

RESEARCH ARTICLE

대규모 언어 모델 기반 농산물 가격 예측: 시계열 데이터의 텍스트 표현 접근법

최상혁[†], 박상춘[†], 김태형^{*}

서울대학교 바이오시스템공학과

Enhancing Agricultural Price Forecasting Based on Large Language Models: A Textual Representation Approach to Time-Series Data

Sang-hyeok Choi[†], Sang-chun Park[†], Tae-hyeong Kim^{*}

Department of Biosystems Engineering, College of Agriculture and Life Sciences, Seoul National University, Seoul, 08826, Republic of Korea

*Corresponding author: taehyeong.kim@snu.ac.kr

[†]These authors equally contributed to this study as first author.

Abstract

Agricultural price volatility poses significant challenges for producers, distributors, and consumers, underscoring the need for improved forecasting methods. Traditional statistical and machine learning techniques often struggle to capture the complex, dynamic and unpredictable nature of agricultural markets, which motivates the exploration of alternative approaches. In this study, we propose a novel method that converts time-series data into text prompts to enhance forecast accuracy by utilizing the language processing capabilities of large language models (LLMs). We analyze daily trading prices for 23 crops over 1,862 days to predict price changes over the subsequent 185 days under stable conditions. Unlike traditional methods, our approach employs customized prompt template to convert numerical data into text, enabling LLMs to more effectively interpret and forecast market trends. A comparative analysis with Holt-Winters, LSTM, GRU, and Transformer models demonstrates that our LLM model (Llama 3) outperforms traditional methods, achieving a MAE of 913.61, RMSE of 1477.74, MAPE of 28.89, and an R^2 of 0.6494. These findings suggest that converting time-series data into text prompts enables the use of LLMs, substantially improving agricultural price forecasting and leading to more accurate and data-driven market decisions.

Keywords: Agricultural Price Prediction, Time Series Forecasting, Large Language Models



OPEN ACCESS

Journal of Agricultural Machinery Engineering
5(1):37-45

DOI: <https://doi.org/10.12972/jame.2025.5.1.4>

Received: March 11, 2025

Revised: March 29, 2025

Accepted: March 30, 2025

Copyright: © 2025 Korean Society for
Agricultural Machinery



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

농산물 가격은 기후 변화, 국제 정세, 정책 변화 등 다양한 요인이 복합적으로 영향을 미치기 때문에 큰 변동성을 나타내며, 이는 농업 생산자뿐만 아니라 유통업자, 소비자 모두에게 경제적 부담과 불확실성을 초래한다(Kim, 2024). 특히 농산물은 소비 빈도가 높고 가계 지출에서 차지하는 비중이 커, 가격 변동이 가계 경제에 직접적인 영향을 미친다. 이에 따라 농산물 가격 변동은 단순한 경제적 문제를 넘어 사회 전반의 중요한 이슈로 부각되고 있다(KREI, 2013). 따라서 농산물 가격 예측의 정확성을 높이는 것은 안정적인 농업 및 식품 공급망 관리와 정책 결정, 그리고 소비자 보호에 필수적이다(KREI, 2024).

전통적으로 농산물 가격 예측에는 통계적 모델이나 회귀 분석 기법이 활용되어 왔으나, 이러한 방법들은 과거 데이터의 경향성을 반영하는 데 초점을 맞추며 비선형적인 요인을 충분히 고려하지 못하는 한계를 가진다. 또한 최근에는 인공지능 기반의 딥러닝 기법이 활발히 사용되고 있지만(Chen et al., 2021), 이는 수치 데이터에만 의존하여 가격 변동의 근본적인 원인을 반영하지 못하고, 단순히 상관관계만을 학습함으로써 예측 정확도의 한계를 보일 수 있다.

최근 대규모 언어 모델(Large Language Models, LLM)은 자연어 처리 분야에서 뛰어난 성능을 보일 뿐만 아니라, 수치 데이터의 추세를 고려하여 패턴을 인식하는 능력을 갖추고 있기 때문에 시계열 예측 분야에서도 새로운 접근법으로 주목받고 있다. 예를 들어, Zhou et al. (2023)은 자연어 분야에서 사전 학습된 대규모 언어 모델을 전력 수요, 변압기 온도, 강수량 예측 등의 시계열 예측에 적용하였으며, Xue et al. (2023)은 프롬프트 템플릿(prompt template)을 구성하여 LLM의 입력으로 활용함으로써 시계열 예측 성능을 향상시켰다. 또한, Gruver et al. (2024)은 텍스트 기반의 Zero-shot 예측을 시도하였고, Jin et al. (2024)은 임베딩 계층과 패치 재프로그래밍 기법, 접두사형 프롬프트 등을 활용하여 시계열 예측에서 LLM의 강력한 성능과 그 확장성을 제시하였다. 이러한 선행 연구들은 LLM이 텍스트를 기반으로 시계열 데이터의 고유한 특성을 효과적으로 포착할 수 있음을 보여주며, 이를 다양한 분야에 확장 적용할 수 있는 가능성을 시사한다.

본 연구에서는 LLM의 자연어 처리 능력을 활용하여 농산물 가격을 예측하는 방법을 제안한다. 해당 접근 방식은 LLM이 언어적 맥락을 바탕으로 작물별 상이한 특성을 효과적으로 인식할 수 있음을 보여주며, 기존 통계 및 딥러닝 모델의 한계를 보완한 더 정확하고 범용적인 농산물 가격 예측 모델을 제시할 것으로 기대된다. 궁극적으로는 예측 정확도 향상을 통해 농산물 시장의 가격 안정 및 농가 수익 변동성 완화에 기여할 수 있을 것으로 전망된다.

Materials and Methods

데이터 수집 및 전처리

한국농수산식품유통공사(aT) 도매시장 통합 홈페이지 API를 통해, 전국 32개 공영 도매시장에서 거래된 농산물의 일별 거래량(kg)과 거래액(원) 데이터를 확보하였다. 수집 기간은 2018년 1월 1일부터 2023년 12월 31일까지이며, 이 기간 동안의 일자별 거래 기록을 32개 도매 시장에 대해 합산하여 해당 일자의 총 거래량과 총 거래액을 구하였다. 다만, 공휴일이나 일요일 등 실제로 거래가 이루어지지 않은 날은 제외하고 집계하였으며, 이후 총 거래액을 총 거래량으로 나누어 kg당 평균 가격(원/kg)을 산출하였다. 계산된 값은 원 단위에 맞추기 위해 소수점 이하는 반올림하여 정수 형태로 변환하였다.

시장에서 활발히 유통되는 품목일수록 경제적 의미가 크다는 점을 고려하여, 전체 품목 중 사과, 호박, 무 등 거래량을 기준으로 상위 30개 품목을 선정하였다. 또한, 수입 과일(오렌지, 바나나, 블루베리, 파인애플)과 계절성이 극단적으로 나타나는 일부 품목(딸기, 복숭아, 단감)은 정책 및 계절적 패턴 등의 이유로 제외하였다. 최종적으로, 23개 작물에 대해 1,862일간의 평균 거래 가격 데이터를 확보하였고, 모든 작물의 시계열 데이터를 통합하여 개별 작물 간 상호작용을 반영하는 범

용 예측 모델을 구축할 수 있도록 데이터를 구성하였다. 해당 전처리 과정을 요약한 도식은 Fig. 1에 제시하였고, 각 품목별 주요 통계치(평균, 표준편차 등)는 Table 1에 정리하였다.

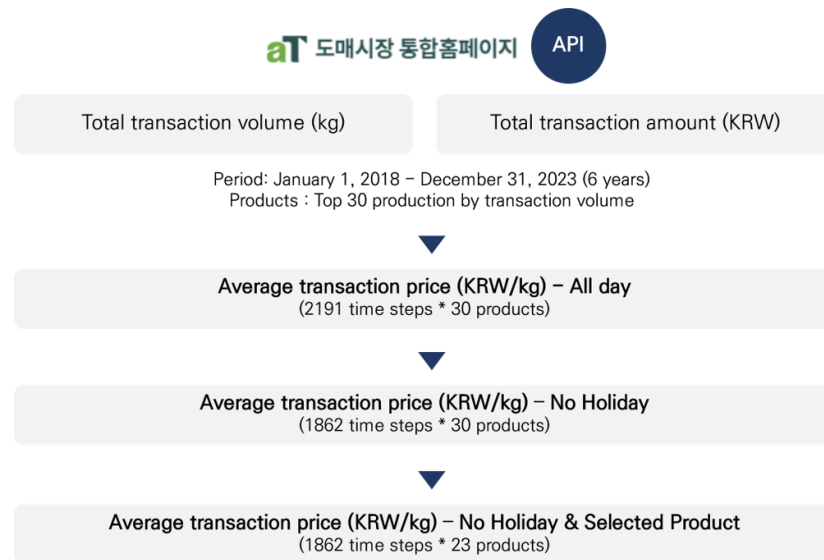


Fig. 1. Overall process of data preprocessing.

Table 1. Statistical characteristics of price data for 23 crops collected over 1,862 days

Crop	Mean	STD	Min	Max
Apple	2413.52	669.79	811	5345
Onion	805.57	344.92	274	1875
Napa cabbage	704.85	320.80	241	3066
Grapes	5432.22	1921.27	2146	31538
Mandarin orange	3777.95	1874.92	498	8984
Green onion	1494.22	762.65	510	5440
Cucumber	2087.50	869.92	742	5868
Cherry tomato	3919.65	1258.91	1479	8871
Watermelon	2001.82	812.30	628	6690
Lettuce	3638.42	2558.49	1007	18409
Tomato	2385.34	874.15	717	7348
Potato	1412.60	619.54	530	5053
Pumpkin	1794.21	734.95	505	5343
Green chili pepper	4048.28	2045.65	1436	16352
Sweet potato	2155.59	691.77	133	4552
Radish	533.92	241.08	208	1801
Cabbage	594.17	258.92	108	1653
Chives	2949.98	1479.11	325	9375
Perilla leaf	6374.42	2572.20	2308	29040
Spinach	3185.55	2213.54	761	15329
Bell pepper	3831.59	1585.54	1144	13059
Carrot	1114.94	270.62	611	3750
Scallion	3541.90	1703.51	838	13887

시계열 데이터의 텍스트 변환

농산물 가격 데이터를 LLM에 입력하기 위해, Fig. 2와 같은 프롬프트 템플릿을 사용하였다.

Prompt template
From 2018-01-05, the daily price of Apple per kilogram is 1878, 1787, 1837, 1804, 1880, 2126, 1885, 1857, 1925, 1828, 1889, 1869, 1844, 1711, 1859 won each day. What will be the price next?

Fig. 2. Prompt template used for LLM tuning

구체적으로, 과거 15일간의 가격 데이터를 텍스트 형태로 변환하여 모델에게 제공하고, 다음날 가격 예측을 자연어로 요청하는 지시문(instruction)을 설계하였다. 모델은 이에 대한 답변을 텍스트로 반환하며, 이 중 수치 정보를 파싱(parsing)하여 LLM의 최종 예측값으로 활용하였다.

LLM 기반 시계열 예측 모델

본 연구에서는 Meta의 오픈 소스 언어 모델인 Llama 3-8B (Dubey et al., 2024)를 활용하였다. 해당 모델은 대규모 텍스트 데이터를 통해 선학습되어 뛰어난 자연어 처리 능력을 보유하고 있다. 특히 오픈 소스(open source)로 제공되기에 접근성과 확장성이 뛰어나며, 연구 및 실제 응용 환경에 맞도록 미세 조정이 가능하다는 점에서 농산물 가격 변동 예측에 적합하다고 판단하였다.

LLM을 농산물 가격 예측에 활용하기 위해 두 가지 학습 기법을 적용하였다. 지시 조정(Instruction tuning; Wei et al., 2022) 기법을 활용하기 위해 농산물 가격 데이터를 프롬프트 템플릿을 이용해 텍스트 형태로 변환한 뒤 모델을 학습시켰다. 이를 통해 LLM이 자연어 형태의 프롬프트를 입력받아 가격 예측을 수행하는 능력이 강화되도록 하였다. 동시에, LoRA (Low-rank adaptation; Hu et al., 2022) 기법을 도입하여, LLM의 모든 매개변수를 업데이트하지 않고도 저차원 행렬만을 학습함으로써 효율적인 학습이 가능하도록 하였다.

이러한 접근을 통해 LLM은 시계열 예측에 필요한 패턴을 효율적으로 학습할 수 있으며, 본연의 자연어 이해 능력을 유지하면서도 농산물 가격 변동이라는 특수한 영역에 효과적으로 적응하도록 설계되었다.

실험 설계

농산물 가격 변동을 예측 성능을 비교하기 위해 LLM뿐만 아니라, 기존 시계열 예측 분야에서 널리 활용되는 다양한 모델들을 선정하였다. 먼저, Holt-Winters (Holt, 1957) 모델은 추세와 계절성을 고려한 지수 평활법을 확장한 방식으로, 계절적 패턴이 뚜렷한 농산물 가격 변동을 효과적으로 예측할 수 있다. 또한, 구조가 비교적 단순하기에 구현과 해석이 용이하며, 시계열 예측 분야에서 널리 활용되는 점을 고려하여 본 실험에 활용하였다.

딥러닝 기반 모델로는 LSTM (Long short-term memory; Hochreiter and Schmidhuber, 1997)과 GRU (Gated recurrent unit; Chung et al., 2014), 그리고 Transformer (Vaswani et al., 2017)를 선정하였다. 두 모델은 기존 순환신경망의 한계를 극복하기 위해 고안되어, 장기적인 패턴을 효과적으로 학습할 수 있다. 또한, Transformer는 Attention 메커니즘을 활용하여 시계열 예측에서 우수한 성능을 보이는 것으로 보고되어 왔으며, 다양한 응용 분야에서 활용되고 있기에 비교 모델로 채택하였다.

비교 모델의 성능은 하이퍼파라미터(hyperparameter) 설정 및 구조에 따라 달라질 수 있으므로, 각 모델에 대해 다양한 설

정을 적용하여 학습을 수행하였다. 이후 검증 데이터셋에서 가장 우수한 성능을 보인 구성을 최종 비교 모델로 선정함으로써, 제안한 모델과의 성능 비교의 타당성을 높였다.

총 1,862일간의 농산물 가격 데이터를 시계열 분석에 활용하였으며, 이 중 약 90%에 해당하는 1,677일은 모델 학습에, 나머지 약 10%인 185일은 모델의 예측 성능을 평가하는 데 사용하였다. 시계열 자료에서 무작위 재배열을 적용할 경우 시간적 순서가 뒤섞일 수 있으므로, 과거부터 미래로 이어지는 시간 순서를 그대로 유지한 채 데이터를 분할하였다.

모든 예측 모델은 15일 단위로 구성된 가격 정보를 입력으로 받아, 다음 날(1일 후)의 가격을 예측하도록 훈련하였다. 학습이 완료된 모델은 평가 구간(185일)에 대해 재귀적 다중 스텝 예측(recursive multi-step forecasting)을 수행한다. 먼저, 초기 15일 동안의 실제 데이터를 입력받아 첫 예측 값을 생성한 뒤, 이후 시점부터는 새로 예측된 값을 입력 데이터에 포함하여 다음 날의 가격을 연쇄적으로 예측한다. 이 과정을 185일 동안 반복하여 전체 평가 구간의 가격 변동을 추정한다.

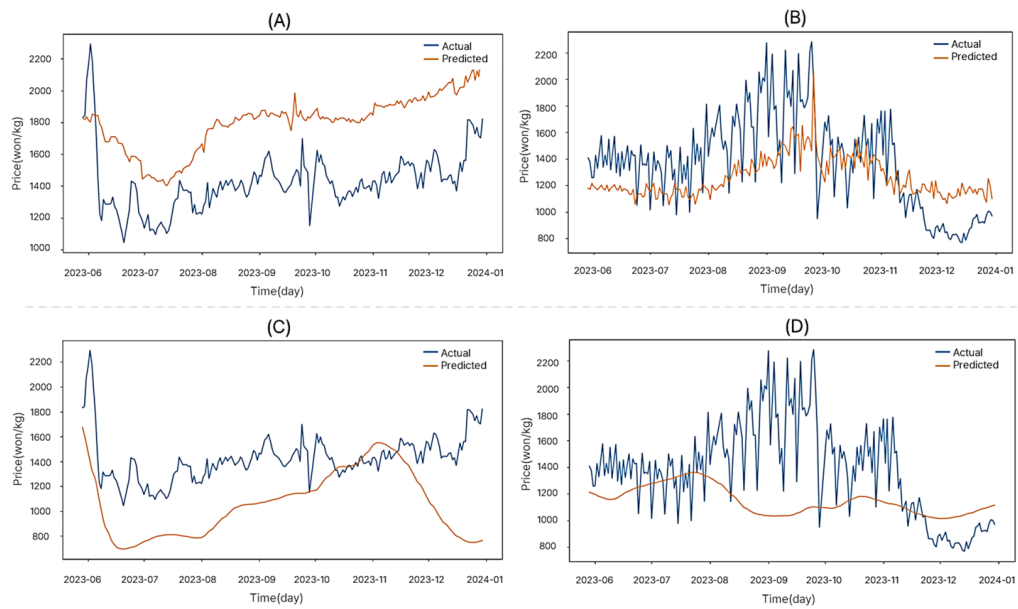
평가 방법

예측 성능에 대한 평가는 크게 네 가지 지표인 평균절대오차(MAE), 평균제곱근오차(RMSE), 평균절대백분율오차(MAPE), 결정계수(R^2)를 이용해 정량적으로 측정하였다. 특히 MAPE는 예측값과 실제값의 비율 차이를 백분율로 나타내어, 작물별 가격 수준이 달라도 상대 비교가 가능하다는 장점 때문에 결과 섹션에서 중점 지표로 활용되었다. 이를 통해 LLM을 포함한 다양한 모델의 농산물 가격 변동 예측 성능을 종합적으로 분석하고 적합성을 비교하였다.

Results and Discussion

농산물 가격 예측 성능 비교

제안한 LLM 기반 방법과 비교 모델들의 가격 예측 성능을 Table 2에 제시하였으며, Holt-Winters, LSTM, GRU, Transformer, LLM 순으로 개선되는 경향을 보였다. 또한 당근과 감자는 농산물 시장에서 경제적으로 중요한 위치를 차지하는 작물로, 계절성, 가격 변동성, 거래 패턴이 서로 다르게 나타나는 대표적인 사례로써 예측 결과를 Fig. 3에 시각화하였다.



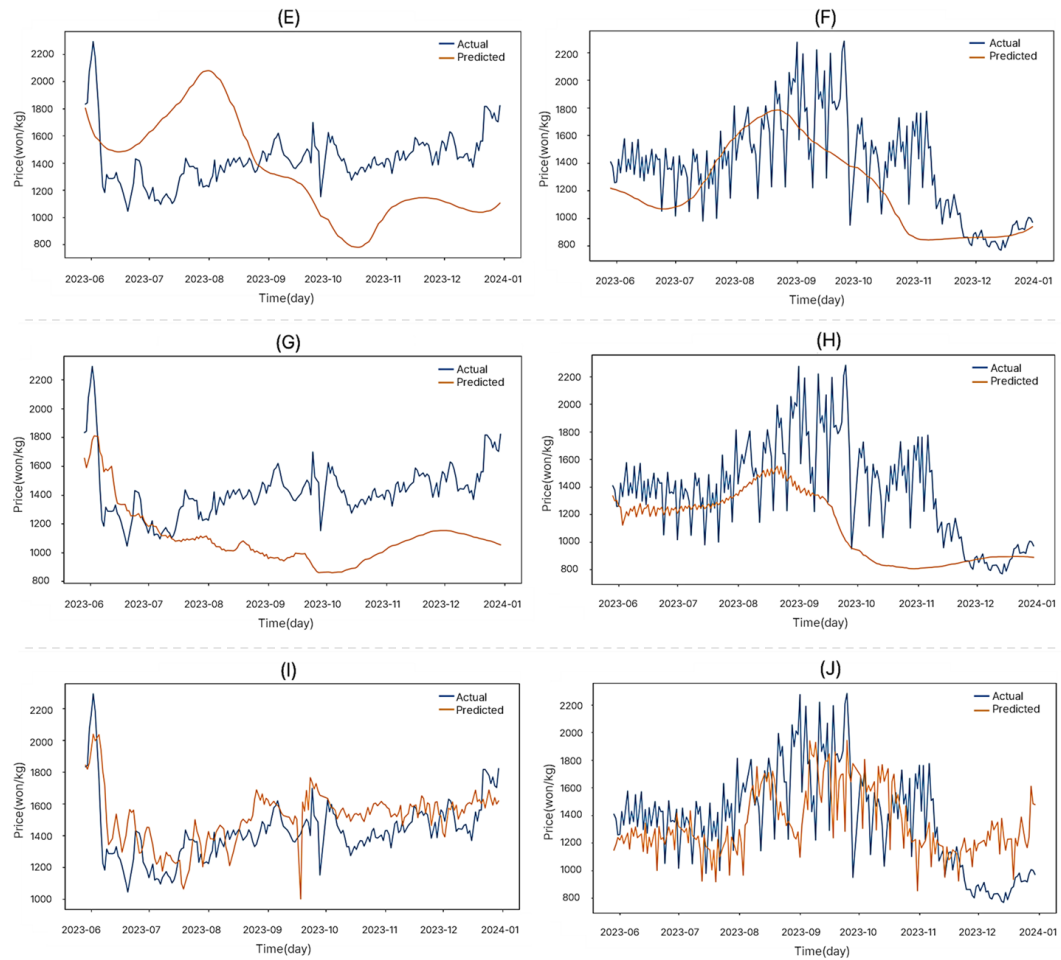


Fig. 3. Visualization of price forecasting results for potatoes (left) and carrots (right) using different models. Plots (A, B) display results from the Holt-Winters model, (C, D) from LSTM, (E, F) from GRU, (G, H) from Transformer, and (I, J) from our proposed LLM-based model. The blue line represents actual prices, while the orange line represents predicted values. The variations in prediction patterns across models illustrate their differing abilities to capture price trends and fluctuations.

Table 2. Evaluation of prediction results across different models for 23 crops

Evaluation metrics	Traditional models				LLM
	Holt-Winter	LSTM	GRU	Transformer	
MAE	1293.62	1096.23	1100.33	1022.83	913.61
RMSE	2270.66	1763.93	1724.47	1640.39	1477.74
MAPE	40.19	32.52	34.18	29.66	28.89
R ²	0.1722	0.5005	0.5226	0.568	0.6494

Holt-Winters 모델의 MAPE 값은 감자에서 27.89, 당근에서 19.95로 나타났으며, 해당 예측 결과는 Fig. 3(A, B)에 제시되었다. Holt-Winters 모델은 일정 수준의 추세와 계절성을 학습하였으나, 실제 가격 변동을 정밀하게 반영하는 데에는 한계를 보였다.

딥러닝 기반 모델인 LSTM과 GRU의 MAPE 값은 각각 감자에서 26.82와 29.82, 당근에서 21.36과 17.14로 나타났으며, Fig.

3(C-F)에서 두 모델의 예측값과 실제값의 비교를 확인할 수 있다. 두 모델 모두 전체적인 추세는 따라가지만, 정밀한 값의 변동을 제대로 반영하지 못하고 지나치게 단순한 예측 곡선을 보이는 경향이 나타났다. Transformer 모델의 MAPE 값은 감자에서 24.6, 당근에서 20.29로 나타났으며, Fig. 3(G, H)에서 해당 모델의 예측 추세를 확인할 수 있다. Transformer는 세 딥러닝 모델 중 가장 높은 성능을 기록했으며, 상대적으로 안정적인 예측 패턴과 함께 약간의 가격 변동성을 반영한 것을 확인할 수 있다.

제한한 LLM은 감자에서 10.48, 당근에서 21.24의 MAPE 값을 기록하며, 다른 예측 모델들과 비교하여 가장 낮은 오차를 보였다. Fig. 3(I, J)에서 확인할 수 있듯이, LLM은 기존 통계 및 딥러닝 기반의 모델들보다 실제 가격 변동 패턴을 더욱 효과적으로 반영하였다. 또한, Fig. 4의 MAPE 및 R^2 분석에서도 LLM 모델은 각각 28.89 및 0.6494의 성능을 기록하며 가장 높은 예측력을 입증하였다.

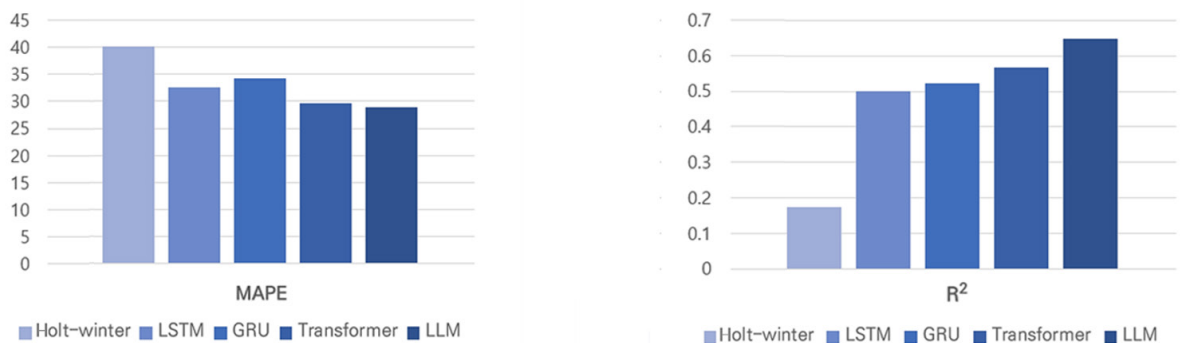


Fig. 4. Comparison of MAPE (left) and R^2 (right) for overall agricultural price prediction by each model

결과에 대한 논의

본 연구에서는 전통적인 통계 기반 모델부터 딥러닝 및 대규모 언어 모델에 이르기까지 다양한 방법을 활용하여 농산물의 가격 변동을 예측하고자 하였다. 이를 통해 각 모델이 포착할 수 있는 시계열 패턴의 범위와 한계를 분석하였다.

전통적인 통계 기반 모델인 Holt-Winters는 계절성과 추세를 어느 정도 따라가는 모습을 보였지만, 농산물 가격에 내재된 비선형적 요인(기후 조건, 시장 충격 등)을 충분히 포착하지 못해 상대적으로 높은 MAPE를 보였다. 특히, Table 2 및 Fig. 4에서 확인할 수 있듯이 해당 모델의 R^2 는 0.1722로 매우 낮아, 전체적인 설명력이 제한적임을 보여준다. 예측 결과에서도 일정한 흐름 내에서는 비교적 정확한 예측을 수행하지만, 급격한 가격 변동 구간에서는 실제 가격 곡선과의 오차가 크게 벌어지는 경향이 관찰되었다. 이는 통계적 가정의 단순성으로 인해 발생한 것으로 보이며, 비정상적 가격 급등·급락 같은 이벤트 대응력 부족을 더욱 부각시킨다.

LSTM과 GRU 모델은 순환신경망 구조를 기반으로 시계열 정보의 장기 의존성을 처리하는 데 강점을 갖는 것으로 알려져 있다. 그러나 본 연구 결과에서 두 모델은 장기 추세를 어느 정도 반영하면서도, 단기 변동성에 대한 학습에서는 예측 곡선이 지나치게 평탄화되는 현상이 나타났다. 이러한 결과는 농산물 가격의 비정형적인 외부 요인이나 갑작스러운 수급 변화와 같은 단기적 변동성을 충분히 학습하지 못했기 때문으로 판단된다.

Transformer 모델은 어텐션 메커니즘을 통해 시계열 전 구간에 대해 병렬적으로 관계를 학습할 수 있기 때문에, 순환신경

망 구조의 모델들이 놓치기 쉬운 장·단기 패턴을 모두 포착하는 데 비교적 우수한 성능을 보였다. 다만, 일부 예측 구간에서 실제 가격과 예측값 간의 오차가 발생했는데, 이는 극단값이나 급격한 변동 구간에서 학습된 어텐션 가중치가 실제 시장 충격과 정확히 일치하지 않았기 때문으로 해석된다. 그럼에도 불구하고 세 딥러닝 모델 중 가장 낮은 평균 오차율을 보임으로써, 농산물 가격과 같은 복잡한 시계열 문제에서 Transformer 구조의 효용을 재확인하였다.

제안된 LLM 모델은 주요 평가 지표 전반에서 기존 통계 기반 및 딥러닝 모델보다 낮은 예측 오차를 기록하였으며, 실제 시장 가격 변동 패턴에 대한 추세 반영이 가장 정확하게 이루어졌다. 또한 LLM은 장기 추세뿐 아니라 단기 변동성도 일정 부분 반영하는 모습을 보였는데, 이는 기존의 시계열 예측 모델들과 달리, 대규모 사전 학습을 통해 방대한 데이터에서 추출한 다양한 패턴을 학습하고 활용할 수 있기 때문으로 생각된다.

기존의 통계 기반 및 딥러닝 모델 대비 LLM이 우수한 예측 성능을 보였으나, 실제 적용을 위해서는 몇 가지 한계를 고려하고 이를 개선하는 방향으로 연구를 확장할 필요가 있다. 우선, 높은 연산 비용과 실무 적용의 효율성 문제가 있다. LLM 모델은 대규모 사전 학습을 통해 복잡한 시계열 패턴을 효과적으로 학습할 수 있지만, 예측 수행에 필요한 계산량이 크고, 연산 자원이 많이 소모된다는 점에서 실무 활용이 쉽지 않다. 특히, 농산물 시장에서는 정기적인 예측 업데이트가 요구되므로, 과도한 연산 비용은 운영 부담으로 이어질 수 있다. 따라서, 향후 연구에서는 모델 경량화 및 최적화를 통해 계산 효율성을 높이고, 더 적은 자원으로도 안정적인 예측이 가능하도록 개선할 필요가 있다.

또한, 예측 결과의 해석 가능성 부족이 있다. 기존에 활용되던 시계열 예측 모델들과 마찬가지로, LLM 역시 블랙박스 모델로서 작동하며 예측 결과의 근거를 명확하게 설명하기 어렵다. 그러나 LLM은 언어 모델이라는 특성을 가지고 있으므로, 모델이 도출한 가격 예측 결과를 단순한 수치가 아닌 자연어 해석을 통한 설명 제공 방식으로 확장할 수 있다. 후속 연구에서는 이러한 자연어 기반 해석 기능을 도입하여, 모델이 학습한 패턴과 시장 요인 간의 연관성을 보다 직관적으로 설명하는 방안을 모색할 필요가 있다.

뿐만 아니라, 모델의 확장성과 일반화에 대한 검증이 필요하다. 본 연구에서는 과거 가격 데이터만을 기반으로 미래 가격 예측을 수행하였으나, 농산물 가격은 기후 변화, 유통망 상황, 소비 트렌드 등 다양한 외부 요인에 영향을 받는다. 따라서, 특정 시점의 가격 패턴만으로 미래 가격을 예측하는 데에는 한계가 있기 때문에 다양한 환경에서도 일관된 성능을 유지할 수 있는 모델 확장성에 대해 추가적인 검증이 필요하다.

Conclusion

본 연구는 LLM 기반의 농산물 가격 예측 방법을 제안하고, 이를 기존의 통계 기법(Holt-Winters), 그리고 인공지능망 모델(LSTM, GRU, Transformer)과 비교하여 성능을 평가하였다. 1,862일간 수집된 23개 농작물의 일별 평균 가격 데이터를 동일한 조건 하에서 실험한 결과, MAE, RMSE, MAPE, R^2 의 모든 평가 지표에서 LLM이 가장 우수한 예측 성능을 보였으며, 이는 대규모 사전 학습을 통해 얻은 비선형성 패턴 반영 능력이 농산물 가격 예측에서도 효과적으로 작동할 수 있음을 시사한다.

그러나 본 연구에는 몇 가지 한계가 존재한다. 첫째, 모델들의 하이퍼파라미터 조정은 완전히 최적화된 조건에서 이루어지지 못했으며, 실험을 통해 정성적으로 설정되었다. 둘째, 공휴일 등에 대한 결측치 처리 과정에서 단순 삭제 방식을 적용하였기에, 시계열 연속성이 일부 훼손되었을 가능성이 있다. 셋째, 실제 농산물 가격은 기후 변화, 유통망, 수출입량, 소비 트렌드 등 복합적 요인에 의해 결정되지만, 단순 과거 가격 데이터만을 기반으로 학습 및 평가하였기에 복합적인 정보 통합이 이루어지지 못했다.

향후 연구에서는 다양한 농산물 및 지역 데이터를 활용하여 모델의 일반화 가능성을 검증하고, 기후, 유통, 정책 등 다양한 데이터를 추가적인 입력 변수로 활용하여 가격 변동 요인을 보다 정밀하게 반영하는 방안을 모색함으로써 예측 성능을

한층 더 향상시킬 수 있을 것으로 기대된다. 또한, 모델 경량화와 자연어 기반 해석 기능을 도입하여 LLM의 실용성을 높인다면, 농산물 시장에서 보다 신뢰성 있는 의사결정 도구로 자리매김할 수 있을 것이다.

Acknowledgements

이 연구는 서울대학교 신입교수 연구정착금, 한국연구재단(과학기술정보통신부 재원, RS-2023-00302123), 농림식품기술기획평가원(농림축산식품부 재원, RS-2024-00398300)의 식물건강 증진 첨단기술 개발 및 융합형 연구인력 양성 사업의 지원을 받아 수행되었습니다.

References

- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.
- Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., ... & Ganapathy, R. 2024. The llama 3 herd of models. Accessed in <https://ai.meta.com/research/publications/the-llama-3-herd-of-models/> on 23 July 2024.
- Gruver, N., Finzi, M., Qiu, S., Wilson, A. G. 2024. Large language models are zero-shot time series forecasters. *Advances in Neural Information Processing Systems*, 36.
- Hochreiter, S. 1997. Long Short-term Memory. *Neural Computation* MIT-Press.
- Holt, C. C. 1957. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5-10.
- Hu, E. J., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. 2022. LoRA: Low-Rank Adaptation of Large Language Models. *International Conference on Learning Representations*.
- Jin, M., Wang, S., Ma, L., Chu, Z., Zhang, J. Y., Shi, X., Chen, P., Liang, Y., Li, Y., Pan, S., Wen, Q. 2024. Time-llm: Time series forecasting by reprogramming large language models. *International Conference on Learning Representations*.
- Kim D. 2024. Climate change and agricultural income – an empirical analysis using farm household economic survey data in Korea. *Journal of Rural Development* 47:143-167.
- Korea Rural Economic Institute (KREI). 2013. Influence on Aggregate General Price Level of Agricultural Pricing and Policy Implications. Policy Research Report, 400 pp. [in Korean]
- Korea Rural Economic Institute (KREI). 2024. Current Status of Digital Technology Utilization and Policy Tasks for Improving Agricultural Product Supply and Demand Management. Policy Research Report, 292 pp. [in Korean]
- Vaswani, A. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*.
- Wei, J., Bosma, M., Zhao, V., Guu, K., Yu, A. W., Lester, B., ... , Le, Q. V. Finetuned Language Models are Zero-Shot Learners. 2022. *International Conference on Learning Representations*.
- Xue, H., Salim, F. D. 2023. Promptcast: A new prompt-based learning paradigm for time series forecasting. *IEEE Transactions on Knowledge and Data Engineering*.
- Zhou, T., Niu, P., Sun, L., Jin, R. 2023. One fits all: Power general time series analysis by pretrained lm. *Advances in neural information processing systems*, 36, 43322-43355.