

ARTICLE

Application of LSTM and CONV1D LSTM Network in Stock Forecasting Model

Qiaoyu Wang* Kai Kang Zhihan Zhang Demou Cao

Electrical and Computer Systems Engineering Department, Monash University, Melbourne, Victoria, Australia

ARTICLE INFO

Article history

Received: 11 January 2021

Accepted: 29 January 2021

Published Online: 30 April 2021

Keywords:

Linear regression

Polynomial regression

Long short-term memory network

One dimensional convolutional neural network

ABSTRACT

Predicting the direction of the stock market has always been a huge challenge. Also, the way of forecasting the stock market reduces the risk in the financial market, thus ensuring that brokers can make normal returns. Despite the complexities of the stock market, the challenge has been increasingly addressed by experts in a variety of disciplines, including economics, statistics, and computer science. The introduction of machine learning, in-depth understanding of the prospects of the financial market, thus doing many experiments to predict the future so that the stock price trend has different degrees of success. In this paper, we propose a method to predict stocks from different industries and markets, as well as trend prediction using traditional machine learning algorithms such as linear regression, polynomial regression and learning techniques in time series prediction using two forms of special types of recursive neural networks: long and short time memory (LSTM) and spoken short-term memory.

1. Introduction

Nowadays, with the rapid development of forecasting technology, forecasting technology can be applied in different fields. For example, the application in the economic field can form the economic forecast, and the application in the securities investment field can become the stock market forecast. The development and change of the stock market are regular, and the stock market forecast is based on the history and reality of the stock price. On the basis of all aspects of comprehensive information, we use the scientific methods of qualitative and quantitative analysis to obtain the objective law of the stock price change. The paper makes a scientific analysis of the relationship between various phenomena and mechanisms in stock forecasting and points out the possible future development

trend and results of stock prices.

As the rapid development of computer and artificial intelligence technology, many new technologies and methods have been provided for the modeling and prediction of stock market. Also, artificial neural network has a wide range of adaptive ability, learning ability, mapping ability. Forecasting models have a strong generalization ability and adaptive ability and have the ability of approximating any nonlinear mapping. The stock prediction methods based on neural network mainly use neural network to train the stock price data, and then use the training model to predict the stock market. In this context, we will use various methods based on machine learning and deep learning algorithms to conduct experiments to improve the accuracy of prediction, and then compare the accuracy of stock motion prediction and goodness of fit score to verify

**Corresponding Author:*

Qiaoyu Wang,

Electrical and Computer Systems Engineering Department, Monash University, Melbourne, Victoria, Australia;

Email: 1643360071@qq.com

the rationality of each method.

The purpose of this paper is to provide a method for time series analysis of the most popular machine learning and deep learning techniques, especially the prediction of stock price movements. By assessing the accuracy of different models in predicting stock price movements, it develops more profitable automated trading strategies for investors, provides risk managers with more accurate forecasts, and provides a deeper understanding of the most commonly used time series models [1].

2. Prediction Tasks

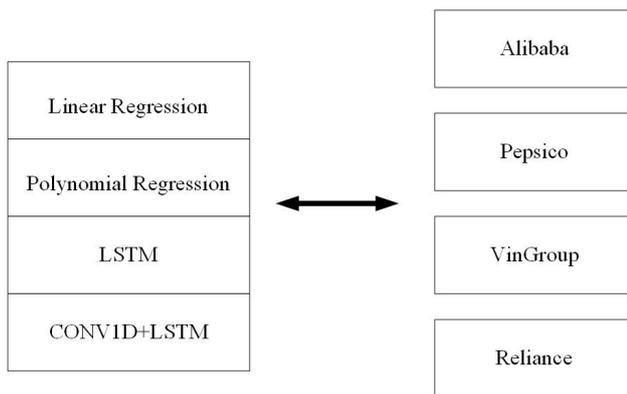


Figure 1. Project Tasks

This paper uses the stock market data of Alibaba, Pepsico, VinGroup and Reliance to make time series prediction. Firstly, linear regression and polynomial regression are used to carry out regression analysis and predictive analysis on stock data. Secondly, due to the excellent performance of LSTM network in sequential data processing, the LSTM model is designed according to the characteristics of stock market data. In this paper, the hidden layer memory cells of the LSTM model replace the artificial neuron cells.

By assigning different weights to each neuron cell, the LSTM has the unprecedented ability to distinguish between early and recent information, while the forgetting structure is helpful to eliminate the memory that is considered unnecessary in the decision-making process. This improves the accuracy of LSTM prediction. Finally, due to the high error rate of LSTM, this paper adopts the LSTM combined model optimized by one-dimensional convolutional neural network for prediction, which improves the accuracy of the prediction model.

3. Stock Forecasting Model

3.1 Data Collection

By collecting historical stock data on Alibaba, Pepsi-

co, VinGroup and Reliance. Historical information about each stock includes its daily opening, closing, high, low and volume. These original features will be served as the basis for further implementation of the feature process and will be discussed for each selected algorithm. We chose the five-year window to ensure that both bullish and bearish trends are included during this period. It is a detailed description of each technical indicator is shown in Table 1.

Table 1. Description of each technical indicator

Data Item	Annotation
Date	Trading in specific days
Opening Price	Opening price in specific trading days
High Price	Highest price in specific days
Low Price	Lowest price in specific days
Closing Price	Closing price for specific trading days
Volume	Trading volume in specific days

3.2 Linear Regression and Polynomial Regression Analysis

3.2.1 Details of Regression Model

This algorithm uses the closing price change of the first N time steps to predict the closing price change at time T. For each algorithm to determine the best N, a for loop will collect the RMSE value and the MAPE corresponding N. Then, the best N, RMSE value and the MAPE reach the minimum function N. In this article we will examine each value for N for 120. Model hyperparameters will be tuned with a validation set size of 0.5, and a test set size of 0.5. Also, it is important that a feature of the performance regression model is that all observations must be independent of each other, which is not the case with time series data. I decided to take the first order of different time series and ensured that the observations were all independent of each other, so the regression analysis was reasonable. I used modules from the sklearn library to do this experiment. For linear regression, the sum of the squared residuals of the sklearn model convergence is minimized, so hyperparameters tuning is not required at this stage. For polynomial regression, gridsearchcv is used to ensure that the best hyperparameter is used to determine the coefficient maximized - R square score in the fitting process for predicting ground live observations. Polynomial kernel functions range in degree from 2 to 8 [2]-[7].

3.2.2 Simulation Results

Table 2. Linear Regression Result

Stock	RMSE	MAPE	Accuracy
Alibaba	1.13	200.36%	55.97%
Pepsico	3.26	150.68%	51.82%
VinGroup	4.87	246.33%	48.67%
Reliance	3.22	294.78%	52.10%

Table 3. Polynomial Regression Result

Stock	RMSE	MAPE	Accuracy
Alibaba	1.36	215.22%	49.66%
Pepsico	2.85	207.46%	47.31%
VinGroup	2.87	152.89%	52.81%
Reliance	3.52	154.79%	50.23%

It can be seen from the experimental results that MSE values are all very small, indicating that there is not much difference between the predicted value and the real value of the model, and the model has a high accuracy. It can be seen from the accuracy of the model that the model has certain rationality. However, the MAPE in both models was large, exceeding 100% and even approaching 300%. Since a MAPE of 0% represents a perfect model, a MAPE of more than 100% represents a bad model. Therefore, even though the accuracy of the model is very high, these two groups of models are inferior models due to their high MAPE values. Due to the high accuracy of the linear regression model, the latter model is improved and enhanced based on the linear regression theory [8]-[9].

3.3 Long Short-Term Memory Model Analysis

3.3.1 Long Short-Term Memory Model

The long short-term memory model is a special RNN model, which is proposed to solve the gradient dispersion problem of the RNN model. In the traditional RNN, the training algorithm uses BPTT. When the time is relatively long, the residual that needs to be returned will decrease exponentially, resulting in the slow update of network weight and unable to reflect the long-term memory effect of RNN. Therefore, a storage unit is needed to store memory, so the LSTM model is proposed. The LSTM acts as a network of memories, and its memory refers to the sequence in which memories are transmitted in different Time steps. At the heart of the LSTM is the cell state "cell state" (the biggest difference from RNN). The focus is on Cell State.

Each cell consists of:

A. Input node (GC): As RNN, the output of the hidden node at the last point in time and the current input are accepted as inputs, then via a tanh activation function.

B. Input gate (IC): It plays the role of controlling input information. The input of the gate is the output of the hidden node at the previous point in time and the current input.

C. Internal state node (SC): The input is the current input filtered by the input gate and the output of the internal state node at the previous time point. A core element introduced by the LSTM is the cell.

D. Forget the gate (FC): It plays the role of controlling the internal state information. The input of the gate is the output of the hidden node at the previous point in time and the current input. The original LSTM in the position is just a value of 1.

E. Output gate (OC): The input of the gate is the output of the hidden node at the previous point in time and the current input [10].

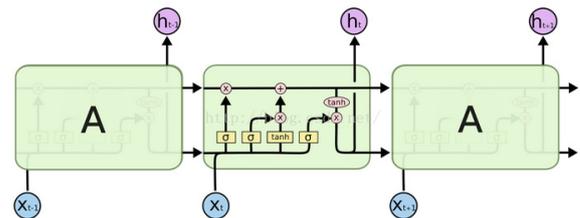


Figure 2. LSTM Schematic Diagram

3.3.2 My Long Short-Term Memory Model

So, I design my LSTM model is here.

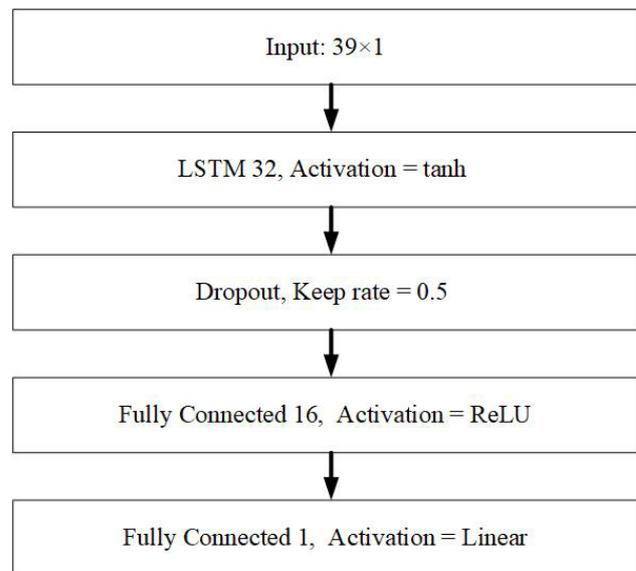


Figure 3. LSTM Model Architecture

After data analysis and screening, the input of our model retains the original characteristics of the closing price to predict the closing price of the next trading day. Through a series of experiments on Alibaba, Pepsico, VinGroup and Reliance stock price prediction under different backtracking days: 9,19,29,39,49,59. The results show that the 39-day window is the most accurate, so we will keep the number of backtracking days in the next section. The LSTM network consists of an input layer, an LSTM layer and a full connection layer.

Instead of using the default weight and bias initializations, this model uses initializations to ensure that the variance of the cross-layer weight gradient remains constant. The activation function switches from ReLU to tangent hyperbolic function (tanh) to prevent the explosive gradient phenomenon observed during the training phase. Mean square error is used as a loss function and L2 regularization technique to prevent potential model overfitting phenomenon ($lr = 0.01$, epoch = 120) [11]-[12]. In this paper, we choose Stochastic Gradient Descent (SGD) and Adaptive Moment Estimation (Adam) as optimizer for my LSTM model.

3.3.3 Simulation Results

Table 4. LSTM Result with SGD Optimizer

Stock	RMSE (SGD)	Accuracy (SGD)	MAPE(SGD)
Alibaba	82.70	47.40%	44.66%
Pepsico	8.52	54.17%	6.84%
VinGroup	22.65	50.99%	33.18%
Reliance	36.71	51.10%	23.81%

Table 5. LSTM Result with Adam Optimizer

Stock	RMSE (Adam)	Accuracy (Adam)	MAPE (Adam)
Alibaba	10.45	51.04%	4.25%
Pepsico	2.62	58.44%	1.84%
VinGroup	7.64	49.40%	7.89%
Reliance	8.20	47.15%	4.80%

It can be seen from the experimental results that when SGD optimizer is used, RMSE and MAPE values are very large, indicating that the predicted value of the model differs greatly from the actual value and the model accuracy is low. However, according to Pepsico’s prediction results, both RMSE and MAPE values

were small and had high accuracy. It can be seen from the precision of the model that the model is reasonable to some extent, so we need to select an appropriate optimizer to optimize the model. So, I chose the Adam optimizer. RMSE and MAPE values are small when using the Adam optimizer. As RMSE and MAPE values are smaller and closer to the perfect model, the LSTM model optimized by Adam has good accuracy and is more reasonable [12].



Figure 4. Alibaba Stock Market Prediction



Figure 5. Pepsico Stock Market Prediction



Figure 6. VinGroup Stock Market Prediction



Figure 7. Reliance Stock Market Prediction

The following figure shows the forecast results. However, when training this model, the training rate is slower, and the accuracy is still some distance from the perfect model. Therefore, I designed an improved LSTM model of one-dimensional convolution model.

3.4 1D Convolutional Long Short-Term Memory Model

3.4.1 CONV1D-LSTM Model

The CONV1D-LSTM network superposed 2 CONV1D layers on the basis of the original architecture.

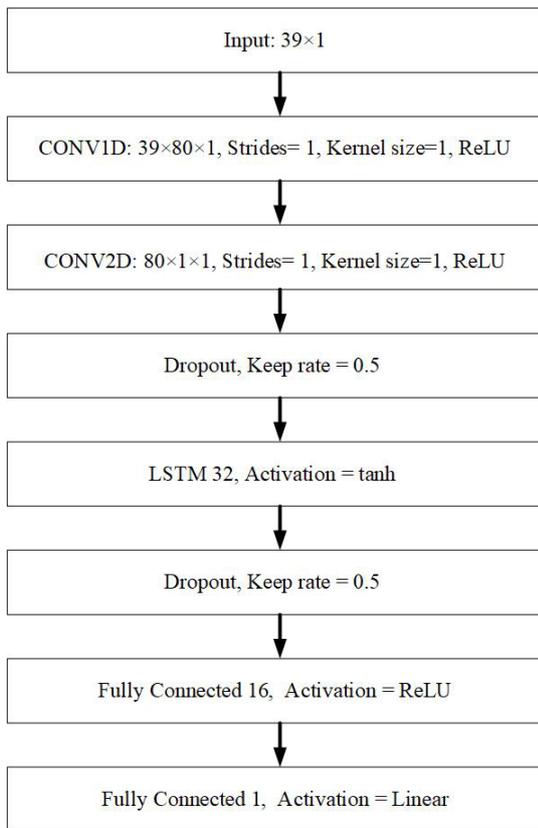


Figure 8. CONV1D-LSTM Model Architecture

3.4.2 Simulation Results

Table 6. CONV1D-LSTM Result with SGD Optimizer

Stock	RMSE (SGD)	Accuracy (SGD)	MAPE (SGD)
Alibaba	83.38	48.91%	44.18%
Pepsico	8.55	51.33%	6.83%
VinGroup	22.66	51.78%	33.04%
Reliance	36.57	49.90%	24.03%

Table 7. CONV1D-LSTM Result with Adam Optimizer

Stock	RMSE (Adam)	Accuracy (Adam)	MAPE (Adam)
Alibaba	14.40	54.17%	3.13%
Pepsico	2.87	51.56%	1.79%
VinGroup	7.53	51.38%	7.11%
Reliance	8.27	50.01%	4.74%

It can be seen from the experimental results that the accuracy of the LSTM model optimized by one-dimensional convolution increases significantly, while the RMSE and MAPE values decrease significantly, and the model is close to perfection. When SGD optimizer is used, RMSE and MAPE values are large, indicating that the predicted value of the model differs greatly from the actual value and the model accuracy is low. Again, I used the Adam optimizer. RMSE and MAPE values are small when using the Adam optimizer. As RMSE and MAPE values are smaller, the model is closer to the perfect model. Therefore, the CONV1D-LSTM model optimized by Adam has good accuracy and is more reasonable. The following figure shows the forecast results ^{[13]-[16]}.



Figure 9. Alibaba Stock Market Prediction



Figure 10. Pepsico Stock Market Prediction



Figure 11. VinGroup Stock Market Prediction

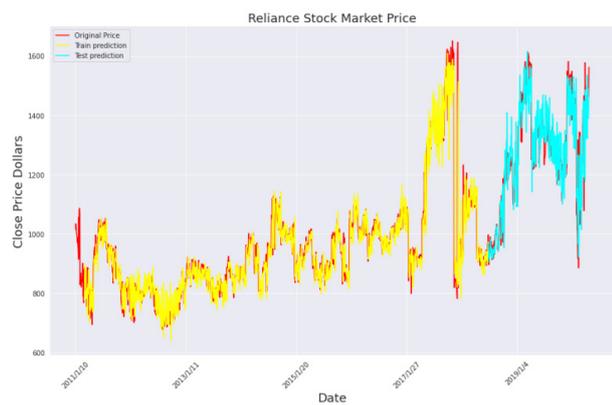
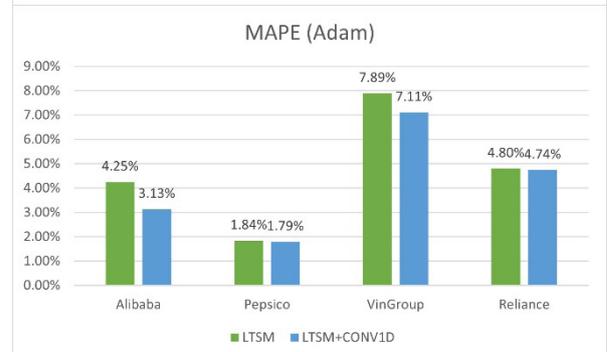
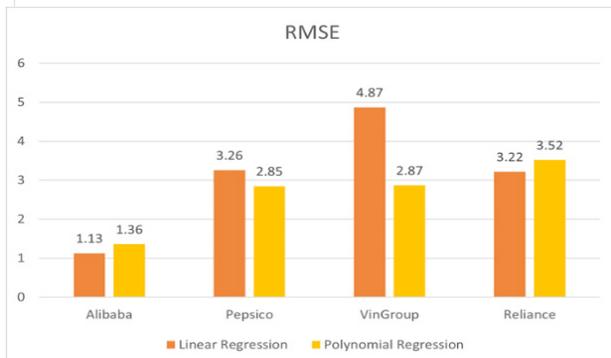
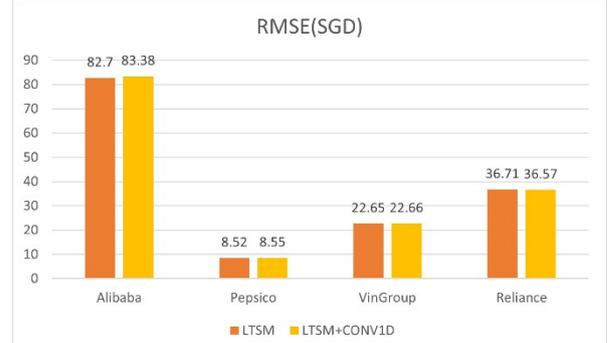
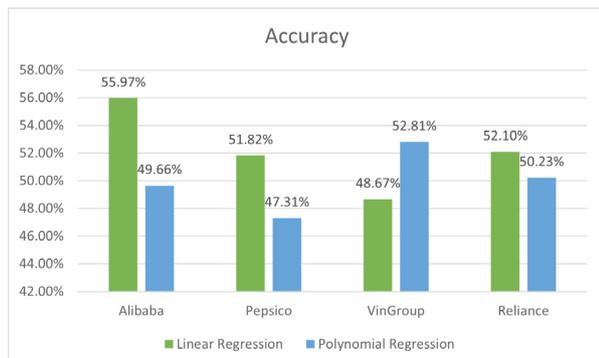
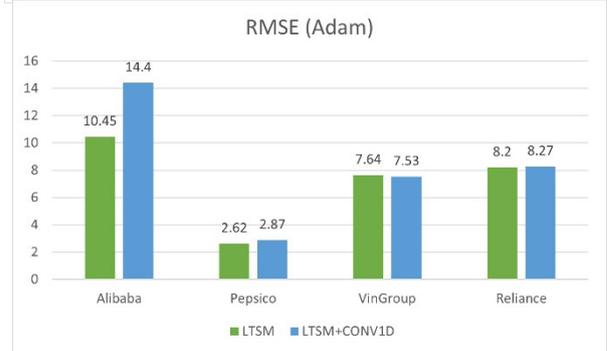
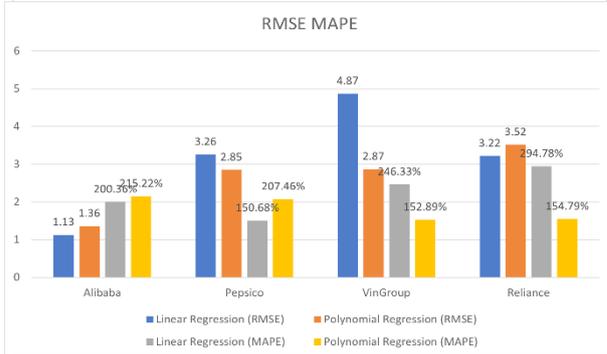
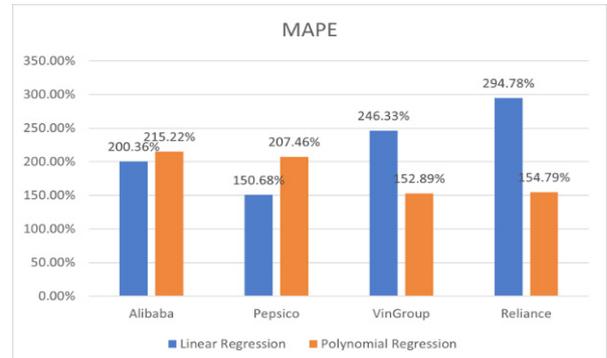
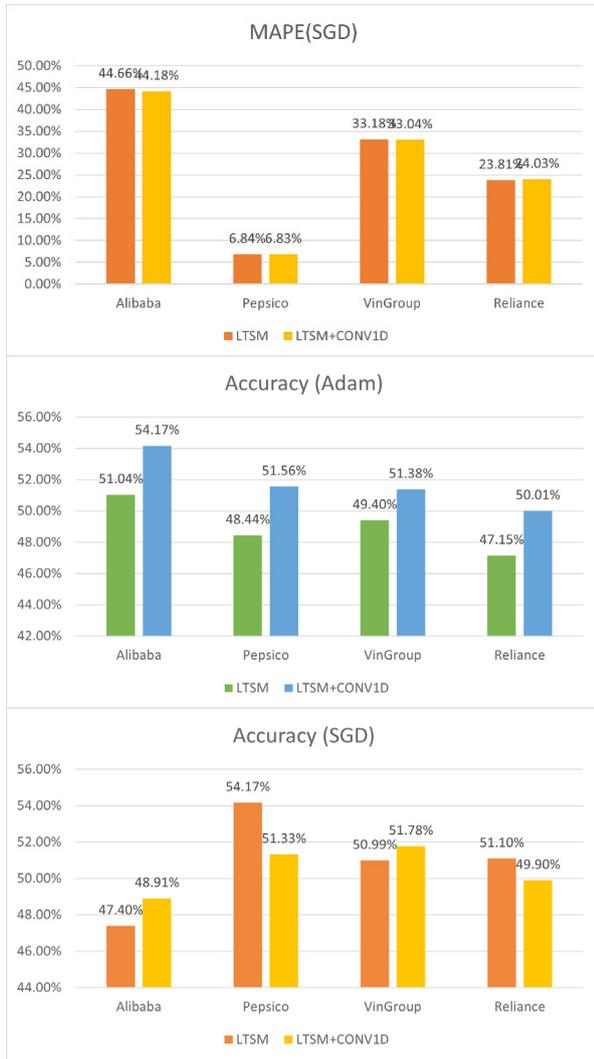


Figure 12. Reliance Stock Market Prediction

4. Comparison over All Algorithms





The experimental results show that the LSTM network model can better predict time series problems, which proves that the MAPE index is always low in the whole experiment. It is obvious from our example that the prediction produced by the LSTM model has the highest degree of fitting. Alibaba's stock is also more influential than the polynomial algorithm, so the addition of polynomial features does not improve the overall accuracy of the algorithm. Also, the goodness of fit of the model generated by polynomial regression is lower than that of the linear regression model. The addition of two CONV1D layers on the original LSTM network was indeed conducive to improving the overall accuracy, and the MAPE index was lower, indicating that the fitting model was better^{[17]-[21]}.

5. Conclusions

In order to verify the applicability of machine learning and deep learning models in stock forecasting models, the study uses two traditional models as baselines: linear

regression and polynomial regression. I evaluated the performance of state-of-the-art deep learning networks in predicting the price and direction of movements of four stocks. It is found that the polynomial features do not improve the MAPE and accuracy of the regression model. As for, accuracy and goodness of fit score (MAPE), the experimental results showed that the CONV1D-LSTM model produced the model with the highest goodness of fit, while the MAPE decreased, indicating that the CONV1D-LSTM model had a high predictive ability overall.

References

- [1] Olive, David J., Olive, David J., & SpringerLink. (2017). Linear regression.
- [2] Pandis, Nikolaos. (2016). Linear regression. American Journal of Orthodontics and Dentofacial Orthopedics, 149(3), 431-434.
- [3] Montgomery, D. C., Peck, Elizabeth A., & Vining, G. Geoffrey. (2013). Solutions manual to accompany Introduction to linear regression analysis (5th ed.).
- [4] Celant, G., & Broniatowski, Michel. (2016). Interpolation and extrapolation optimal designs 1 : polynomial regression and approximation theory.
- [5] Nye, C. D., Prasad, J., Bradburn, J., & Elizondo, F. (2018). Improving the operationalization of interest congruence using polynomial regression. Journal of Vocational Behavior, 104(C), 154-169.
- [6] Weidmann, R., Schönbrodt, F. D., Ledermann, T., & Grob, A. (2017). Concurrent and longitudinal dyadic polynomial regression analyses of Big Five traits and relationship satisfaction: Does similarity matter? Journal of Research in Personality, 70(C), 6-15.
- [7] Lee, K., Woo, H.-G., & Joshi, K. (2017). Pro-innovation culture, ambidexterity and new product development performance: Polynomial regression and response surface analysis. European Management Journal, 35(2), 249-260.
- [8] Mishra, P. (2019). PyTorch Recipes A Problem-Solution Approach (1st ed. 2019.).
- [9] Huang, K., Hussain, Amir, Wang, Qiu-Feng, & Zhang, Rui. (2019). Deep Learning: Fundamentals, Theory and Applications (1st ed. 2019.).
- [10] Zhang, Jianfeng, Zhu, Yan, Zhang, Xiaoping, Ye, Ming, & Yang, Jinzhong. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. Journal of Hydrology (Amsterdam), 561, 918-929.
- [11] Liu, J., Zhang, T., Han, G., & Gou, Y. (2018). TD-LSTM: Temporal Dependence-Based LSTM Networks for Marine Temperature Prediction. Sensors (Basel, Switzerland), 18(11).

- [12] Coto-Jiménez, M., & Goddard-Close, J. (2018). LSTM Deep Neural Networks Postfiltering for Enhancing Synthetic Voices. *International Journal of Pattern Recognition and Artificial Intelligence*, 32(1).
- [13] Zang, Haixiang, Liu, Ling, Sun, Li, Cheng, Lilin, Wei, Zhinong, & Sun, Guoqiang. (2020). Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renewable Energy*, 160, 26-41.
- [14] Chahkandi, V., Fadaeieslam, M., & Yaghmaee, J. (2018). Improvement of image description using bi-directional LSTM. *International Journal of Multimedia Information Retrieval*, 7(3), 147-155.
- [15] Zhu, Yonghua, Gao, Xun, Zhang, Weilin, Liu, Shenkai, & Zhang, Yuanyuan. (2018). A Bi-Directional LSTM-CNN Model with Attention for Aspect-Level Text Classification. *Future Internet*, 10(12).
- [16] Li, P., Abdel-Aty, M., & Yuan, J. (2020). Real-time crash risk prediction on arterials based on LSTM-CNN. *Accident Analysis and Prevention*, 135, 105371.
- [17] Kim, Taewook, & Kim, Ha Young. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PLoS ONE*, 14(2), e0212320. <https://doi.org/10.1371/journal.pone.0212320>.
- [18] Quan-Hoang Vo, Huy-Tien Nguyen, Bac Le, & Minh-Le Nguyen. (2017). Multi-channel LSTM-CNN model for Vietnamese sentiment analysis. 2017 9th International Conference on Knowledge and Systems Engineering (KSE), 2017, 24-29.
- [19] Chen, N., & Wang, P. (2018). Advanced Combined LSTM-CNN Model for Twitter Sentiment Analysis. 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 684-687.
- [20] Wu, Y., Zheng, B., & Zhao, Y. (2018). Dynamic Gesture Recognition Based on LSTM-CNN. 2018 Chinese Automation Congress (CAC), 2446-2450.
- [21] Chen, N., & Wang, P. (2018). Advanced Combined LSTM-CNN Model for Twitter Sentiment Analysis. 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 684-687.