Long-term Pattern Forecasting in Stock Market using Multivariate LSTM with Expectation-Biased

Tam T. Vo
Faculty of Engineering, Science and Technology
Swinburne University of Technology
Ho Chi Minh City, Vietnam
Email: 103487596@student.swin.edu.au

***Abstract*—Stock Prediction using Deep Learning models such as Recurrent Neural Networks, or its advanced architecture such as Long Short-term Memory or Gated Recurrent Units is challenging since the market is affected by social and polical events. This paper aims to fulfill the task of stock patterns forecasting in long-term period regardless of any factors but five essential features in the Open-High-Low-Close (OHLC) Chart: Open, High, Low, Adjusted Close and Volume using the proposed model from A. Ismail, Timothy Wood, H. C. Bravo.**

***Keywords*: Deep Learning, LSTM, Multivariate Prediction, Long-term, Patterns Forecasting**

I. Introduction

The application of deep learning varies in different fields such as: Medical, Computer Vision, Electricity, even in Economics. In Medical Study, detecting Lung Cancer using Support Vector Machine (SVM) and Logistic Regression [1] requires very basic algorithms in Machine Learning. Moving to the era of deep learning, the very first Deep Learning model, which is Artificial Neural Networks (ANNs), is fed forward by the Histogram of Oriented Gradient (HOG) Feature Vector [2] to solve the problem of Lung Cancer Prediction. However, the impact of ANNs in the real world is limited since it could only handle structured datatype such as tabular data, images, numbers. In fact, scientists usually deal with unstructured datatype: audio, video, and text, etc. Hence, Sequence Model arrived as a solution to deal with those datatypes: Recurrent Neural Networks (RNNs). Then, a new architecture of RNNs was introduced to solve the problem of Gradient Vanishing: Long Short-term Memory (LSTM) [3]. LSTM is usually applied in the study of Sign Language Recognition (SLR), which combines 2 essential fields in Deep Learning: Computer Vision (CV) and Natural Language Processing (NLP). In SLR, a Convolutional Neural Networks (CNNs) is used to extract the feature of a video frame (which is an image) and feed it to the LSTM model in a single state. Doing so until reaching the last video frame, we will get the classification result in SLR. A group of authors proposed a customized architecture of LSTM to predict the electricity consumption [4].

Without the consideration in any polical events or social activities, data scientists dedicated their effort to solve the problem of stock prediction. In [5], five factors (Open, High, Low, Close, and volumes) are considered to make the stock prediction. The sequence of stock price is fed into the LSTM model to evaluate the output results. In this paper, we tried to implement the new proposed LSTM architecture from [4] to forecast the long-term patterns/trends of Apple Inc. in the next k sequential days based on five factors: Open Price, High Price, Low Price, Adjusted Close, Volume. Then, the Apple Stock Data is retrieved from Yahoo Finance is used as the dataset to be fed into our model and 3 built-in sequence models in Keras: RNNs, LSTM, and Gated Recurrent Units (GRUs) for comparison.

II. Methodology

A. LSTM Proposed Architecture

In [4], the authors introduced two advanced LSTM architectures: Multivariate LSTM with bias and single/multiple output(s). In this paper, we decided to implement the Multivariate LSTM with single output and bias to solve the case of stock paterns forecasting given a sequence of observations. The Figure 1 below is the implemented architecture:

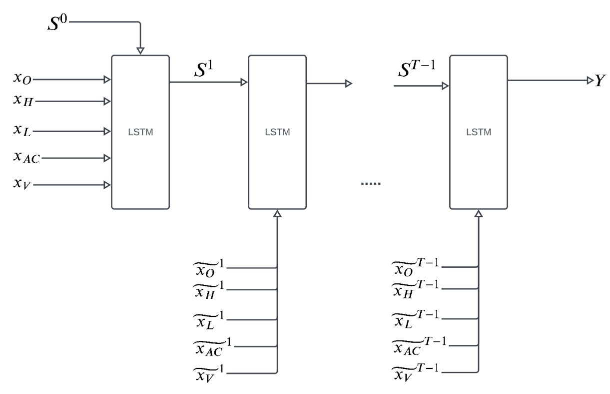

Fig. 1. Multivariate LSTM with Single Output

In Figure 1:

- T is the length of observed sequence
- k is the length of predicted sequence
- Y the predicted sequence of Close Price
- S^t is the cell state at state $t, t \in [0, T-1]$
- $x_O, x_H, x_L, x_{AC}, x_V$ are the Open Price, High Price, Low Price, Adjusted Close Price, and Volume at Stage 0, respectively

- $\widetilde{x_O}^t, \widetilde{x_H}^t, \widetilde{x_L}^t, \widetilde{x_{AC}}^t, \widetilde{x_V}^t$ are Open Price, High Price, Low Price, Adjusted Close Price, and Volume at Stage t after applying the Bias Function $F(x_i{}^t)$, respectively

B. Bias Function

Given an observed input matrix $\mathbf{A}, \mathbf{A} \in \mathbb{R}^{T\times 5}$ as below:

$$\mathbf{A} = \begin{bmatrix} x_O{}^0 & x_H{}^0 & x_L{}^0 & x_{AC}{}^0 & x_V{}^0 \\ x_O{}^1 & x_H{}^1 & x_L{}^1 & x_{AC}{}^1 & x_V{}^1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_O{}^{T-1} & x_H{}^{T-1} & x_L{}^{T-1} & x_{AC}{}^{T-1} & x_V{}^{T-1} \end{bmatrix} \tag{1}$$

Then the Bias Function is defined as below:

$$F(x_i{}^t) = \widetilde{x_i}^t = \beta(t)x_i{}^t + (1-\beta(t))\mu_i \tag{2}$$

Where:

$$i = \{\text{O, H, L, AC, V}\}, t \in [0, T-1] \tag{3}$$

$$\beta(t) = \frac{1}{t} \tag{4}$$

$$\mu_i = \frac{1}{T} \times \sum_{j=0}^{T-1} x_i{}^j \tag{5}$$

C. Loss Function and Optimization Algorithm

The goal is using the Adam Optimization Algorthm to minimize the Huber Loss Function as below:

$$\text{AdamOptimze(HuberLoss}(Y, \hat{Y}))$$

D. Dataset

The stock data of Apple Inc. is collected via Yahoo Finance.

- Training Set: Records from January 1^{st}, 2010 to December 31^{st}, 2022
- Testing Set: Records from January 1^{st}, 2024 to March 31^{st}, 2024

E. Training Process

1) Retrieving Apple Inc. Stock Data from Yahoo Finance
2) Using MinMax Scaler to scale the training dataset into the range of $[0, 1]$
3) Training 4 models (RNNs, LSTM, GRUs, and our model). Each model is trained in 30 epochs. In each epoch, 20% of the training set will be used for validation.

III. Results

There are 4 models in this experiment:

- RNNs
- LSTM
- GRUs
- Our built model (AdvanceLSTM)

A. Observing the sequence of 20 records to forecast the pattern in the next 5 days

The Figure 2 shows the predicted pattern in the next 5 days after observing the last 20 records of Open Price, High Price, Low Price, Adjusted Close Price, and Volume.

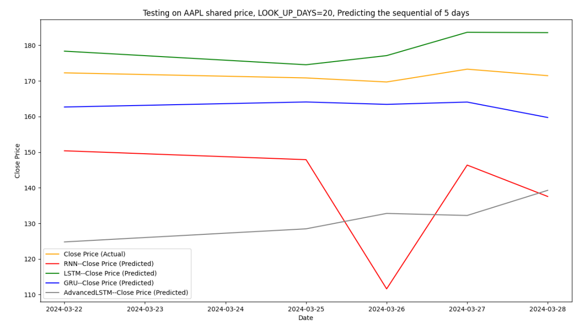

Fig. 2. Predicted Pattern in the next 5 days

TABLE I
PREDICTED ERRORS IN THE NEXT 5 DAYS

	MSE	MAE	R2 Score
RNNs	2.3203e-04	0.0113	0.9734
LSTM	6.6643e-05	0.0056	0.9924
GRUs	4.6058e-05	0.0046	0.9947
AdvanceLSTM	9.1251e-05	0.0064	0.9896

B. Observing the sequence of 20 records to forecast the pattern in the next 10 days

The Figure 3 shows the predicted pattern in the next 10 days after observing the last 20 records of Open Price, High Price, Low Price, Adjusted Close Price, and Volume.

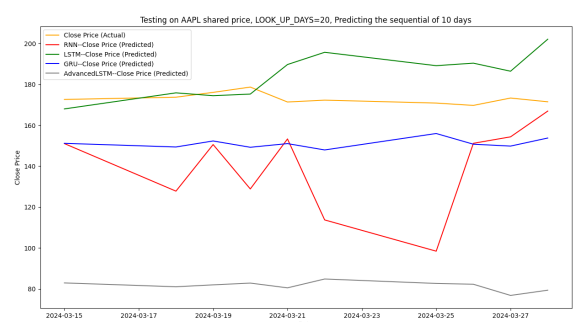

Fig. 3. Predicted Pattern in the next 10 days

TABLE II
PREDICTED ERRORS IN THE NEXT 10 DAYS

	MSE	MAE	R2 Score
RNNs	1.5992e-04	0.0096	0.9816
LSTM	9.4293e-05	0.0064	0.9892
GRUs	8.0342e-05	0.0060	0.9908
AdvanceLSTM	3.4027e-04	0.0118	0.9608

C. Observing the sequence of 20 records to forecast the pattern in the next 15 days

The Figure 4 shows the predicted pattern in the next 15 days after observing the last 20 records of Open Price, High Price, Low Price, Adjusted Close Price, and Volume.

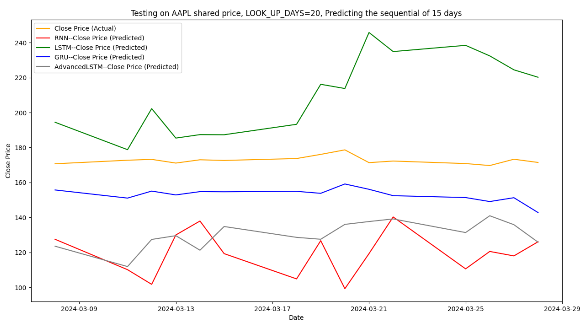

Fig. 4. Predicted Pattern in the next 15 days

TABLE III
PREDICTED ERRORS IN THE NEXT 15 DAYS

	MSE	MAE	R2 Score
RNNs	1.6359e-04	0.0093	0.9811
LSTM	1.1241e-04	0.0071	0.9870
GRUs	1.2479e-04	0.0075	0.9856
AdvanceLSTM	1.4668e-04	0.0082	0.9831

D. Observing the sequence of 20 records to forecast the pattern in the next 20 days

The Figure 5 shows the predicted pattern in the next 20 days after observing the last 20 records of Open Price, High Price, Low Price, Adjusted Close Price, and Volume.

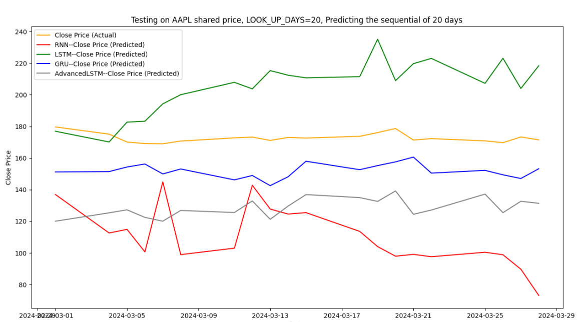

Fig. 5. Predicted Pattern in the next 20 days

TABLE IV
PREDICTED ERRORS IN THE NEXT 20 DAYS

	MSE	MAE	R2 Score
RNNs	2.2368e-04	0.0106	0.9741
LSTM	1.4406e-04	0.0080	0.9833
GRUs	1.4804e-04	0.0079	0.9829
AdvanceLSTM	1.7967e-04	0.0090	0.9792

IV. CONCLUSION AND FUTURE WORK

In these above experiments, it is obvious that the AdvanceLSTM performs the pattern prediction better than the RNN. In Tables I, II, III, and IV, sometimes the error from the AdvanceLSTM is greater than LSTM's, but observing the Figure 2, 3, 4 and 5, the AdvanceLSTM performs the pattern of the actual stock pattern better than the LSTM and RNNs. The result from GRUs has the most fitted pattern compared to others. The more length in long-term forecast, the more fluctuations we get. In this research, we do not cover the stock price prediction but only pattern forecasting. Hence, in the near future, long-term stock price prediction will be the topic.

REFERENCES

[1] A. Hazra, N. Bera, and A. Mandal, "Predicting lung cancer survivability using svm and logistic regression algorithms," *International Journal of Computer Applications*, vol. 174, no. 2, p. 19–24, Sep 2017.

[2] E. Adetiba and O. O. Olugbara, "Lung cancer prediction using neural network ensemble with histogram of oriented gradient genomic features," *The Scientific World Journal*, vol. 2015, p. 1–17, 2015.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, p. 1735–1780, Nov 1997.

[4] A. A. Ismail, T. Wood, and H. C. Bravo, "Improving long-horizon forecasts with expectation-biased lstm networks," *ArXiv*, vol. abs/1804.06776, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:4951181

[5] A. Jagtap, Y. Patil, and D. Oswal, "Visualizing and forecasting stocks using machine learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 4, p. 2562–2566, Apr 2022.