Stock Prediction with Machine Learning Methods

Zhuo Diao[*] **Zhihao Yang**[*] **Chaoer Ren**[*]
School of Information Science and Technology
ShanghaiTech University
{diaozhuo, yangzh4, renche}@shanghaitech.edu.cn
2021533{046,185,192}

Abstract

This study investigates the application of machine learning techniques, including LSTM, SVR, Attention Mechanism, and CNN, for predicting stock prices. Subsequently, it conducts a comprehensive analysis of the predictive outcomes. The aim of this research is to provide investors with precise and reliable stock price predictions, offering valuable decision support to navigate the complexities of the financial market.

1 Introduction

The financial market, characterized by complexities and uncertainties, poses challenges for predicting stock price trends. Despite some success with machine learning models, issues like dynamic financial data, noise, irrational behavior, and unexpected events persist. This study analyzes Google's stock data using advanced algorithms (LSTM, SVR, Attention Mechanism, and CNN) to uncover patterns and provide investors with precise predictions, enhancing decision support in the complex financial landscape.

Our research focuses on analyzing Google's stock data from the past two years, given its relevance as a major player in the tech industry influenced by various market dynamics. The dataset includes crucial financial features: Open, High, Low, Close, Adjusted Close, and Volume. Open and Close prices indicate the trading period's start and end, High and Low prices reflect stock fluctuations, Adjusted Close accounts for dividends and stock splits, offering a more accurate value, and Volume provides insights into market interest and participation.

2 Related Works

The application of machine learning in predicting stock prices continues to be a focal point of research in both academic and financial domains. Some research has introduced innovative hybrid algorithms, such as CEEMD-CNN-LSTM and EMD-CNN-LSTM, with the aim of extracting deep features and time series information.(1) Others have sought to enhance the prediction accuracy of gold volatility by integrating two deep learning approaches—Long Short-Term Memory networks (LSTM) and Convolutional Neural Networks (CNN), notably leveraging the pre-trained VGG16 network.(2) Certain studies advocate for the construction of a CNN-based framework capable of processing data from diverse sources, to consider the potential wealth of information in the correlation among different markets and extract features conducive to predicting future trends in these markets.(3) Additionally, some research introduces a novel Recurrent Neural Network Attention model. (4)This Attention Technique can play a part in predicting the stock price.

[*] All authors contributed equally to this project.

NeurIPS 2023 Generative AI and Biology Workshop.

3 Proposed Methods

3.1 Ridge Regression and Lasso Regression

When there are numerous features but relatively few samples, linear regression is prone to overfitting. To alleviate this overfitting issue, a regularization term is added to the linear regression.

$$\min_{\omega} \sum_{i=1}^{m} (y_i - \omega^T x_i)^2 + \lambda ||\omega||_2^2 \quad (\text{or } \lambda ||\omega||_1)$$

We use 4 features: closing price (Adj. Close), percentage change between high and close price of a stock (HL_PCT $=$ (High $-$ Adj. Close)$/$Adj. Close $*$ 100), the percentage change between a stock's closing price and its open price (PCT_change $=$ Adj. Close $-$ Open)$/$Open $*$ 100), and volume to predict the next 30 days' stock price.

3.2 Support Vector Regression

Support Vector Regression aims to minimize the distance from the farthest data points to the hyperplane.

$$\min_{\omega,b} \frac{1}{2} ||W||^2$$
$$s.t. \quad |y_i - (\omega x_i + b)| \leq \epsilon$$

In SVR, losses are only calculated for data points within a certain margin, specifically when the absolute difference between the predicted value $f(x)$ and the true value y exceeds a predefined tolerance ϵ. The optimization objective is to maximize the width of the margin while minimizing the total loss for these selected points.

The same as ridge and lasso regression, SVR uses 4 features: Adj. Close, HL_PCT, PCT_change and volume.

3.3 Long Short-Term Memory

The provided LSTM architecture in the code adeptly captures long-term dependencies and subtle temporal nuances in financial time series. It employs a multi-layered LSTM structure, stacking LSTM layers together to effectively learn from sequential stock price data. The model is dynamically constructed with a user-specified LSTM cell size, utilizing dropout in the cells to prevent overfitting during training. Trained on historical stock price data, the model generates the final prediction through a dense output layer after 300 training iterations.

3.4 Convolution Neural Network

A 1D CNN, focusing solely on the width dimension, is employed for stock price prediction, excluding convolution along the height dimension. Given the intricate factors influencing stock prices, leveraging deep learning proves advantageous. Our approach utilizes a single-layer 1D CNN with a size of 128 and a kernel size of 3. Using the closing price as input, the model forecasts prices for the next 30 days through 1000 training iterations for optimal performance. Post-training, the model predicts prices for successive days, repeating the process 10 times to obtain average accuracy.

3.5 Attention Technique

From a mathematical standpoint, attention mechanisms enable the model to dynamically assign varying degrees of importance to different elements in the input sequence. This adaptability allows the model to focus on crucial information, capturing temporal dependencies and relevant patterns without being constrained by fixed weights.

Attention mechanisms provide a seamless integration of external factors, such as macroeconomic indicators or sentiment analysis. By dynamically adjusting attention based on the perceived importance of these external factors, the model can enhance its predictive accuracy in response to the broader financial landscape.

Algorithm 1 Attention Core Mechanism

```
Get $queries, keys, query\_masks, future\_binding$
split $Q\_head$ and $K\_head$ with $queries$ and $keys$
$alignment \leftarrow$ calculate_attention($Q\_head, K\_head$) / sqrt($dimension\_of\_keys$)
$alignment \leftarrow$ apply_masks($alignment, query\_masks, future\_binding$)
$attention\_weights \leftarrow$ softmax($alignment$)
$weighted\_sum \leftarrow$ weighted_sum_with_values($attention\_weights, K\_head$)
$output \leftarrow$ linear_transform_and_normalize($weighted\_sum + queries$)
```

4 Evaluation

We use RMSPE to calculate our algorithms‘s accuracies, which is widely used in the field of financial.

$$RMSPE = 100\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\frac{P_t - P_t'}{P_t})^2}$$

where P_t stands for real value and $P_t^{'}$ stands for predict value.

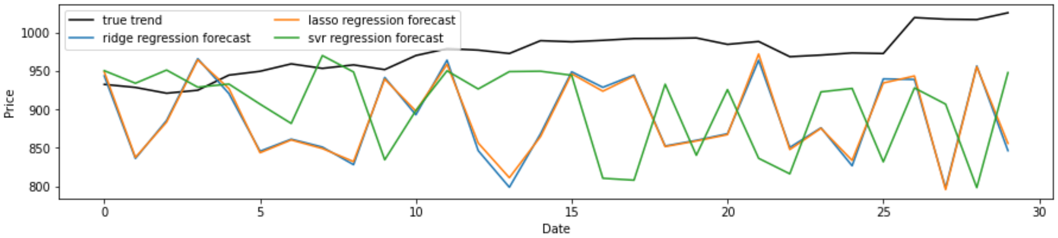

Figure 1: Ridge regression, lasso regression and SVR

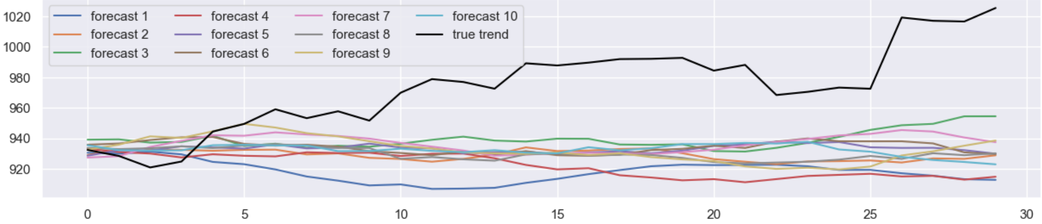

Figure 2: Prediction of Long Short-Term Memory

4.1 Ridge Regression, Lasso Regression and SVR

Figure 1 indicates the results of Ridge Regression, Lasso Regression and SVR. Ridge regression and lasso regression predict almost the same. Performance of SVR is slightly better than linear regression. Ridge regression‘s accuracy is 89.4822%, lasso regression‘s accuracy is 89.6312% and SVR‘s accuracy is 90.3083%. All the algorithms predict a similar price at the first day and is close to the true value while none of them shows an upward trend.

4.2 Long Short-Term Memory

Figure 2 indicates the results of LSTM. The average accuracy rate stood at 94.8147%. With multi-layered structure and dynamic construction of cells. The LSTM’s inherent capacity to capture and retain long-term dependencies enabled it to adapt effectively to varying conditions. It is worth noting that LSTM has a good performance in predicting stock prices on first day , it is useful sometimes.

4.3 Convolution Neural Network

After 1000 times training, we forecast the next 30 days‘ stock price 10 times. The average accuracies are 94.6940% . Figure 3 indicates that forecast 7 is the best, whose accuracy reaches 97.6409%. It has a upward trend. The figue also indicates that most of the forecasts predict a similar price at the first day and is close to the true value.

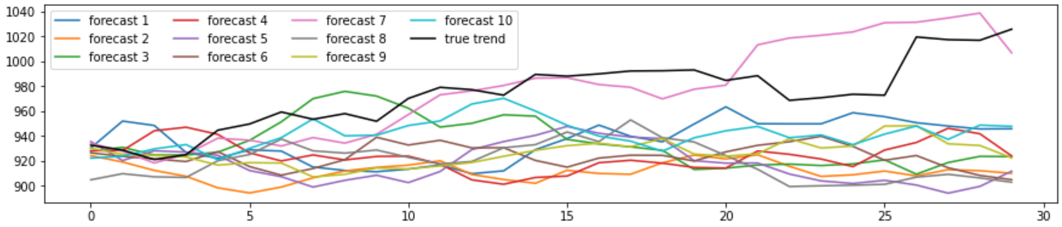

Figure 3: Prediction of convolution neural network

4.4 Attention Technique

Single training iteration repeated ten times, utilizing a single-layer attention mechanism. The model incorporates a specified number of units or hidden units within each attention layer. Across ten testing instances, the average accuracy rate stood at 94.8284%.

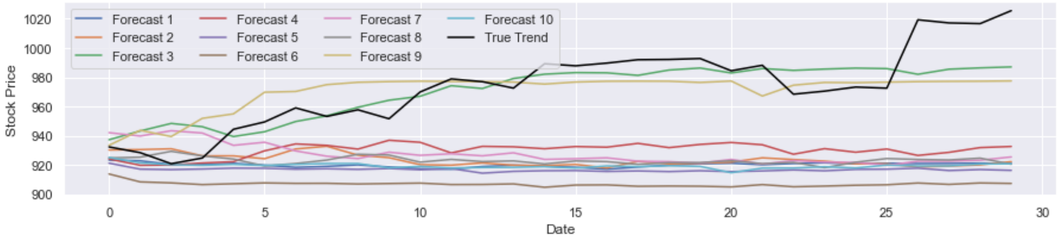

Figure 4: Prediction of attention

5 Conclusion

We utilize Ridge Regression, Lasso Regression, SVR, LSTM, CNN and Attention technique to predict stock prices. Among these algorithms, LSTM, CNN, and attention technique have demonstrated superior performance, showing effective fitting of stock trends.

Despite the promising results, there remains some disparity between our predictions and the actual trends. Predicting each day's price based on the preceding day may contribute to a gradual amplification of errors. The accuracy assessment of our model has limitations, particularly with the use of percentage error based on stock prices. This approach may lead to an overestimation of predictive capabilities. Each algorithm in our study has inherent limitations: LSTM may have prolonged training times, Attention mechanism lacks interpretability. Additionally, the dataset's limited scope, encompassing only a year, may be insufficient for robust model training.

In our future work, we plan to explore combining Attention technique with CNN or integrating CNN with LSTM to improve prediction accuracy. We anticipate that these hybrid approaches will yield more favorable outcomes.

References

[1] H. Rezaei, H. Faaljou, and G. Mansourfar, "Stock price prediction using deep learning and frequency decomposition," *Expert Systems with Applications*, vol. 169, p. 114332, 2021.

[2] A. Vidal and W. Kristjanpoller, "Gold volatility prediction using a cnn-lstm approach," *Expert Systems with Applications*, vol. 157, p. 113481, 2020.

[3] E. Hoseinzade and S. Haratizadeh, "Cnnpred: Cnn-based stock market prediction using a diverse set of variables," *Expert Systems with Applications*, vol. 129, pp. 273–285, 2019.

[4] V. Mnih, N. Heess, A. Graves, and K. Kavukcuoglu, "Recurrent models of visual attention," in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, (Cambridge, MA, USA), p. 2204–2212, MIT Press, 2014.