Stock Prediction Based on LSTM Under Different Stability

Mrs. T.Jeevitha[1], A.Deva Nivas[2], S.Gokulprasanth[3], K.Jothi Prakash[4], K.Kamalesh[5]
[1]Professor, Dept of Technology in Information Technology
[2, 3, 4, 5]Dept of Technology in Information Technology
[1, 2, 3, 4, 5]Info Institute of Engineering, Coimbatore, India

Abstract- *The boom of Big Data has made the development of prediction algorithms more intelligent, so the studies have gradually shifted from the traditional linear prediction algorithm to the popular deep learning prediction algorithm. Long Short Term Memory network (LSTM) is a special algorithm for processing time-series problem. In this work, we conducted a stationary analysis of the stock's time-series data and then used the LSTM neural network algorithm to predict stock data under different stationary conditions, and performed statistical analysis on multiple experimental data. A large number of experimental results show that the LSTM neural network prediction algorithm has higher prediction accuracy and is not sensitive to the stability response.*

I. INTRODUCTION

Intuitively, the stationary time-series is a sequence whose statistical characteristics will not change with time shifting (such as mean and covariance), or means there is no obvious tendency or periodicity . Stock price forecasting is the most common forecast based on time-series data in financial market. Stock price data are noisy, complex and non-linear which is easy to be affected by many factors such as policy, economy and psychology. we need to conduct data pre-processing and analysis before the prediction Traditional prediction methods cannot accurately capture historical information, which often rely on near regression and parameter estimation. models Deep learning is a method of machine learning which developed from the original neural network. Recurrent Neural Network (RNN) adds the concept of time to its network structure, making it dedicated to processing time-series data.If the sequence is too long, the vanishing gradient problem will occur during optimization proposed a Long Short Term Memory network in The bidirectional LSTM and the stacked LSTM with the simple LSTM were compared, and the results showed that the performance of the bidirectional LSTM is the best in stock prediction in . In these models, scholars have confirmed the superiority of the LSTM prediction model over traditional prediction models, which can be applied to various nonlinear time-series data, but have not studied the influence of time series stationary difference on prediction results. This Project mainly compares the influence of stock data under different stability conditions on the prediction results of LSTM model.The financial market is a dynamic and composite system where people can buy and sell currencies, stocks, equities and derivatives over virtual platforms supported by brokers.The stock market allows investors to own shares of public companies through trading either by exchange or over the counter markets. This market has given investors the chance of gaining money and having a prosperous life through investing small initial amounts of money, low risk compared to the risk of opening new business or the need of high salary career. Stock markets are affected by many factors causing the uncertainty and high volatility in the market. Although humans can take orders and submit them to the market, automated trading systems (ATS) that are operated by the implementation of computer programs can perform better and with higher momentum in submitting orders than any human. However, to evaluate and control the performance of ATSs, the implementation of risk strategies and safety measures applied based on human judgements are required. Many factors are incorporated and considered when developing an ATS, for instance, trading strategy to be adopted, complex mathematical functions that reflect the state of a specific stock, machine learning algorithms that enable the prediction of the future stock value, and specific news related to the stock being analysed.Time-series prediction is a common technique widely used in many real-world applications such as weather forecasting and financial market prediction. It uses the continuous data in a period of time to predict the resultin the next time unit. Many time- series prediction algorithms have shown their effectiveness in practice. The most common algorithms now are based on Recurrent Neural Networks(RNN), as well as its special type- Long-short Term Memory (LSTM) and Gated Recurrent Unit (GRU). Stock market is a typical area that presents time-series data and many researchers study on it and proposed various models. In this project, LSTM model is used to predict the stock price.

II. LITERATURE SURVEY

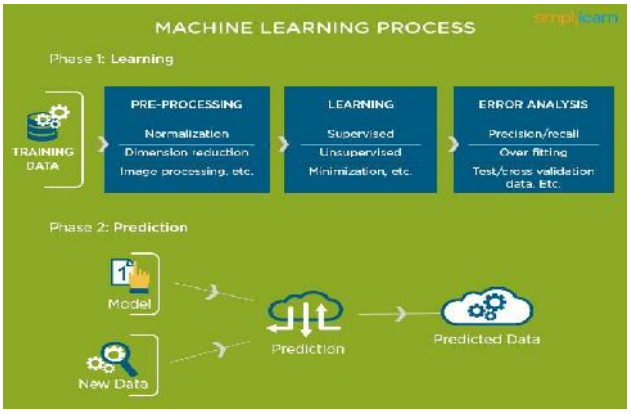
(1) Automated Stock Price Prediction Using Machine Learning

The research work done by Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut. Traditionally and in order to predict market movement, investors used to analyse the stock prices and stock indicators in addition to the news related to these stocks. Hence, the importance of news on the stock price movement. Most of the previous work in this industry focused on either classifying the released market news as (positive, negative, neutral) and demonstrating their effect on the stock price or focused on the historical price movement and predicted their future movement. In this work, we propose an automated trading system that integrates mathematical functions, machine learning, and other external factors such as news' sentiments for the purpose of achieving better stock prediction accuracy and issuing profitable trades. Particularly, we aim to determine the price or the trend of a certain stock for the coming end-of day considering the first several trading hours of the day. To achieve this goal, we trained traditional machine learning algorithms and created/trained multiple deep learning models taking in to consideration the importance of the relevant news. Various experiments were conducted, the highest accuracy (82.91%) of which was achieved using SVM for Apple Inc. (AAPL)stock.

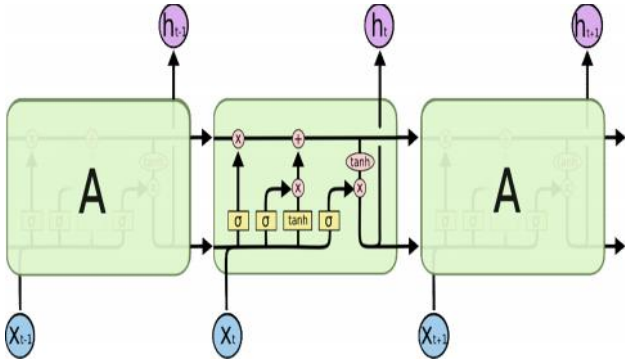
(2)Stock Price Correlation Coefficient Prediction with ARIMA- LSTM Hybrid Model

The research work done by Hyeong Kyu Choi, B.AStudent Dept. of Business Administration Korea University Seoul, Korea. Predicting the price correlation of two assets for future time periods is important in portfolio optimization. We apply LSTM recurrent neural networks(RNN) in predicting the stock price correlation coefficient of two individual stocks. RNN's are competent in understanding temporal dependencies. The use of LSTM cells further enhances its long-term predictive properties. To encompass both linearity and nonlinearity in the model, we adopt the ARIMA model as well. The ARIMA model filters linear tendencies in the data and passes on the residual value to the LSTM model. The ARIMA-LSTM hybrid model is tested against other traditional predictive financial models such as the full historical model, constant correlation model, single-index model and the multi-group model. In our empirical study, the predictive ability of the ARIMA-LSTM model turned out superior to all other financial models by a significant scale. Our work implies that it is worth considering the ARIMALSTM model to forecast correlation coefficient for portfolio optimization.

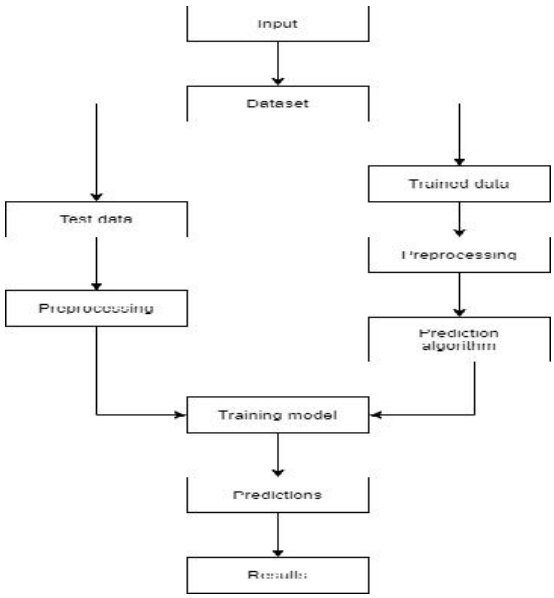
Machine Learning Process:



LSTM Architecture:



BLOCK DIAGRAM:



III. WORKING OF LSTM

LSTM is a special network structure with three "gate" structures. Three gates are placed in an LSTM unit,

called input gate, forgetting gate and output gate. While information enters the LSTM's network, it can be selected by rules. Only the information conforms to the algorithm will be left, and the information that does not conform will be forgotten through the forgetting gate. The experimental data in this paper are the actual historical data downloaded from the Internet. Three datasets were used in the experiments. It is needed to find an optimization algorithm that requires less resources and has faster convergence speed.

- Used Long Short-term Memory (LSTM) with embedded layer and the LSTM neural network with automatic encoder.
- LSTM is used instead of RNN to avoid exploding and vanishing gradients.
- In this project python is used to train the model, MATLAB is used to reduce dimensions of the input. MySQL is used as a dataset to store and retrieve data.
- The historical stock data table contains the information of opening price, the highest price, lowest price, closing price, transaction date, volume and soon.
- The accuracy of this LSTM model used in this project is57%.

IV. DATA COLLECTION

In this Project, we center around the preliminaries of stocks structure AAPL. By beginning, we brought the data legitimately from Y finance site, which has data on each stock value that is in the securities exchange for as long as years.

V. DATA PRE-PROCESSING

Pre-processing refers to the transformations applied to our data before providing the data to the algorithm. Data Preprocessing technique is used to convert the raw data into an understandable data set. In other words, whenever the information is gathered from various sources it is collected in raw format that isn't possible for the analysis.

VI. TRAINING DATA AND TEST DATA

- For choosing a model we split our dataset into train and test
- Here data's are split into 3:1 ratio that means
- Training data having 70 percent and testing data having 30 percent
- In this split process preforming based on train_test_split model
- After splitting we get xtrain xtest and ytrain ytest

VI. MODEL CREATION

- Contextualise machine learning in your organisation.
- Explore the data and choose the type of algorithm.
- Prepare and clean the dataset.
- Split the prepared dataset and perform cross validation.
- Perform machine learning optimisation.
- Deploy the model.

VII. MODEL PREDICTION

Predictive modeling is a statistical technique using machine learning and data mining to predict and forecast likely future outcomes with the aid of historical and existing data. It works by analyzing current and historical data and projecting what it learns on a model generated to forecast likely outcomes.

VIII. OUTPUTES

Figure : 1

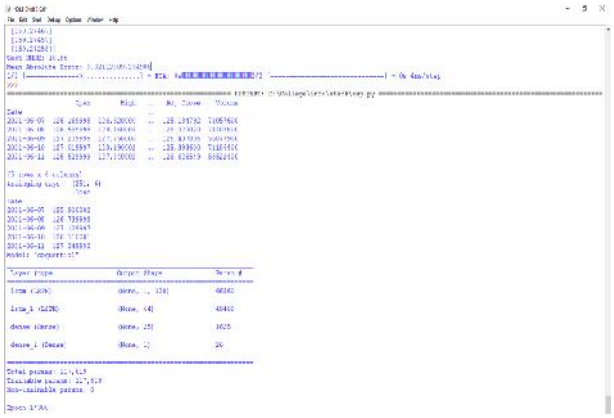


Figure : 2

Figure : 3

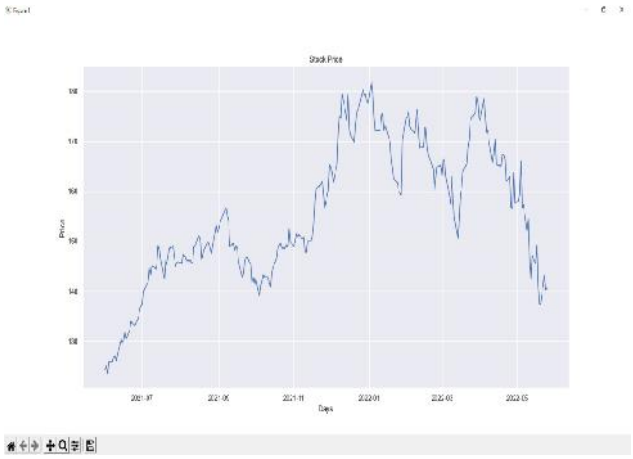
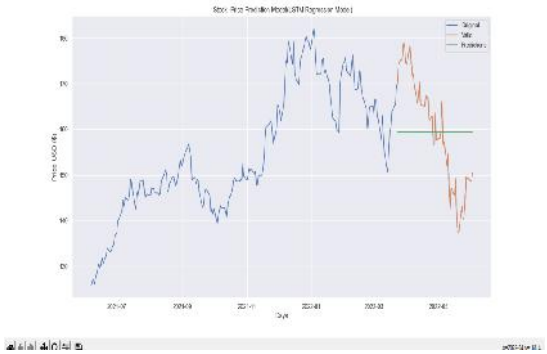


Figure : 4



IX. CONCLUSION

In this Project, we have studied the influence of stability's difference on LSTM stock price prediction. The stability of experimental results has little effect on the prediction results, but it has a slight influence on the convergence speed of the algorithm. The greater the P value is, the faster the convergence speed will be. At last, we conducted a comparative experiment with the LSTM model and the model, and the results showed that the error rate of the LSTM algorithm was 66.78% lower than that of. Therefore, we can draw a conclusion that LSTM algorithm performs better in prediction and has smaller errors. However, the disadvantage of LSTM algorithm is that it takes a lot of time to train the model and requires large sample of data.

X. FUTURE WORK

- We want to extend this application for predicting cryptocurrency trading.
- We want to add sentiment analysis for better analysis.

REFERENCES

[1] Jian feng ZHAO, "Modeling of Software Aging Based on Nonstationary Time Series," in International Conference on Information System and Artificial Intelligence, 2016, pp. 176-180

[2] Box, G. E. P. And Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco

[3] Javier Contier, "ARIMA Models to Predict Next-Day Electricity Price," in IEEE Transactions on Power Systems 2003, 18(3): pp. 1014-1020

[4] Gododfellow I, Bengio Y, Courville A. Deep learning [M]. The MIT Press, 2016