

농산물 가격 예측을 위한 기계학습 모델 학습 연구

연구 배경 및 목표

농산물 가격의 가격 변동성을 예측하는 것은 농업인과 정책결정자 들에게 매우 중요합니다. 시계열 분석 방법으로 잘 알려진 ARIMA 모델을 기반으로 농산물 가격 예측 모델을 개발하여 복잡한 요인들 간의 예측 정확도를 높이고자 합니다.

공공데이터를 활용한 정확한 농산물 가격 예측은 다음과 같은 이유로 필요합니다

- 농업인의 소득 안정화: 예측된 가격 정보를 바탕으로 적절한 시기와 방법을 통해 작물을 출하하거나 재배 계획을 수립할 수 있음.
- 정책 결정 지원: 정부의 농업 정책과 정책 수립에 있어, 정확한 데이터를 제공하여 보다 체계적이고 실효성 있는 정책을 가능하게 함.
- 소비자 보호: 가격 예측을 통해 물가 변동에 따른 소비자 부담을 최소화할 수 있음.

정량목표

- 농산물 가격 예측을 위한 ARIMA 모델 기반 기술 확보.
- 공공데이터를 바탕으로 기계학습 기술을 결합하여 신뢰도 높은 예측 결과 도출.

연구 방법 및 결과

데이터 수집 : 농넷 공공데이터 (2018 ~ 2021년의 모든 농산물의 상세 정보 및 가격, 전국 도매 경락 정보 & 산지 공판장 경락 정보)

데이터 전처리 : 품목 가격의 시계열 데이터를 추세(Trend), 계절성(Seasonal) 그리고 '잔차(Residual)' 성분으로 분해합니다.

이를 통해 해당 품목 가격의 장-단기 패턴과 예측에 활용할 수 있는 feature를 추출할 수 있습니다.

누락된 날짜와 기존의 train 데이터를 합친 후, 누락된 값을 채우고, 모든 컬럼의 o값을 앞의 값으로 채워 보정함.

이를 기반으로 각 주차별로 예측할 가격에 대한 컬럼을 추가하여 학습 데이터를 생성합니다.

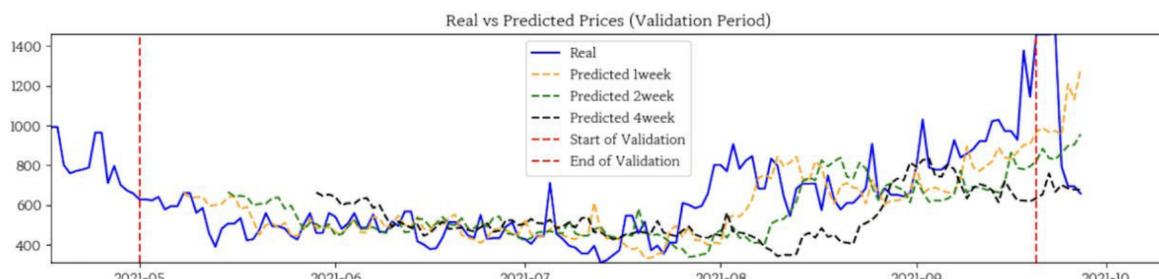
활용 모델과 연구 개발

공공 데이터 플랫폼에서 농수산물 가격 및 관련 데이터 수집을 바탕으로 데이터 정제 및 전처리 작업을 수행합니다.

ARIMA 모델을 위한 데이터를 준비하며, 모델 설계 및 개발, 초기 학습 및 검증 데이터셋을 활용하여 실험을 수행합니다.

모델 성능 평가 및 파라미터 튜닝을 진행하고, 모델 최적화 작업을 수행합니다.

테스트 데이터셋을 통한 성능 평가 및 개선을 하며, 최종 모델 성능 확인 및 주요 성과를 도출합니다.



ARIMA 모델로 확인한 해당 데이터셋 학습 결과 - 예측 거동 시각화 함수를 통한 예측 값과 실제 가격의 비교

결과물의 활용 방안 및 기대 효과

정책 수립 지원 도구: 농업 정책 결정 시, 가격 예측 데이터를 활용하여 가격 안정화 대책 및 긴급 조치 마련에 도움을 줄 수 있음.

농업인 대상 예측 시스템: 농업인들이 이용할 수 있는 가격 예측 플랫폼을 제공하여 재배 및 출하 계획 수립에 실질적인 도움을 줄 수 있음.

소비자용 물가 정보 서비스: 일반 소비자들에게 실시간으로 농산물 가격 변동 정보를 제공하여 합리적인 구매 결정을 지원할 수 있음.

정책적 결정 지원: 정책 결정자들이 신속하고 효과적인 농업 지원 및 보상 정책을 수립하는 데 중요한 근거를 제공.

농업 생산성 및 효율성 증가: 농업인들이 정확한 예측 데이터를 바탕으로 재배 및 출하 시기를 조절하여 경제적 이익을 극대화할 수 있음.

소비자 물가 부담 경감: 가격 변동에 대비한 정보를 통해 소비자들이 농수산물 구매 시기를 조절할 수 있어 물가 부담을 줄일 수 있음.

향후계획

ARIMA 모델을 기준으로 Neural ODEs 모델과 같은 다양한 기계학습 모델을 학습할 예정입니다.

새로운 기계학습 모델을 통해 기존의 모델과 성능 평가를 비교하여 시각화하고, 해석 가능한 결과를 도출할 것입니다.