A Study of Daily K-Level Quantitative Trading Based on Deep Learning

Hao Cheng.

chlwr1997@163.com

School of Information and Electronic Engineering

Abstract: This paper proposes a Daily K-Level Quantitative Trading Strategy (DKTS) based on deep learning, which aims to predict the future stock price trends in the daily K-level and conduct corresponding quantitative trading using historical data and LSTM networks. We use Long Short-Term Memory (LSTM) for feature extraction and predictive modeling, transforming historical data into multidimensional time series data to adapt to the input format of deep learning models. Through the analysis of experimental results on various indexes in the A-share market, this method shows good prediction accuracy and stability. To verify the practical effect of the model, we conducted trading backtests in the actual stock market. The experimental results show that using a deep learning model for trading yields significant returns on multiple indexes and individual stocks, while also demonstrating better risk control and drawdown rates. The research results of this paper indicate that deep learning has broad application prospects in the field of quantitative trading. In the future, we will continue to explore and optimize deep learning models to improve their reliability and stability in practical trading.

Keywords: Quantitative Trading, LSTM, Daily K-Line, Quantitative backtesting

1. LSTM STOCK PREDICTION MODEL

1.1 LSTM

LSTM, short for Long Short-Term Memory, is a type of recurrent neural network (RNN) commonly used for processing sequential data. Unlike traditional RNNs[1], LSTM[2] introduces three gating units (input gate, forget gate, output gate) and a memory cell to control the flow of information input, forgetting, and output. These gating units can adaptively determine which information needs to be passed and which needs to be forgotten based on the input data. The memory cell can retain important historical information and can be modified as needed[3], making it well-suited for modeling long sequences. LSTM has been widely used in natural language processing, speech recognition, stock prediction, image recognition, and other fields, achieving many outstanding results.

LSTM has several advantages for stock prediction. First, it can handle long sequences of historical data, which is necessary for capturing complex patterns in stock prices[4]. Second, its ability to selectively remember or forget past information is useful for filtering out irrelevant noise in the data. Third, the input, forget, and output gates in LSTM allow for better control over the information flow, allowing the model to focus on the most relevant data. Finally,

ICBBEM 2023, May 19-21, Hangzhou, People's Republic of China

DOI 10.4108/eai.19-5-2023.2334298

LSTM can handle non-linear relationships between inputs and outputs, making it suitable for modeling the complex and dynamic nature of stock prices. These advantages have made LSTM a popular choice for stock prediction and have led to many successful applications in the field.

1.2 LSTM prediction process

Assuming the training dataset is x_{Pre}, the pre-training model F is trained usingx_{Pre} data and the LSTM method.

$$F = LSTM(x_{Pre})$$

The stock prediction model is trained using stock prediction training data, which is fed into the LSTM prediction model to train and obtain the model F.After the data training stage of the LSTM stock prediction model, the model preparation work is completed. In the quantitative backtesting stage, first, n pieces of data before the t-th trading day are extracted from the quantitative trading database to predict the data for the t+1-th trading day.

$$\bar{x}_{t+1} = F(x_{t-n} - x_t)$$

After completing the training of the LSTM stock prediction model using the stock prediction training data, the model preparation work is completed. In the quantitative backtesting phase[5], first, data from the n days before the t-th trading day is extracted from the quantitative trading database for predicting the data of the t+1-th trading day. Once the data is passed through the LSTM stock prediction model, the predicted results are sent to the quantitative simulation decision-making stage for generating specific operations on the stocks, and then the backtesting phase of quantitative trading is entered.

2. THE OVERALL ARCHITECTURE OF DEEP LEARNING-BASED QUANTITATIVE DAILY K TRADING

2.1 Quantitative Trading System Architecture

The overall architecture of the deep learning-based quantitative stock daily K trading method consists of three parts, of which the first part is the LSTM stock pre-training model, The second part is on quantitative trading strategies, and the third part is the data backtesting method which will be introduced later.

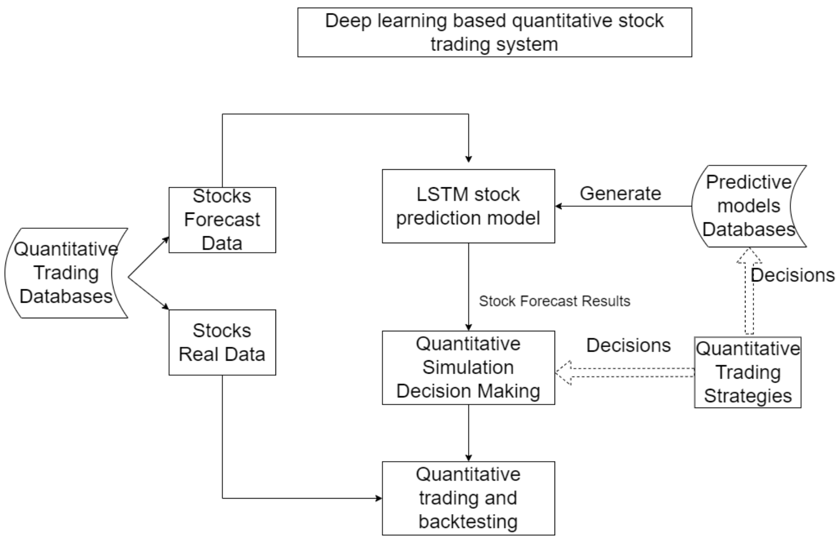

Figure 1. The overall architecture of a deep learning based quantitative daily K trading method for stocks

As shown in Figure 1, the model training data and prediction data are two different types of data[6]. Model training data is only used during model training and is a preparatory work in the early stage. The data is first extracted from the database and divided into prediction data and real data. The prediction data is first preprocessed[7], such as standardization and normalization, before being sent into the LSTM model to obtain prediction results. The prediction results will generate trading instructions according to the quantitative trading strategy. The trading instructions can only be calculated with real data to form a quantitative trading backtest.

2.2 Quantitative Trading System Process

The quantitative trading process is illustrated here for a stock, x_t represents the stock's data for the t trading day, $x_{t-n} - x_t$ is the stock's forecast data, x_{t+1} is the real data for the next trading day, and x'_{t+1} the forecast result.

Assuming a simulated capital size of b, to simulate the trading situation on the t+1 trading day in the quantitative trading system, we need to input $x_{t-n} - x_t$ into the prediction model to get the prediction x'_{t+1} and get the specific action (e.g. buy or sell). Afterwards, the new amount of money b after the transaction is obtained according to the x_t operation price and the x_{t+1} operation price.

$$b = b * \frac{(x_t - x_{t-1})}{x_{t-1}} \tag{1}$$

As the A-share market implements the T+1 (buy on the day, sell on the next day at the earliest) trading mechanism, to achieve daily buy and buy operations, we use the A-share market's opening bid mechanism to carry out a buy operation at the closing bid time of the tth trading day and a sell operation before the close of the t+1 trading day, thus achieving a pseudo-t+0 transaction. The details are as follows.

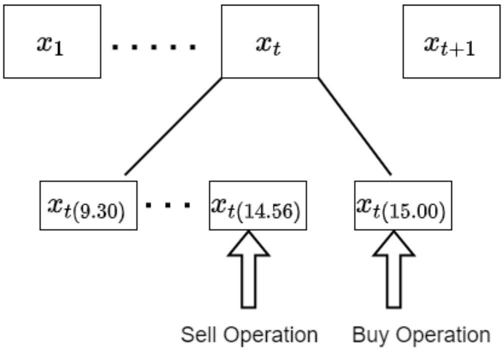

Figure 2. Stock buying and selling operation point diagram

As shown in Figure 2, this trading strategy is used to perform a buy operation during the closing bid phase[8] of t trading day on the first trade, followed by a sell operation at 14:56 hours before the closing bid and a buy operation at 15:00 hours during the closing bid phase for all subsequent trading operations on the next day or one trading cycle apart. For the operation to be successful, the buy is submitted at the stop price and the sell is submitted at the stop price. When we put the $x_{t-n} - x_t$ forecast intox_t, it is clear that x_t is missing the last three minutes of bid data, so we use the $x_{t(14:56)}$ data from x_t to fill in the$x_{t(15:00)}$ .data, which is done using$fill(x_t)$. This implements a pseudo-t+1 operation and avoids quantitative trading strategy bias caused by stock traders going back to the market at night intervals on trading days and by certain financial information that may cause the daily opening price of a stock to be different from the previous day's closing price.

3. QUANTITATIVE TRADING STRATEGIES

Quantitative trading strategies refer to a series of trading rules and strategies designed through the use of mathematical models and algorithms[9], analyzing historical market data and other relevant information to guide the process of making trading decisions. Their formulation is often based on statistical principles, such as buying when the 5-day moving average crosses above the 10-day moving average and selling when the 10-day moving average crosses below the 5-day moving average. Specific quantitative trading strategies may utilize different levels of candlestick charts, indicators, and factors, and generally consist of two parts: data feature selection and specific operational decision-making. These trading strategies are usually interpretable. However, deep learning is a black box model that can only obtain prediction results through the training objectives during training, and the process of obtaining prediction results is not interpretable. Therefore, quantitative trading strategies based on deep learning need to be formulated from three aspects: model and data type selection, model prediction task selection, and specific operational decision-making. In this chapter, we choose to use daily K-level stock data as the data type selection. The quantitative trading strategy section mainly introduces task selection and specific operational decision-making[10] to better implement the daily K quantitative trading method based on deep learning.Here is a sample illustration and caption for a multimedia file:

3.1 Quantitative trading strategy for stock pool selection based on daily K

The main idea of the quantitative trading strategy based on stock pool selection using daily K-line data is to select a stock pool (such as the constituents of the CSI 300 index), and then predict the price change of each stock in the pool for the next trading day's closing price[11]. The strategy involves selecting some of the stocks with the highest predicted increase and buying them. Figure 3 illustrates this thinking:

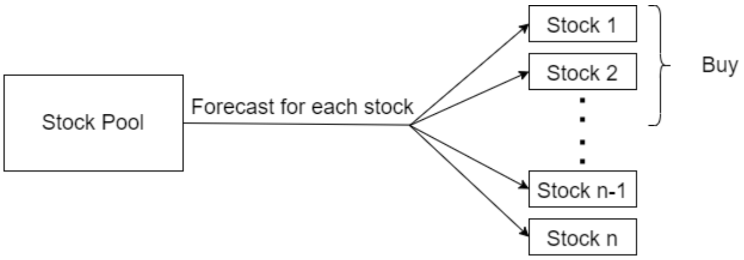

Figure 3. Stock pool selection diagram

The main idea of this quantitative trading strategy based on daily K-line stock pool selection is to select a stock pool (such as the constituents of the CSI 300 index) and predict the next day's closing price change for each stock in the pool, and then select some stocks with the highest predicted increase for buying. The schematic diagram is as follows: In this trading strategy, the goal is to predict the price change of stocks for the next trading day. The specific operation is to use a stock prediction model to predict each stock in the stock pool on the t+1 trading day, sort the prediction results according to the magnitude of the increase or decrease, and select the top m stocks for buying. On the next trading day, sell the previously purchased stocks and predict each stock in the stock pool again, selecting the stock with the highest predicted increase for buying.

3.2 Single Stock Daily K to T Quantitative Trading Strategy

The main idea of the single stock day K doing T quantitative trading strategy is to select a stock based on personal subjective preference, which remains unchanged in the strategy. Then, according to the stock prediction model's prediction of the rise or fall, buy or sell operations are conducted on the selected stock. This trading strategy adopts a more aggressive full position buying and selling strategy[12]. The prediction network is used to predict the stock's rise or fall for the next trading day and classify it. If the prediction result is a rise, a full position buy operation is conducted at the closing auction today, and the full position is sold at 14:57 the next day. For ease of reading, we refer to this method as the "daily K lines strategy" in this article.

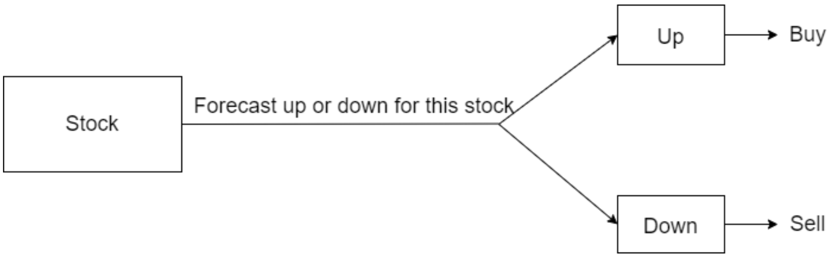

Figure 4. Single Stock Daily K for T Quantitative Trading Strategy Diagram

As shown in Figure 4, when the forecast result is up, continue to hold the stock if you are holding it at this time, otherwise proceed to buy. When the forecast result is down, if you do not hold the stock at this time, you will continue to hold a short position, otherwise you will perform a sell operation.

4. QUANTITATIVE TRADING EXPERIMENT WITH STOCK POOL SELECTION BASED ON DAILY K

4.1 Quantitative trading experiment on stock pool selection based on daily K

In this section, experiments are conducted using a quantitative trading method for stock pool selection based on the daily K. In order to verify the validity of the model in the real world, pre-training data for the model and model training data are used before the backtest date. The pre-training, training and back-test data for all experiments in this section were collected at the same time.Pre-training data collection time interval 2019.1.1-2021.1.1 Training data collection time interval 2019.1.1-2021.1.1. backtesting interval 2021.6.16-2022.2.27 starting capital 1,000,000 transaction stamp duty 1/1000 transaction fee 1/10000 number of stocks selected for the pool 3, 10. The results of the experiment were as follows.

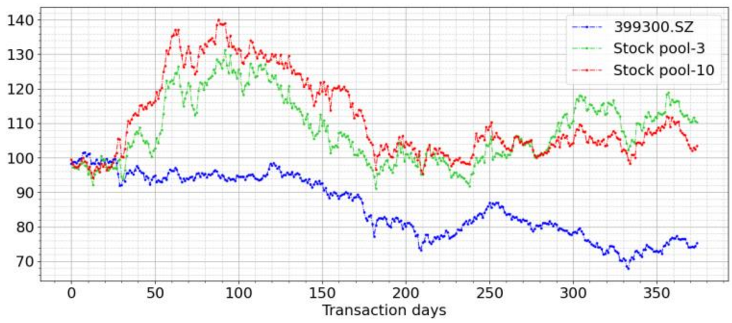

Figure 5. CSI 300 Backtest Experiment Results

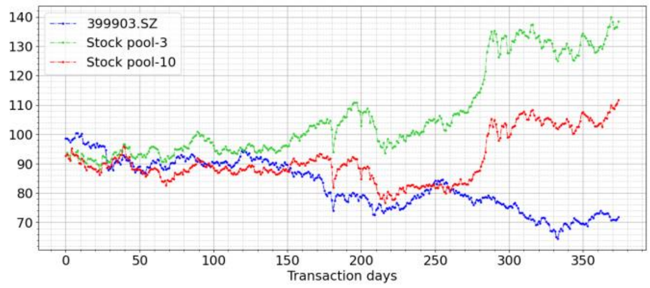

Figure 6. CSI 100 Backtest Experiment Results

As can be seen in Figures 5 to 6, in the backtest interval, the DKTS quantitative trading strategy achieved good return results during the drawdown interval even though the original indices were in the backtracking phase.

4.2 Quantitative trading experiments with single stock daily K-lines

In this section, we use a quantitative trading approach based on the daily K-line T of a single stock for our experiments. To verify the effectiveness of the model in practical applications, both pre-training data and model training data were obtained from data prior to the backtest date. The training and backtest data are collected at the same time for all experiments in this section.

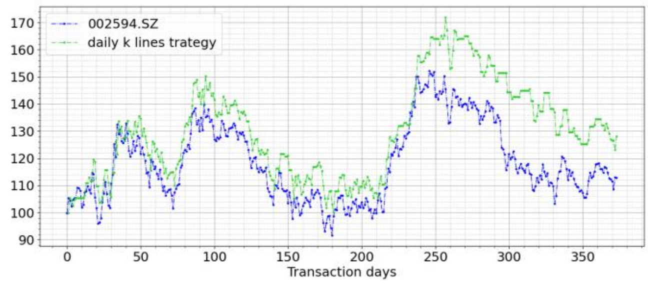

Figure 7. BYD Stock Quantitative Experiment Results

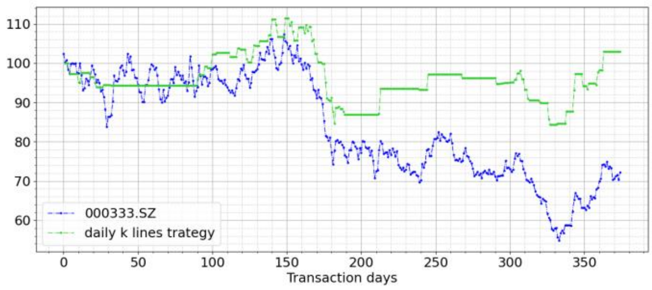

Figure 8. Midea Stock Quantitative Experiment Results

As can be seen in Figures 7 to 8, the Single Stock Daily K to T Quantitative Trading Strategy escaped many of the downside ranges during the backtesting interval and ultimately achieved good returns.

5. CONCLUSION

In this study, we proposed a day-level quantitative trading strategy based on LSTM and backtested it on historical data. By comparing the experimental results, we found that this strategy achieved superior trading performance on indexes such as the SSE 300 and individual stocks, demonstrating the effectiveness and practicality of our proposed method. Moving forward, we will continue to explore the knowledge in the field of quantitative trading, optimize the parameters and models of the strategy, and improve the returns and stability of trading.

REFERENCES

[1] Elman J L . Finding Structure in Time[J]. Cognitive Science, 1990, 14(2):179-211.

[2] Hochreiter S , Schmidhuber J . Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[3] Chang P C , Liu C H . A TSK type fuzzy rule based system for stock price prediction[J]. Expert Systems with Applications, 2008, 34(1):135-144.

[4] Lu C J . Integrating independent component analysis-based denoising scheme with neural network for stock price prediction[J]. Expert Systems with Applications, 2010, 37(10):7056-7064.

[5] Fu Y, Cao S, Pang T. A Sustainable Quantitative Stock Selection Strategy Based on Dynamic Factor Adjustment[J]. Sustainability, 2020, 12(10): 3978.

[6] Liu C, White M, Newell G. Selecting thresholds for the prediction of species occurrence with presence-only data[J]. Journal of biogeography, 2013, 40(4): 778-789.

[7] Soni J, Ansari U, Sharma D, et al. Predictive data mining for medical diagnosis: An overview of heart disease prediction[J]. International Journal of Computer Applications, 2011, 17(8): 43-48.

[8] Wellman M P, Wurman P R, O'Malley K, et al. Designing the market game for a trading agent competition[J]. IEEE Internet Computing, 2001, 5(2): 43-51.

[9] Nuti G, Mirghaemi M, Treleaven P, et al. Algorithmic trading[J]. Computer, 2011, 44(11): 61-69.\

[10] Thechatakerng P. Immigrants as Entrepreneurs: From Mynmar to Thailand[J].

[11] Serbera J P , Paumard P . The fall of high-frequency trading: A survey of competition and profits[J]. Research in International Business & Finance, 2015, 36(jan.):271-287.

[12] Atak A . The Future of Computer Trading in Financial Markets. 2011.