#101

BICA VPS 2023

Application and Modeling of LLM in Quantitative Trading Using Deep Learning Strategies.

Tiejun Pan M.I. Author[1], *Jinjie Yu Presenter*[*2], *Jinjie Yu Corresponding-Author*[3]. [1]First affiliation including address and email. Asterisk in the list of authors above should indicate the presenter. [2]Second affiliation. Please replace the number 000 in the top left corner with your submission number. [3]Third affiliation. Please include at least one email address.

First Author

Second Author

SUMMARY

- The newly emerged Large Language Model can give the general framework of the code.This project asked LLM to train a stock price prediction model using LSTM, then use backtrader to perform backtest.

INTRODUCTION

After more than 100 years of development, with the breakthrough of computer technology, deep learning and big data industry, the quantitative trading market has gradually matured, and more and more investors have begun to use quantitative trading to invest. Quantitative trading automatically executes transactions through written programs, eliminating the interference of human subjective factors on transaction execution. But the threshold for quantitative trading is high, requiring researchers to have a deep understanding of mathematics, statistics, finance, and computer technology. The newly emerged Large Language Model (LLM) can help users get started to a certain extent, by giving the general framework of the code, so that users can have a preliminary understanding of the countermeasures faster and more accurately.

APPROACH

In terms of model training and testing, this paper adopts the CSI 300 index obtained from tushare platform to study the results of daily data, weekly data and monthly data after training.

This project trained a stock price prediction model using long short-term memory (LSTM) methods. Then, the backtest model was established with the classic double moving average strategy in quantitative trading, and the backtrader platform was used to visualize the return results simulated by the backtest.

METHODS

Table 1. CSI 300 Data

Trading Date	Open	High	Close	Low	Volume
20230609	3822.4088	3836.7026	3836.7026	3811.2189	134319863
20230608	3791.0826	3834.3888	3820.1867	3777.7752	118700934
20230607	3815.4823	3823.1124	3789.3418	3780.1021	98862432

The project uses the data of the CSI 300, we focus on 399300.SZ, which is obtained from tushare platform. The obtained values include time, opening price, highest price, lowest price, closing price and trading volume, as shown in Table I.

The time range of this project is between 2010/01/01 and 2023/06/10. The model was trained using daily, weekly and monthly data respectively using LSTM strategy. The target of this test is to compare the results and find the best dataset for model training. The model adopts mean square error (MSE) as the model evaluation index, and the formula is as follows,

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{1}$$

Where, m is the number of samples, y_i is the stock price, and $\hat{y}_i$ is the model forecast stock price. After training, MSE for each dataset are shown below. We also plot the predict value of stock price and the true value. By comparing those plots (see Fig. 1), the accuracy of the models can be found directly by eyes.

Similar to the previous example, we start our project by asking LLM ‘Could you please show me an example of using double average strategy to build a backtest model on backtrader platform and visualize the results, using the CSI 300 data obtained from tushare? ’ LLM provides the simplest version of double average strategy, but the model is not matching to real world trading case. When a trade occurs, the trader needs to pay a transaction fee of 0.01% to 0.30%. In this model, we set the proportion of the transaction fee to 0.05%. We also need to determine the position when the ‘gold cross’ or ‘dead cross’ appears.

In this project, we use the data of the CSI 300 stock dataset again and focus on 399300.SZ, which is obtained from tushare platform. Total initial funding is set at 100,000. After performing the backtest using daily data, weekly data and monthly data, we will compare the profits and get a conclusion.

RESULTS

Fig. 1. Training results of the stock price prediction models recommended by LLM using LSTM method. From left to right are the results from daily, weekly and monthly data

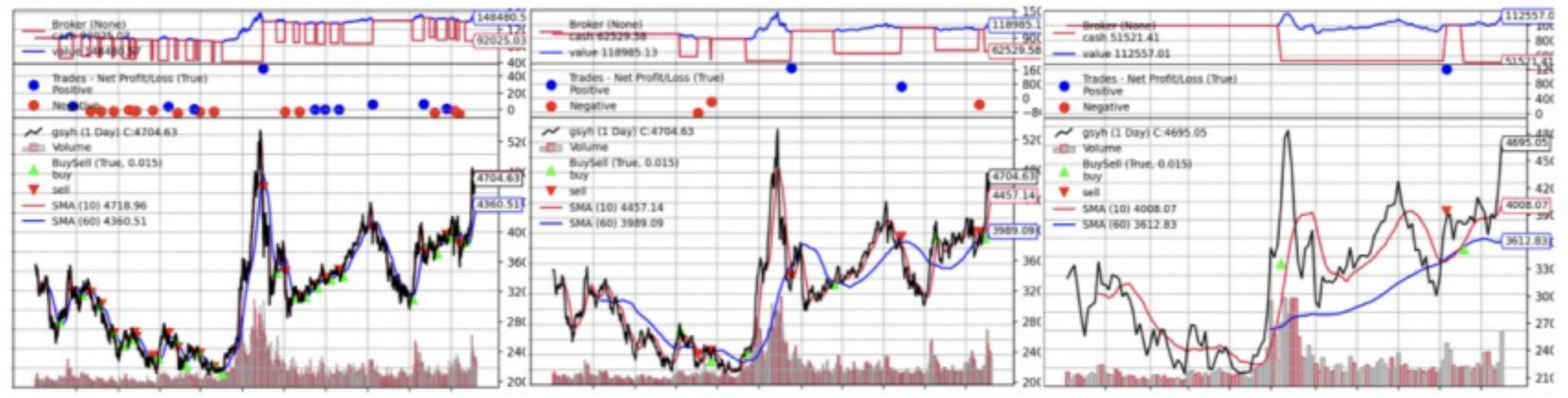

Fig. 2. Backtest results of the quantitative trading models recommended by LLM using double moving average strategy. From left to right are the results from daily, weekly and monthly data.

ANALYSIS

For LSTM strategy, the best training data for stock prices prediction model depends on the need. If users want to see the trend as accurate as possible, they may need to take costs for training models with daily data. If users want to see the general trend in short time or with a little training cost, weekly data could be better. In this example, LLM helps to construct the basic structure of codes, though it cannot do all works for training a prediction model, it is a beginner-friendly tool in quantitative trading. The only thing the user has to do is adjust the codes to fit their training demands.

DISCUSSION

There are many factors affecting the stock market, such as changes in political situation, sudden natural disasters and artificial manipulation, which can cause stock prices not to change according to the expected trend. Therefore, there are certain risks in trading with the model recommended by the LLM.

CONCLUSIONS

LLM, as a newly emerging chat artificial intelligence, is of great help to beginners to understand the basic knowledge related to quantitative trading. However, its ability to write specific strategies is limited, for example, it can provide the basic framework code using long short-term memory modeling, but the improvement and optimization of specific details still need researchers' own understanding of quantitative transactions to achieve.

ACKNOWLEDGMENTS

This paper is supported by Zhejiang Province's 14th Five Year Plan Teaching Reform Project (jg20220738), Ningbo Science and technology innovation Fund (2020Z008), Zhejiang College Student Innovation and Entrepreneurship Training Program (S202310876001), Zhejiang Provincial Basic Public Welfare Fund Research Project (LGF20G020002), Zhejiang Provincial Philosophy and Social Science Planning Project (22NDJC127YB), Ningbo Municipal Basic Public Welfare Fund Research Project (2021S070).