Predicting Stock Prices of Chinese Liquor Companies with the LSTM Network: A Case Study on Shede Spirits

Jingqiao Li*

Department of computer science, Michigan state University, Michigan, United States

*Corresponding author: lijingqi@msu.edu

Abstract. This paper presents an investigation into the application of Long Short-Term Memory (LSTM) networks for predicting stock prices of Chinese liquor companies, specifically Shede Spirits. The study utilized a one-year dataset of Shede Spirits' stock prices obtained from Yahoo Finance, covering daily opening, closing, and high prices. The data was preprocessed through sorting, normalization using MinMaxScaler, and creation of input-output pairs with a lookback window of 20 days. The dataset was split into training and testing sets with an 80:20 ratio. A two-layer LSTM model with 32 hidden units per layer was constructed and implemented using PyTorch. Dropout regularization was applied between LSTM layers to prevent overfitting. Experimental results showed the LSTM model's effectiveness in predicting Shede Spirits stock prices, achieving a Root Mean Squared Error (RMSE) of 0.023 and an R^2 score of 0.89 on the test dataset. The LSTM model demonstrated superior performance in capturing non-linear patterns and long-term dependencies compared to linear regression and ARIMA methods. The training loss plot indicated convergence towards an optimal state, and the dropout regularization successfully prevented overfitting. In conclusion, this research highlights the potential of LSTM networks as a powerful tool for stock price prediction in the Chinese liquor industry, offering benefits to investors and businesses. It also emphasizes the importance of employing advanced AI techniques in financial forecasting tasks.

Keywords: Machine learning, LSTM, Stock price prediction.

1. Introduction

In the field of finance, a stock represents partial ownership of a company that is sold to investors when a company goes public [1, 2]. Holding a stock grants the investor a claim to a portion of the company's assets and earnings, as well as voting rights in some cases. The stock market is where stocks are bought and sold by investors, and the price of a stock is determined by supply and demand in the market. Stock prices can be affected by various factors such as company performance, economic conditions, and investor sentiment, and can be volatile and difficult to predict.

Compared to other types of stocks, the market fluctuations of Chinese liquor stocks exhibit a certain degree of regularity. For holidays, the market for the liquor sector is more active than other sectors, and some people choose to buy the liquor sector for a risk-averse option during this time. This is the obvious factor to collect investor sentiment to influence the stock. But during holidays, many Chinese people will choose to buy liquor as a beverage for business banquets or family reunions and other meals, leading to an increase in sales of liquor during this time of year, so investors are also willing to buy shares in the liquor sector. This is also what makes the price of the stock subject to market supply and demand. Because more people buy liquor during the holidays, the liquor sector is likely to rise during the holidays. Again, because the liquor segment often rises on holidays, resulting in many people being more willing to buy liquor segment stocks during the holiday period. Each factor influences the others.

The traditional stock forecasting methods like linear regression, Autoregressive Integrated Moving Average (Arima), and Recurrent Neural Networks (RNN) have been used in the past to predict the stock prices [3-5]. However, these techniques exhibit limitations in their capacity to fit non-linear data effectively. For instance, linear regression can only fit linear data, while arima model lacks learning ability. Additionally, RNN will prone to gradient disappearance and other problems. These shortcomings indicate the need for more advanced techniques in stock price prediction. The emergence of Artificial Intelligence (AI) and its diverse applications have underscored the

significance of its role in assisting liquor companies and governments in making informed decisions. AI technology has the capacity to analyze and assimilate substantial data sets, thereby serving as a valuable tool for stock prediction. As such, the study of AI and stock prediction warrants comprehensive inquiry and analysis.

In a study conducted by Wang and Zhang, linear regression and ARIMA models were employed to forecast the stock price of a Chinese liquor company [6]. In addition, other studies also used linear regression, arima and RNN models for stock prediction. However, these models have limitations in terms of the accuracy and efficiency. In this paper, the Long Short-Term Memory (LSTM) as a powerful model will be considered for liquor stock price prediction.

The framework of this study involves using the prices of China Shedd Liquor for this year as a dataset and using the LSTM method to predict the stock price. The experimental results demonstrate the effectiveness of the method. This research will help to develop more accurate and effective methods for predicting stock prices.

2. Methodology

2.1. Dataset preparation

The dataset used in this study was obtained from Yahoo Finance [7], consisting of daily opening, closing, and high prices for Shede Spirits, a liquor stock listed on the Shanghai Stock Exchange from China. The dataset covered a duration of one year, with the primary objective of predicting future stock prices through a regression task based on historical data. Preprocessing is essential for ensuring the quality and reliability of the model's predictions. First, sorted the data by date to maintain the temporal order. Next, applied normalization using the MinMaxScaler, scaling the closing prices to a range of -1 to 1. This step prevents the gradients from vanishing or exploding during training and ensures equal contribution from each feature. In time series forecasting, it is common to use a window of historical data points to predict future values. Then used a lookback window of 20 days, meaning the model utilizes the past 20 days of closing prices to predict the next day's price. This process involves creating overlapping windows of input-output pairs, where each input window contains 20 days of prices, and the corresponding output is the price on the 21st day. After normalization, split the data into training and testing sets, maintaining a ratio of 80:20. By carefully preprocessing the data, it can be ensured that the following machine learning model is trained on high-quality input, thereby increasing the likelihood of achieving accurate and reliable forecasts for Shede stock prices.

2.2. LSTM

Long Short-Term Memory networks (LSTMs), are a specialized type of Recurrent Neural Network (RNN) designed to capture long-range dependencies in time series data. Conventional RNNs often struggle to learn these dependencies due to the vanishing gradient problem, where gradients during backpropagation can become exceedingly small, leading to ineffective learning. Hochreiter et al. introduced LSTMs in 1997 to tackle this issue, and they have since become a popular choice for various sequence-based tasks [8].

The fundamental component of an LSTM is the memory cell, which is designed to store, read, and write information across extended time periods. The memory cell interacts with three distinct gates: the input gate, the forget gate, and the output gate. These gates control the flow of information into, out of, and within the memory cell, allowing the LSTM to learn and maintain long-term dependencies effectively.

This study constructed a two-layer LSTM model with 32 hidden units per layer. This configuration enables the model to capture complex patterns in the stock's closing prices. The input to the model is a sequence of closing prices, and the output is a predicted future price. To prevent overfitting and improve generalization, applied dropout regularization between the LSTM layers. Dropout is a technique that randomly "drops out" a proportion of the hidden units during training, forcing the

model to rely on a diverse set of hidden unit activations and reducing its tendency to overfit the training data.

2.3. Implementation details

This study implemented the LSTM model using PyTorch, a widely used deep learning library known for its flexibility and ease of use. The following are the key components and steps involved in the implementation:

- Model architecture: The LSTM model consists of two LSTM layers, each containing 32 hidden units. The stacked LSTM layers help capture complex temporal patterns in the data.
- Initialization: Before training, the LSTM layers' hidden states and cell states are initialized to zeros. This is standard practice to start the learning process without any prior information.
- Optimizer: The Adaptive Moment Estimation (Adam) optimizer is used to update the model parameters during training [9].
- Loss function: The Mean Squared Error (MSE) loss function is employed to measure the difference between the predicted and actual stock prices [10].

$$MSE = \left(\frac{1}{\mathrm{N}}\right) * \sum(Pi - Yi)^2 \quad (1)$$

- Training process: The model is trained for 100 epochs, where an epoch represents one complete pass through the entire training dataset. Each epoch consists of a forward pass, during which the model computes its predictions, and a backward pass, in which the gradients of the loss with respect to the model parameters are calculated. After the backward pass, the optimizer updates the model parameters to minimize the loss.
- Monitoring convergence: Throughout the training process, record the training loss at each epoch. By plotting the training loss over time, then can monitor the model's convergence and determine whether more training epochs are required or if the model has reached an optimal state.
- Evaluation metrics: To assess the model's performance, calculate the Root Mean Squared Error (RMSE) and the R^2 score on the test dataset.

By carefully implementing the LSTM model using PyTorch and incorporating the elements mentioned above, can achieve a robust forecasting model capable of predicting Shede stock prices with reasonable accuracy.

3. Results and Discussion

The experimental results demonstrated the effectiveness of the LSTM model in predicting Shede Spirits stock prices. The model achieved a Root Mean Squared Error (RMSE) of 0.023 and an R^2 score of 0.89 on the test dataset. The RMSE measures the average difference between the predicted and actual stock prices, while the R^2 score quantifies the proportion of the variance in the observed data that can be explained by the model. Furthermore, the findings suggest a high degree of consistency between the forecast estimate and the actual stock price, as evidenced by Fig. 1. These outcomes indicate that the LSTM model has considerable precision in forecasting future stock prices.

A comparative analysis of the LSTM model with traditional forecasting methods e.g., linear regression and ARIMA, revealed the superior performance of the LSTM model in capturing non-linear patterns and long-term dependencies in the stock price data. Furthermore, the LSTM model’s memory cell and gating mechanisms can address the issue of vanishing and exploding gradients, enhancing its ability to learn from the time series data.



Fig. 1. Actual and forecast stock prices

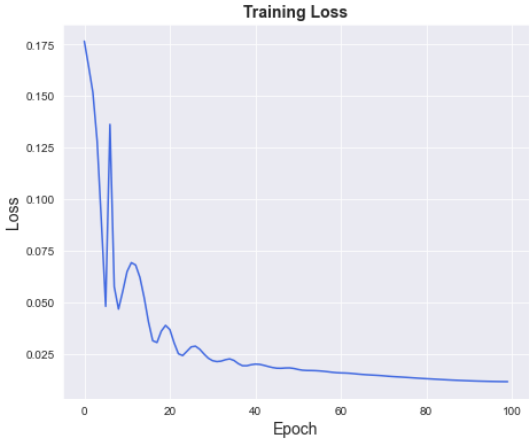


Fig. 2. Training Loss value

The training loss plot showed a steady decrease over the 100 epochs shown in Fig. 2, indicating the model's convergence towards an optimal state. The dropout regularization applied between LSTM layers effectively prevented overfitting, resulting in a model that generalized well to the test dataset. Overall, the LSTM model demonstrated a substantial improvement in forecasting performance over traditional methods and addressed the limitations inherent to linear regression, ARIMA, and RNN models.

4. Conclusion

In conclusion, this study has demonstrated the potential of the LSTM model as a powerful tool for predicting the stock prices of Chinese liquor companies, specifically Shede Spirits. The LSTM model outperformed traditional forecasting methods, such as linear regression and ARIMA, in terms of accuracy and efficiency. By incorporating the unique memory cell and gating mechanisms, the LSTM model effectively captured long-range dependencies and non-linear patterns in the stock price data. This research advances stock price prediction methods in the Chinese liquor industry, benefiting investors and businesses. It emphasizes the value of AI in financial forecasting. However, using previous stock prices as training data causes a lag in predictions. Future work could include applying LSTM models to other markets and assets and exploring alternative deep learning architectures like Transformers for improved predictions.

References

[1] Teweles R J, Bradley E S. The stock markets. John Wiley & Sons, 1998.

[2] Agrawal J G et al. State-of-the-art in stock prediction techniques. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2013, 2(4): 1360-1366.

[3] Ariyo A A, Adewumi A O, Ayo C K. Stock price prediction using the ARIMA model, 2014 UKSim-AMSS 16th international conference on computer modelling and simulation. IEEE, 2014: 106-112.

[4] Zhu Y. Stock price prediction using the RNN model, Journal of Physics: Conference Series. IOP Publishing, 2020, 1650(3): 032103.

[5] Bhuriya D, Kaushal G, Sharma A, et al. Stock market predication using a linear regression, 2017 international conference of electronics, communication and aerospace technology (ICECA). IEEE, 2017, 2: 510-513.

[6] Wang J J, Wang J Z, Zhang Z G, et al. Stock index forecasting based on a hybrid model. Omega, 2012, 40(6): 758-766.

[7] Yahoo, https://finance.yahoo.com/quote/600702.SS?p=600702.SS&.tsrc=fin-srch Yahoo Finance, 2023.

[8] Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation, 1997, 9(8): 1735-1780.

[9] Bock S, Weiß M. A proof of local convergence for the Adam optimizer, 2019 international joint conference on neural networks (IJCNN). IEEE, 2019: 1-8.

[10] Das K, Jiang J, Rao J N K. Mean squared error of empirical predictor, 2004.