Forecasting Stock Prices with Artificial Intelligence

Danxuan Zhao

Course of Professional Study (CPS), Northeastern University, Boston, MA 02115, U.S.A.

Abstract. The purpose of this study is to investigate the closing prices of stocks in Artificial intelligence. The objective is to enhance the accuracy of future stock price Prediction to support investment or trading decisions. The models used in this paper include Simple Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), LSTM with peephole connectivity, and Gated Recurrent Unit (GRU). To conduct the study, Wal-Mart stock data is utilized to accurately predict future stock prices. The results show that the MSE for the SimpleRNN test is higher, indicating weaker generalization. The MSE of the basic LSTM test is lower than that of the RNN, indicating stronger generalization. The validated and tested MSEs of LSTM with peephole connectivity are higher than the basic LSTM and GRU. GRU performs as well as the basic LSTM but has the lowest LSTM training MSE. This stock prediction task requires the GRU model, which is the most suitable choice based on the training time.

1 Introduction

The stability and development of financial markets are supported by stock forecasting, which aids investors in making better investment decisions. The problem of forecasting itself is very difficult due to the non-static nature of stock prices [1].

A comprehensive review and taxonomy of stock market prediction techniques has been conducted by researchers, which encompasses both traditional statistical methods and modern machine learning approaches [2]. However, there are limitations to traditional forecasting methods. The SMA method, for instance, is simple but slow to react to new information. The WMA and exponential smoothing methods are relatively complex and require analysts to carefully select parameters. The naïve method, while straightforward, is often too simple to make accurate long-term stock price predictions. Since statistical methods are linear, an obvious limitation is shown in the way in which these methods deal with sudden rises or falls in stock prices. Stock data are usually non-stationary, chaotic and random, which makes it difficult for linear methods to capture the complexity of these fluctuation characteristics [3]. As a result, statistical methods often fail to accurately predict stock prices due to their inability to effectively deal with the nonlinear characteristics of stock data and the effects of multiple technical parameters [4].

Corresponding author: zhao.danx@northeastern.edu

To compensate for the limitations of the traditional methods, researchers have come up with machine learning methods to predict stocks. A method combining convolutional neural networks (CNNs), bi-directional long short-term memory (LSTMs), and attention mechanisms (AMs) was proposed by Wenjie Lu and colleagues to predict the closing price of the next day's stock [5]. Stock market prediction techniques are covered comprehensively by Gandhmal et al., who include Bayesian models, fuzzy classifiers, artificial neural networks, and support vector machines [6]. Decision trees, random forests, and deep learning models like LSTMs are explored by Nabipour et al [7]. Yu et al. talk about how Deep Neural Networks (DNNs) can be used to predict stock prices, with a particular emphasis on addressing the nonlinearity and time dependence of financial data [8]. Zhang et al. have developed a new system for predicting stock price trends and growth intervals based on data, called Xuanwu, which utilizes unsupervised morphological pattern recognition and random forest models with feature selection to predict stock price trends and growth intervals. The system processes raw trading data into categorical segments (up, down, flat, unknown) and trains models that are robust to market fluctuations, with an accuracy that outperforms existing methods [9]. Vijh and colleagues investigated the use of machine learning algorithms, including Artificial Neural Networks (ANNs) and Random Forests, to predict the closing prices of stocks of companies in different industries [10].

This paper aims to use machine learning algorithms to predict the stock price of Walmart and improve it for investment or trading decisions. To accomplish this objective, this paper constructs a range of neural network models, including basic RNN, LSTM (with or without peephole connections), and GRU units for stock price prediction, and provides potential directions for improvement in future research.

2 Methodology

The historical stock data of Walmart used in this paper is obtained from kaggle.com. This dataset, including 6118 stock price samples, ranges from 2000 to 2024. The data is characterized as a time series and contains multiple price-related features such as open, close, high, low, volume and adjusted close prices. Non-numeric columns including date and volume columns are dropped. Subsequent data preprocessing, including de-duplication, missing value processing and normalization was performed to make it suitable for use in machine learning models.

To predict Walmart stock prices, this paper utilizes different machine learning models. Dedicated to the study of Recurrent Neural Networks (RNNs), which include the fundamental RNNs, LSTMs, LSTMs with peephole connectivity, and GRUs. The objective is to provide accurate predictions of future stock prices based on historical data. MSE, which measures the predictive accuracy of the models, is the primary evaluation metric used for all four of the models in this paper. A more sensitive picture of the model's predictive performance and an easier time calculating the gradient are provided by the MSE measure, which is the mean of the squared difference between predicted and actual observations. This measure amplifies the prediction error and makes the effect of larger errors on the overall error more significant.

3 Analysis of results

This section focuses on evaluating each model's performance, advantages, and disadvantages using metrics like mean square error (MSE) and additional observations from training and validation. This paper uses the correct symbolic prediction to illustrate the relationship between the closing price and the opening price. A model's ability to accurately predict the direction of price movement—that is, whether the closing price on a given day will be higher or lower than the opening price—is known as correct sign prediction. This metric is critical for traders because it can be more valuable to predict the direction of price movement than the exact price.

3.1 Analysis of the SimpleRNN model

In the SimpleRNN model, the test target and test predictions did not overlap highly, as shown in Fig.1. Regarding the training data, the model does reasonably well, correctly predicting the direction of change in price in 72% of the cases. This indicates that the model effectively learns the patterns in the training data. The high percentage of correct symbol predictions in training suggests that the model can capture specific trends and patterns in the training dataset. However, if the validation and testing results are significantly low, this may also indicate possible overfitting. The model's accuracy in symbol prediction falls to 54% when applied to the validation dataset. This sharp drop compared to the training data indicates the model's limited ability to generalize to unseen data. The model's accuracy in symbol prediction falls even further to 47%, below the 50% level of the random guess. This confirms the model's poor generalization ability and highlights how difficult it is to predict the direction of price movement for newly discovered data.

The initial loss of 0.0169 is relatively high, but the loss improves significantly as training progresses, indicating good learning. The learning rate remains decreasing during the training process, which usually helps to fine-tune and stabilize the training process as the model approaches convergence. The model fits the training data, as evidenced by the very low final training loss of 2.95× 10^-5. The gradient vanishing problem, which is well known to cause problems for simple RNNs when handling long-term dependencies, could account for the relatively high validation error when compared to other models. The test MSE value is 9.83× 10^-4. This slightly higher test loss indicates that the model performs well, but slightly worse when exposed to brand new data compared to the validation set. The training logs show that the model learns well and reduces losses significantly from the start. The use of the culling layer helped prevent overfitting, while the learning rate scheme allowed the model to converge effectively. The final evaluation metrics show that the model fits well and generalizes well.

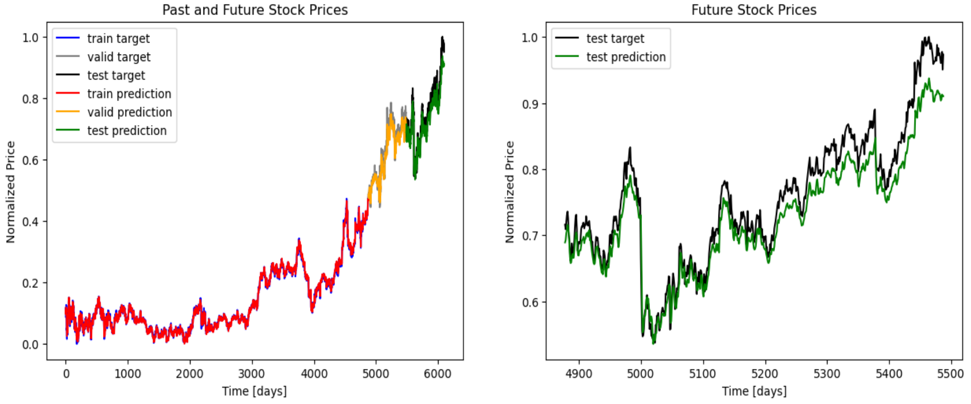

Fig. 1. The correct sign prediction results for the closing-opening price with the RNN model (Photo/Picture credit: Original)

3.2 Analysis of the LSTM model

Compared to SimpleRNN, the LSTM model predicts the correct sign prediction accuracy of the close-open price of 74% in training, 45% in validation, and 52% in testing, all of which are better than SimpleRNN MODEL shown in Fig 2. The performance of LSTM is improved and the training and validation MSE values are significantly lower. The model exhibits a good fit to the training data, as evidenced by the extremely low training MSE of 3.005 × 10^-5. Furthermore, the test MSE of 1.47× 10^-4 and validation MSE of 1.17× 10^-4 are marginally higher than the training MSE, suggesting that the model can satisfactorily fit the training set. The training time was 65 out of 100 periods, which suggests that improvement reaches a steady state at epoch 65 and training is stopped early. The learning rate starts at 0.001 and then systematically decreases to 1.00e-06 at the end of training. The initial loss of the model is 9.42× 10^-4. There is a significant reduction in both the training and validation losses, indicating a significant fit between the model and the data. Starting at time 14, the learning rate was halved several times, allowing the model to converge more accurately as it approached the optimal parameters. After a small fluctuation as the learning rate is lowered, the loss values stabilize in epochs 14 to 65. This indicates that the model is being fine-tuned, with smaller adjustments to the weights. In the final epochs, there is minimal change in the losses, indicating that the model has found a satisfactory solution.

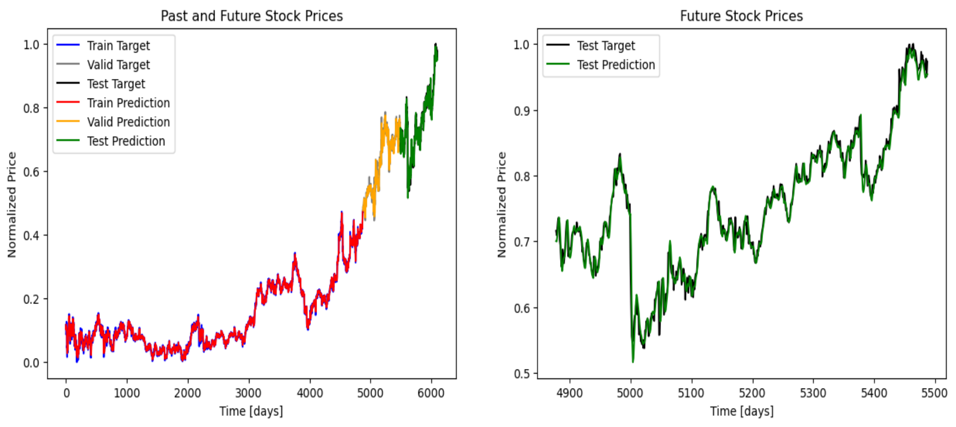

Fig. 2. The correct sign prediction results for the closing-opening price with the LSTMs model (Photo/Picture credit : Original)

3.3 Analysis of the LSTM model with peephole

The LSTM model with peephole connections is a variant of the LSTM network that includes “peephole connections”. These connections allow the gate layer to view the state of the cells, thus potentially improving the capabilities of the model. In order to aid in convergence and avoid overfitting, the model was trained by reducing the learning rate and stopping early. The final training results show a very low value of 3.01× 10^-5 for the training MSE. This indicates a good fit for the training data. The actual and expected values for the training, validation, and test sets are displayed in the first subplot. It is clear that there has been good learning when the training predicted values (red) closely match the actual target values (blue). However, there may be some difficulties with generalization, as evidenced by the stark difference between the test's (green) predicted values and the validation's (orange) predicted values, as well as the test's actual target values, which are black.

The model's performance on raw data, where the differences are even more noticeable, is highlighted in the second subfigure, which concentrates on the test data in Fig. 3. It is possible that overfitting occurred or that the extra complexity of the peephole connections had no discernible impact on the model's performance on this dataset, as the test MSE for the peephole LSTM model is marginally higher than that of the standard LSTM. It is possible that overfitting occurred or that the extra complexity of the Peephole connection did not appreciably improve the model's performance for this dataset, as indicated by the slightly higher test MSE for the Peephole LSTM model when compared to the standard LSTM. Despite the higher-than-expected test error, the Peephole LSTM model demonstrated effective learning with low training and validation errors. The model may benefit from additional tuning or regularization to prevent overfitting, according to the final test mean square error (MSE). The use of peephole connections, while theoretically beneficial, did not show a clear advantage over standard LSTMs in this particular task. Trying different architectures or additional regularization techniques may lead to improvements.

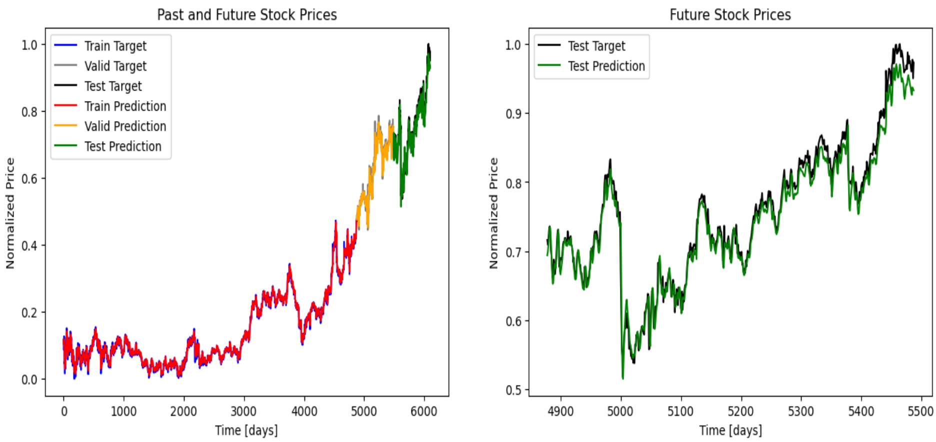

Fig. 3. The correct sign prediction results for the closing-opening price with LSTMs model with peephole connection(Photo/Picture credit : Original)

3.4 Analysis of the GRU model

The GRU model uses a stack of GRU units with leaky ReLU activations to predict stock prices. An input layer, a GRU layer, and a dense output layer make up the architecture. As shown in the first subplot of Fig. 4 the training predictions (red) overlap with the actual target values (blue), indicating good learning. There is no significant difference between the validation prediction (orange) and test prediction (green) and their actual target values (grey for validation and black for test) indicating good generalization. The model outperforms the earlier models on unseen data, as evidenced by the second subplot.

The model is trained over 100 epochs during the training process, with 98 batches making up each calendar element. The model's training loss and validation loss both get smaller as training goes on and the learning rate is changed. The training loss decreases significantly from the initial 0.0012 to 2.43 × 10^-5. The validation loss decreases from 4.98 × 10-4 to 1.24 × 10^-4. Since the model is applied to unseen data during validation, it is expected that the validation MSE will be slightly higher than the training MSE. The final test MSE is 1.47 × 10^-4. The relatively small difference indicates that the model generalizes well and is not overfitted. The test MSE is slightly higher than the training and validation MSEs, reflecting the model's performance on the test dataset that was kept separate. This increase is common and acceptable, indicating that the model maintains consistent performance across datasets.

Additionally, by limiting the learning rate, the loss is stabilized, overfitting is avoided, and more nuanced patterns in the data are captured. The model's learning rate starts at 0.001 for the first 11 epochs, and then the learning rate decreases systematically in a process called learning rate decay. At the 17th epoch, it falls to 0.00025, and at the 22nd epoch, it rises to 0.000125. It is finally set to 0.0000625 at the 27th epoch. On the training, validation, and test datasets, the GRU model achieves low error rates, making it a good fit for this task. GRU's performance is comparable to that of the LSTM, with the validation MSE being slightly higher, but the training MSE is similar. However, the training MSE is similar, suggesting that although GRU learns efficiently, it may not be able to capture as much complexity as the LSTM model in some cases.

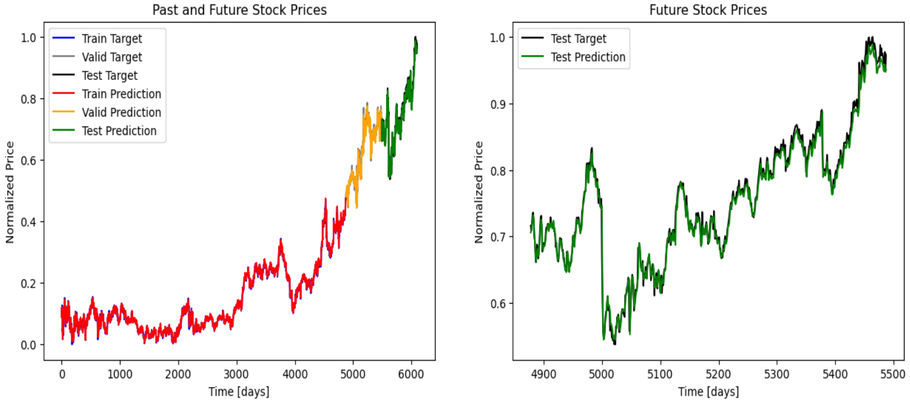

Fig. 4. The correct sign prediction results for the closing-opening price with the GRU model (Photo/Picture credit : Original)

4 Conclusion

In order to increase the accuracy of stock price forecasts, this study investigates the application of artificial intelligence in stock forecasting. The study used a machine learning algorithm to analyze Walmart's Stock Price data to more accurately predict future stock prices. This paper concludes that the LSTM model has the best test mean square error (MSE) with a value of 0.000147, followed by the GRU model (0.000147). In comparison, the MSE of the basic RNN is significantly higher (0.000983), which indicates its weak generalization ability. The MSE of basic LSTM is lower than SimpleRNN, showing its advantage in generalization ability. However, the complex structure of the basic LSTM results in a long training time. The MSE of LSTM with peephole connections is higher than that of basic LSTM and GRU, which may originate from the overfitting phenomenon or insufficient learning efficiency. In comparison, the GRU model has a smaller number of parameters, is faster to train, and has a lower risk of overfitting. Its overall performance is comparable to the basic LSTM, but the training MSE is slightly higher than the LSTM. If further improvements are desired, consideration can be given to exploring hybrid models, ensemble methods, or additional hyperparameter tuning.

References

1. D. Shah, H. Isah, F. Zulkernine, Stock market analysis: A review and taxonomy of prediction techniques. International Journal of Financial Studies, 7(2), 26 (2019)

2. H. Grigoryan, A stock market prediction method based on support vector machines (SVM) and independent component analysis (ica), *Database Systems Journal*, vol. 7, no. 1, pp. 12-21 (2016).
3. Bhattacharjee and P. Bhattacharja. Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach, 4th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, pp. 1-6 (2019) doi: 10.1109/EICT48899.2019.9068850.
4. Mintarya, Halim, J.N., C.Angie, S. Achmad, A.Kurniawan, Machine learning approaches in stock market prediction: A systematic literature review. Procedia Comput. Sci. 216, 96–102 (2023)
5. W. Lu, J. Li, J. Wang, L. Qin, A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications, 33, 4741–4753 (2021).
6. Dattatray P. Gandhmal, K. Kumar, Systematic analysis and review of stock market prediction techniques. *Computer Science Review, 34*,100190 (2019)
7. M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, & S. Shahab. Deep learning for stock market prediction. *Entropy, 22*(8), 840 (2020)
8. Yu, P., Yan, X. Stock price prediction based on deep neural networks. *Neural Comput & Applic* **32**, 1609–1628 (2020).
9. J. Zhang, S. Cui, Y. Xu, Q. Li, T.Li, A novel data-driven stock price trend prediction system. *Expert Systems with Applications, 76*, 20982118 (2020)
10. M.Vijh, D. Chandola, T. A.Vonary, A. Kumar, Stock closing price prediction using machine learning techniques. *Procedia Computer Science, 167*, 599-606 (2020)