

# ***Stock Market Prediction Method Based on Simple Market Feature Data and the Attention Framework***

**Yiran Ding<sup>1,a,\*</sup>**

<sup>1</sup>*School of Information Management, Wuhan University, Hubei, 430072, China*

*a. yrding@whu.edu.cn*

*\*corresponding author*

**Abstract:** As financial markets become increasingly complex and volatile, traditional statistical and machine learning methods reveal limitations in accurately predicting market trends. The introduction of deep learning techniques, particularly models like Long Short-Term Memory (LSTM) networks and Convolutional Recurrent Neural Networks (CRNN), has significantly enhanced the precision of market forecasts. However, these models often fail to account for the inherent uncertainties of financial markets, leading to poor predictive performance under extreme market conditions. To address this issue, this paper proposes a stock market prediction method that integrates attention mechanisms, bi-directional LSTM, and generative Transformers. By leveraging the attention mechanism and bi-directional LSTM to learn complex patterns within time series data, a model for forecasting stock market trends is constructed. This model is further combined with a generative Transformer to introduce a method for predicting future stock market movements. This paper conducts a performance analysis of the stock market prediction models using Attention as the framework, as well as the combined model (prediction model + generative Transformer), based on selected metrics. The experimental results demonstrate that the stock market prediction model exhibited a good fit across most subjects, with exceptionally strong performance on certain specific cases. The research aims to contribute a novel modeling framework to the field of financial market prediction, analyzing the advantages and disadvantages of this approach and providing insights and guidance for future research.

**Keywords:** Attention, Simple feature, Stock market prediction, Transformer.

## **1. Introduction**

In recent years, the financial market has witnessed a growing integration of deep learning techniques to enhance forecasting accuracy and risk management. While traditional methods such as statistical models (e.g., ARIMA, GARCH) and fundamental analysis have long been applied to price forecasting, their ability to capture the complex, nonlinear patterns within financial data remains limited [1,2]. Meanwhile, machine learning approaches like support vector regression and artificial neural networks provide improvements but often fall short due to their dependence on feature engineering and limitations in handling large-scale, unstructured data [3]. These constraints have led to the rise of more sophisticated methods that can learn complex temporal dependencies and adapt to market dynamics, particularly during periods of high volatility.

The application of deep learning architectures, such as Long Short-Term Memory (LSTM) networks and hybrid models like Convolutional-Recurrent Neural Networks (CRNNs), has gained prominence in financial time series forecasting [4,5]. These models, powered by advances in computational capacity (e.g., GPUs) and high-level programming frameworks (e.g., TensorFlow, PyTorch), demonstrate a superior ability to capture both linear and nonlinear relationships within financial data [3,6]. However, despite their efficacy, these models often overlook the uncertainties inherent in financial markets, leading to suboptimal predictions during rare or extreme events such as market crashes.

To address these challenges and resolve existing problems, the Attention model has emerged as a valuable tool. The attention mechanism, widely applied in natural language processing, has increasingly been adopted in time-series data analysis. It enables models to focus on key features, thereby enhancing the interpretability and accuracy of predictions. On the other hand, deep learning architectures can effectively learn from simple feature information to forecast the market, paving the way for a promising direction in market prediction based on simplified market characteristics.

Experiments show that an Attention model framework, based on simple market information, can achieve relatively accurate predictions of financial markets. By integrating Transformer networks equipped with residual mechanisms and convolutional layers, the model can learn simple market feature information, allowing for the prediction of future features and supporting price predictions under unknown future conditions.

The contributions of this paper can be summarized as follows: In this paper, we introduce an Attention-based framework for effectively handling time-series data and propose an innovative approach to market prediction. By simplifying the process of information extraction, our method bases market forecasts on straightforward feature inputs. Additionally, we present a novel technique that integrates two deep learning frameworks, resulting in an Attention-based model specifically designed for future market prediction applications.

## **2. Related Work**

### **2.1. Attention Model**

The attention mechanism has been widely applied in natural language processing and computer vision, significantly improving models' ability to handle complex relationships and long-range dependencies [6]. In financial market prediction, attention mechanisms have also been increasingly employed to address the dynamic and nonlinear nature of financial time series [4]. By assigning weights to different time points, attention mechanisms can capture key features in time series, particularly showing promising results in stock price prediction and trading behavior analysis [5]. Furthermore, attention models can be integrated with deep learning techniques to enhance the representation of long-term dependencies in time series, thereby improving market volatility prediction capabilities [7].

To tackle uncertainty in the market, Monte Carlo simulations provide robust support by facilitating multi-scenario analysis. Combined with attention mechanisms, they generate market prediction outcomes across various scenarios, assisting investors in making multidimensional decisions [8]. This integrated approach not only improves prediction accuracy but also enhances the interpretability of the results, offering more robust analytical tools to address market uncertainties. Additionally, multi-task learning and transfer learning show potential in financial forecasting by integrating multi-source data, further enhancing the model's generalization ability [9].

### **2.2. Deep Learning: Market Prediction Field**

Deep learning has gained significant traction in financial market forecasting due to its ability to model complex patterns and nonlinear relationships in financial time series data. Various deep learning

architectures have been explored, including Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and hybrid models like Convolutional-Recurrent Neural Networks (CRNNs). These models have demonstrated improved performance in capturing temporal dependencies and spatial features, making them well-suited for price forecasting tasks with spatiotemporal dynamics [5]. However, challenges such as computational expense and hyperparameter tuning persist, and the interpretability of these models remains a concern.

Generative Adversarial Networks (GANs) have also been applied to price forecasting. In this context, GANs typically consist of a generator (often an LSTM or its variants) and a discriminator (commonly a CNN or feed-forward neural network) trained in a competitive setting. GANs can generate synthetic price sequences that closely resemble real market data, although they require large datasets and can be unstable during training [10]. Despite these limitations, GANs represent a promising approach for capturing the complexities of financial markets.

Another emerging method is Deep Quantum Neural Networks (DQNNs), which combine principles of quantum mechanics with classical deep learning models. For example, the QuantumLeap model utilizes quantum states to encode multivariate financial data and predicts future market behavior using a hybrid quantum-classical architecture [11]. While this approach shows potential, it faces practical challenges such as the need for specialized hardware and complex training processes.

Ensemble methods, including stacking and decomposition ensembles, have been widely used to improve the accuracy and robustness of financial forecasts. Stacking ensembles combine predictions from diverse base models (such as LSTM, GRU, and CNN) using meta-learning techniques, enhancing adaptability to market conditions [12]. Decomposition ensembles, on the other hand, first break down time series data into intrinsic mode functions (IMFs) before applying deep learning models to each component. This approach improves prediction accuracy and offers better interpretability of financial data, but risks data leakage due to the inclusion of future information in the training process [13].

In conclusion, while deep learning techniques offer substantial improvements in market prediction, challenges such as computational complexity, data requirements, and model interpretability remain. The feature data used is relatively complex, and the model's parameters are large in scale, leaving room for simplification and reduction in time costs.

### 3. Method

#### 3.1. Overview

This paper employed a bidirectional LSTM and multi-head attention framework (see Section 3.2) to learn feature data, utilizing the bidirectional LSTM as the hidden layer of the entire deep learning model to enhance the model's ability to process time series data. In supplementary experiments, a Transformer model is used to predict feature data that was input into the model designed to forecast market price changes. This approach, which integrates a model built on an attention framework, proposes a method for predicting future market price trends solely based on historical fundamental feature information from the financial market (see Section 3.3). No additional special processing of the feature data was performed, and the use of non-public or easily accessible information was also avoided.

#### 3.2. Market Prediction Model

Utilizing the fundamental structure of neural networks, the basic stock market data obtained from public datasets serves as the input layer data. This data is processed through hidden layers, culminating in the final linear layer that provides the output, specifically the predicted target closing

price. In this process, the hidden layer consists of a bidirectional LSTM layer and a multi-head attention layer. The output from the bidirectional LSTM layer undergoes matrix operations to be transmitted as keys, queries, and values to the multi-head attention layer for subsequent computations. The final fully connected layer primarily serves to reduce dimensionality, ensuring that the output dimension is 1. For the basic information obtained from public datasets, no special processing is required; after dividing the features and targets, the model computation can proceed directly.

### 3.3. Transformer Model

Due to the foundational information about the stock market in financial public datasets being historical data, and since models based on attention and bidirectional LSTM are not generative models, predicting future market trends using an attention-based framework requires additional forecasts of future feature data. Therefore, in the supplementary experimental section of this article, a model is constructed based on a generative architecture of the Transformer framework. This model learns from the input foundational stock market information (such as the opening price and the highest price of the day) to forecast the basic stock market information for the next trading day. Experiments demonstrate that incorporating an LSTM layer enhances the model's capability to handle time series data, while utilizing convolutional layers to extract local features and implementing a residual mechanism to prevent gradient vanishing significantly improves the model's accuracy. The output results from this Transformer-based model are then used as inputs for the attention-based stock market prediction model, thereby linking the two models to realize the application of the stock market prediction model proposed in this article.

## 4. Experiment

### 4.1. Experimental Setup

**Data Acquisition and Processing** All data used in this article is sourced from publicly available financial market websites (e.g., Yahoo Finance). This study focuses on historical data from five well-known publicly traded companies, including Apple, NVIDIA, covering the past two years. The selection is based on the notable fluctuations in stock prices for certain companies during this period, as utilizing data with high market uncertainty can enhance the model's ability to learn from market volatility, thereby improving its predictive accuracy under high fluctuation scenarios. Additionally, comparing multiple companies aims to validate the model's generalizability, ensuring effective predictions for both volatile and relatively stable firms. No special processing was applied to the collected data; it was handled according to standard practices for processing basic feature data in deep learning models.

### 4.2. Stock Price Prediction

In this study, the processed data of five companies were divided into features and targets, which were placed into two separate lists. Subsequently, these lists were split into training and testing sets in a 90:10 ratio for model training and evaluation. The true values of the targets in the test set were compared against the model's predicted target values, allowing for the calculation of the overall  $R^2$  (Coefficient of Determination) as an indicator of the model's predictive capability. The choice of this metric is justified as it eliminates potential distortions in the results due to the non-standardization of price data, which could affect metrics like MSE and MAE.  $R^2$  provides a more accurate reflection of the model's fit regarding stock price trends, supported by standardized statistical formulas and interpretations, and has been used in previous experiments as one of the metrics for assessing model performance. Figure 2 illustrates the model's predictions for selected stocks in this experiment. The

results indicate that the model exhibits a commendable ability to predict stock price movements, although it still shows a certain preference towards specific data distributions. Under identical conditions, the stock trend predictions for some companies, such as NVIDIA (Figure 2A), are notably more accurate than those for others, such as APPLE (Figure 2B). All five companies achieved  $R^2$  values exceeding 0.7, indicating a good fit, with NVIDIA's trend prediction particularly outstanding, reaching over 0.96 (It is generally considered that a coefficient of determination above 0.7 indicates a good fit, while a value above 0.95 represents an ideal fit.), categorizing it as an excellent fit. The experimental results reveal that the model is particularly adept at capturing upward and downward trends, effectively determining whether the stock price will rise or fall the following day.

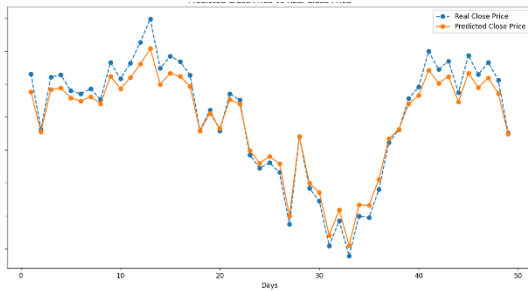


Figure 1A

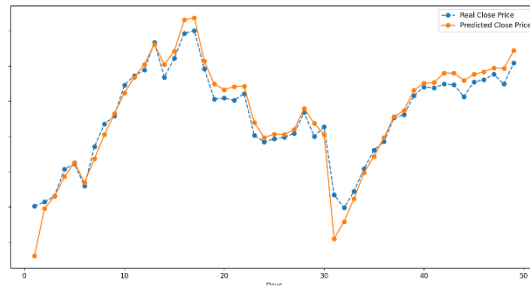


Figure 1B

Figure 1: The experimental results of the stock market prediction model on selected test subjects

Figure 1 illustrates the trend fitting results of the stock price prediction model based on the Attention framework for select experimental subjects. In panel 1A, the subject is NVIDIA stock, which achieves an  $R^2$  value exceeding 0.96. In panel 1B, the subject is Apple stock, with an  $R^2$  value exceeding 0.85. The blue line represents the actual stock price fluctuations, while the orange line indicates the predicted stock price movements.

### 4.3. Supplementary Experiment

In the supplementary experiment, this paper configured a pre-training data learning model based on the Transformer framework, consisting of 12 heads, with both the encoder and decoder layers set to 12 and a hidden layer size of 256. The training and testing sets were retained in a 90:10 split, excluding close price data, and the remaining five categories of feature data were input into the model for training. The model's predictions were then compared to those from the attention-based model trained in section 4.2, ultimately yielding forecasts of future stock price trends and completing the application of the model discussed in section 4.2. Figure 2 illustrates the visual results of the combined model on selected experimental subjects. The experimental findings indicate that this combined model exhibits significant performance differences compared to standalone stock valuation prediction models, with  $R^2$  values hovering around 0.7, indicating a less favorable fit. Additionally, there is a notable discrepancy between the predicted and actual trends. However, the model demonstrates a strong accuracy in predicting whether the stock price will rise or fall in the next time series. This suggests that while the combined model is highly accurate in forecasting overall upward and downward trends, the precision of predicting the magnitude of these movements still requires further improvement.

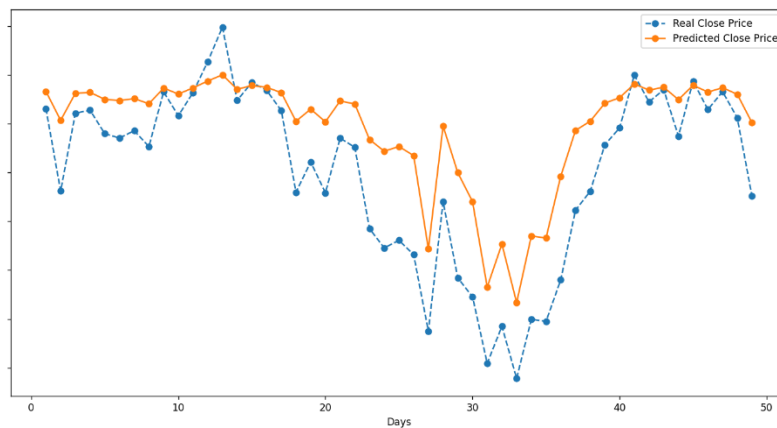


Figure 2: The experimental results of the combined model on selected test subjects

Figure 2 presents the prediction trend results of the combined model for selected experimental subjects, showcasing the data forecast for NVIDIA stock. The blue line represents the actual stock price fluctuations, while the orange line indicates the predicted stock price movements. The significant discrepancy between the two is reflected in the  $R^2$  value, which deviates notably from 1.

## 5. Conclusion

This paper introduces a model based on simple foundational data to predict stock price trends, providing an efficient and straightforward method for such predictions. This model employs several common optimization techniques, including the use of bi-directional LSTM to enhance the processing capabilities for time series data and a residual mechanism to improve model performance. Additionally, the paper proposes a fully automated combined model that incorporates future data for predictions and utilizes these forecasts to estimate stock prices, further demonstrating the functionality of an attention-based framework for stock price prediction.

**Discussion and Limitations:** the primary limitation of this experiment lies in the supplementary experiment section, where the combined model's performance was suboptimal. This was mainly due to device constraints that led to mediocre training outcomes for the future data generation model based on the Transformer framework. Moreover, combining the two models amplified the errors, resulting in poor predictive performance. Future researchers may consider enhancing model complexity and strengthening the operations of local features to optimize the earlier model, though this would significantly increase the required computational resources. Nevertheless, the results from the supplementary experiment are not without merit; they demonstrate that the combined model can accomplish predictive tasks. While the predictions under this method may not accurately reflect the magnitude of changes, the trends are highly precise, making it a potentially effective approach for forecasting future fluctuations in stock value. However, the high computational demands of the model may pose cost-related challenges.

## References

- [1] Adhikari R, Agrawal R K. (2014) A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications*, 24(6): 1441-1449.
- [2] Cheng C, Sa-Ngasoongsong A, Beyca O, et al. (2015) Time series forecasting for nonlinear and non-stationary processes: A review and comparative study. *IIE Transactions*, 47(10): 1053-1071.
- [3] Jiang Z.(2021) Deep learning-based stock price prediction with attention mechanism.
- [4] Sezer O B, Gudelek M U, Ozbayoglu A M. (2020) Financial time series forecasting with deep learning: A systematic literature review. *Applied Soft Computing*, 91: 106181.



- [5] Tsantekidis A, Passalis N, Tefas A, et al. (2020) *Using deep learning for price prediction by exploiting stationary limit order book features*. *Applied Soft Computing*, 93: 106401.
- [6] Goodfellow I, Bengio Y, Courville A. (2016) *Deep Learning*. Cambridge: MIT Press.
- [7] Liu S, et al. (2018) *Visual interrogation of attention-based models for natural language inference and machine comprehension*. Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States).
- [8] Dong L, Lapata M. (2016) *Language to logical form with neural attention*. *arXiv:1601.01280*.
- [9] Liu Y, et al. (2016) *Learning natural language inference using bidirectional LSTM model and inner attention*. *arXiv:1605.09090*, 2016.
- [10] Staffini A. (2022) *Price Forecasting with GANs for Financial Markets*. *Applied Artificial Intelligence*, 36(1): 64-83.
- [11] Paquet J, Soleymani M. (2022) *QuantumLeap: A Deep Quantum Neural Network for Financial Market Forecasting*. *Quantum Computing and Finance*, 45: 1-15.
- [12] Patel J, Shah S, Thakkar P, et al. (2020) *Predicting stock and commodity prices using deep learning, convolutional neural networks, and long short-term memory*. *Financial Engineering*, 34(2): 204-215.
- [13] Wu H, Zhang D, Liu H. (2022) *A new decomposition ensemble model for price forecasting of financial time series*. *Journal of Forecasting*, 41(1): 20-35.