Stock Price Prediction System*

Chanyoung Lee[1]2019313902, Yujin Seo[2]2018314589, and Donghun Jung[3]2020312141

[1] Sungkyunkwan University, Computer Science and Engineering, Republic of Korea
[2] Sungkyunkwan University, Department of Mathematics, Republic of Korea
[3] Sungkyunkwan University, Department of Physcis, Republic of Korea

Abstract. This project aims to predict the stock price of SAMSUNG Electronics utilizing advanced machine learning techniques. Our primary objective is to develop a predictive model tailored to the stock prices of SAMSUNG Electronics. Leveraging data sourced from the Korea Exchange (KRX), we will undertake the training and evaluation of a diverse set of machine learning models, including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN), and Transformer. This accessibility contributes significantly to the domains of investment decision-making, lowering the entry barrier for users interested in this field. Given the computationally intensive nature of our endeavor, our primary focus remains on constructing an accurate predictive model specifically designed for the stock prices of SAMSUNG Electronics. SAMSUNG Electronics stands as one of the most popular and actively traded stocks in South Korea, making it an ideal candidate for our predictive model. Furthermore, as a testament to our commitment to accessibility, our model will be deployed as a publicly accessible web service. This ensures that users can harness the power of our predictive model conveniently.

Keywords: Stock Price Prediction · Machine Learning · Decision Making

1 Introduction

In today's world of capitalism, stocks offer an attractive way to make money, potentially leading to significant profits. Investors are always eager to predict what will happen to stock prices because their financial well-being depends on it. Investors must make prediction about future stock prices, even if they don't have a deep understanding of the stock market.

Predicting stock prices essentially involves trying to figure out the real value of a company. If you know a company's true value, you can decide whether its stock price will go up or down from its current level. However, determining a company's true value is a challenging task. Many factors, both within and beyond a company's control, influence stock prices. Some of these factors include global events and economic conditions.

* Supported by LINC

Historically, professional fund managers have played a crucial role in helping people invest wisely and profitably in the stock market. However, because fund managers are human, their predictions and strategies can be influenced by their personal biases. This has led to a growing interest in using machine learning, to predict stock prices. Machine learning is believed to offer a more objective and unbiased approach to forecasting.

These days, numerous services utilize machine learning to predict stock prices. For instance, in Korea, a service known as 'IAM CHART' employs machine learning for this purpose. However, this service comes at a cost and is not open source. With a monthly fee of $150, it is beyond the means of many potential users.

In light of this, our intention is to harness machine learning techniques to offer a more accessible solution, thereby contributing to the realms of financial analysis and investment decision-making at a lower price point. Our project's primary objective is to construct a predictive model for the stock prices of SAMSUNG Electronics.

2 Background and Related Works

2.1 Background and Related Works

Stock Price Prediction Stock Price Prediction is the task of forecasting future stock prices based on historical data and various market indicators. It involves the use of statistical models and machine learning algorithms to analyze financial data and make predictions about a stock's future performance. Stock price trends are influenced by numerous factors, including interest rates, inflation rates, and financial news. To make accurate stock price predictions, one must leverage this diverse set of information. In the banking industry and financial services sector, teams of analysts are dedicated to scrutinizing, analyzing, and quantifying qualitative data from news sources. A substantial amount of information regarding stock trends is extracted from the extensive corpus of textual and quantitative data involved in such analysis.

Research Trends Recent research trends in stock price prediction include advancements in deep learning-based regression models. [4] These models often utilize Long Short-Term Memory (LSTM) networks and innovative validation techniques, such as walk-forward validation, to enhance their predictive capabilities.

In addition, some researchers have explored Particle Filter Recurrent Neural Networks (PF-RNNs), a new RNN family explicitly designed to model uncertainty within their internal structure. [5] Unlike traditional RNNs that rely on a deterministic latent state vector, PF-RNNs maintain a latent state distribution approximated as a set of particles. To enable effective learning, researchers have introduced a fully differentiable particle filter algorithm that updates the PF-RNN latent state distribution based on Bayes' rule. Experimental results have shown that PF-RNNs can outperform conventional gated RNNs across various domains, including synthetic robot localization datasets and real-world sequence prediction tasks, which is stock price prediction.

Furthermore, recent studies have proposed novel approaches, such as the development of a sentiment-ARMA model, which combines the autoregressive moving average model (ARMA) with sentiment analysis of financial news articles. [6] This model is integrated into an LSTM-based deep neural network consisting of three components: LSTM, VADER model, and a differential privacy (DP) mechanism. The proposed DP-LSTM scheme has demonstrated the potential to reduce prediction errors and enhance model robustness. Extensive experiments conducted on S&P 500 stocks have indicated promising results, including a 0.32% improvement in mean Mean Percentage Absolute Error (MPA) and a significant up to 65.79% reduction in Mean Squared Error (MSE) for the prediction of the market index S&P 500.

3 Problem Statement and Proposed Solution

3.1 Problem Statement

Determining Data Set Data collection will be conducted using the FinanceDataReader package in Python[3]. This package serves as a data crawler, extracting information from Korea Exchange[1] for Korean stock data and NASDAQ[2] for US stock data. It provides access to historical and up-to-date data, making it suitable for both prediction and training/validation/testing purposes.

Determining Model Stock prices exhibit one-dimensional, non-stationary, time-series characteristics. In consideration of these unique traits, the selection of an appropriate model is crucial.

Determining Loss Function and Evaluation Metric The choice of a loss function and evaluation metric hinges on the nature of our stock price prediction task. If we are predicting only daily price movements (rise/fall), the appropriate loss function will be Cross Entropy, and the evaluation metric will be Accuracy. Conversely, if we aim to predict specific values such as closing prices, the preferred loss function will be Huber Loss or Mean Squared Error (MSE), and the evaluation metric will be Mean Absolute Error (MAE) or Root Mean Square Error (RMSE).

3.2 Proposed Solutions

LSTM Long Short-Term Memory(LSTM) is an advanced Recurrent Neural Network (RNN) architecture as shown in the Figure. LSTMs were introduced to address some of the limitations of traditional RNNs, which struggle with capturing long-range dependencies in sequential data due to the vanishing gradient problem.

LSTM has three gates: input gate, forget gate, and output gate.

1. Input gate: Input gate is denoted by orange box. It decides what new information should be stored in the cell.
2. Forget gate: Forget gate is denoted by blue box. It determines what information from the previous state should be discarded or reflected.
3. Output gate: Output gate is denoted by gray box. The actual outputs are h_i and y_i, which are same and c_i represents the status of the cell. It specifies what information from the cell should be used to generate the output.

Compared to the traditional RNN, LSTM performs various mathematical operations, including including element-wise multiplication and addition, to control the flow of information and perform updates to the memory cell and hidden state.

Through this architecture and characteristics, LSTM can handle the long sequential data by maintaining a memory cell with gates to control information flow, making it capable of capturing long-term dependencies and patterns in the data.

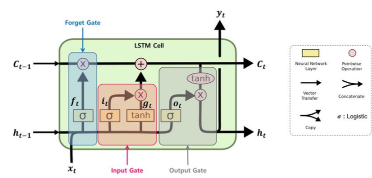

Fig. 1. LSTM model structure

GRU A Gated Recurrent Unit(GRU) is another type of recurrent neural network (RNN) architecture, similar to the Long Short-Term Memory (LSTM) network. GRUs are simpler in structure compared to LSTMs but have been found to be highly effective in various applications. GRU is also designed to address the vanishing gradient problem and enable RNNs to better capture long-range dependencies in sequential data. Compared to LSTM, GRU does not distinguish between cell status and the output.

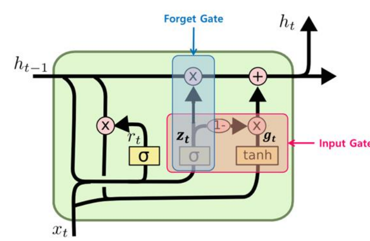

Fig. 2. GRU model structure

GRU has two gates: reset gate and update gate.

1. Reset gate: Reset gate is denoted by color box. It decides how much of the past information to forget.
2. Update gate: Update gate is denoted by color box. It decides how much of the past information to remember.

GRUs also perform mathematical operations, including element-wise multiplications and additions, to control the flow of information and update the hidden state.

Through this architecture and characteristics, GRU can also handle the long sequential data by maintaining a memory cell with gates to control information flow, making it capable of capturing long-term dependencies and patterns in the data.

One-dimensional CNN A Convolutional Neural Network(CNN) is a neural network architecture widely employed for processing and analyzing one-dimensional data sequences. In the context of stock price prediction, which inherently involves one-dimensional data, the utilization of a one-dimensional CNN is particularly relevant and effective.

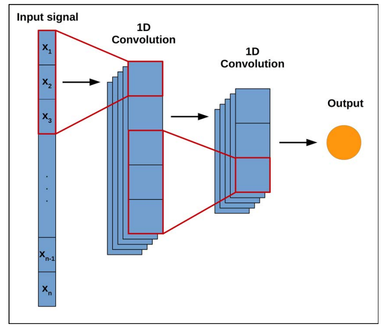

Fig. 3. CNN model structure

Compared to other recurrent neural network (RNN) variants like LSTM and GRU, CNNs offer a notably simpler structural design. Also it seems possible to analyse the various patterns of stock price data through the convolutional layers. As the stock price data is characterized by its non-stationary nature, exhibiting evolving trends and patterns over time, it is probable that CNN excels the performance of LSTM, and GRU.

Transformer The Transformer is a neural network architecture that has gained significant prominence in natural language processing (NLP) tasks due to its remarkable effectiveness. Transformers have exhibited promising potential in the domain of time series prediction.

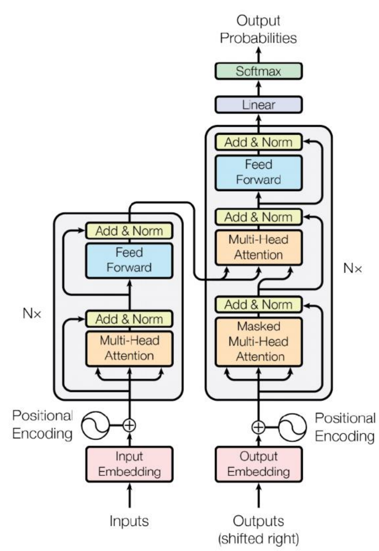

Fig. 4. Transformer model structure

Transformer operates as an Encoder-Decoder model, leveraging an attention mechanism. In the Encoder-Decoder architecture, the Encoder takes an input sequence and encodes the information into a single context vector. Conversely, the Decoder utilizes this context vector to generate an output sequence. Within the model structure, a crucial component is the "embedding" process. This process involves the conversion of input values into a unified vector representation.

The attention mechanism employed by the Transformer is a key feature. It assigns varying weights to elements within the input sequence, placing greater emphasis on pertinent information. This emphasis is then reflected in the model's output. The Transformer employs this mechanism to comprehensively evaluate the significance of the entire input sequence when generating the output.

The utilization of the Transformer model holds the promise of delivering superior performance. It has the capacity to incorporate diverse sources of information such as news, stock indices, and corporate disclosures, distinguishing it from other models.

4 Planning in Detail

4.1 Roles

Our team is organized as follows: Donghun Jung serves as the team leader and is responsible for the development of the web front end. Chanyoung Lee and Yujin Seo share responsibilities for developing the back end and the AI model.

4.2 Tentative Schedule

Week	3	4	5	6	7	8	9	10	11	12	13	14	15
정동훈 UI/UX Front-end	Preliminary Study		UI/UX Design		Front-end Development							Testing	
					AWS Server Building								
서유진 AI Back-end			Data Collection & Distillation		Model Implementation (GRU, CNN)		Performance Improvement						
					Back-end API Development								
이찬영 AI Back-end					Model Implementation (LSTM, Transformer)		Performance Improvement						
					Back-end API Development								

Our project timeline is outlined as follows: In the first three weeks, we conducted preliminary studies and research to establish the project's foundation. From the fourth to the sixth week, the frontend team will focus on designing the user interface(UI) and user experience(UX), while the backend team will collect data and prepare the necessary infrastructure. In the subsequent weeks up to the thirteenth week, the frontend team will proceed to develop the web page and set up the AWS server, while the backend team will be responsible for implementing the AI model, designing backend APIs, and continually improving the model's performance. Finally, in the last three weeks leading up to the presentation, we will rigorously test the AI model and the web application before deploying the web page for the final presentation.

References

1. Korea exchange. `http://krx.co.kr/`

2. Nasdaq: Stock market, data updates, reports & news. `https://www.nasdaq.com/`
3. FinanceData: Financedatareader. `https://github.com/FinanceData/FinanceDataReader` (2022)
4. Li, X., Li, Y., Yang, H., Yang, L., Liu, X.Y.: Dp-lstm: Differential privacy-inspired lstm for stock prediction using financial news. arXiv preprint arXiv:1912.10806 (2019)
5. Ma, X., Karkus, P., Hsu, D., Lee, W.S.: Particle filter recurrent neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34, pp. 5101–5108 (2020)
6. Mehtab, S., Sen, J., Dutta, A.: Stock price prediction using machine learning and lstm-based deep learning models. In: Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, October 14–17, 2020, Revised Selected Papers 2. pp. 88–106. Springer (2021)