**이력서**

**[지원분야: AI 개발자 / 신입]**

**▶ 인적 사항**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **성명** | 정종혁(鄭宗赫) | **영문** | Jonghyeak, Jeong |  |
| **생년월일** | 1997. 01. 11 (만 28세) | | |
| **주소** | 서울특별시 은평구 은평로21길 20 | | |
| **연락처** | 010-7624-5728 | | |
| **이메일** | devna0111@gmail.com | | |

**▶ 핵심역량(Skills) :** [**https://github.com/devna0111**](https://github.com/devna0111)

|  |  |
| --- | --- |
| **언어/프레임워크** | Python, TensorFlow, Flask, Django |
| **AI/ML 기술** | 머신러닝, 딥러닝(CNN, RNN, Transformer), LLM, NLP, 생성형AI기반 RPA, RAG |
| **플랫폼/도구** | Hugging Face, Ollama(Open source활용), Git |
| **DBMS/데이터** | Oracle, MySQL, SQLite, Qdrant, Pandas, NumPy |
| **Web** | HTML5, CSS3, Bootstrap, Jinja2, JavaScript, jQuery |

**▶ 프로젝트 수행사항**

|  |
| --- |
| **2025.07.10 ~ 07.23 / 1차 팀 프로젝트 – 서울시 민간 PM 재배치 (MBC 아카데미)** |
| * 소스 : <https://github.com/devna0111/FlowCast> * 서울시 공공자전거의 대여량 예측과 이를 통해 민간 PM을 재배치하여 소비자/기업 양측 효율성을 추구 * 주 기술 스택   - python, tensorflow(LSTM,DNN), LightGBM, CatBoost, RandomForest, flask, OpenAI API   * 기여 부분   - 데이터 수집 및 분석, 모델의 설계 및 조정, 웹 서비스 개발 |
| **2025.08.04 ~ 08.22 / 2차 팀 프로젝트 – RAG를 활용한 \*\*\*\* (MBC 아카데미)** |
| * 소스 : [**https://github.com/본인개별프로젝트git주소**](https://github.com/본인개별프로젝트git주소) * 스마트 지역 상권을 분석한 내용을 기반으로 창업 컨설팅을 쳇봇으로 제공 |
| * 주 기술 스택 : * 기여 부분 : 단순 기능 나열은 피하자 |

**▶ 학력 사항**

|  |  |
| --- | --- |
| **2015.03 ~ 2019.02** | 서울과학기술대학교 건설시스템공학과 졸업(졸업 성적 : 3.98/4.5) |
| **2012.03 ~ 2015.02** | 서울 현대고등학교 졸업 |

**▶ 교육 및 연수 사항**

|  |  |
| --- | --- |
| **2025.04.07 ~ 2025.09.02**  **(824시간)** | AI+X를 활용한 기업형 인재 양성 [MBC 아카데미 신림점] |

**▶ 자격 사항**

|  |  |
| --- | --- |
| **2025. 07. 04** | 정보처리기능사(한국 산업 인력 공단) |

**▶ 병역사항**

|  |  |
| --- | --- |
| **2019.03 ~ 2021.06** | 공병 장교(ROTC) 만기 전역  - 소대장 : 병력관리, 중대 행정 업무, 훈련 관리  - 공병단 정보장교 : 부대보안관리(문서, 인트라넷), 부대 첩보 관리 등 |

**▶경력 사항**

|  |  |
| --- | --- |
| **2019.03 ~ 2021.06 (2년 4개월) / 육군본부** | |
| **19.03 ~ 19.06** | (소위) 공병학교 신임 장교 지휘자 과정  - 리더십 및 공병훈련 교육 |
| **19.06 ~ 19.12** | (소위) 공병 소대장  - 병력 중심 지휘 및 관리  - 중대 문서 행정 전담 |
| **20.01 ~ 20.12** | (중위) 정보장교 : 보안 관리 및 부대 첩보 수집 등  - 공병단 보안 관리 총괄  - 작전사령부 연계 보안 관리 활동  - 정작과 문서 행정 보조 |
| **21.01 ~ 21.06** | (중위) 정비소대장 : 중대장 대리 임무 수행  - 작전계획 수정 및 조정  - 중대 보안 관리  - 중대 문서 행정 전담 |
| **2021.12~ 2025.03 (3년 8개월) / 주식회사 래버커스** | |
| **21.12 ~ 22.04** | (팀원) ckbglab 본점 매장 운영  - 매장 매출 데이터화  - 간이 회계, 노무, 물류 관리 |
| **22.04 ~ 23.03** | (실장) ckbglab 압구정 지점 매장 총괄  - 매장 매출 데이터화 및 매출 분석, 운영 방안 제시  - 회계, 노무, 물류, 본점 연계, 타 매장 런칭 관리 |
| **23.04 ~ 23.07** | (팀원) 유용욱 바비큐 연구소 FOH  - 매장 운영  - 확장 이전 대비 필요 행정 서류 구비 |
| **23.08 ~ 25.03** | (실장) 이목 스모크 다이닝 총괄  - 매장 운영 및 인력 관리  - 매출 데이터 분석 및 인사이트 도출, 운영 방안 제시  - 회사 운영 매뉴얼 제작 및 각종 행정 업무 일괄  - 케이터링 행사 간 타 회사와 네트워킹, 조정, 견적 및 세금처리 등에 기여 |

**서울 PM 재배치를 위한 수요 예측 모델 개발**

|  |  |  |
| --- | --- | --- |
| **프로젝트 명** | | 서울 PM 재배치를 위한 수요 예측 |
| **참여 인원** | | 4명 |
| **기간** | | 2025.07.10 ~ 2025.07.23 |
| **개발 환경** | Language | Python 3.10 |
| Framework | Flask (API), Jinja2 (Template) |
| IDE Tool | VSCode, Jupyter Notebook, WorkBench |
| DB | MySQL |
| Open Source | TensorFlow, scikit-learn, lightgbm, pandas, numpy, matplotlib, seaborn, seleninum, 공공데이터포털 Open API(data.go.kr), OpenAI\_API, Prophet, openmeteoAPI |
| 데이터 출처 | 서울시 월별 공공자전거 대여 – 서울시 열린데이터광장  기상 및 기후 – 기상청 통계 포털  PM 견인률 및 운영 현황 – 공공데이터 포털 |
| 프로젝트 기여 | 1. 데이터 분석(인구, 인프라, 시계열 속성)   1. 인구 데이터의 경우 3종의 데이터를 확인(행정인구, 생활인구, 유동인구) : 데이터가 갖는 의미에 대해 분석하고 프로젝트의 적합성(실제 PM을 이용할 인구) 체크 2. 인프라 데이터의 경우 자전거 노선의 길이와 그 수에 따라 대여량과의 상관관계를 확인하고 중요도를 체크, Feature 사용여부 결정 3. 외부 활용 자원에 따라 계절성, FM/LM의 활용에 따른 시간에 따른 영향 체크를 통해 시계열 속성 체크   2. 데이터 전처리  가. 데이터의 결측치를 확인  : 공공 제공 데이터이며 5분 간격의 데이터: 결측치 0건  나. 정류소 중 일반인이 사용 가능한 정류소 + 수리 등을 위한 정비정류소를 확인, 프로젝트 목적에 적합한 일반인 사용 정류소만 추출  다. 데이터 이상치 처리 : 수요가 많을 것으로 예상되는 데이터는 이상치로 선 처리가 아닌 모델 체크 후처리로 대응  3. 모델의 설계  가. 필요 모델 체크 : 최초 가설 설정 시 민간 PM의 경우 공공자전거의 대여량 영향을 설정, 이에 따라 필요 모델 최소 2개 분류(공공자전거 대여량 예측 / 민간 PM 대여량 예측)  나. 공공자전거 대여량 예측 모델  - 시계열 속성과 타겟 변수(대여량)의 불균등 분포, 극단적 값에 따라 모델을 설계  - 최초 Classification 계획하였으나 타겟 변수 분포에 따라 PM수요 예측 모델에 부적합 판단, Linear Regression 모델로 변경  - LSTM(중장기 데이터 예측에 강점), DNN(시계열 속성을 작년도 대여량을 Feature로 활용하여 해소), 앙상블 모델(LightGBM(작년도 대여량 Feature의 활용에 따라 복잡한 Feature 분석에 강한 모델 선정))   1. PM 대여량 예측 모델   - 공공자전거 대여량 예측 모델의 결과를 Feature 활용하는 모델의 설계  - 공공자전거 대여량 예측 모델과 달리 재배치가 필요한 시간대를 예측함에 큰 목표로, 사분위수 Q3 초과에 대해 재배치가 가능하게 Classification 모델 설계  4. 웹서비스 구현  가. 웹서비스가 가능하게끔 기능들의 함수화 진행  - 전년도 데이터 MySQL load  - 실시간 기상 예보(openmeteoAPI)  - 공공자전거 대여량 예측 모델 예측  - 민간 PM 대여량 예측  - 행정구별 주요 역들에 대해 대여량 가중치에 따라  상위 5개 역에 대한 대여량 예측  - 시각화 함수 등  나. flask와 Jinja2 템플릿을 통한 웹서비스 내용 구현  다. openAI API를 활용한 재배치 필요 시간대 확인과 데이터 분석 기능 구현 |
| 협업도구 | GitHub, Google Drive, Notion |

**▶ 프로젝트 연관 자료**

|  |  |
| --- | --- |
| **Github Address** | <https://github.com/devna0111/FlowCast> |
| **프로젝트 발표 동영상 및 실행 화면 시연** | [https://youtu.be/](https://youtu.be/@jongjong__) |

**▶ 1. 주제 및 프로젝트의 정의**

1. 주제 : 개인형 이동장치(Personal Mobility; 이하 PM)의 수요 대비 공급의 불균형 문제를 해소하기 위해 기상, 인구, 인프라 등을 고려하여 행정지역별 · 시간별 공급의 부족과 과다를 시각화 및 분석하는 웹서비스

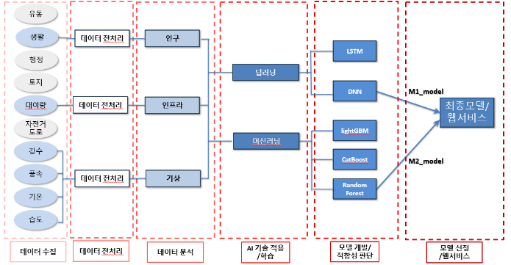


그림 1 워크플로우 설정

1. 문제 정의 : PM의 공급을 충분히 해결하지 못해 불편함을 겪는 기사, 민원을 확인

텍스트, 지도, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

그림 2 문제 배경 제시

1. 프로젝트 목표 설정 : 주제와 부합하는 프로젝트의 목표와 목표물 계획

(웹서비스=>스마트폰용 확장)

텍스트, 스크린샷, 폰트, 디자인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 폰트이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

그림 3 프로젝트 목표 설정 그림 4 목표물 설정

1. 프로젝트 수행을 통한 기대효과 및 응용분야 선정을 통한 정확한 목표 지정

텍스트, 의류, 스크린샷, 사람이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 신발류, 스크린샷, 지도이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

그림 5 프로젝트 기대 효과 그림 6 기술 응용 및 확장 가능 분야

▶ 2. 일정관리 및 업무분담

1. 일정 관리 : Gantt Chart 작성 및 책임자 정 · 부 부여를 통해 전체 프로젝트 관리

- 제한된 시간과 한정 자원을 효율적으로 운영하기 위해 소목적선정과 책임자 지정을 통한 효율성 추구

- 업무 전반에 걸쳐 각자의 역량으로 가능한 프로젝트 책임 부여로 효율성 극대화

텍스트, 스크린샷, 소프트웨어, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

그림 7 각자의 역량에 따른 프로젝트 책임자 선정 및 Gantt Chart 작성

1. 업무 분담 : 총괄, 서브 기획, 데이터 분석, 모델 엔지니어로 구분하여 프로젝트 책임 분장

- 각자 역량에 따라 서로 상호보완을 통해 프로젝트를 진행

**-** 데이터 활용 분야와 데이터 분석 분야 상호 커뮤니케이션을 통해 소목적에 대한 데이터 토의

및 분석 · 활용

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

그림 8 팀 구성 및 역할 부여

▶ 3. 데이터 수집 : 서울시 열린 데이터 광장, 공공 데이터 포털, 기상청 통계 포털

|  |  |
| --- | --- |
| **참고자료 1인구 · 인프라 · 기상 데이터 수집** | **인구 데이터 : 서울시 열린 데이터 광장**  **(행정, 생활, 유동 3개 개념)**  **인프라 :서울시 열린 데이터 광장(자전거 노선길이와 수)**  **공공 데이터 포털(PM 견인률, 운영현황)**  **기상 : 기상청 통계 포털**  **(기온, 습도, 강수, 풍속)** |

가. 여러 기관의 공개 데이터를 활용하여 기본 데이터 셋을 구축

나. 결측치는 없었으나 시계열 해상도(단위 : 분 / 시 / 일 / 연 단위) 차이 확인

▶ 4. 데이터 분석

1. 수집 데이터 분석 : 인구, 인프라, 기상

|  |  |
| --- | --- |
| **참고자료 1 인구데이터 분석 결과** | **인구 데이터 분석 결과**   1. 생활 인구수 분포는 우측으로 꼬리가 긴,   왜도 1.416의 비균형 분포   1. 생활 인구수와 대여량의 상관관계를 확인하기 위한 도수분포표에서 양의 상관관계를 시각화로 확인하였으며 일정 인구수 이상부터 대여량이 급증함을 확인 2. 위 결과에 따라 인구 밀집 지역을 중심으로 수요 예측의 필요성 점검 |
| **참고자료 2 인프라 데이터 분석 결과** | **인프라 데이터 분석 결과**   1. 노선길이에 따른 대여량의 형태를 시각화해 단순 길이의 증가와 대여량의 상관관계가 미비함을 확인 2. 행정구별 대여량 분포를 시각화하여 지역별 수요의 차이를 확인 3. 필요에 따라 자전거 인프라의 경우 feature 사용 여부 재확인(feature importances 및 모델 학습결과 기반 의사결정) |
| **참고자료 3 기상 데이터 분석 결과** | **기상 데이터 분석 결과**   1. PM 대여는 적정 온도(15~25도)에서 활발 2. 습도가 90%를 넘을 때 대여량 급감 3. 강수 시 대여량 급감 4. 풍속이 4m/s 이상 시 대여량 급감   => 계절성 요인 확인  => 계절성 및 시계열 분석 필요 |

1. **시계열 분석 : 계절성 요인을 기상 데이터에서 확인하였으며 FM/LM으로 활용도가 높은 PM의 특성을 고려 시계열 분석을 실시**

|  |  |
| --- | --- |
| **참고자료 1 시계열 속성 분석 결과** | **시계열 분석 결과 : EDA 및 Prophet**   1. 일간 시간대에 따른 대여량의 증감을 확인, 특히 출퇴근 시간과 평일/주말 간 차이 확인 2. 7월을 기준으로 대여량의 파동을 확인   봄(4~5월)과 가을(9~10월)에 대여량이 급증하며 여름 폭염기와 겨울 한파기 대여량 급감 |

▶ 5. 모델 가설 및 Feature 추가 전처리

1. 최초 계획 시 필요 데이터를 도메인 분석을 통해 구상 후 이에 따라 데이터 분석과 동시에 모델에 적용할 가설을 설정
2. 민간 PM 예측 모델을 설계하기 위해서 Feature로 활용할 “공공자전거 수요 예측 모델”의 필요성을 확인, 2개 모델을 설계

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

그림 9 모델 가설 선정 의사결정 과정

1. 학습 전 데이터 전처리를 위한 분포 체크 : 전반적인 데이터의 이상치, 비정규성, 극단값, 불균등분포 등 데이터 특성 체크 및 변수별 특성 고려 맞춤 스케일 조정과 라벨인코딩

텍스트, 스크린샷, 도표, 그래프이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 도표, 평면도, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 소프트웨어, 웹 페이지이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.**

**텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.**

**텍스트, 스크린샷, 소프트웨어, 웹 페이지이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.그림 10 공공자전거 대여량 수요 예측 모델 Feature 및 종속변수 분석**

**텍스트, 스크린샷, 폰트, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 디자인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.**

**텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 그래프이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.**

**텍스트, 스크린샷, 소프트웨어, 웹 페이지이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.그림 11 민간 PM 대여량 수요 예측 모델 Feature 및 종속변수 분석**

1. 데이터 점검 결과에 따라 독립변수는 MinMaxScaler와 StandardScaler를 활용해 스케일 조정, 종속변수는 프로젝트 목표에 따라 “공급 부족”과 “공급 절대 부족” 범주를 맞출 수 있게 이상치 조정 없이 예측 정확도를 높이기 위해 MinMaxScaler, RobustScaler, Log Scaling을 고려, 학습

**▶ 6. 모델의 설계와 프로젝트 적합성 체크**

1. 모델은 상위 분석한 데이터를 토대로 총 2개 ; 공공자전거 수요 예측 모델(이하 M1)과 PM 수요 예측 모델(이하 M2) 설계 실시
2. M1 모델을 설계할 때 **1) Classification 모델**과 **2) Linear Regression 모델**을 최초 고려하였지만 타겟 변수의 범주와 이상치 특성에 따라 향후 M2 모델의 feature로 활용을 고려하였으며 최종적으로 Linear Regression 모델을 선택
3. Classification : 모든 모델에서 accuracy가 90% 이상으로 높은 성능을 기록하였으나 M2 모델의 feature로 사용할 때, Cluster의 범주가 너무 광범위해 M2 샘플 모델의 accuracy가 35% ~ 40% 수준으로 한계

텍스트, 스크린샷, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다. 텍스트, 스크린샷, 번호, 도표이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다. 텍스트, 스크린샷, 번호, 폰트이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 12 M1 Classification 모델 체크 및 한계 확인**

1. Linear Regression : LSTM과 DNN 2가지 방식을 고려하였으며 **DNN의 MAE가 0.32**로 **LSTM의 1/6 수준으로 높은 성능**을 보이며 **DNN 모델을 사용**

텍스트, 스크린샷, 번호, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다. 텍스트, 스크린샷, 번호, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 13 M1 Linear Regression 모델 확인**

1. M2 모델의 경우 M1 모델의 예측 값과 일시, 기상, 1년전 인구데이터, 1년전 대여량, 계절, 주말구분, 출퇴근 시간 여부의 독립변수를 사용, 프로젝트의 목적에 적합하게 4분위 수를 기준으로 Q3 : 75% 이상의 수요를 공급의 부족과 공급의 절대부족의 Class로 예측할 수 있도록 설계 및 학습

텍스트, 스크린샷, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 14 M2 모델의 선정 과정**

1. DNN, LSTM 두 모델과 앙상블 모델 LightGBM, CatBoost, RandomForest 3가지 모델, 총 5가지 모델을 점검하였으며 accuracy 목표 75%와 공급 부족과 공급 절대 부족 class를 예측하여 사업성을 높일 수 있게 두 class에 대한 Recall Score가 85% 이상인 모델 체크
2. DNN

- accuracy 74%로 목표 accuracy 미도달

- 공급 부족에 대한 recall score : 0.31

- 공급 절대 부족에 대한 recall scorea : 0.08

- 원하는 수준에 도달하지 못해 다른 모델을 체크

1. LSTM

- **accuracy 76%로 목표 accuracy 도달**

- 공급 부족에 대한 recall score : 0.34

- 공급 절대 부족에 대한 recall score : 0.31

- 원하는 수준에 도달하지 못해 다른 모델을 체크

1. 사전 샘플 앙상블 모델(LightGBM)에서 확인한 성능과 비슷한 수준의 딥러닝 모델

: **두 딥러닝 모델에서 해결하지 못한 recall score에 대해 앙상블 모델을 점검**

텍스트, 스크린샷, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다. 텍스트, 스크린샷, 도표, 직사각형이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 15 M1 딥러닝 Classification 모델 체크**

1. 앙상블 모델의 고려는 가) 시계열 속성 부여에 대한 강점을 갖고 나) 범주형 변수 처리 및 불균형 데이터에 강점을 갖으며 다) 갑작스러운 변화에 대한 강점을 갖는 모델을 체크

: LightGBM, CatBoost, RandomForest

1. 앙상블 모델의 경우 최적의 하이퍼 파라미터를 사용해 성능의 향상 도모, sklearn의 GridSearchCV 활용

텍스트, 스크린샷, 폰트, 도표이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 16 GridSearchCV 탐색 과정**

1. LightGBM

**- accuracy 78%로 목표 accuracy 도달**

- 공급 부족에 대한 recall score : 0.79

**- 공급 절대 부족에 대한 recall score : 0.89**

- 원하는 수준(공급 부족에 대한 recall score 0.85 이상)에 도달하지 못해 다른 모델을 체크

1. CatBoost

**- accuracy 78%로 목표 accuracy 도달**

- 공급 부족에 대한 recall score : 0.70

**- 공급 절대 부족에 대한 recall score : 0.90**

- 원하는 수준(공급 부족에 대한 recall score 0.85 이상)에 도달하지 못해 다른 모델을 체크

1. **RandomForest**

**- accuracy 84%로 목표 accuracy 도달**

**- 공급 부족에 대한 recall score : 0.88**

**- 공급 절대 부족에 대한 recall score : 0.90**

- **목표치 도달 및 수요가 많아 재배치가 필요한 지역을 가장 잘 예측하는 모델로 최종 모델**

**선정**

텍스트, 스크린샷, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 직사각형이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 17 앙상블 모델 점검 및 모델 선정**

1. 최종 모델 선정은 M1 : DNN Linear Regression 모델과 M2 : RandomForest Classification 모델을 활용

**텍스트, 도표, 번호, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.**

**그림 18 최종 모델의 선정**

**▶ 7. 웹서비스 구현**

1. 웹서비스의 경우 Flask와 Jinja2 템플릿을 통해 구현하였으며 민간 사업자가 1주일의 데이터를 체크하여 사업성을 극대화할 수 있도록 시간대별로 공급 부족 여부를 체크할 수 있도록 하였으며 참고자료로 공공자전거의 수요를 체크 가능하도록 설계

1. 일자별 수요가 높은 5개의 역을 예측하고 이에 대해 점검하여 재배치를 유도

1. 예측 결과에 대한 출력화면에서 LLM을 활용하여 그래프를 분석하고 이해를 용이하게 설명

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 평행이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.텍스트, 스크린샷, 도표, 폰트이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 19 웹서비스의 소개**

**▶ 6. 결론**

1. 서울 PM 수요 분석을 통해 1) **외부에서 활용하는 교통수단으로 기온, 강수, 습도, 풍속에 영향을 받으며 악기상에 수요량 급감함을 증명**하였으며 2) **출퇴근 시간에 활용하는 First Mile, Last Mile 교통수단**임을 또한 증명, 3) **행정구와 생활 인구수에 따라 수요가 결정되므로 아직 PM사업이 활성화되지 않았거나 사업자가 적은 특정 행정구에서 신사업에 대한 가능성을 점검할 수 있는 인사이트 도출용 프로그램** 개발

1. 이러한 수요 예측 서비스는 **1) 자원의 최적화를 통해 시민의 불편함 감소와 사업 이익의 극대화**, **2) 불필요한 재배치 작업을 감소시키고 운영을 최적화하여 비용의 절감 효과**와 **3) 향후 기술 응용을 통해 PM 대여 신사업 및 공공 정책과 도시계획에 활용**에 큰 기여 기대

텍스트, 의류, 스크린샷, 신발류이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다. 텍스트, 지도, 신발류, 스크린샷이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

**그림 20 서비스의 기대효과**

**2차 팀 프로젝트 포트폴리오**

**교육 이수 및 기술내역서**

|  |  |
| --- | --- |
| **교육 과정 명** | (KDT) 기업맞춤형 AI+X 융복합 인재 양성 교육 |
| **교육 기간** | 2025. 04. 07. ~ 2025. 09. 02 (20.6주, 824시간) |

|  |  |  |
| --- | --- | --- |
| **교과목** | **교육일정** | **이수 기술** |
| Python | 6주 | * Python 기본 문법, 자료형과 데이터 구조, 함수, * 모듈 패키지 제작 및 사용, OOP(클래스, 상속), 예외처리 * 파일 I/O, DB 연동, Numpy 배열, Pandas 라이브러리 * 데이터 시각화(matplotlib, seaborn, folium) * 공공데이터를 이용한 빅데이터 분석(부동산, 상가정보 : 데이터셋 분리 및 스케일 조정, 결측치 및 이상치 처리, 시각화) * 정적, 동적 웹 데이터 수집(날씨, 환율, 쇼핑몰 베스트셀러 정보) * Open API(공공 Open Api, 네이버 Open Api) |
| Database | 1주 | * 데이터베이스 개념과 이해, Oracle 및 MySQL * DML과 DDL(테이블 생성과 조작 언어), DCL, DBMS 권한제어 * DB설계(ERD) |
| 웹표준 | 2주 | * HTML5, CSS, CSS3, 반응형 웹 * JavaScript, JQuery, Ajax, fetch API, bootstrap 5.0 |
| 머신러닝 | 1주 | * 통계분석, 데이터 EDA 가속화, 데이터 전처리, 데이터 분리, 군집분석, 분류분석, 모형 최적화를 위한 변수 선택과 차원 축소, 파라미터 탐색, 자료 불균형 처리, 앙상블모형 |
| 딥러닝 | 1.5주 | * Tensorflow v1과 Tensorflow v2 * DNN(회귀분석, 이진분류, 다중 분류, 모델 저장 및 다양한 callback, 성능지표), CNN * 모델의 성능을 올리는 기법들 * 통계적 기법 : 가설설정, 통계처리에 대한 해석 |
| 자연어처리 | 1주 | * 자연어 처리(NLTK, KoNLPy 형태소 분석 및 자연어 시각화) * RNN, LSTM, GRU(번역, 영화감상평 감성 분석), 시계열분석 * Sequence to sequence, Transformer |
| LLM | 1.6주 | * Hugging Face, Ollama Open source를 통한 LLM * Open ai API * LangChain, RAG, Streamlit을 통한 개인화 쳇봇 구현 * 생성형 AI 기반 RPA(로봇 프로세스 자동화) |
| 웹 프레임워크 | 2주 | * Flask(route, render, 전송방식, Jinja2, fileupload, orm) * Django (Model, ORM, Tempalte, View, Form, File, 사용자인증) |
|  | 4.5주 | * 1, 2차 팀 프로젝트, 협업 관리 툴로 Github, Branch전략 사용 |

**▶수상 경력**

|  |  |
| --- | --- |
| **2020. 12** | 대령급(연대장급) 개인표창 : 모범 및 우수 간부 유공 |
| **2020. 11** | 대령급(연대장급) 개인표창 : 전술훈련평가 유공 |
| **2019. 12** | 대장급(군사령관급) 개인상장 : 최정예전투원 |
| **2019. 06** | 대령급(연대장급) 개인상장 : 성적 우수 |
| **2015.03 ~ 2019.02** | 성적 우수 장학금 4회 수여(서울과학기술대학교) |

**▶어학 능력**

|  |  |
| --- | --- |
| **2025.01.01** | OPIc / IM2 |

**자기소개서**

**[지원분야: AI 엔지니어 / 신입]**

**지원동기**

|  |
| --- |
| **[문제 정의부터 서비스화까지, 전방위 AI 역량을 갖춘 엔지니어로]** |

AI 기술은 데이터 속에 숨어 있는 패턴을 발견하고, 이를 통해 세상을 더 나은 방향으로 변화시키는 힘이 있다고 믿습니다. 저는 데이터를 기반으로 한 문제 해결에 깊은 흥미를 가지고 있으며, 근거 있는 가설을 세우고 이를 분석과 모델링을 통해 검증하는 과정에서 큰 보람을 느낍니다.

최근에는 서울시 민간 퍼스널 모빌리티(PM) 수요 예측 프로젝트를 수행하며, 공공 데이터를 분석하고 LSTM 기반 수요 예측 모델을 설계·조정하였습니다. 또한 Flask를 활용한 웹 서비스 프로토타입 구현까지 직접 수행하며, 데이터 기반 AI 모델이 어떻게 서비스로 확장될 수 있는지를 실무적으로 경험하였습니다.

TensorFlow, Flask, Django, Pandas, Hugging Face, Qdrant 등 다양한 프레임워크와 툴을 다루며 머신러닝·딥러닝(CNN, RNN, Transformer), 자연어 처리(NLP), 생성형 AI 기반 RAG, RPA 등 최신 AI 기술에 대한 이해도를 쌓았습니다. 항상 성실하고 적극적인 태도로 업무에 임하며, 기술력과 태도 모두에서 신뢰받는 엔지니어가 되기 위해 노력하고 있습니다.

앞으로도 데이터로부터 가치를 발견하고, 이를 실질적인 서비스로 구현하는 AI 엔지니어로 성장해 나가고 싶습니다.

**입사 후 포부**

|  |
| --- |
| **[ 아래의 내용을 대표하는 제목 언급 ]** |
| 입사 후에는 **데이터 수집부터 분석, 모델링, 서비스 구현까지 전 과정을 주도적으로 수행할 수 있는 AI 엔지니어**로 성장하고자 합니다. 특히 문제 정의 단계에서부터 사용자 관점의 인사이트를 반영하고, 이를 데이터 기반 예측 및 자동화 시스템으로 연결하는 **End-to-End 설계 역량**을 더욱 고도화할 계획입니다.  업무에 있어서는 단순한 기술 구현을 넘어, **프로젝트의 전체 흐름을 이해하고 책임지는 자세**를 갖추는 것을 가장 중요하게 생각합니다. 여기에 그치지 않고, **기회가 주어진다면 배움을 확장하여 회사의 전략과 방향에 기여할 수 있는 AI 관련 연구**까지도 도전하고자 합니다.  변화하는 AI 기술 환경 속에서도 끊임없이 학습하고 실험하며, 실무적 기여는 물론 **중장기적으로는 조직의 기술적 깊이를 더할 수 있는 구성원**이 되겠습니다. 또한 다양한 직무와의 협업을 통해, AI 기술이 실제 비즈니스 가치를 창출할 수 있도록 적극적으로 기여하겠습니다. |

**직무 관련 경험 및 경력**

|  |
| --- |
| **[ 아래의 내용을 대표하는 제목 언급 ]** |
| 1. 어울   서울시 공공데이터 및 민간 PM 운영 데이터(자치구별 기기 수, 견인 현황, 시간대별 기상 정보 등)를 기반으로, **PM 수요와 공급의 불균형을 파악하고 예측하는 모델**을 개발하였습니다. 데이터 수집 과정에서는 이질적인 형식과 해상도를 가진 데이터를 통합해야 했으며, **시간 단위 해상도의 차이**, **범주형 변수와 수치형 변수 혼합 문제** 등을 해결하기 위해 데이터 정제 및 전처리를 적용하였습니다.  또한, 시계열적 특성을 반영하기 위해 모든 데이터를 기준 시간대를 중심으로 정렬하고, **특정 주기성과 계절성을 분석**하여 예측 모델 설계의 기초를 마련했습니다.  문제 정의 단계에서는 “**시간대별로 특정 자치구에 수요가 몰리는 현상을 예측할 수 있을까?**”라는 가설을 세웠고, 이를 바탕으로 다양한 예측 모델을 비교 설계하였습니다.  **선형 회귀 모델**을 먼저 적용하여 변수 간 상관관계를 확인하고, 전체적인 흐름을 이해하는 데 활용했습니다. 이후 **RandomForest, LightGBM, LSTM** 등 분류 및 시계열 모델들을 적용하여 수요를 6단계 범주로 분류하고 예측하였으며, 각각의 모델이 갖는 **설명력, 복잡도, 연산 속도, 과적합 위험** 등 장단점을 직접 비교하고 실험해 보았습니다.  이 과정에서 **모델의 성능에 영향을 주는 주요 feature**들을 탐색하기 위해 **상관관계 분석**, **피처 중요도 분석(Feature Importance** 등을 활용하였고, 이를 통해 불필요한 변수를 제거하고 모델을 간결하게 구성할 수 있었습니다.  모델의 성능 향상을 위해 다양한 **하이퍼파라미터 튜닝**을 시도했습니다. 예를 들어, LightGBM 모델에서는 num\_leaves, max\_depth, learning\_rate, class\_weight 등 주요 파라미터를 그리드 서치와 수동 튜닝을 병행하여 조정했습니다.  또한 **클래스 불균형 문제**를 해결하기 위해 class\_weight를 조정하거나 **언더 샘플링** 기법, 이상치를 사분위수로 조정하는 등 해결방법을 강구하였습니다.  LSTM 모델에서는 과적합 방지를 위해 **Dropout, EarlyStopping, Validation split** 등 다양한 기법을 적용하며 모델 일반화 성능을 높이고자 했습니다.  모델의 결과를 단순한 수치 예측에 그치지 않고, **의사결정에 실질적인 도움이 되는 서비스로 구현**하기 위해 Flask 기반의 웹 프로토타입을 개발하였습니다. 사용자는 행정구, 시간대, 날씨 조건 등을 입력하면 예측 결과를 시각화된 형태로 확인할 수 있도록 하였으며, **공급 부족 시간대에 대한 대안 자치구 추천 기능**도 탑재하였습니다.  이 과정을 통해 **단순한 모델 개발을 넘어, 사용자와 소통할 수 있는 인터페이스 설계**, **데이터 기반의 직관적인 정책 도출**까지 경험하였습니다.  이 프로젝트를 통해 저는 단순히 모델을 ‘적용하는’ 단계를 넘어서, 데이터를 **분석하고 가설을 설정**하고 **모델을 직접 설계 및 평가**하며 **최적화와 피처 중요도 해석**, **웹 인터페이스를 통한 실제 활용**까지 **End-to-End 전체 흐름을 이해하고 책임지는 역량**을 기를 수 있었습니다.  이러한 실전 경험을 바탕으로, 앞으로 더욱 복잡한 AI 문제 해결과 실제 서비스 설계에 주도적으로 참여할 수 있는 엔지니어로 성장해가고자 합니다. |

**성장과정**

|  |  |
| --- | --- |
| **[ 애피소드를 대표할 제목 ]** | |
| “왜?”라는 질문을 멈추지 않았던 어린 시절이, 지금의 저를 만들었습니다.  저는 어릴 적부터 현상에 대한 의문을 가지고 스스로 설명을 찾으려는 습관이 있었습니다. 장난감이 어떻게 움직이는지 분해해보고, 수학 문제의 공식을 단순히 외우기보다는 그 원리를 납득하려 노력하곤 했습니다. 이처럼 원인과 결과 사이의 논리를 연결하는 것에 재미를 느끼는 성향은, 이후 제가 데이터를 다루고 분석하는 데 깊은 흥미를 갖게 된 계기가 되었습니다.  대학교에서는 공학적 사고와 논리적 문제 해결 능력을 바탕으로, 다양한 실습과 팀 프로젝트에 적극 참여했습니다. 비록 전공은 AI나 데이터와 관련이 깊지는 않지만 공학적인 사고를 중요시 여기는 교수님들께 “왜” 이렇게 문제를 해결해야 하는지 모델링과 설계를 배웠으며 “어떻게” 사람을 이롭게 할 수 있는 지에 대해 끊임없이 고민하도록 배웠습니다. 대학시절 기억에 가장 남는 과제는 팀프로젝트에서 상수도의 물이 흐를 때 어떤 외력들이 작용할 지 고민하고 어떤 부가 압력이 발생하는 지 토의하고 결과를 산출하며 개인이 놓치기 쉬운 작용력을 팀으로 접근해 풀어나가는 과제입니다. 조장으로서 프로젝트를 주도하며 공학적인 사고를 팀원과 공유함에 따라 어떤 문제든 처음부터 끝까지 흐름을 파악하고 해결하려는 태도는, 함께한 팀원들로부터 신뢰를 받는 기반이 되었습니다.  저는 이러한 성장 배경 속에서 기술과 공학은 원리를 이해하고, 가치를 만들어내는 도구라고 배웠습니다. 그래서 단순히 모델을 구현하는 데 그치지 않고, 데이터의 흐름을 읽고 의미를 해석하며, 그것이 실제 사회에서 어떤 영향력을 가질 수 있는지를 고민하는 AI 엔지니어로 성장하고자 합니다. |

**성격의 장단점을 보여줄 만한 구체적 사례 (단점 작성 시, 보완노력 포함)**

|  |
| --- |
| **[End-to-End, 시작과 끝은 우리가]** |
| 처음 외식업 매장의 운영 책임을 맡았을 때, 대표님께서 하신 말씀이 아직도 기억에 남습니다. “계획은 잘하지만, 결과가 약하다.” 이 말은 곧 **업무의 시작은 능동적이나 끝맺음에 대한 고민이 부족하다는 지적**이었습니다.  사실 이와 유사한 피드백을 과거 군 복무 중 정보장교로 참모 임무를 수행하던 시절에도 들은 적이 있습니다. “스스로 업무를 잘 시작하지만, 결과는 아직 미흡하다”는 평가였습니다. 당시에는 경험 부족 때문이라고 생각하고 지나쳤지만, 사회에서도 같은 지적을 다시 듣게 되자, 비로소 **제 업무 스타일을 돌아보고 반드시 개선해야 할 과제**임을 깨달았습니다.  그 후 저는 피드백을 단순한 지적으로 넘기지 않고, 이를 성장의 기회로 삼았습니다. **업무의 목적을 근본부터 이해하려 노력했고**, 어떤 결과를 도출해야 하는지 목표 지점을 명확히 한 뒤, 필요한 자원과 시간 분배, 업무 흐름을 전체적으로 재설계했습니다. **매뉴얼화와 시스템화를 통해 업무의 효율성과 일관성을 높였고**, 점차 마무리까지 책임지는 습관과 실행력을 갖추게 되었습니다.  이듬해, "끝이 약하다"는 말은 사라졌고, 오히려 “성과 중심으로 전 과정을 책임지는 사람”이라는 긍정적인 평가를 받았습니다.  저는 단점이 없다고 생각하지 않습니다. 하지만 **피드백을 회피하지 않고, 적극적으로 수용하여 발전의 동력으로 바꾸는 태도**는 누구보다 자신 있습니다.  이 경험을 통해 저는 단순히 주어진 업무를 수행하는 것이 아니라, **문제의 흐름을 처음부터 끝까지 이해하고 완성하는 End-to-End형 인재로 성장할 수 있었습니다.** |

**지원서 상의 모든 기재사항은 사실과 다름이 없음을 확인합니다.**

**작성일 : 2025년 09월 03일**

**작성자 : 정 종 혁 ( 인 )**