



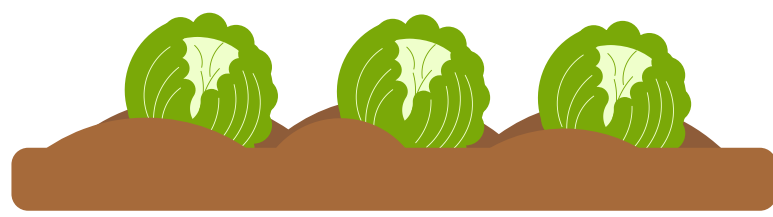
# 기상변화에 따른 배추 가격 예측

빅데이터융합전공 박소현 · 안현정 · 장원혁 / 통계데이터사이언스전공 정지원

## 1.분석 배경 및 목적

파이낸셜뉴스 “폭염, 가뭄에 배춧값 급등”...정부, 중국산 배추 직접 수입한다.” 24.09.24

- 정부가 기록적 폭염과 가뭄의 여파로 가격에 크게 오른 배춧값 안정을 위해 중국산 배추를 긴급 수입한다.
- 현재 여름 배추는 재배면적이 감소한 데다 가뭄과 고온의 장기화로 작황이 부진한 상태다.



배추는 우리나라 주요 식재료 중 하나로, 특히 김장철과 명절 기간 동안 수요가 급증한다. 그러나 최근 기후 변화와 지구 온난화로 인해 배추 재배량과 재배면적이 지속적으로 감소하며, 배추의 수급 불안정 문제가 심화되고 있다. 이러한 수급 불안정은 김장철 배추 가격의 급등으로 이어져 소비자와 농가 모두에게 부정적인 영향을 미친다. 실제로 2024년 9월에는 비축 물량이 전량 소모된 가운데 배추 가격이 전월 대비 48% 상승하는 사례도 나타났다.

배추 가격은 기상 요인, 물가지수, 수입량, 무 가격 등 다양한 요인의 영향을 받는다. 기존 연구에서는 배추와 무의 대체재 관계를 기반으로 가격 변동의 상관성을 분석한 바 있다. 하지만 생산량 예측의 진동성과 복잡성으로 인해 단순 통계나 머신러닝만으로는 한계가 있었다. 이에 본 연구는 RNN, LSTM, BiLSTM 등의 시계열 분석 모델을 활용해 기상 요인이 배추 공급량에 미치는 영향을 체계적으로 분석하고, 이를 기반으로 배추 가격을 예측함으로써 소비자 부담 완화와 농산물 시장 안정에 기여하고자 한다.

## 2.데이터 전처리

변수명	데이터 내용	연도	출처
배추 공급량	전국 도매시장 경락 데이터	2020-2024 일별 데이터	농산물유통 종합정보시스템
기상 요인	평균/최저/최고기온, 일강수량, 강수 계속시간, 최대/평균 풍속, 평균 상대습도, 일사량, 일조시간, 일최심신적설	2018-2024 일별 데이터	기상청 기상자료개발포털
배추 가격	주요 5개 도시 상품 중도매가(단위:10kg)	2020-2024 일별 데이터	한국농수산식품 유통공사
무 가격	주요 5개 도시 상품 중도매가(단위:20kg)		
배추 수입량	배추 수출입 데이터(단위:톤)	2020 - 2024 월별 데이터	농산물유통 종합정보시스템
소비자물가지수	2020년도 100 기준		

### 공급량 예측 데이터

2020년 1월부터 2024년 11월까지의 전국 도매시장 경락 데이터 사용  
주요 도매지역(서울, 부산, 대구, 대전, 광주)의 일별 배추 총거래물량 산출  
거래단위(kg)와 총거래물량(개)을 곱하여 총거래물량 계산  
음수 값은 절댓값으로 변환하고, 상위 1% 이상치를 제거한 뒤 선형보간으로 결측값 처리  
품종은 시장의 전반적 공급량을 반영하고 모델의 복잡도를 줄이기 위해 통합  
기상 데이터는 2018년부터 2024년까지의 전국 평균값(16개 도별 평균) 사용  
지역과 날짜 기준으로 배추 공급량 데이터와 기상 데이터를 병합하여 데이터 프레임 구성

### 가격 예측 데이터

공급량 예측 데이터를 바탕으로 주요 5개 도시의 중도매가 데이터 병합  
배추와 무의 가격 데이터는 한국 농수산식품유통공사에서 제공한 일별 데이터 활용  
주말 및 공휴일에 가격이 책정되지 않은 경우 선형보간으로 결측값 처리  
배추 수입량 데이터와 소비자물가지수는 월별 데이터를 일별로 변환하여 포함  
배추 가격, 배추 공급량, 무 가격, 수입량, 물가지수 변수를 포함한 데이터 프레임 구성

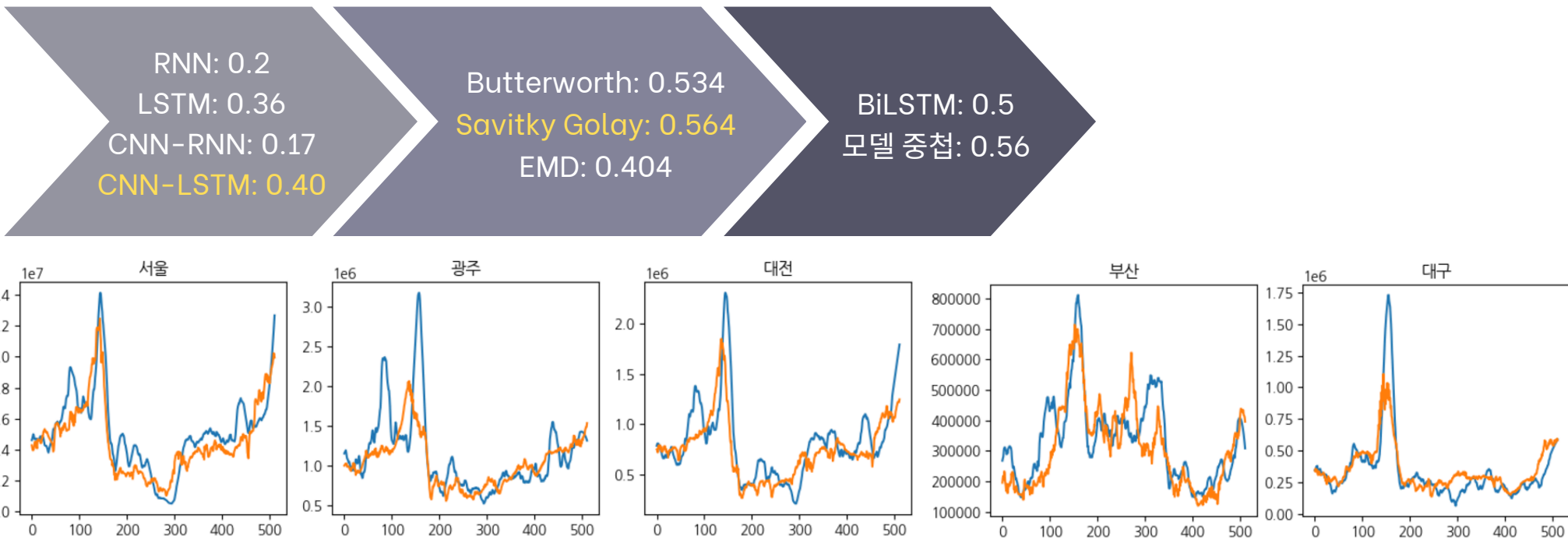
## 4. 데이터 분석

### 배추 공급량 예측

2020년~2024.10월까지의 서울, 대전, 대구, 부산, 광주의 일별 배추 공급량 데이터와 기상 데이터를 바탕으로 시계열 분석을 진행하여 각 지역의 배추 공급량을 예측했다.

-> RNN, LSTM, CNN-RNN, CNN-LSTM 4가지 모델을 이용한 결과 CNN-LSTM이 결정계수(R<sup>2</sup>) 0.40으로 가장 우수함을 확인했다.

이후 진동제어를 위해 저주파수 통과필터인 버터워스필터(Butterworth Filter)와 사비츠키-골레이필터(Savitzky-Golay Filter), 고주파에서부터 저주파 순서로 시계열을 분해하는 EMD를 적용해보았고, 모델 정밀도 개선을 위해 BiLSTM과 모델 중첩을 이용해 성능을 높이기 위한 시도를 했다.



사용 모델 : CNN-LSTM  
목적 시계열: 개별 지역의 배추 공급량  
공변량: 평균/최고/최저기온, 일강수량, 강수 계속시간, 최대/평균 풍속, 평균 상대습도, 일사량, 일조시간, 일 최심신적설  
train&test: 2020.01.02-2023.06.02. & 2023.06.03-2024.11.18.  
Scaler: Min-Max Scaler  
Filter: Savitzky-Golay  
평가지표: MAE, MSE, RMSE, R2

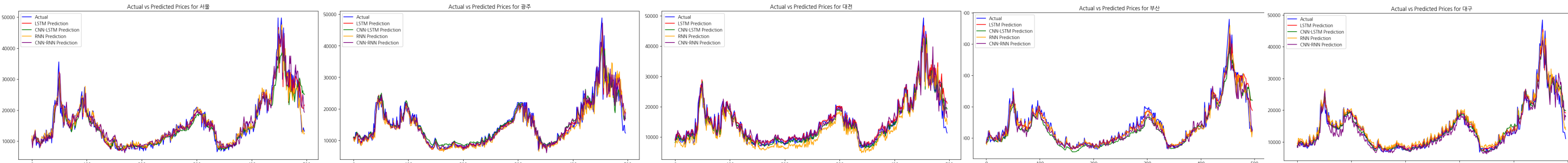
지역	MAE	MSE	RMSE	R2
서울	1.17e+06	2.00e+12	1.42e+06	0.68
광주	2.26e+05	1.39e+11	3.73e+05	0.42
대전	1.84e+05	6.55e+10	2.56e+05	0.54
부산	7.37e+04	8.88e+09	9.42e+04	0.51
대구	8.65e+04	2.29e+10	1.51e+05	0.69

최종적으로 CNN-LSTM에 timestep=80에 Savitzky-Golay 필터를 쓴 결과가 가장 우수함을 확인했다.

### 배추 가격 예측

#### 각 지역에 대한 로컬 모형

사용모델: RNN, LSTM, CNN-RNN, CNN-LSTM  
목적 시계열: 배추의 가격  
공변량: 물가지수, 배추 수입량, 공급량, 무가격  
train&test: 2020.03.22-2023.06.13. & 2023.06.13-2024.10.31  
Scaler: Min-Max Scaler  
평가지표: MAE, MSE, RMSE, R2



분석을 진행한 5개 지역(서울, 부산, 대구, 대전, 광주) 모두 가격 예측 변동성에서 유사한 추세를 보였다.

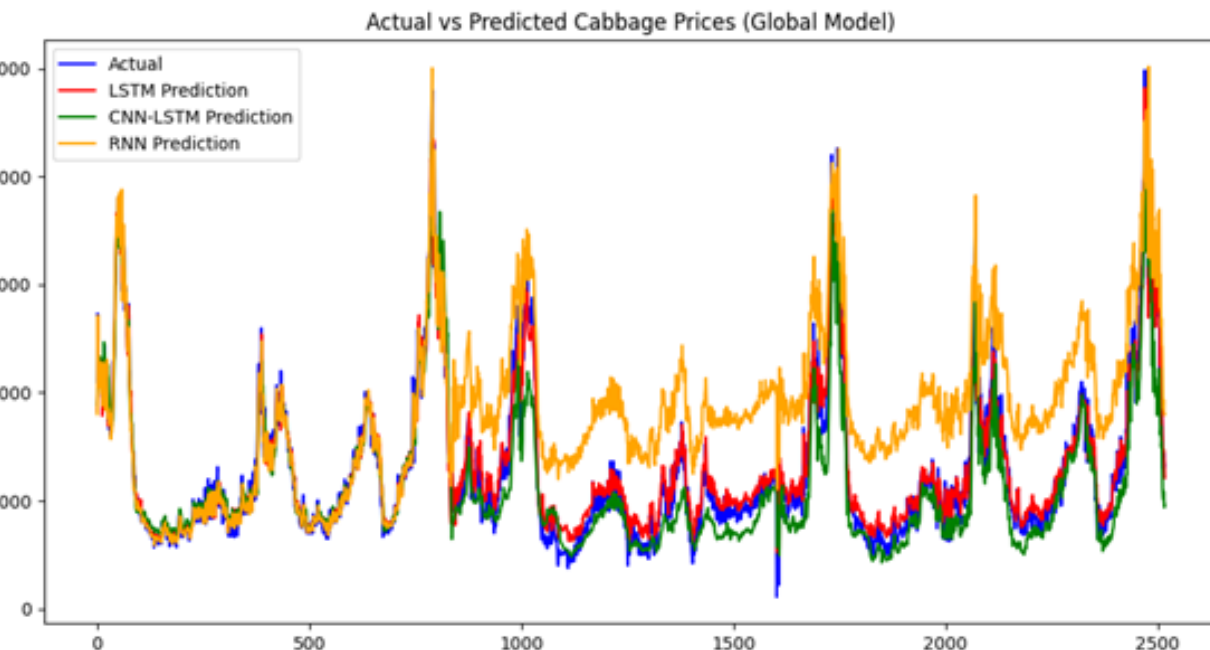
이에 5개 지역을 통합하여 글로벌 모델을 활용한 시계열 예측을 진행하였으며, 그중에서 LSTM 예측 모델의 설명력이 약 94%로 가장 우수한 성능을 보였다.

#### 모든 지역에 대한 글로벌 모형

사용모델: LSTM  
목적 시계열: 배추의 가격  
공변량: 물가지수, 배추 수입량, 공급량, 무가격

train&test: 2020.03.22-2023.06.13. & 2023.06.13-2024.10.31.  
Scaler: Min-Max Scaler  
평가지표: MAE, MSE, RMSE, R2

## 5. 분석 결과



### 공급량 예측

기상 데이터를 활용하여 배추 공급량과 가격을 예측하였으며, 주요 기상 요인들이 배추 생산량 변화의 핵심 동인으로 작용함을 확인하였다. Savitzky-Golay 필터를 적용한 CNN-LSTM 모델은 데이터의 노이즈를 제거해 평균 절대 오차(MAE)와 평균 제곱 오차(MSE)를 크게 감소시키고, 결정계수(R<sup>2</sup>)를 0.6 이상으로 향상시켰다. 이는 필터를 통해 데이터의 주요 패턴에 모델이 집중할 수 있었음을 시사한다.

### 가격 예측

글로벌 모델을 사용한 배추 가격 예측에서는 지역별 모델 대비 성능이 유의미하게 개선되었으며, 결정계수(R<sup>2</sup>)가 0.94를 기록하며 높은 설명력을 보였다. 이는 글로벌 모델이 지역 간 상호작용과 전반적인 공급량 변화를 통합적으로 반영함으로써, 특정 시점의 급등락 현상을 보다 정확히 해석할 수 있었기 때문이다.

결과적으로, 기상 요인을 고려한 공급량 예측과 글로벌 모델의 통합 분석을 통해 배추 가격 변동성을 효과적으로 제어하고, 농산물 시장의 안정성을 높이는 데 기여할 수 있음을 보여주었다.

### 평가지표

Model	MAE	MSE	RMSE	R2
LSTM	1242.3	2992358	10486070.0	0.94
CNN-LSTM	1951.5	7881176	2807.3	0.85
RNN	5547.3	44393450	6662.8	0.18
CNN-RNN	2391.5	10486070	3238.2	0.80

## 6. 기대효과 및 활용 전략

- 농업 생산성 향상과 경제적 안정화에 기여할 수 있는 중요한 기반을 마련함.
- 농민과 소비자, 정책 입안자 등 다양한 이해관계자들에게 실질적인 도움을 제공함.
- 농업 정책 수립 및 공급관리에서의 의사결정을 지원하는 데 활용될 수 있으며 농업 분야의 전반적인 발전에 기여할 수 있음.
- 본 연구에서 제시된 접근법은 배추뿐만 아니라 다른 농작물로의 확장 가능성을 가지고 있어 농업 분야의 지속 가능한 발전과 효율성을 높이는 데 기여할 수 있음.

- 농민: 수익 안정화, 생산비 절감
- 유통업체: 물류비 및 재고비 최적화
- 정부: 가격 안정화 정책 효율성 강화, 소비자 물가 관리, 위기 대응 체계 개선
- 소비자: 가계 부담 경감, 가격 정보 제공으로 소비자 만족도 향상, 다양한 대체 소비 전략 개발

## 7.한계점 및 개선방안

### 지역별 생산 특성과 계절적 변동 반영하지 못함

- 데이터 부족으로 예측에 활용되는 데이터의 기간이 짧았고, 배추 생산량 데이터가 연 단위로 제공되어 이를 예측에 이용했다.
- 일별 생산량 데이터 및 품종별 데이터를 포함한 다양한 시점의 생산량 데이터를 확보한 후 이를 바탕으로 생산량 예측을 진행하고 이후 공급량 예측으로 이어지는 접근 방식을 이용하면 가격 메커니즘에 부합하는 연구가 가능할 것이다.

### 정책 변수를 고려하지 못함

- 배추 수출입 데이터만 반영했는데, 정부의 농업 보조금, 수급 조절 정책 등의 데이터를 더 확보해 정책 요인을 추가해 수요 변화를 분석하면 예측력을 강화시킬 수 있을 것이다.

### 발전가능성

- 예측 모델의 적용 범위를 배추 외 다른 농작물로 확장하여 농업 분야의 다양한 문제를 해결할 수 있는 일반화된 예측 시스템을 구축할 수 있다.
- 클라우드 기반의 실시간 데이터 수집 및 분석 플랫폼을 구축하여 다양한 이해 관계자들이 실시간으로 데이터를 활용하고 변화에 신속히 대응할 수 있는 체계를 마련할 수 있다.