| **주 간 업 무 일 지(1조)**  **2024년 11월 25일 (월)** | |
| --- | --- |
|  | 업 무 내 용 |
| 금  주 | * 자료조사   + 수집 데이터(파일)   + 물가지수와 농산물 가격은 연관성이 없음   + 유통 비용이 50% 정도 차지함 * 데이터 확보   + 독립변수 : 지역별 날씨, GDP, 한국은행 금리, 시간급, 경유가격   + 종속변수 : 농산품 일별 소매가(쌀, 배추, 무, 사과(후지), 배(신고) , 건고추(화건), 등 총 17 품목) * 서버   + 모바텀으로 개발서버 접속 IP, user, P/W 확인 * 데이터 통합   + 각 농산물 별로 독립변수와 병합해서 예측 * EDA   + 일별로 corr() 상관분석 한 결과는 유의미하지 않음. 0.4 정도   + 평균기온은 5개 지역 차이 거의 없음   + 농산물 대부분은 2021년 이후로 가격이 상승함 * 전처리   + 결측치     - 시장가 조사가 없는 공휴일 : 전에 있는 데이터 값으로 ffill()     - 년별, 분기별 값 : 일별로 변환 후 ffill()   + 피처링1     - 지역별 ‘일 평균기온’, ‘일교차’, ‘일 평균 상대습도’, ‘일 강수량’     - 동두천, 인제, 대전, 광주광역시, 밀양   + 피처링2     - ‘재배기간’, ‘재배기간 내 강수량’ 등을 시도 중 * 시계열을 통한 예측 모델링   + 일별 예측은 런타임 에러   + 주변 예측을 통해서는 향후 1년에 대한 예측 그래프 확인 * PowerBI   + 데이터 간의 상관관계를 시각화를 통해 확인 |
| 차    주 | * 자료조사   + 부대조달 시 농가 계약 가격 확인?     - 국방전자조달시스템([www.d2b.go.kr](http://www.d2b.go.kr/)) - 사업자 필요     - 나라장터([www.g2b.go.kr](http://www.g2b.go.kr/)) * 서버   + GPU 환경 확인 * 데이터베이스   + DB 구축(csv, 엑셀 등 파일을 오라클 SQL로 변환) * 전처리   + 재배기간을 반영할 컬럼 생성   + 딥러닝시 필요한 컬럼 추가, 변형(ex- 딥러닝시 계절성을 학습시키기 위하여 sin\_month,cos\_month,sin\_day,cos\_day 컬럼 추가) * 시계열을 통한 예측 모델링   + 과적합 방지를 위해 모델 수정   + GPU서버 환경에서 일별 예측 실행 * 독립변수를 통한 예측 모델링   + 다양한 모델을 사용하여 모델링 생성(ex-LSTM,Transformer)   + 하이퍼파라미터 수정을 통한 모델 최적화 * PowerBI   + 구축된 DB에서 데이터를 불러와 시각화 * 웹페이지   + 웹사이트 디자인 구상 |
| 비        고 | 멘토 의견   * 기존 주제에서 크게 변경하지 말고 국방과 연결하여 수정하면 좋을 것 같다 * 군부대에서 농산품 구입 시 상(上)품의 국산품을 우선적으로 구입한다 * 중앙조달과 지역조달(부대조달)이 있고 지역조달은 부대 근처 농가와 계약해서 조달한다   이슈 사항   * 시계열 머신러닝 예측결과의 과적합   + (해결책) 모델을 수정해서 과적합 되지 않은 모델 생성 * 독립변수(ex-날씨, 금리 등)를 통한 예측시 환경변수와 타겟값(농산물 일별 소매가)을 시각화 하였을 때 연관성을 크게 발견하지 못함   + 독립변수의 발생일이 타겟값(일변 소매가)에 영향을 주기까지의 시간적 간격이 발생하고 품목마다 조금씩 다름   + (해결책) 딥러닝시 시간적인 컬럼을 추가하여 학습시켜 품목마다 독립변수가 영향을 주는 시간적 간격 또한 학습시킴 * 학습시 독립변수가 얼마나 정확히 영향을 줄 수 있을지 미지수   + (해결책) 하이퍼파라미터를 수정하고, 변수들을 수정해가면서 학습시켜보고 결과를 보면서 변수들을 수정 * GPU서버에서 CUDA버전과 torch버전이 호환되지 않음   + (해결책) 버전 업그레이드를 시도해보고 해결되지 않을 시 CUDA말고 다른 방법을 시도   지원 필요 사항   * 딥러닝 학습 할 수 있게 GPU를 사용할 수 있는 환경 지원 요청 |