**PIAI 인공지능연구원 19기 인턴 김재희 인수인계 설명서**

**0. 세부과제명**

**[1차년도]**

**1) 방문 및 소비 데이터 패턴 분석을 위한 데이터베이스 설계 및 개발**

- 방문자 추천 동선 도출을 위한 매장 별 방문자, 방문시간 등의 정보기반 데이터베이스 설계 및 개발

- 소비자 맞춤형 상품 추천을 위한 소비자 별 주문정보 기반 데이터베이스 설계 및 개발

**[2차년도]**

**2) 데이터 학습 및 RNN 기반 예측 알고리즘을 적용한 장래 유동인구 증감 및 수요예측 기술 개발**

- 점포 주변 방문인구 및 매출, 상품 판매 데이터를 활용한 수요 예측 모형 개발

- (경로 분석) 클러스터링 된 상권내 방문인구 이동 흐름 분석

- (상권 상세 정보) 방문인구 정보(방문인구량, 방문패턴, 주요 판매 상품)

- (소비자 객층 분석) 주변 방문인구 객층 및 상품 구매객층 데이터 분석

- (추세 예측) 방문 및 매출 데이터, 공공 데이터 분석을 통한 방문자 증감 추세 예측 및 예상 매출액 추정

**1. 참고논문**

1) Forecasting Daily Visitors and Menu Demands in an Indonesian Chain Restaurant using Support Vector RegressionMachine – Makmur A Zhao.(2020)

· 인도네시아 체인 레스토랑의 일일 고객 수(Number of Visitor Forecasting) 및 메뉴 수요량(Menu Demand Forecasting) 예측

2) 텐서플로우 LSTM 모델을 활용한 음식점 방문자수 예측 – 권혁민, 김화종(2018)

· LSTM모델을 Tensorflow 라이브러리를 활용하여 시계열 데이터로 제공되는 음식점 방문자 수

3) [논문 독후감] LSTM을 이용한 주가예측 모델의 학습 방법에 따른 성능분석

<https://blog.naver.com/id8436/222608855502>

**2. 데이터셋**

- recruit-restaurant-visitor-forecasting

텐서플로우 LSTM 모델을 활용한 음식점 방문자수 예측 – 권혁민, 김화종(2018)

위 국내논문에서 LSTM모델 성능 확인을 위해 사용한 데이터

<https://www.kaggle.com/competitions/recruit-restaurant-visitor-forecasting/code>

|  |  |  |
| --- | --- | --- |
| air | | AirREGI  포스 기능을 수행하는 스마트 앱 |
| hpg | | Hot Pepper Gourmet  레스토랑 찾기, 예약 웹사이트 |
| 데이터셋 | air\_visit\_data.csv | air restaurants의 과거 방문 data |
| air\_reserve.csv / hpg\_reserve.csv | air & hpg 레스토랑 예약 시스템 |
| air\_store\_info.csv / hpg\_store\_info.csv | air & hpg 레스토랑 세부 정보 |
| store\_id\_relation.csv | air & hpg id |
| date\_info.csv | 일본 공휴일 정보 |
| sample\_submission.csv | test set으로 사용. ‘id’는 air id와 방문 날짜를 결합하여 생성 |

- tokyo\_2016\_weather

- tokyo\_2017\_weather

\* kaggle 내 weather 데이터를 검증하기 위한 통계청 도쿄 날씨 데이터(2016, 2017년)

**2. 사용코드**

1) LSTM\_2T.ipynb: LSTM, GRU를 적용한 코드

2) ARIMA\_1T.ipynb: ARIMA 분석을 진행한 코드

데이터 전처리 내역은 주석 및 마크다운을 참고하며, 2022.12.24 개인업무보고 파일 참고 시 더 상세히 확인 가능

**3. 추가상황변수**

상권의 방문자 수 예측에 영향을 끼칠 추가 변수 탐색 과정으로, 일본의 1인당 국내총생산(GDP) 지표 추가 첨부

(분석에 적용은 아직 x)