비스 1 팀, 퇴사사유 : 계약만료

회사소개: KT 그룹 정보시스템 구축과 노하우 기술을 바탕으로 고객의 가치를 실현하는 IT 서비스 전문 기업

**데이터 수집, 분석**

**KT AI 상담센터 보이스봇 언어모델링(LM)**

- 단어 간 확률 기반의 N-gram 언어모델 방식을 이해하고 코퍼스(말뭉치) 제작
- 기업 고객의 시나리오 분석을 토대로 도메인별 예상 발화 수집
- 완성된 코퍼스(말뭉치)로 리눅스 셸 스크립트 수정 및 명령어 실행을 통해 언어 모델링 작업
- 언어 모델링 작업 후 기업 고객 서비스에 인입된 음성파일로 인식률 점검 및 서버 배포
- 배포 이후 정상 작동 확인 및 점검 결과 사업부에 공유
- 상시 축적되는 음성파일 및 엔진인식 결과 데이터 분석을 통해 인식률 품질 유지
- 데이터 축적 경로 파악 및 추출로 고객 VOC 발생 시 대응 지원

**CJ 바이오사이언스**

2017.11~ 2021.03 (3 년 5 개월)

정규직 대리, 재경팀, 퇴사사유 : 업직종 전환

회사소개: 장내미생물 유전정보 분석 기업, 헬스케어 서비스 및 신약개발

**재무회계**

- 자금관리 : 외부 투자금 사용용도별 관리 및 보고자료 작성
- 매입관리 : 재고자산/매출원가, 기타 매입비용 관리
- 회계감사자료 작성 : 분기/반기/연간 감사자료 작성
- 한국거래소 상장 심사자료 작성 : 연도별 비교식 자료 작성

## 자기소개

---

### • 지원 동기

AI서비스팀에서 언어모델의 데이터 분석 업무와 자연어처리 단에서 의도분류 및 개체명 인식 업무를 하며 AI기술이 어떻게 접목되고 있는지 경험하였습니다. 2년이 넘도록 AI상담센터에서 유입되는 음성데이터와 AI가 인식한 문구 및 처리 내용들을 분석하며 어떻게 하면 자연스러운 서비스를 고객에게 제공하고, 직원들이 더 효율적으로 업무에 임할 수 있을까 고민하였습니다. 저의 이러한 경험과 고민은 회사가 제공하는 서비스 품질을 한층 더 끌어올릴 수 있다고 생각합니다. 특히 재경팀에서 근무하며 쌓았던 엑셀 스킬은 이후 업무에서 보고서 작성 및 자료정리에 큰 도움이 되었고 앞으로도 유용하게 활용될 것이라고 생각하며, 아울러 교육기관 및 대학원에서 공부한 파이썬 프로그래밍을 통해 해당 분야의 데이터를 수집, 정제한 후 분석 및 활용하여 회사의 성장에 이바지하도록 하겠습니다.

### • 본인의 강점

저는 개발언어로 파이썬을 공부하였습니다. 우선은 대학원 커리큘럼 상 인공지능 알고리즘을 구현하기 위해 처음부터 텐서플로, 케라스 등의 딥러닝 프레임워크로 접근하기 보다 파이썬만을 이용하여 단계별로 직접 알고리즘을 구현하여 동작 원리를 기초부터 자세히 마스터 하였습니다. 다만, 딥러닝의 CNN, RNN 동작 알고리즘은 상대적으로 이론이 복잡하고 코드가 많기 때문에 텐서플로로 구현하는 연습을 하였습니다. 또한 학부에서 경제수학과 통계학을 심도있게 학습하여 인공지능 이해에 필요한 확률 및 통계, 선형대수, 미분 등의 개념을 수월하게 이해할 수 있었습니다. 이러한 인공지능과 머신러닝, 딥러닝의 전반에 대한 이론 및 코드 이해를 기반으로 AI학습 모델-딥러닝 연구 개발의 업무를 담당하고 싶습니다.

아울러, 저는 대학원 졸업을 위해 논문 작성을 완료한 경험이 있습니다. 업무 여건 상 회사 서비스 문제점 파악 및 개선방안 기획의 형태로 논문 작성을 완료하였습니다. 이 기간동안 AI 상담센터에 대한 전체적인 구조(STT-NLP-TTS)를 넓은 시야로 파악할 수 있었고, 각 단계별로 타 회사 서비스 및 관련 논문들을 찾아가며 현 기술 수준 및 신 기술에 대한 탐색 활동을 열심히 하였습니다. 이러한 경험을 바탕으로 AI신기술 기반 응용 모델 연구개발 및 AI 신기술 대응에 필요한 인력이 되도록 하겠습니다.

부가적으로 이전 경력인 재무팀에서 쌓은 수치를 기반으로 한 엑셀 보고서 작성 능력과 영문으로 의사소통하고 문서를 읽는데 적합한 영어 실력은 원활한 업무 흐름에 좋은 요소로 작용할 것입니다.

### • 업무 경험 및 역량

초기 AICC 구독형 사업 진행 시 식당 예약과 관련된 언어모델링 고도화 업무를 맡았습니다. 처음 고객사는 VIPS였는데, 식당 예약과 질의에 대한 비정형 데이터들을 기본 시나리오 흐름을 토대로 수집하기 시작했습니다. 초기 인식률은 70점대였는데, 추가적으로 예상되는 발화와 날짜, 요일, 시간 등의 특정 문구들을 엑셀 및 리눅스 쉘 스크립트로 증폭시키는 작업을 통해 서비스 개시 전 인식률을 90점대 초반으로 올렸습니다. 본격적인 서비스가 개시된 이후로 VIPS의 전 지점으로 회선은 확대되었고, 하루에 8,000~10,000 콜 정도의 음성데이터가 유입되었습니다. 서비스 품질 유지를 위해 인입된 음성파일 및 전사데이터를 엑셀을 활용하여 분석하였고, 오인식된 키워드 및 기타 어미 등을 추가하여 실제적으로 개선해 나갔습니다. 결국 인식률은 97점으로 상승하였고, 이 언어모델을 시작으로 다른 예약과 관련된 식당

및 카페 업체들에게도 서비스를 제공할 수 있었습니다.

NH투자증권 음성자산화 프로젝트 진행 중에서는 지식구축업무를 담당하며 주식 주문 및 금융상품 관련 고객 발화의 원문 데이터를 분석하였습니다. 고객 발화의 의도를 분류하고, 문장 속에 종목명과 같은 개체명을 인식하는 작업을 진행하였습니다. 의도분류와 개체명 인식이 제대로 되지 않으면 분석단계에서 미추출로 떨어지는 경우가 많았습니다. 프로젝트를 인수 받았을 때 데이터 추출률은 89점이었는데, 이후 의도분류에 혼동이 오는 부분을 제대로 분류하고 미추출된 문장들을 대거 수집하여 학습을 진행하여 94~95점으로 추출률을 끌어 올렸습니다. 문제점의 정확한 원인을 데이터를 통해 분석하고 발견하여 그 해결점을 찾아 학습을 진행하였더니 품질이 좋아지는 결과를 얻을 수 있었습니다.

이러한 경험은 유입되는 데이터들의 유형을 파악하고, 의사결정에 필요한 데이터를 툴을 통해 수집 및 정제하여 인사이트를 제공하는데 도움이 될 것입니다.