

장단기 기억 신경망을 사용한 다변수 데이터 농산물 가격 예측 모델

Agricultural Product Price Prediction Model Using Multi-Variable Data Long Short Term Memory

강 동 곤*, 장 영 민*, 이 주 석*, 이 성 수*

Donggon Kang*, Youngmin Jang*, Joosock Lee*, and Seongsoo Lee*

Abstract

This paper proposes a method for predicting agricultural product prices by utilizing various variables such as price, climate factors, demand, and import volume as data, and applying the Long Short-Term Memory (LSTM) model. The analysis of prediction performance using the LSTM model, which learns the long-term dependencies of time series data, showed that integrating diverse data improved performance compared to traditional methods. Furthermore, even when predicting without price data as a dependent variable, meaningful results were achieved using only independent variables, indicating the potential for further model development. Moreover, it was found that using a multi-variable model could further enhance prediction performance, suggesting that this complex approach is effective in improving the accuracy of cabbage price predictions.

요 약

본 논문에서는 가격, 기후 요인, 수요, 수입량 등 다양한 변수를 데이터화한 후, LSTM(Long Short-Term Memory) 모델을 활용하여 농산물 가격을 예측하는 방법을 제안하였다. 시계열 데이터의 장기 의존성을 학습하는 LSTM 모델을 통해 예측 성능을 분석한 결과, 다양한 데이터를 통합함으로써 기존 방법보다 성능이 향상되었음을 확인하였다. 또한, 종속 변수인 가격 데이터 없이 독립 변수들만을 활용한 예측에서도 의미 있는 성과를 거두어, 모델의 발전 가능성을 확인할 수 있었다. 더 나아가, 다변수 모델을 사용할 경우 예측 성능이 더욱 개선될 수 있음을 알게 되었으며, 이러한 복합적인 접근이 배추 가격 예측의 정확도를 높이는 데 효과적임을 시사한다.

Key words : Time Series Analysis, Long Short Term Memory, Multi-Variable Data, Price Prediction

* School of Electronic Engineering and Department of Intelligent Semiconductor, Soongsil University (Student, Student, Professor, Professor)

★ Corresponding author

E-mail : sslee@ssu.ac.kr, Tel : +82-2-820-0692

※ Acknowledgment

This work was supported by the R&D Program of the Ministry of Trade, Industry, and Energy (MOTIE) and Korea Evaluation Institute of Industrial Technology (KEIT). (RS-2022-00154973, RS-2023-00232192, RS-2024-00403397). It was also supported by MOTIE and Korea Institute for Advancement of Technology (KIAT) (P0012451). The authors wish to thank IC Design Education Center (IDEC) for CAD support.

Manuscript received Sep. 23, 2024; revised Sep. 24, 2024; accepted Sep. 25, 2024.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. 서론

농산물 가격 예측은 농업 경제 분야에서 중요한 연구 주제 중 하나로, 특히 배추와 같은 주요 농산물의 가격 예측은 농가의 수익 안정화와 소비자의 합리적 소비 결정에 중요한 영향을 미친다[1]. 배추는 한국 식단에서 필수적인 식재료로, 가격 변동이 사회 경제적으로 큰 파급 효과를 가져온다. 이러한 배추의 가격 변동은 계절성이 존재하지만, 가격 변동 패턴이 매우 복잡하고 비선형적이다. 따라서 단순한 통계적 방법만으로는 이러한 변동성을 정확히 예측하는 데 한계가 존재한다.

이전 연구에서는 과거 배추 가격 데이터를 기반으로 예측을 수행하였으나, 가격 데이터만을 사용한 모델은 기후 변화나 시장 상황 등 중요한 외부 요인의 영향을 충분히 반영하지 못해 예측 정확도가 떨어지는 한계가 있었다. 특히 배추의 생육과 수확량은 기후 조건에 민감하게 반응하므로, 이를 고려하지 않는 예측은 실제 가격 변동성을 반영하기 어렵다. 이를 해결하기 위해 본 연구는 기존의 가격 데이터뿐만 아니라, 온도, 강수량, 일조 시간 등 배추의 생육에 직접적인 영향을 미치는 기후 조건과 수입량, 수요와 같은 환경 요소를 추가하여 예측 모델의 성능을 개선하고자 한다. 이러한 요소를 모델에 통합함으로써 가격 변동의 원인을 정밀하게 파악하고, 예측의 신뢰성을 높일 수 있을 것으로 기대된다.

이러한 다양한 요인들을 효과적으로 학습해 가격을 예측하기 위해, 본 연구에서는 딥러닝 기법 중 하나인 LSTM(Long Short-Term Memory) 네트워크를 활용하고자 한다. LSTM은 복잡한 시계열 데이터를 학습하는데 강점을 지니며, 시간적 패턴과 장기 의존성을 효과적으로 학습할 수 있어 비선형적이고 다차원적인 농산물 가격 변동을 예측하는 데 적합하다. 이에 따라 본 연구에서는 LSTM 모델을 통해 기존의 가격 데이터와 더불어 추가된 기후 조건 및 시장 요인들을 통합적으로 고려한 배추 가격 예측 모델을 제안한다.

II. LSTM

1. LSTM의 구조

LSTM(Long Short Term Memory)은 RNN(Recurrent Neural Network)에서 발전된 구조로, 장기기억(Long-term memory)과 단기기억(Short-term memory)을 효과적으로 처리할 수 있는 능력을 갖추고 있다. 이러한 특성 덕분에 LSTM은 시계열 데이터 분석, 자연어 처리,

음성 인식 등에서 활용되고 있다[2].

LSTM의 구조는 그림 1과 같이 망각 게이트(Forget gate), 입력 게이트(Input gate), 출력 게이트(Output gate)로 이루어진 세 개의 게이트와 셀 상태(Cell state), 히든 상태(Hidden state)로 구성된다.

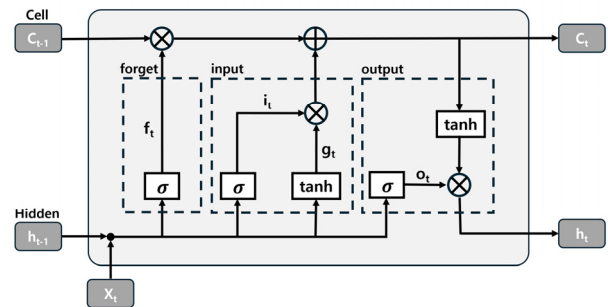


Fig. 1. LSTM network architecture.

그림 1. LSTM 네트워크 구조

각각의 구성 요소는 LSTM이 데이터를 처리하는 과정에서 중요한 역할을 수행한다. 망각 게이트는 이전 타임스텝에서 전달된 셀 상태의 정보를 얼마나 유지할지를 결정한다. 이 게이트의 활성화와 정도에 따라, 과거의 정보가 잊혀지거나 유지된다. 입력 게이트는 새로운 입력 정보가 얼마나 셀 상태에 반영될지를 결정한다. 이 게이트는 입력되는 정보를 선택적으로 받아들여, 현재 상태에 필요한 정보만을 반영하도록 한다. 출력 게이트는 현재 타임스텝의 셀 상태를 기반으로 최종 출력 값을 생성하며, 이를 다음 타임스텝으로 전달하거나 네트워크의 출력으로 사용한다. 이 세 개의 게이트는 각각 특정한 역할을 수행하며, 이러한 구조 덕분에 LSTM은 시간에 따라 변화하는 데이터를 효율적으로 처리할 수 있으며, 장기적인 의존성을 유지하는 데 뛰어난 성능을 발휘한다.

2. LSTM의 수식과 메커니즘

앞서 LSTM의 기본 구조와 각 구성 요소의 역할에 대해 개략적으로 설명하였다. 본 절에서는 LSTM이 정보를 처리하고 갱신하는 과정을 수식적으로 구체화하고, 각 수식이 네트워크 내에서 어떻게 작동하는지 설명한다[3].

2.1. 망각 게이트(Forget gate)

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (1)$$

여기서 f_t 는 망각 게이트의 출력값으로, 이전 시간의 셀 상태 C_{t-1} 에서 얼마나 많은 정보를 잊을지 결정하는 역할을 한다. f_t 는 시그모이드(Sigmoid) 함수를 통해 계

산되며, 이 값은 0과 1사이로 조정된다. 이 값이 1에 가까울수록 이전 셀 상태의 정보를 더 많이 유지하고, 0에 가까울수록 정보를 더 많이 잊어버린다. 수식에서 x_t 는 현재 타임스텝의 입력값이고, h_{t-1} 은 이전 타임스텝의 은닉상태 값이다. 또한 W_f 는 망각 게이트의 가중치 행렬이며, b_f 는 편향 값이다.

2.2. 입력 게이트(Input gate)

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

i_t 는 입력 게이트의 값으로, 시그모이드 함수를 적용해 0과 1 사이의 값을 가지도록 한다. 이 값은 새로운 정보가 셀 상태에 얼마나 반영될지를 결정한다. 입력 게이트의 값이 1에 가까울수록 셀 상태 c_t 에 새로운 정보가 더 많이 반영되며, 0에 가까울수록 반영되는 정보의 양이 적어진다. 수식에서 W_i 는 가중치 행렬이며 b_i 는 편향 값이다.

$$g_t = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \quad (3)$$

g_t 는 셀 상태 후보로, 현재 타임스텝에서 입력된 정보가 셀 상태에 반영되기 전에 계산된 값이다. g_t 는 하이퍼볼릭 탄젠트 함수 \tanh 를 적용해 -1과 1 사이의 값으로 변환된다. g_t 는 셀 상태에 추가될 정보의 크기와 방향을 나타낸다. 특히 \tanh 를 사용하는 이유는 셀 상태가 정보의 증가(양수) 또는 감소(음수)를 모두 표현할 수 있어야 하기 때문이다. g_t 가 양수일 경우 셀 상태를 증가시키고, 음수일 경우 셀 상태를 감소시켜, LSTM이 데이터를 더 유연하게 조정할 수 있도록 한다.

최종적으로 입력게이트 i_t 와 셀 상태 후보 g_t 의 원소별 곱(element-wise multiplication)을 통해 셀 상태에 새로운 정보가 얼마나 반영될지 결정된다.

2.3. 셀 상태(Cell state)

셀 상태는 LSTM이 장기적으로 기억해야 하는 정보를 유지하는 경로로, 시간에 따라 정보가 손실되지 않도록 도와준다. 셀 상태는 각 타임스텝에서 게이트 연산을 통해 필요없는 정보를 잊고, 중요한 정보를 추가함으로써 정보의 장기적인 의존성을 보장한다. LSTM의 셀 상태는 다음과 같은 수식을 통해 갱신된다.

$$c_t = f_t \circ c_{t-1} + i_t \circ g_t \quad (4)$$

망각 게이트의 출력 값 f_t 와 이전 셀 상태 c_{t-1} 의 연산을 통해 불필요한 정보는 제거되고, 입력게이트 값 i_t 와 셀 상태 후보 g_t 의 원소별 곱을 통해 새로 입력된 정보를 얼마나 반영할지 결정된다.

2.4. 출력 게이트(Output gate)

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (5)$$

출력 게이트 값인 o_t 를 연산하는 수식이다. 현재 타임스텝의 입력값 x_t 와 이전 타임스텝의 은닉 상태 h_{t-1} 에 출력 게이트 가중치 행렬 W_o 를 곱한 후 편향 값 b_o 를 더해준다. 이후, 시그모이드 함수를 적용하여 o_t 를 계산할 수 있다. 계산된 출력 게이트 값 o_t 는 LSTM 셀이 내보내는 은닉상태 h_t 에 얼마나 많은 정보를 내보낼지 조절한다.

2.5. 은닉 상태(Hidden state)

은닉 상태는 단기적인 정보를 저장하며, 이는 현재의 출력과 다음 타임스텝으로 전달될 정보를 결정한다. 은닉 상태는 셀 상태와 출력 게이트를 결합하여 계산되며, 다음 수식으로 표현된다.

$$h_t = o_t \circ \tanh(c_t) \quad (6)$$

출력 게이트 o_t 는 시그모이드 함수를 통해 계산된 값으로, 0과 1 사이의 값을 가진다. 출력 게이트가 1에 가까울수록 셀 상태의 정보가 은닉 상태에 더 많이 반영되며, 0에 가까울수록 정보의 반영이 적어진다.

3. 성능 지표

RMSE(Root Mean Squared Error, 평균 제곱근 오차)는 예측값과 실제값 간의 차이를 평가하는 성능 지표 중 하나로, 모델의 예측 정확도를 평가할 때 자주 사용된다. RMSE는 실제값과 예측값의 차이를 제곱한 값들의 평균을 구한 후, 그 평균에 제곱근을 취하여 계산된다[4].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

위 식은 RMSE의 공식을 나타내며, 여기서 y_i 는 실제 값을, \hat{y}_i 는 예측값을 의미하고, n 은 데이터의 개수를 나타낸다.

RMSE의 주요 장점은 제공된 오차를 사용하기 때문에,

큰 오차에 더 민감하게 반응한다는 점이다. 이는 모델이 큰 오차를 갖지 않도록 유도하는 데 유리하다. 또한 RMSE는 원래 데이터와 동일한 단위를 사용하기 때문에 해석이 용이하고 직관적이다. 본 연구에서는 배추 가격 예측에서 큰 오차를 줄이는 것이 중요하기 때문에, RMSE를 성능 평가 지표로 채택하였다. RMSE를 통해 예측 모델이 실제 데이터와 얼마나 일치하는지 평가함으로써, 모델의 전반적인 성능을 객관적으로 측정하고 개선 방향을 모색했다.

III. 데이터 전처리

본 연구에서는 배추 가격 예측을 위해 농산물 유통정보 시스템인 KAMIS에서 제공하는 배추 가격 데이터를 사용하였으며, 기상청에서 수집한 기후 데이터(기온, 강수량, 일조시간)와 수입량 데이터를 포함하였다. 또한 도매시장 통합 홈페이지에서 전국 도매시장의 배추판매량을 수요데이터로 통합하여 분석을 수행하였다.

데이터 전처리는 다음과 같은 단계로 진행되었다. 먼저, 2020년부터 2023년까지의 배추 가격, 기후 데이터, 수요 및 수입량 데이터를 수집하였다. 배추 가격 데이터와 기후, 수요 데이터는 일별로 제공되었으나, 수입량 데이터는 월별로 제공되었기 때문에 월별 데이터를 일별 데이터로 변환하여 통합하였다. 수집된 데이터 중 일부 기간에는 결측치가 존재하였으며, 배추 가격 데이터의 경우 인접한 기간의 데이터를 활용한 결측치 보완 방법을 통해 결측값을 보완하였다. 기후 및 수요 데이터의 결측치는 평균값으로 대체하였다.

또한, 전체 데이터 간의 단위 차이를 보정하기 위해

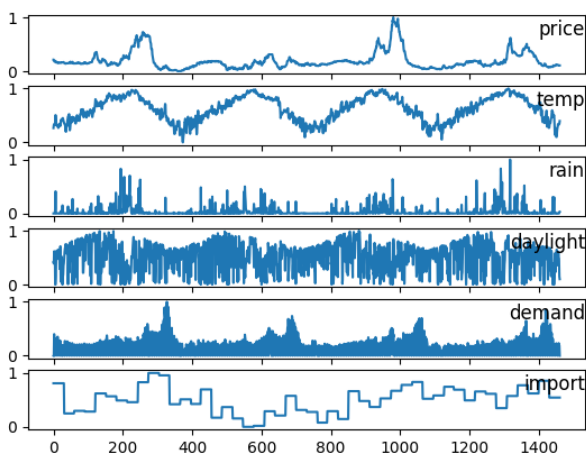


Fig. 2. Normalized time series data.

그림 2. 정규화된 시계열 데이터

Min-Max 정규화를 적용하여 모든 데이터 값이 0과 1 사이에 위치하도록 조정하였으며, 그림 2에서 정규화된 전체 데이터의 모습을 확인할 수 있다. 이를 통해 모델 학습 시 각 변수의 균형을 유지하였다. 다음 수식은 사용한 Min-Max 정규화 공식이다.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (8)$$

정규화된 데이터는 시계열 분석을 용이하게 하기 위해, 각 타임스텝별로 분할되어 학습이 가능하도록 구성하였다[5]. 여기서 타임스텝은 모델이 예측을 수행할 때 참조하는 이전 시점의 데이터 포인트 개수를 의미한다. 예를 들어, 타임스텝을 5로 설정하면, 모델은 현재 시점의 값을 예측하기 위해 직전 5개의 시점 데이터를 활용하게 된다. 이를 설정함으로써, 모델은 과거 데이터를 활용하여 단기 및 장기 패턴을 학습할 수 있다.

전체 데이터 중 앞의 1000개 샘플을 학습용 데이터로, 나머지 데이터를 테스트용 데이터로 분할하여 예측 모델의 일반화 성능을 평가할 수 있도록 하였다. 학습 데이터셋인 'train_x'의 형태는 (sample, time step, feature) 형태의 3차원 배열로, time step을 5로 설정한 경우 (1000, 5, 6)의 차원을 가진다. 목표 변수 데이터셋인 'train_y'는 (sample, feature) 형태의 2차원 배열로, (1000, 1)의 차원을 가지며, 이는 1000개의 샘플 각각에 대해 하나의 출력값을 예측하도록 설계되었음을 나타낸다. 테스트 데이터의 샘플 수는 전체 데이터에서 학습용 데이터를 제외한 나머지 데이터로 구성된다. 마찬가지로 time step이 5인 경우 test_x는 (431, 5, 6), test_y는 (431, 1)의 차원을 갖는다. 이와 같은 전처리 과정을 통해 배추 가격 예측을 위한 데이터셋을 구축하였으며, 이를 바탕으로 모델을 학습 및 평가하였다.

IV. 모델 설계

모델의 구조는 그림 3에 제시된 바와 같이, 3개의 LSTM 층, Dropout 층, 그리고 Dense 층으로 구성되었다[5]. 3개의 LSTM 층을 사용하는 이유는 모델의 일반화 성능을 향상시키기 위함이다. 일반화는 모델이 다양한 데이터에 대해 일관된 성능을 유지할 수 있는 능력을 의미하며, 이는 복잡한 비선형 데이터를 효과적으로 학습하기 위해 필수적이다. 그러나 LSTM 층의 수가 증가할수록 모델이 과적합될 가능성도 높아지므로, 이를 방지하기 위한 정규화 기법이 필요하다. 이러한 이유로,

각 LSTM 층에 L2 정규화(L2 Regularization)를 적용하고, LSTM 층 다음에는 Dropout 층을 배치하였다.

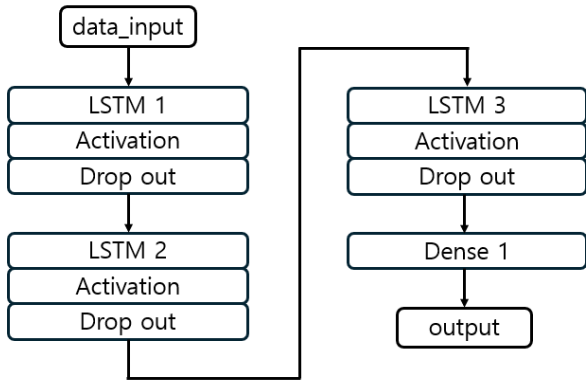


Fig. 3. Model architecture.
그림 3. 모델 아키텍처

L2 정규화는 각 LSTM 층에서 가중치의 크기를 제어하여 모델이 과도하게 복잡해지는 것을 방지하고, 일반화 성능을 향상시키는 역할을 한다. 이는 손실 함수에 가중치의 제곱 합을 추가함으로써, 큰 가중치가 발생하지 않도록 하여 모델의 복잡성을 줄이는 방법이다. 이와 함께, Dropout 기법은 학습 과정에서 무작위로 일부 뉴런을 비활성화하여, 모델이 특정 뉴런에 지나치게 의존하지 않도록 함으로써 과적합을 효과적으로 방지한다. 이를 통해 모델의 일반화 성능을 더욱 강화하였다.

첫 번째 LSTM 층은 입력 시퀀스를 처리하여 시간적 패턴을 학습하며, 두 번째와 세 번째 LSTM 층은 앞서 학습된 정보를 기반으로 장기적인 의존성을 강화한다. 각 LSTM 층에 활성화 함수로 tanh를 선택하여, 입력값을 -1과 1 사이로 조정하였으며, 앞서 설명한 L2 정규화와 Dropout 기법을 통해 과적합을 방지하고, 모델의 성능을 최적화하였다. 출력층으로는 Dense 층을 사용하여 최종 예측값을 생성하도록 하였다.

모델의 최적화 함수로는 Adam Optimizer를 사용하였다. Adam Optimizer는 학습률을 자동으로 조정하면서 빠르고 안정적인 수렴을 보장하는 최적화 알고리즘으로, 다양한 상황에서 우수한 성능을 발휘하는 대표적인 최적화 방법이다.

학습 과정에서의 손실 함수로는 MAE(Mean Absolute Error)가 사용되었다. MAE는 예측값과 실제값 간의 절대 오차를 평균하여 계산되며, 이를 최소화하는 방향으로 모델이 학습된다. 이로 인해 모델이 예측값을 실제값에 가깝게 조정할 수 있도록 돕는다.

V. 실험 결과

1. 학습 결과

표 1은 Dropout 비율 0.1, Regularization 강도 0.001, Batch Size 50, Epoch 100을 기준으로 100번의 학습을 통해 얻은 RMSE의 변화를 time step 별로 나타낸 것이다.

Table 1. RMSE variation vs. time step.

표 1. time step 별 RMSE 변화

	5	10	15	20	25	30
AVG	1257.6	1359.4	1065.2	1050.7	1063.4	1102.2
MIN	1195.4	1228.8	974.7	962.3	984.3	973.2
MAX	1469.4	1560.1	1267.3	1225.7	1226.1	1411.5

각 time step에 대해 평균(AVG), 최소(MIN), 최대(MAX) RMSE 값을 계산하였으며, 이를 통해 모델의 성능이 time step에 따라 어떻게 변하는지 분석했다. 실험 결과, time step이 20일 때 모델이 가장 작은 RMSE 값을 기록한 것으로 나타났다. 이는 time step이 20인 경우가 다른 설정에 비해 예측 성능이 뛰어나다는 것을 의미하며, time step 20을 선택하여 추가적인 훈련을 진행했다.

그림 4는 추가적인 훈련 결과를 시각적으로 보여주는 그래프이다. 상단의 그래프는 학습 과정에서의 손실 함수값(Training Loss)과 검증 손실 함수 값(Validation Loss)을 나타내어 모델이 얼마나 잘 학습되었는지를 확인할 수 있다. 하단의 그래프는 실제 값(Actual)과 예측 값(Pred)을 비교한 가격 예측 결과를 나타낸다. 학습 과정 동안의 손실 함수값과 검증 손실 함수값을 분석한 결과, 학습 손실 값은 초기 0.4367에서 최종 0.0259으로 감소하였고 검증 손실 함수값은 초기 0.3273에서 최종 0.0194으로 감소했다. 이를 통해 검증 손실 함수값이 학습 손실 함수값보다 낮아 과적합이 발생하지 않았음을 확인할 수 있다. RMSE 값은 672.46으로 기존 1050.7보다 향상된 성능을 확인할 수 있다. 하지만 가격 변수를 포함한 훈련에서는 이전 연구와 마찬가지로 모델이 예측하는 것이 아닌 그 전 시점의 데이터를 반영하는 것을 볼 수 있었다. 이러한 결과를 고려하여, 가격 변수를 제외한 데이터로 모델을 다시 훈련하여 가격 변수에 대한 의존성을 줄이고, 다른 변수들이 가격 예측에 미치는 영향을 보다 명확하게 분석하고자 하였다.

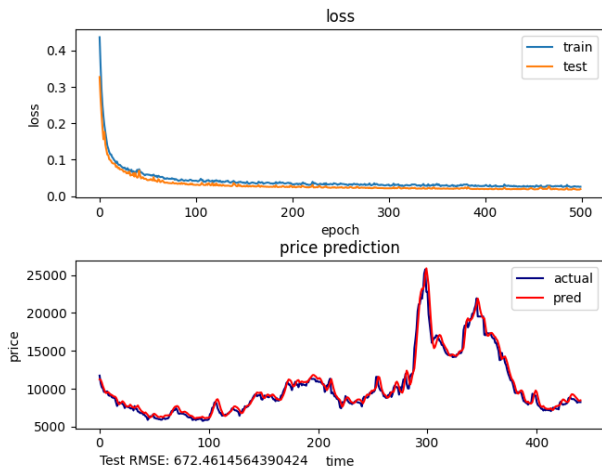


Fig. 4. Price Prediction Graph and Loss Function Graph.
그림 4. 가격 예측 그래프와 손실함수 그래프

2. 비가격 데이터만을 사용한 학습 결과

그림 5는 가격 변수를 제외한 비가격 변수들만을 사용하여 훈련된 LSTM 모델의 예측 결과를 나타낸 것이다. Loss 그래프에서 확인할 수 있듯이, 훈련 손실 값은 초기 0.4328에서 최종 0.0839로 감소하였고, 검증 손실 값도 초기 0.3590에서 최종 0.0614로 감소했다. 이를 통해 모델이 과적합 없이 안정적으로 학습되었음을 확인할 수 있다.

그림 5의 그래프를 보면, 그림 4와 달리 가격 예측 그래프와 실제 가격 그래프가 일치하지 않은 모습을 확인할 수 있다. 이는 예측된 가격이 과거 가격을 반영한 것이 아니라, 모델이 데이터를 학습하여 예측한 결과임을 보여준다.

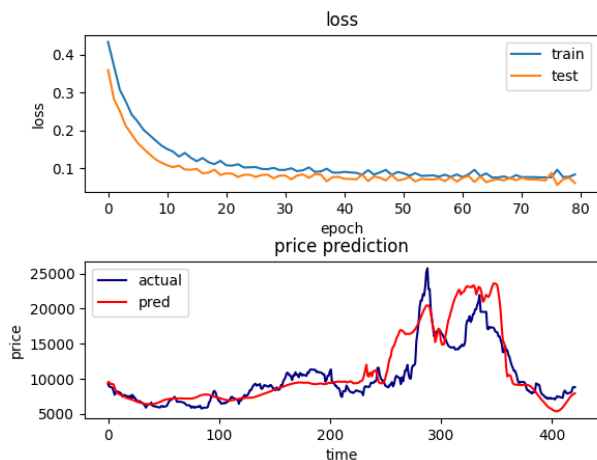


Fig. 5. Price Prediction Graph Using Non-Price Data.
그림 5. 비가격 데이터만을 사용한 가격 예측 그래프

RMSE 값이 약 2300으로 나타나, 예측된 값과 실제 값 사이에 오차가 존재함을 보여준다. 이러한 오차로 인

해 예측된 가격 그래프가 실제 가격 그래프와 일치하지 않는 부분이 있지만, 전반적인 추세는 대체로 유사하게 나타났다.

VI. 결론

본 연구에서는 2020년부터 2023년까지의 배추 가격에 영향을 미치는 데이터들을 바탕으로, 미래의 배추 가격을 예측하기 위해 LSTM(Long Short-Term Memory) 모델을 설계하였다. 이전 연구에서는 배추 가격 데이터만을 사용하여 예측을 수행했으나, 이는 가격 변동의 다양한 외부 요인을 충분히 고려하지 못하는 한계가 있었다. 이러한 한계를 극복하기 위해, 본 연구에서는 배추 가격 뿐만 아니라 기후 조건(기온, 강수량, 일조 시간), 수요 및 수입량 데이터를 추가적으로 반영하여 모델을 개선하였다.

연구 결과, 다양한 외부 요인을 통합한 LSTM 모델이 기존의 가격 데이터만을 사용한 모델보다 예측 성능이 향상되었음을 확인할 수 있었다. 특히, 기후 조건, 수요, 수입량과 같은 변수들이 배추 가격 변동에 미치는 영향력을 반영하여, 더 정교한 예측 결과를 도출할 수 있었다. 이는 배추와 같은 농산물의 가격 예측에 있어 다양한 변수 요인들을 고려하는 것이 중요함을 시사한다.

또한 중속 변수인 배추 가격 데이터 없이 독립 변수들만을 사용해 학습한 결과도 어느 정도 의미 있는 예측 성능을 보여주었다. 비록 세부적인 가격 변동에 대한 예측 정확도는 다소 떨어졌지만, 이러한 연구를 통해 비가격 변수만을 사용한 예측 모델이 가격 변동의 큰 흐름을 효과적으로 학습할 수 있음을 확인했다는 점에서 의미가 있다.

향후 연구에서는 모델의 예측 정확도를 더욱 향상시키기 위해 다양한 방법을 모색할 계획이다. 특히, 배추의 생육 기간 등 추가적인 데이터를 수집하여 비가격 데이터만을 활용한 학습을 진행할 예정이다. 또한, 하이브리드 모델의 도입과 데이터 전처리 과정을 더욱 정교하게 수행함으로써, 배추 가격 예측의 정확도를 한층 더 높일 수 있을 것으로 기대된다.

References

[1] K. Kurumatani, "Time Series Prediction of Agricultural Products Price Based on Time Alignment

of Recurrent Neural Networks,” *Proceedings of IEEE International Conference on Machine Learning and Applications*, 2018.

DOI: 10.1109/ICMLA.2018.00020

[2] V. Sarika, G. Kamal, S. Pratham, N. Deepak, and T. Veneela, “An LSTM-Based Model for Stock Price Prediction,” *Proceedings of Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems*, 2023.

DOI: 10.1109/aicera/icis59538.2023.10420270

[3] B. Ren, X. Xu, and H. Yu, “Research of LSTM-RNN Model and Its Application Evaluation on Agricultural Products Circulation,” *Proceedings of IEEE Eurasia Conference on IOT, Communication and Engineering*, 2021.

DOI: 10.1109/ECICE52819.2021.9645687

[4] K. Moharm, M. Eltahan and E. Elsaadany, “Wind Speed Forecast using LSTM and Bi-LSTM Algorithms over Gabal El-Zayt Wind Farm,” *Proceedings of International Conference on Smart Grids and Energy Systems*, 2020.

DOI: 10.1109/SGES51519.2020.00169

[5] S. Selvin, R. Vinayakumar, E. Gopalakrishnan, V. Menon, and K. Soman, “Stock price prediction using LSTM, RNN and CNN-sliding window model,” *Proceedings of International Conference on Advances in Computing, Communications and Informatics*, 2017.

DOI: 10.1109/ICACCI.2017.8126078

BIOGRAPHY

Donggon Kang (Member)



2019~ : Candidate for BS degree in Electronic Engineering, Soongsil University.
 <Main interest> Automotive SoC, AI SoC, Processor SoC

Youngmin Jang (Member)



2019~ : Candidate for BS degree in Electronic Engineering, Soongsil University
 <Main Interest> Automotive SoC, AI SoC, Processor SoC

Joosock Lee (Member)



1983 : BS degree in Electronic Engineering, Sogang University.
 1985 : MS degree in Electronic Engineering, Korea University.
 1999 : PhD degree in Electrical Engineering, Korea University.
 1985~1995 : Senior Engineer, LG Central Laboratory

2004~2005 : CTO, MtekVision Ltd

2006~2010 : Chief of SoC Center, Chungbuk Technopark

2022~Now : Professor in School of Electronic Engineering, Soongsil University

<Main Interest> AI SoC, Power Management SoC, Battery Management SoC

Seongssoo Lee (Life Member)



1991 : BS degree in Electronic Engineering, Seoul National University.

1993 : MS degree in Electronic Engineering, Seoul National University.

1998 : PhD degree in Electrical Engineering, Seoul National University.

1998~2000 : Research Associate, University of Tokyo

2000~2002 : Research Professor, Ewha Womans University

2002~Now : Professor in School of Electronic Engineering, Soongsil University

<Main Interest> AI SoC, Automotive SoC, Security SoC, Processor SoC, Power Management SoC, Battery Management SoC, Reliability and Safety