

STOCK PREDICTION BY LSTM ALGORITHM

B. TECH CSE WITH ML AND AI,

GRAPHIC ERA HILL UNIVERSITY

ABSTRACT

The financial market's movement is always stochastic and unexpected, and the return on a security is thought to be uncertain. At this time, analysts are attempting to integrate modelling methods from Natural Language Processing into the resemblance of the data's sequential quality in the realm of finance.

The Long Short Term Memory Model (LSTM), Stacked LSTM, and Attention-Based LSTM state-of-the-art deep learning sequential models, as well as the conventional ARIMA model, have been developed and implemented in this study to make predictions about stock prices for the following day. Additionally, based on our prediction, we developed two trading strategies and evaluated them against the benchmark. In addition to the customary end-day price and trading volumes, our input data also includes corporate accounting.

This article introduces the use of recurrent neural networks (RNN) and long short-term memory cells (LSTM) for stock market forecasting in portfolio management while taking into account time series historical stock data for the portfolio's holdings. Regression, Support Vector Machine, Random Forest, Feed Forward Neural Network, and Back propagation are some of the more well-known machine learning algorithms that have been compared to the model. Numerous metrics and architectural designs for the LSTM RNN model have been taken into consideration, tested, and analysed. There is discussion regarding the impact of changing trends and customer sentiment on stock prices.

INTRODUCTION

As the future trading strategy is always implemented and created based on our view of the financial market in the future, predicting the future security returns is at the core of the quantitative trading industry. Quantitative trading and fundamental analysis are the two main methods used in the trading world. The fundamental method, which primarily relies on publicly available information like market news, corporate statistics, and a string of financial statement releases, bases trading decisions on a subjective assessment of an industry's or company's future direction. On the other hand, the quantitative trading strategy avoids the disruption of human subjectivity and emotion by using mathematical models to make the decision. Traditionally, the quantitative strategy methodology entails

Stock market has gained extensive interest from investors. Investors and researchers have long been interested in learning how to understand the shifting regularity of the stock market and forecast the trajectory of stock prices. Politics, the economy, society, and the market all have an impact on how stock prices grow and decrease. The stock market's trend projection is directly tied to the acquisition of profits for stock investors. The forecast's ability to effectively avoid dangers increases with its accuracy. For publicly traded companies, the stock price not only represents current operating conditions and aspirations for future development, but also serves as a crucial technical gauge for corporate analysis and study. Research on stock forecasts is equally crucial to the major goal of this study is to develop a deep network model that can concurrently forecast a stock's starting price, lowest price, and highest price the next day based on the stock's history price as well as other technical parameter data. In order to forecast the three associated values, a deep recurrent neural network model based on LSTM is suggested (so it is called the associated neural network model, and abbreviated as associated net model). By contrasting the accuracy of the three models, the associated net model is put up against LSTM and LSTM-based deep recurrent neural networks, demonstrating the model's viability.

LITERATURE REVIEW

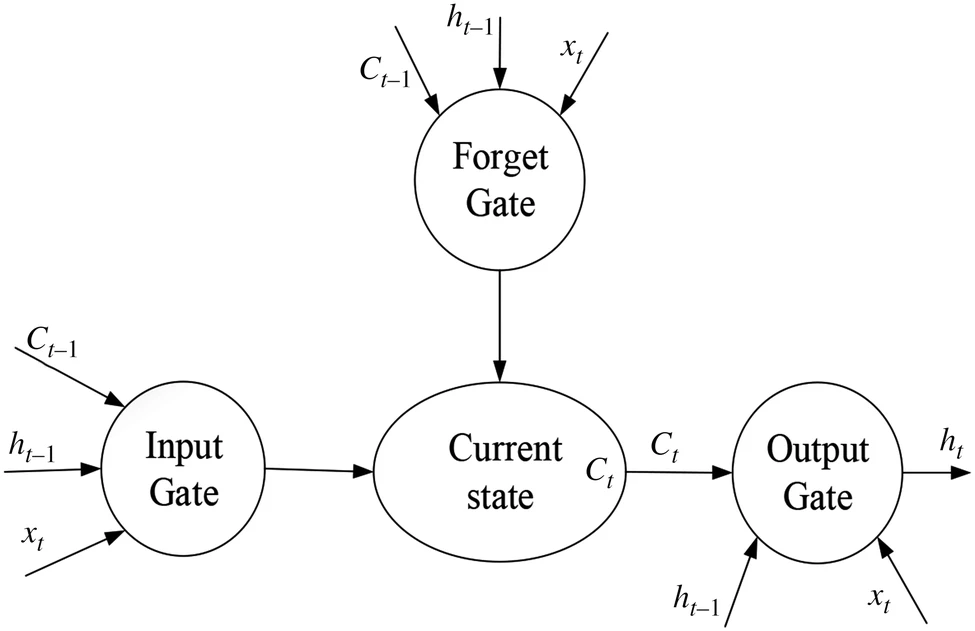
1. The classical time series model has been shown to have good predictive power up to a point in the pre-deep learning era, where it was mostly focused on the field of ARIMA and any changes on this. For instance, the asymmetric Min young Kim has substituted an asymmetric loss function for the conventional Maximum Likelihood Estimation in the context of financial time series return. The performance of ARIMA and artificial neural networks in forecasting the Korean stock price index was compared by C.K. Lee et al. The research demonstrated that ARIMA forecasts were more precise than those produced by a back-propagation neural network. Deep learning techniques have recently performed better because of increased computer power and the capacity to discover non-linear correlations included in a variety of financial characteristics. For the prediction of NSEI listed equities, Selvin etc al. compared three different deep learning architectures, including RNN, LSTM, and CNN-sliding window models. They came to the conclusion that CNN architecture outperforms other models and is capable of detecting changes in stock trend. Yan and Ouyang demonstrated that the wavelet transform of the financial time series paired with the LSTM outperformed classical SVM and K-nearest Neighbors in terms of performance. In order to increase forecast accuracy, Thien Hai Nguyen et al. integrated sediment features that were taken from social media. By supplying pertinent data based on financial domain expertise, the performance of LSTM-RNN would be significantly improved. In addition, Kim Won created a hybrid strategy that combines LSTM and GARCH models, and the resulting model has significantly lower prediction errors
2. Stock price prediction has been the subject of numerous related studies. To create a regression model using historical stock data and forecast the trajectory of stocks, support vector machines were used. The support vector machine's parameters are optimised using the particle swarm optimization algorithm, which can accurately predict stock value . The support vector machine method is improved by this study, although the particle swarm optimization algorithm is computationally expensive. In order to increase the effectiveness of prediction, LSTM and naïve Bayesian technique were coupled . This approach can be used to forecast financial markets using additional factors across completely different time horizons. To create a reliable time series model, the emotional analysis model was merged with the LSTM time series learning model. Jia talked on how well LSTM predicts stock price, and the research indicated that LSTM is a good way to anticipate stock profits. To solve several logical flaws in earlier studies, real-time wavelet denoising and an LSTM network were coupled to predict the east Asian stock index. This combination model is far better than the original LSTM, with high prediction accuracy and minimal regression error. Each neural network was trained using the back propagation method and the Adam optimization algorithm, and the results show that the method has varying degrees of accuracy for predicting different stock indices, but the prediction on the Shanghai composite index and the Shenzhen component index was the most accurate.
3. The findings indicate that this strategy outperforms the technical analysis method in terms of daily stock price prediction accuracy. For the Dhaka Stock Exchange (DSE), an efficient soft computing method was developed to forecast closing prices. This strategy is more successful, according to a comparison experiment with artificial neural networks and adaptive neural fuzzy reasoning systems. For stock price forecasting, the artificial bee colony technique was integrated with wavelet transforms and recurrent neural networks. The Dow Jones industrial average (DJIA), the London FTSE 100 index (FTSE), the Tokyo Nikkei-225 index (Nikkei), and the Taiwan stock exchange Capitalization Weighted Stock Index were among the many foreign stock indices that were mimicked for evaluation (TAIEX). The simulation outcomes demonstrate that the system can be used and has a good prediction performance.
4. Based on the preceding research, and given that some parameters and indicators of a stock are related, it is necessary to develop a multi-value associated neural network model that can handle multiple associated prices of the same stock and output these parameters and indicators at the same time. For this purpose, an associated neural network model based on LSTM deep recurrent network is proposed for predicting the opening price, lowest price, and highest price of the stock on the following day using historical data
5. A number of attempts have been made to predict the trend of stock prices using textual information . For example, Lavrenko etc combined stock price and financial news article trends and predicted the trends using news article content before the trends appeared. The predicted trends in their simulation were profitable. Schumaker and Chen compared several textual representations for stock price prediction, including Bag-of-Words, Noun Phrases, and Named Entities. They demonstrated that Bag-of-Words is insufficient, and that support vector machines (SVM) with proper noun features outperform in predicting stock price trends. Hagenau et al. predicted the difference between a stock's open and close prices using data from DGAP and Euro Adhoc, which are corporate announcements from Germany and the United Kingdom, respectively. They used bigram, a sequence of two adjacent words, and a 2-word combination as features, and selected features based on the 2 statistic for each brand of stock prices. Their experiment demonstrated that the Chi-Square-based feature selection method improved classification accuracy while reducing over fitting. These works, however, share a common issue: the inefficiency in dealing with large amounts of textual information. Ding et al used Open IE (Information Extraction) techniques to predict the S&amp P 500 index by extracting the actor and object of events from news article titles. In their experiments, they used the deep neural network model as a classifier and outperformed SVM.

MODEL DESIGN

Long short-term memory network

The reason behind using this because this will give us one of the best result and have accurate result as comparison to other LSRM part in this part I will explain the LSTM and its working , later I will explain its algo and parameter.

Long short-term memory networks (LSTM) are a subset of recurrent neural networks (RNNs), which are a class of neural networks capable of processing sequential data. LSTM is a network structure that consists of three "gate" structures. An LSTM unit contains three gates: an input gate, a forgetting gate, and an output gate. Rules can be applied to information as it enters the LSTM network. Only information that conforms to the algorithm will be retained, while information that does not conform will be erased via the forgetting gate.



The gate allows information to be passed selectively, and depicts the sigmoid activation function of the LSTM network. Through the gating unit, the LSTM can add and delete information for neurons. It consists of a Sigmoid neural network layer and a pair multiplication operation to determine whether information is passed or not. Each Sigmoid layer element is a real number between [0, 1], representing the weight through which the corresponding information passes. There is also a layer in the LSTM neural network that contains the tanh activation function, as shown in. It is used to maintain the state of neurons.

σ(x)=1/1+e−x

tanh(x)=e^x−e^−x/e^x+e^−x

The LSTM neural network's forgetting gate determines what information should be discarded, which reads ht1 and xt and gives the neuron state Ct1 a value of 0-1. The calculation method for forgetting probability is

ft=σ(Wf⋅[ht−1,xt]+bf)

where ht1 is the previous neuron's output and xt is the current neuron's input. The sigmoid function is. The amount of new information added to the neuron state is determined by the input gate. The input layer with the sigmoid activation function first determines which information needs to be updated, and then a tanh layer generates candidate vectors ct, and the neuron's state is updated.

Ct=ft∗Ct−1+it∗C^t

2 Time Series Model:

The Auto Regressive Integrated Moving Average (ARIMA) model is a popular statistical method for forecasting time series (equation 1). We used the Box-Jenkins Methodology to create an ARIMA model as a baseline to compare to Deep Learning models in this work. To fit the ARIMA model, only "adjusted close price" was used. To identify data trends and the parameters (p, d, and q) of the ARIMA model, we used summary statistics and functions such as moving average and autocorrelation function.

## **Design of algorithm and experiments**

The regression method is used to predict a specific value, which is an arbitrary real number rather than a pre-defined category. In general, a regression problem has only one output, which is the predicting value. The mean square error (MSE) is a common loss function used in regression problems . It is the expected value of the square of the difference between the estimated and actual parameters. MSE can assess the function's degree of change. The lower the MSE value, the more accurate the prediction model describing the experimental data. As a result, during the training phase, MSE is used as a criterion to assess the quality of a network model. MSE(y,y′)=∑ni=1(yi−y′i)^2/n

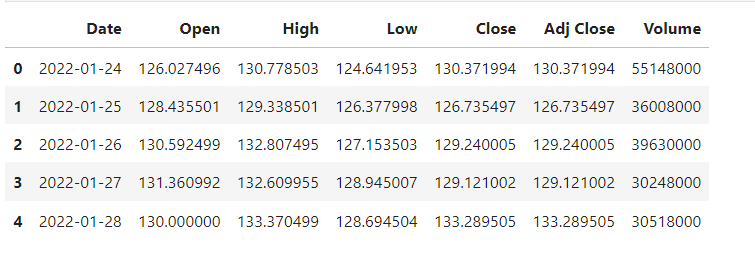
### Algorithm

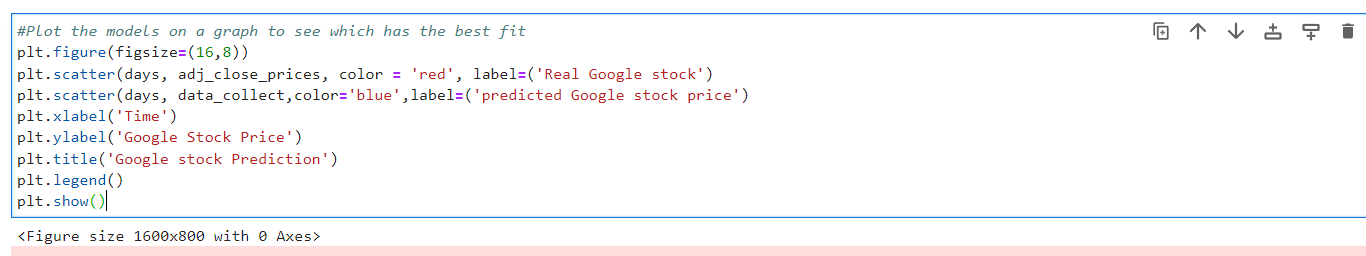
Deep learning often necessitates a significant amount of time and computational resources to train. It is necessary to develop an optimization algorithm that uses fewer resources and has a faster convergence speed. The Adam optimization algorithm is a stochastic gradient descent extension that has significant advantages in solving non-convex optimization problems. The Adam optimization technique is employed in the model during the training phase, and L total is used as the evaluation function. The initial input sequence data to the related net model has three DRNN networks. It has many values associated with the neural network model method framework. Each of the three DRNN networks results in a loss, and the total loss is the sum of the losses from all three DRNN networks. The total loss is then optimised using the Adam algorithm. The training will continue to lower the overall loss if the number of iterations falls short of the model's predetermined number; else, it will end.

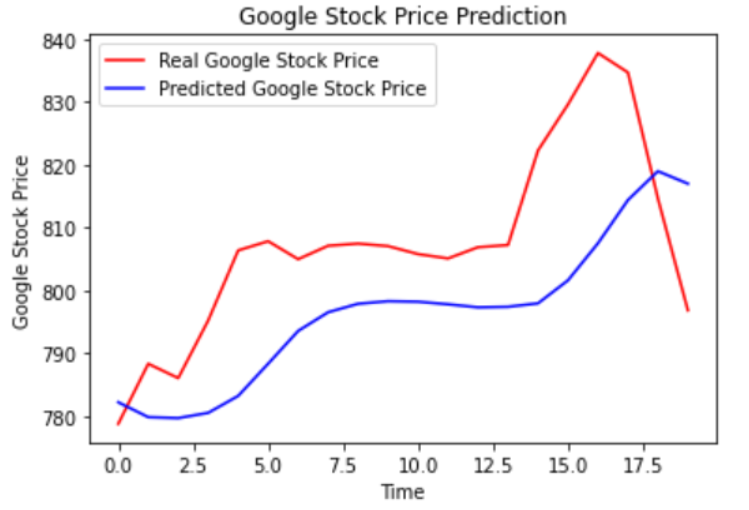
DATASET AND FEATURES

We used two types of data to train, create, and test our model: 1. From 2004 to 2013, the daily prices and volumes for each SP 500 stock From 2004 to 2013, the accounting and business statistics for the SP 500 stocks. By combining two sets according to release dates, missing statistics between the two sets were filled in. The inputs we utilise are "adjusted close price," "trading volume," "Debt-to-Equity Ratio," "Return on Equity," "Price-to-Book" ratio, "Profit Margin ,"Diluted Earnings Per Share," and "Company Beta" after a significant number of iterations in determining the right input.

In order to standardise the data between 0 and 1, we used the min-max scale, prevent others from being overwhelmed by the size of some features. We have the daily information provided above for each stock in the SP 500 from 2004 to 2013. We divide the data for each stock into training, development, and testing data using a rough ratio of 70-15-15. In other words, we use 2013 to 2011 data as training data, 2012 to 2013 data for development, and 2013 to 2013 data for testing. We therefore base our success measures and trading plans on the data from 2013.







This is the graph of result and this show the prediction value and purpose of using LSTM so that accuracy of result will be good or close to real price.

Proposed Work:

To create a algorithms program that give a stock prediction for an institution or for an individual traders. Usually we have to get information from paid platform and have to depend on some third party company. The main problem with them is that some time they are incorrect and we get late data.so I decide to build my own program to get accurate and early data of stocks price . I use LSTM for model features for better results and more accuracy**.**

Results

In this project, we will forecast the return on the following trading day using historical data. Our main model performance statistic is Mean Squared Error, which we use to evaluate the model (MSE). The MSE will be run on test data, which is completely different from the development and training phases. The difference between our forecasted price and the actual price is determined.

Our trading approach and the secondary model performance metric are connected. Our two intraday trading methods are developed based on our performance forecast for 2013:

🡪Long-Short Strategy: If a stock's prediction for the following day is favourable, we buy the stock at its open price and sell it at its closure price on the same day. If the forecast is unfavourable, we short the stock at market open and close the short position at market close.

Keep in mind that our trading method is an intraday trading strategy, which means that we only act on the same day and do not hold positions overnight.

All 500 stocks in the SP universe make up our initial sample. After conducting various tests, we made the decision not to employ all 500 stocks and instead focused on the top 10 SP 500 stocks by market size. Due to the fact that higher market capitalization typically indicates more reliable financial standing, predictable growth, and comprehensive public information from the companies Tanh activation function for the LSTM layers and sigmoid activation function for the output layers were selected for the LSTM models since these activation functions have been shown to produce better outcomes than other activation functions. Additionally, in addition to using the standard regularisation procedure for each hidden layer, a dropout probability of 20% is also applied to address the over fitting problem. The mean squared error is employed as the loss function and the optimization is used to learn the parameters.

We employ mini-batch in our training procedure to quicken the training process and improve the likelihood of convergence. After testing a wide range of mini-batch sizes, we ultimately settled on using the size of 30, as larger quantities do not always guarantee convergence and can occasionally imprison the training process in the saddle point. On this study, we experimented with a huge number of look back days and ultimately opted to utilise look back days of 20 days, which means that we will be using the data from the previous 20 days to predict the stock prices in the following day.

Conclusion:

In order to forecast stock price movements, a forecasting framework is established in this study. As input data, we made use of the combinations of price, volume, and company information. Four models were proposed, created, tested, and trained: the ARIMA, LSTM, Stacked-LSTM, and Attention-LSTM models. Based on the model predictions, we designed long-only and long-short trading strategies. Due to its capacity to apply various weights to the input features and hence automatically select the most pertinent features, the attention-LSTM exhibits better outcomes than previous models. As a result, the Attention-LSTM is better able to detect long-term dependence in time series and is better suited to forecast financial time series. The Attention-exceptional LSTM's trading return further supports our experimentation outcomes In addition, we have demonstrated that, due to the risk of over fitting, stacked-LSTM models do not perform any better than single LSTM models, despite having more complex model structures. The volatility of stock time series will be one area of future research. The non-stationary nature of the stock market makes predictions challenging. Observing Attention-performance LSTM's on data would be fascinating. Even though our suggested solution produced a respectable result, this research provides more room for further investigation. Additionally, we discovered that the RFE algorithm is not sensitive to term lengths other than 2-day, weekly, and biweekly during the review process. A potential direction for future research is to do more in-depth analysis of the technical indices that might affect the irregular term lengths. Additionally, there is a great potential to create a more thorough prediction system that is trained using a variety of data sources including tweets, news, and other text-based data by merging the most recent sentiment analysis techniques with feature engineering and deep learning models.

**REFERENCES**

**1.** Li, R., Fu, D., & Zheng, Z. (2017). An Analysis of the Correlation between Internet Public Opinion and Stock Market. In 2017 4th International Conference on Information Science and Control Engineering (ICISCE) (pp. 150–153). IEEE.

**2.** Tiwari, V., Gupta, S., & Tiwari V. (2010). Association rule mining: A graph based approach for mining frequent itemsets. International Conference of Networking and Information Technology (ICNIT) (pp. 309–313). IEEE.

**3.** Kunal, S., Saha, A., Varma, A., & Tiwari, V. (2018). Textual Dissection of Live Twitter Reviews using Naive Bayes. *Procedia Computer Science, 132*, 307–313. Elsevier.

**4.** Khaidem, L., Saha, S., & Dey, S. R. (2016). Predicting the direction of stock market prices using random forest. arXiv preprint arXiv:[1605.00003](https://arxiv.org/pdf/1605.00003)

5. Al-Nasseri, A., & Ali, F. M. (2018). What does investors' online divergence of opinion tell us about stock returns and trading volume?. *Journal of Business Research, 86*, 166–178.

6. Lin, Y., Guo, H., & Hu, J. (2013). An SVM-based approach for stock market trend prediction. In Neural Networks (IJCNN), The 2013 International Joint Conference on (pp. 1–7). IEEE.

7. Gupta, A., & Dhingra, B. (2012). Stock market prediction using hidden markov models. In Engineering and Systems (SCES), 2012 Students Conference on (pp. 1–4). IEEE.

8. Bhuriya, D., Kaushal, G., Sharma, A., & Singh, U. (2017). Stock market prediction using linear regression. International conference of electronics, communication and aerospace technology (ICECA), Coimbatore, India. IEEE.

9. Yang, J., Rao, R., Hong, P., & Ding, P. (2016). Ensemble model for stock price movement trend prediction on different investing periods. In Computational Intelligence and Security (CIS), 2016 12th International Conference on (pp. 358–361). IEEE.

10. Labiad, B., Berrado, A., & Benabbou, L. (2016). Machine learning techniques for short term stock movements classification for Moroccan stock exchange. In Intelligent Systems: Theories and Applications (SITA), 2016 11th International Conference on (pp. 1–6). IEEE.

11. Mingyue, Q., Cheng, L., & Yu, S. (2016). Application of the Artifical Neural Network in predicting the direction of stock market index. In 2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS) (pp. 219–223). IEEE.

12. Mithani, F., Machchhar, S., & Jasdanwala, F. (2016). A modified bpn approach for stock market prediction. In Computational Intelligence and Computing Research (ICCIC), 2016 IEEE International Conference on (pp. 1–4). IEEE.

13. Sharma, A., Bhuriya, D., & Singh, U. (2017). Survey of stock market prediction using machine learning approach. International conference of electronics, communication and aerospace technology (ICECA), Coimbatore, India. IEEE.

14. George, S., & Changat, M. (2017). Network approach for stock market data mining and portfolio analysis. In Networks & Advances in Computational Technologies (NetACT), 2017 International Conference on (pp. 251–256). IEEE.

15. Dewan, A., & Sharma, M. (2015). Prediction of heart disease using a hybrid technique in data mining classification. In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on (pp. 704–706). IEEE.

16. Sak, H., Senior, A., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In Fifteenth annual conference of the international speech communication association.

17. Heaton, J., Polson, N., & Witte, J. (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, *33*(1).

18. Mrcela, L., Mercep, A., Begusic, S., & Kostanjcar, Z. (2017) Portfolio optimization using preference relation based on statistical arbitrage. In Smart Systems and Technologies (SST), 2017 International Conference on (pp. 161–165). IEEE.

19. Fang, J., & Xiaoyun, M. (2017) Dynamic Multi-Mode Portfolio Optimization Strategy for Markovian Arrival Process. In Robots & Intelligent System (ICRIS), 2017 International Conference on (pp. 139–142). IEEE.

**20.** Das, S., & Goyal, M. (2012) Rebalancing a two-asset Markowitz portfolio: A fundamental analysis. In Computational Intelligence for Financial Engineering & Economics (CIFE), 2012 IEEE Conference on (pp. 1–8). IEEE.