

# ***Leveraging LSTM Neural Networks for Stock Price Prediction and Trading Strategy Optimization in Financial Markets***

**Chengru Ju<sup>1,a,\*†</sup>, Qi Shen<sup>2,†</sup>, Xin Ni<sup>3</sup>**

<sup>1</sup>*Master of Public Administration, Columbia University, New York City, USA*

<sup>2</sup>*Master of Business Administration, Columbia University, NY, USA*

<sup>3</sup>*Business Analytics and Project Management, University of Connecticut, CT, USA*

*a. rexcarry036@gmail.com*

*\*corresponding author*

*†These authors are co-first authors.*

**Abstract:** In the Artificial Intelligence development market in recent years, the atmosphere felt by the financial field will be relatively strong, and some, especially in the deep learning model, widely exist in the performance of complex financial data, which has certain advantages. Therefore, as one of the most prominent models of deep learning, LSTM neural network models are just good at processing some complex financial data of the rest of the time series, such as stock price prediction and trading strategy, optimization, and so on. However, in the actual application process, the stock price prediction of such models still has certain data quality, historical data market fluctuations, complex, non-linear data and other related factors, so there are certain challenges and development space in the process of processing. Nevertheless, by properly addressing these issues and combining them with best practices, LSTM algorithms are a powerful tool to help uncover underlying patterns in financial markets and optimize trading decisions.

**Keywords:** Long short-term memory Network (LSTM), Stock price forecasting, Trading strategy optimization, Artificial Intelligence (AI).

## **1. Introduction**

AI is growing rapidly in every industry, and the widely known financial sector is no exception. Banks and other financial institutions are currently looking for ways to take full advantage of both generative and traditional AI while maintaining high compliance and legal standards. While AI in finance is certainly a dynamically growing trend, there are hardly any solutions that rely solely on the technology. AI is still seen more as part of the solution than the solution itself. This chart 1 shows last year (2022) and forecast (2025) AI adoption rates among financial companies: As early as 2022, 54% of financial firms are either extensively adopting AI or viewing it as a key asset. It is worth mentioning here that since ChatGPT was released on November 30, 2022, in 2022, solutions such as GenAI have not been that popular. The prediction is that by 2025, a higher percentage of companies will see AI as critical to their business, rather than just a supporting factor.

Algorithmic trading has been around for decades. [1] Fifteen years ago, some New York traders contacted me about their new machine learning, algorithmic trading (precursor) AI Trading) marketing campaign. They were way before their time. It's too early. The technology really isn't there, but we're very close now, and these AI stock investment vehicles below are just the beginning. More to come. In this early research, it can be seen that many investments only mainly rely on relevant data indicators and basic parameters to predict certain short-term or long-term stock prices. Therefore, for these reasons, many researchers begin to adopt artificial intelligence-based deep learning methods to realize the trend of stock prices. This research is to carry out more detailed literature research and more extensive understanding of the prediction of related fields of financial market, multiple technologies and multiple researchers' research, and finally conduct some practical experimental research through LSTM model and draw the final conclusion.

## 2. Literature Review

It can be seen from both the early experiments and the related experiments on stock price prediction in recent years that many experts related to finance hope to obtain more information and value of the financial market by predicting the volatility of the stock market. However, there are not many successful cases. The stock market discussed in this paper belongs to the crucial financial field in some developing countries, so if the stock market declines, it will affect the development and growth rate of the whole country. Therefore, the study of stock price prediction is one of the problems and challenges that people have long pursued.

### 2.1. Traditional Financial Predict Stock Market

In the early story forecasting process, the main traditional techniques are based on technical analysis and fundamental analysis, two kinds of technical analysis, generally by predicting the short-term trend of stock prices, through charts and bar charts and other patterns to carry out technical analysis, but these contain basically simple data indicators, such as the basic moving average SMA [2], Lin Wiring, Lin Wiring, and Lin Wiring. Indicators such as the Relative Strength Index (RSI) [3], overbought and oversold, but almost all of these techniques are indicators that are considered lagging. Therefore, these technologies have certain shortcomings in the analysis process. For example, the analysis process does not consider the key data factors of stock price but refers to the analysis of basic aspects. This technology is based on the form of RNN and has certain advantages in the solution of long-term dependence problems. In the early experimental process, it can be seen that ANNS and RN NS are essentially implemented by machine learning, and machine learning requires artificial intelligence assistants to conduct data analysis without repeated execution in a wider range of financial market experiments.

However, this is not a comprehensive analysis, as traditional machine learning prediction methods include backpropagation algorithms. This algorithm usually belongs to the process of ship loss data, so many researchers have found that high prices and delays in the process of using complex time series data to predict stock price trends have a certain framework for predicting price trends using delay and forecast slowness.

### 2.2. Review of traditional SVM prediction techniques

Support Vector Machine (SVM) is a very widely used algorithm, which belongs to supervised learning algorithm, can be used for both classification and regression. There are two common applications of SVM: SVC (support vector classification) and SVR (support vector regression). They are based on the same support vector machine algorithm but are used for two different types of tasks: classification and regression. Because existing research distil an overall stock representation,

obscuring time-specific details of stock sequences, it leads to weaknesses in modeling de facto stock correlations, which often occur in an instantaneous and cross-time manner (Bennett, Cucuringu, and Reinert 2022). Specifically, stock correlations are highly dynamic and may exist over inconsistent time steps rather than through the entire review period [4].

This is because the dominant factors in stock prices are constantly changing, and different stocks may react to the same factors with different delays. For example, upstream companies' share prices may react more quickly to raw material shortages than downstream companies, with individual stocks exhibiting a lot of catch-up and lagging behavior. The main goal of support vector machines is to find an optimal decision boundary (hyperplane) that separates different categories of data points. The choice of this decision boundary is based on the maximum margin, that is, the distance from the point closest to the boundary (the support vector) to the decision boundary.

Support Vector Machines (SVM) [5] offer potential advantages in time series analysis due to their use of structural risk minimization (SRM), which focuses on reducing generalization errors rather than just training errors. This makes SVMs generally superior to traditional neural networks in terms of generalization performance. However, standard SVM models struggle with dynamically changing, complex financial data. Modified SVMs, combined with self-organizing feature maps, show improved predictive accuracy and convergence speed. Additionally, wavelet neural networks (WNN) and Cuckoo Search-based WNN models further enhance predictive capabilities, particularly in capturing stock price correlations. With the rise of algorithmic trading, developing sophisticated algorithms to predict financial markets is increasingly feasible, allowing investors to make informed decisions in real-time.

### 2.3. Technical stock market risk prediction based on deep learning

In the research related to analyzing the deep learning model neural network in artificial intelligence attributes to predict stock prices, it can be seen that most of such prediction models will use single or multi-layer actuators, that is, artificial neural networks such as RBF and SVM for analysis. In recent years, with the development of machine learning technology, more and more models have been applied to the field of stock price prediction. Research has shown that deep learning models, such as deep neural networks (DNN) and artificial neural networks (ANN), excel at processing large amounts of data, providing faster and more accurate predictions. In particular, combining technical indicators such as moving averages (10-day and 50-day), relative strength index (RSI), rate of price change (RoC), etc., can help models better capture market trends. In addition, characteristics such as volatility, Williams's indicator (%R), Commodity Channel index (CCI) [6] are also widely used in stock price forecasting.

Nevertheless, for smaller data sets, some traditional machine learning methods (such as naive Bayes) still have an advantage in accuracy [7]. Therefore, future research could further explore how to combine multiple deep learning models and optimize them for different data sizes to improve the overall performance of stock price predictions.

### 2.4. The role of LSTM model in stock market prediction

LSTM (Long Short-Term Memory Network) is a specially designed recurrent neural network (RNN) architecture designed to solve the "long-term memory" problem that RNNS face when dealing with long time series. Traditional RNNS tend to remember short-term information, and when dealing with longer sequences, it is difficult for the model to effectively retain early critical information due to the gradient disappearance problem. [8] LSTM enhances the memory capacity of RNNS through its unique "memory units" and "gate mechanisms" (input gates, forget gates, and output gates), allowing them to efficiently retain long periods of historical data, thereby improving the processing of long

sequences. This makes LSTM widely used in a number of applications, including video analytics, natural language processing (NLP), geospatial data modeling, and financial time series forecasting.

In LSTM, the gate mechanism can intelligently select the information to be retained or discarded, thus effectively avoiding the gradient disappearance problem. In contrast, traditional RNNs reuse the same parameters in each time step, and the gradient gradually decreases over time, making it difficult for the model to learn effective information when dealing with long time series. LSTM solves this problem by introducing different parameters in each time step and filtering the information through a gate mechanism. Because LSTM can maintain high accuracy and stability in long-term time series modeling, it is an ideal tool for dealing with time series forecasting, such as stock price forecasting, and other financial applications.

Traditionally, our understanding of time series data has been more static, focusing on daily temperature fluctuations or the opening and closing prices of the stock market. However, with the power of LSTM, we can now move beyond this static perspective and explore more dynamic aspects. At this point, we will transition to the coding part, where we will implement LSTM on the stock dataset to demonstrate its ability to analyze time series data.

Thus, in essence, LSTM provides a powerful tool for building predictive models for time series data such as stock prices by overcoming the limitations of traditional methods and standard RNNs. They can capture complex patterns and long-term dependencies in data, making them a valuable method for stock forecasting despite inherent limitations and the ever-present volatility of the market.

### 3. Methodology

In financial markets, stock price forecasting has always been a challenging task due to the nonlinearity and high volatility of its data. In recent years, deep learn-based models have shown strong performance in this field, especially long short-term memory (LSTM) networks, which have significant advantages in stock price prediction tasks due to their ability to capture long-term dependencies in time series data. This experiment aims to build a model that can effectively predict stock prices and optimize trading strategies by using the LSTM network implemented in Python. Through the analysis of historical trading data, we will explore the performance of LSTM in predicting stock price trends and assisting in making trading decisions, in order to quantify the effectiveness of its application in financial markets.

#### 3.1. Dataset

The dataset contains 14 columns associated with time series such as date and different variables such as closing, high, low and volume. We will experiment with the LSTM time series using the opening and closing values. (Table 1)

In this methodology part, we carry out model feature extraction steps by separately extracting date components from date variables of stock data. In this extraction process, researchers can pay more attention to date information. In this case, we are interested in understanding how price value changes over time.

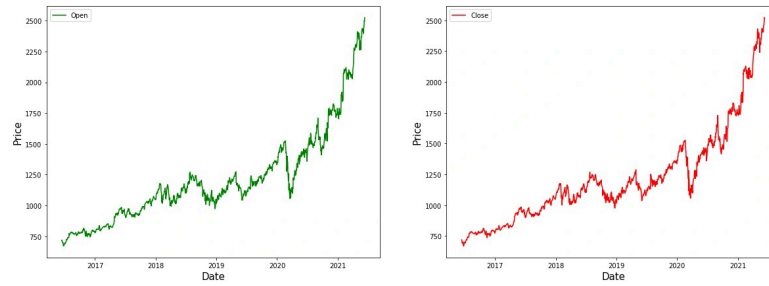


Figure 1: Stock market data Create price date plots

In this part, we first need to create a stock price date chart, which can be represented by different colors, and different variables are represented by green in this paper. The opening price can be more intuitive to see the initial price of each different date and the final price of each corresponding date by comparing the red curve, that is, the closing price. Using different colors to distinguish open and close prices makes it easier to visualize price movements in the same time period or in different time periods. (Figure 1)

### 3.2. Data preprocessing

Before applying LSTM to stock price prediction, the raw data must be preprocessed. First, the data is transformed through the 'fit\_transform' function to ensure that the data fits the input model. In this process, we use Min-Max Scaler, which can scale stock price data to a uniform range, so as to avoid adverse effects of different numerical ranges on model training. This can not only improve the stability of the model, but also speed up the convergence of the model.

We then split the entire data set proportionally into a training set and a test set, typically using 80% of the data for model training to allow the LSTM to learn trends and patterns in stock prices from historical data[9]. The remaining 20% of the data is reserved for testing to evaluate the model's predictions on unseen data. This partitioning of data ensures that the model not only performs well during training, but also demonstrates reliable predictive capabilities in practical applications.

### 3.3. LSTM model model building

In this experiment, we imported the original sequence model from keras and realized the data through two LSTM levels in the database. During the construction of the LSTM model, the setting of parameters has an important impact on the prediction performance. First, in the architecture of LSTM, we assign 50 units to each layer. These units are the number of neurons in the LSTM network, which determines the complexity and memory ability of the model. A higher number of cells can often capture more complex patterns, but also increases computational costs and the risk of overfitting. In order to prevent the model from overfitting, we introduced dropout mechanism between the two layers of LSTM and set a dropout rate of 10%. This means that in each training iteration, 10% of the neurons are randomly ignored to enhance the generalization ability of the model.

In the selection of the loss function, we use the mean square error (MSE) [10-11] to measure the difference between the predicted value and the real value. Mean square error is more sensitive to large deviations, so it can effectively capture large errors in the forecast. To optimize the model, we used the commonly used adam optimizer, which adaptively adjusts the learning rate to help accelerate the convergence process of the model. In evaluating the model performance, we use the mean absolute error (MAE) as a metric. Because MAE can directly reflect the average error margin in time series prediction, it is especially suitable for continuous time series data such as stock price. Through these



optimization and measurement choices, the LSTM model can better capture the long-term trend of the stock price and improve the accuracy of the forecast.

### 3.4. Experimental Results

In the analysis of the experimental results, we further evaluate the effectiveness of these trading strategies based on LSTM predictions. Backtest historical data to see how different strategies perform in real markets.

Strategy backtest: Based on the predicted buy and sell signals, we conducted several simulated trades to verify the effectiveness of the trading strategy. Specifically, when the LSTM model predicts the upward trend accurately, the buying strategy can obtain significant returns; When the model successfully predicts a price decline, the sell signal can effectively hedge against risk.

Benefit-risk ratio analysis: In order to quantify the performance of a trading strategy, we analyze the benefit-risk ratio of a forecast trade. By introducing a Sharpe Ratio or a Calmar Ratio, the overall performance of the strategy can be better assessed. The results show that the strategy combining LSTM prediction with stop-loss and stop-profit mechanism is significantly better than the traditional Buy and Hold strategy, especially in the volatile market environment.

Future directions for strategy optimization: Although LSTM-based trading strategies perform well in experiments, the model still has certain limitations. In the future, the precision of trading strategies can be further improved by introducing more market sentiment data or combining with other machine Learning models, such as Reinforcement Learning.

## 4. Conclusion

Future research can be combined with LSTM neural networks for stock price prediction to enhance the application of text learning and sentiment analysis. By analyzing social media, news reports, and market commentary, researchers are able to capture changes in sentiment among market participants, allowing them to more accurately assess the potential impact of those sentiments on stock prices. Building a comprehensive model that takes stock market sentiment analysis, national economic policies and related indicators into consideration can not only improve the accuracy of economic judgment, but also significantly improve the forecasting effect of stock prices. In addition, the integration of macroeconomic data and technical analysis indicators can help investors make more informed investment decisions in a volatile market environment.

Further exploration of different deep learning models, such as the combination of convolutional neural network (CNN) and long short-term memory network (LSTM), can provide a more reliable method for time series data processing. By employing advanced algorithms or combinations of algorithms, researchers can improve the predictive power of their models and better cope with non-linear features and unexpected events in the stock market. These explorations will lay the foundation for a more robust stock price forecasting framework, drive the development of fintech, and provide effective decision support for investors and institutions.

In summary, this paper explores stock price prediction based on LSTM neural networks and its application in trading strategy optimization, highlighting the potential of deep learning in capturing market dynamics and trends. Although LSTM model has excellent performance in processing time series data, its prediction effect is still affected by external factors, so future research should focus on integrating sentiment analysis and macroeconomic data to improve the accuracy and reliability of the model. Through these comprehensive methods, researchers can provide deeper insights into the financial market, help investors make better decisions in a complex and volatile environment, and ultimately promote the healthy development of the stock market.

## References

- [1] Ta, V. D., Liu, C. M., & Tadesse, D. A. (2020). Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading. *Applied Sciences*, 10(2), 437.
- [2] Namdari, A., & Li, Z. S. (2018). Integrating Fundamental and Technical Analysis of Stock Market through Multi-layer Perceptron. 2018 IEEE Technology and Engineering Management Conference (TEMSCON) .
- [3] Singh, P., Jha, M., Sharaf, M., El-Meligy, M. A., & Gadekallu, T. R. (2023). Harnessing a hybrid CNN-LSTM model for portfolio performance: A case study on stock selection and optimization. *Ieee Access*, 11, 104000-104015.
- [4] Koo, E., & Kim, G. (2022). A hybrid prediction model integrating garch models with a distribution manipulation strategy based on lstm networks for stock market volatility. *IEEE Access*, 10, 34743-34754.
- [5] Wu, J. M. T., Sun, L., Srivastava, G., & Lin, J. C. W. (2021). A long short-term memory network stock price prediction with leading indicators. *Big Data*, 9(5), 343-357.
- [6] Zhu, Y., Yu, K., Wei, M., Pu, Y., & Wang, Z. (2024). AI-Enhanced Administrative Prosecutorial Supervision in Financial Big Data: New Concepts and Functions for the Digital Era. *Social Science Journal for Advanced Research*, 4(5), 40-54.
- [7] Zhao, Fanyi, et al. "Application of Deep Reinforcement Learning for Cryptocurrency Market Trend Forecasting and Risk Management." *Journal of Industrial Engineering and Applied Science* 2.5 (2024): 48-55.
- [8] Wang, S., Zhu, Y., Lou, Q., & Wei, M. (2024). Utilizing Artificial Intelligence for Financial Risk Monitoring in Asset Management. *Academic Journal of Sociology and Management*, 2(5), 11-19.
- [9] Shen, Q., Wen, X., Xia, S., Zhou, S., & Zhang, H. (2024). AI-Based Analysis and Prediction of Synergistic Development Trends in US Photovoltaic and Energy Storage Systems. *International Journal of Innovative Research in Computer Science & Technology*, 12(5), 36-46.
- [10] Zhu, Y., Yu, K., Wei, M., Pu, Y., & Wang, Z. (2024). AI-Enhanced Administrative Prosecutorial Supervision in Financial Big Data: New Concepts and Functions for the Digital Era. *Social Science Journal for Advanced Research*, 4(5), 40-54.
- [11] Li, H., Zhou, S., Yuan, B., & Zhang, M. (2024). OPTIMIZING INTELLIGENT EDGE COMPUTING RESOURCE SCHEDULING BASED ON FEDERATED LEARNING. *Journal of Knowledge Learning and Science Technology* ISSN: 2959-6386 (online), 3(3), 235-260.