LSTM-Based Quantitative Trading Using Dynamic K-Top and Kelly Criterion

Binjing Li, Keli Xie, Siyuan Lu, Jun Lin, and Zhongfeng Wang
School of Electronic Science and Engineering, Nanjing University, Nanjing, China
Email: {bjli, kelixie, sylu}@smail.nju.edu.cn, {jlin, zfwang}@nju.edu.cn

***Abstract*—With the strong capability of modeling time sequence, long short-term memory (LSTM) networks have been widely applied to predicting financial time series. This has attracted tremendous attention in the quantitative trading area. A complete quantitative trading system usually has three tasks, including market timing, stock selection, and portfolio management. In this paper, we present an LSTM-based quantitative trading system and optimize this system from the following two aspects. Firstly, in the process of stock selection, we first introduce the dynamic K-top method in the LSTM-based quantitative trading system to follow the market change. Secondly, concerning portfolio management, we further incorporate the Kelly Criterion to attain an appropriate position ratio. Taking CSI300 constituent stocks as the study example, extensive experiments have been carried out to show the superiority of the proposed method. In comparison with the straight forward LSTM-based trading strategy, the improved LSTM-based trading strategy with the dynamic K-top method and the Kelly Criterion can achieve an increase of 44.97% over ten days in terms of accumulative return. In addition, our novel method can gain a win ratio of 55.95%, a monthly alpha of 0.16, a monthly Sharpe ratio of 2.17, and a monthly Sortino ratio of 2.96 disregarding the transaction costs.**

***Index Terms*—long short-term memory (LSTM), market timing, portfolio management, Kelly Criterion.**

I. Introduction

A complete quantitative trading system is composed of at least three tasks, including market timing, stock selection, and portfolio management. These three tasks have been identified as important but challenging tasks for researchers and investors due to the high volatility of the market.

Market timing is the strategy of making buying or selling decisions of financial assets by attempting to predict future market price movements. Its purpose is to capture stock movements accurately so that investors can grasp the inflection point and earn excess money. According to behavioral finance, market underreaction and overreaction to information are caused by unreasonable behaviors of the investors [1], [2]. The studies of behaviorial finance indicated that the future trend of the market can be reflected through historical data analysis. Technical analysts believe past trading activity and price changes of security can be valuable indicators of the security's future price movements [3]. Numerous comprehensive studies have provided solid evidence that it is possible to achieve trading success through technical analysis (TA). By behavioral finance, TA, and other related theories, great progress has been achieved in the market timing task.

Machine learning (ML) methods have been validated superior to traditional methods in forecasting financial time series [4], [5]. With strong data-driven and non-linearity capability, ML approaches have been demonstrated specialized in classification and regression tasks. They are well-suited for finding hidden patterns in large amounts of financial data [6]. As a variant of recurrent neural network (RNN), long short-term memory (LSTM) [7]–[9] and its hybrid models [10], [11] almost outperform all other models in terms of stock prediction accuracy.

The aim of the stock selection task is to choose the most suitable stocks from a certain stock pool. In the stock selection task, evaluation and scoring models are utilized to distinguish more 'promising' stocks from less 'promising' stocks [12], [13]. Stock scoring and stock ranking are two main procedures of a stock selection process. With respect to stock ranking, the problem of how many stocks shall we choose is still a great challenge for investors. It has a close relationship with the final returns of a quantitative trading system. Previous works tended to go long the most promising stocks for the whole trading period. However, the numbers of stocks chosen by them are fixed. A fixed number of stocks could not adapt to the market change since opportunities are changing in the volatile market.

The ultimate goal of portfolio management [14], [15] is to maximize the expected return given an appropriate level of risk exposure. In the high-frequency trading fields, positions are established and liquidated in a very short period. As part of portfolio management, the choice of the position ratio is a very important aspect. By optimizing position ratio, excess profit can be attained compared with no or unwise adjustment to the position ratio. Therefore, how much money should be invested on each stock is another problem that needs special attention.

To address these two problems, we present an LSTM-based quantitative trading system that combines these three tasks together effectively. Besides, we make some reasonable optimizations to the last two tasks of the trading system. In our design, the LSTM model is used for market timing task. The results of LSTM are used for stock scoring and dynamic K-top is utilized for stock ranking. For portfolio management, the Kelly Criterion are employed to adjust the position ratio reasonably, making the whole trading strategy more reasonable and scientific. The overall framework of the proposed method is shown in Fig. 1. Our contributions are summarized as

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

follows:

• To solve the problem of how many stocks to choose for trading, we introduce the dynamic K-top method in LSTM-based quantitative trading system in replacement of traditional fixed K-top for the first time.

• To solve the problem of how much money to invest in each stock, we further incorporate the Kelly Criterion to attain an appropriate position ratio, greatly exploiting the possibility of earning more money.

• We combine market timing, stock selection, and portfolio management to improve the quantitative trading system to a large extent. From prediction, ranking, and portfolio management aspects, extensive comparisons and analyses have been conducted to validate the effectiveness and reasonableness of our methods in an all-around way.

The remainder of the paper is organized as below: Section II presents related works and background of this article, which lays a foundation for carrying out our work. Section III elaborates on our model. Then, comparisons and results are listed in Section IV. Finally, we will reach a conclusion to summarize all the works.

II. Background And Related Work

A. Financial Time Series Prediction

Yuan Song [16] chose stock price indicators from 20 well-known public companies and found that RNN works better than support vector machine (SVM) and extreme gradient boosting in terms of prediction accuracy. Svetlana Borovkova and Ioannis Tsiamas [17] trained 12 stacked LSTM models for 20 large-cap stocks. The weight assigned to each model is evaluated by the area under the curve score. The final prediction model is the weighted combination of the 12 models. Siyu Yao, Linkai Luo, and Hong Peng [18] took transaction costs into account. In their paper, only when the fluctuation amplitude of the stock price is more than transaction fees, it will be marked as the up or down signal. In day trading, [19] first applied LSTM to large-scale market Standard and Poor's 500 Index constituent stocks dating from 1990 to 2015 in order to predict whether a certain stock's price change is more than the median price change of all stocks. The method has won 0.46 percent daily profit.

B. Stock Selection

There are diverse methods about stock selection, which range from statistical approaches (like fuzzy quantitative analysis [20], ordered weighted averaging operator [21], and cluster analysis [22]) to ML approaches (deep neural network (DNN) [23], SVM [24], and differential evolution (DE) [13]). By using technical and fundamental indicators to characterize each stock, the main point of stock selection is to construct the scoring and ranking mechanism.

With respect to stock ranking, the problem of how many stocks shall we choose is still a great challenge for researchers, which has a close relationship with the final returns of a quantitative trading system. Previous works tended to go long the most promising stocks for the whole trading period. However, the numbers of stocks chosen by them are fixed. It is not reasonable since opportunities are changing in the volatile market.

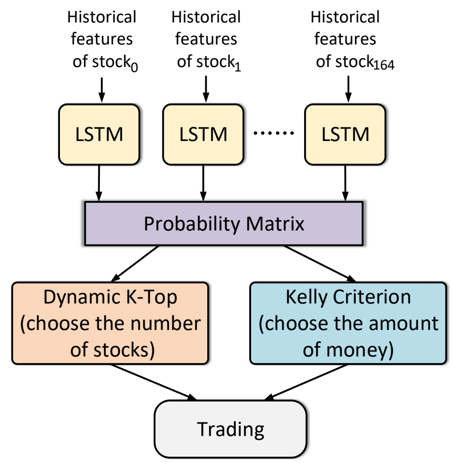


Fig. 1. Overall framework of the proposed model.

C. Kelly Criterion

Kelly Criterion, first proposed in gambling game by J.L. Kelly, has been successfully adopted in financial markets [25], because it suggests a wiser position ratio, and provides a promising method of investment. It is meant to gain maximum expectation value of profit rate. It can be described as follows:

$$f = \frac{p_{win} * b - p_{loss}}{b}, \tag{1}$$

where f denotes the position ratio. More specifically, f stands for the proportion of the money for bets in total capital at a time. p_{win} represents the probability of wins while p_{loss} means the probability of losses. b denotes odds, specifically the ratio of profits (excluding capital) to capital under the circumstance of win.

Emil Ohlsson and Oskar Markusson [26] evaluated the Kelly formula on the Swedish stock market between 2005 and 2015. It used the last three months of data to establish the Kelly Criterion. It is concluded that the Kelly strategy produces five times higher returns than traditional portfolio optimization methods.

In [27], the advantages and disadvantages of Kelly Criterion are investigated in detail. As is concluded in it, the main advantage of the Kelly Criterion is that it maximizes the limiting exponential growth rate of wealth. While the main drawback is that its suggested wagers may be very large.

To improve the large wager problem, [28] shrink the size of the bet by a shrinkage factor. From a simulation study and analysis, it is shown that the shrunken Kelly approaches can greatly improve the 'raw' Kelly Criterion.

III. Methodology

In this section, our proposed method is presented in detail. It is elaborated from the following three perspectives: market timing, stock ranking, and portfolio management.

A. Market Timing

1) Three Classification Task: In technical analysis, the future trend of stock movements can be extracted through processing and analyzing historical financial data. For all candidates in a stock pool, their previous data can be sorted as a financial time series. Assuming that the current time is T. The aim is to estimate the stock movements at time w ($w = T + k$) over k periods. Rather than directly forecast the price change, we do the following three classification task:

$$F(Q_{i,1}^{w-s*k}, ..., Q_{i,j}^{t}, ..., Q_{i,12}^{T}) = \begin{cases} 1 & R_i^w > \gamma \\ 2 & R_i^w < -\gamma \\ 0 & elsewhere \end{cases} \tag{2}$$

, where s stands for timesteps and R_i^w denotes the price change of stock i at time w over k periods. $Q_{i,j}^t$ represents the jth ($j \in \{1, 2, ..., 12\}$) feature of stock i at time $t \in \{w - s*k, w - s*k + k, ..., T\}$. A total of 12 technical indicators are introduced to represent the state of stock at current time T. They are R_i^{T-k}, R_i^{T-2k}, R_i^{T-3k}, open, close, high, low, volume, money, open price of CSI300 index, close price of CSI300 index, and average price. In the above formula, F can be a linear or non-linear function. γ is related to the transaction costs.

The classification task has already taken the transaction costs into account. In fact, even if we grasp the right trend, it may not be profitable due to the high stock trading fees and the low fluctuation of the stock price. Only when the fluctuation amplitude of the stock price is more than transaction fees, it will be marked as the up or down signal.

To be more specific, with the price change more than a certain value γ, the stock is considered to have an up signal. The circumstances of stock's upward movement are classified into category '1'. With the price change less than $-\gamma$, the stock is considered to have a down signal. The circumstances of stock's downward movement are classified into category '2'. The price change between γ and $-\gamma$ stands for no fluctuation of the stock.

The ground truths of the three classification task are based on the one-hot encoding.

2) Preprocessing: In the preprocessing step, the financial data of each stock is normalized using the z-score normalization to transform data of different magnitudes into uniform metrics:

$$\tilde{Q}_j = \frac{Q_j - \mu_j^{train}}{D_j^{train}}, \tag{3}$$

where μ_j^{train} denotes the average value of jth feature during the training period, and D_j^{train} stands for the standard deviation of jth feature during the training period. As stated in [19], it is the key to obtain mean and standard deviation from the training set only. Only in this way can look-ahead bias be avoided.

3) LSTM-individual: The entire data are split into training samples and trading samples. 20 percent of training samples are used as a validation set. The rest of the training samples are used as a training set. Rolling window method [19] is utilized in the processes of training, validating, and trading. Each time the rolling window moves forward k periods, as presented in Fig. 2.

Each stock has its individual LSTM model, the size of which is the same as others. The input layer has 12 features and 100 timesteps. The LSTM model in our trading system contains one single standard LSTM layer with 32 neurons. Its last layer is one dense layer with 3 neurons and a softmax activation function. By applying dropout regularization [29] within the recurrent layer, 0.05 of units are randomly dropped both at the input gates and the recurrent connections. This can avoid overfitting and facilitate generalization. Considering the problem of sample imbalance [18], random under-sampling is employed to improve the training process.

B. Stock ranking: Dynamic K-top

The outputs of our stock prediction model are the probabilities of the price change of stock more than γ, below $-\gamma$, and between γ and $-\gamma$. Thus, associated with the price change, these probabilities can represent the predicted future performance of the stocks.

Since each stock's financial data are fed into its individual model, there are $3N$ outputs (N represents the number of stocks in a stock pool, e.g. $N = 165$). Those stocks with a higher probability of up signal are predicted to have upward movements, while those stocks with a higher probability of down signal are predicted to have downward movements. Previous works tended to go long the most promising stocks for the whole trading period. However, the numbers of stocks chosen by them are fixed, which cannot follow the market change.

From an economic perspective, K represents the balance of return and risk. If K is small, it brings many risks and extreme returns at the same time. As K increases, the volatility of returns will be smaller and the returns will be more stable, which means lower returns and lower risks.

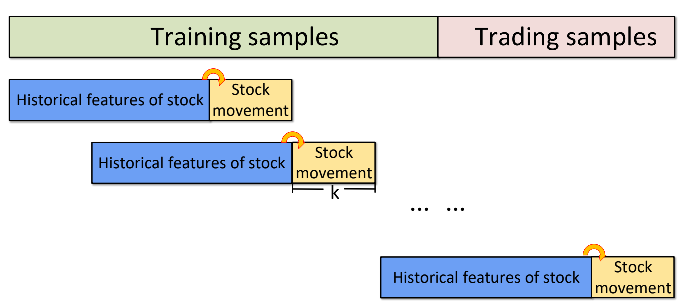


Fig. 2. The rolling window method.

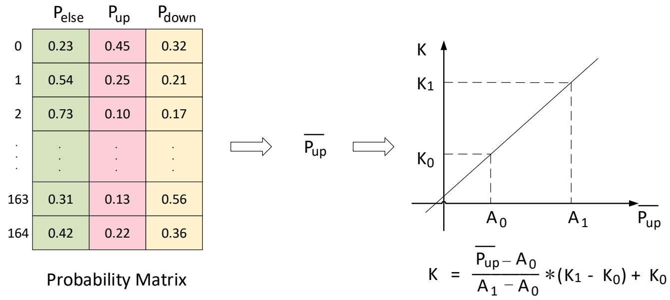


Fig. 3. The procedure of dynamic K-top.

We regard the average value of the predicted probability of 165 stocks' upward movements as the stock pool's predicted trend. The standard deviation of the stock pool's predicted trend is small here. If the average value is high, many stocks tend to move up. So more stocks should be bought. If the average value is low, fewer stocks should be bought. The relationship between stock movements and K can be represented by the linear transformation.

Supposing average value $\overline{P_{up}}$ ranges from A_0 to A_1 and the range of K is from K_0 to K_1:

$$K = \frac{\overline{P_{up}} - A_0}{A_1 - A_0} * (K_1 - K_0) + K_0, \tag{4}$$

where A_0 and A_1 are statistical data based on training period data. We will choose at least one stock, so K_0 is set to 1.

C. Portfolio Management: Kelly Criterion

To some extent, stock trading is just like gambling. The difference is that gambling has a fixed winning rate and fixed odds. While in the stock market, it requires traders to analyze and predict future situations by utilizing relevant information.

In the environment of stock trading, we have to use the general form of the Kelly Criterion:

$$f = \frac{p_{win}}{c} - \frac{p_{loss}}{b}. \tag{5}$$

The detailed descriptions of basic parameters in the Kelly Criterion are listed in Table I.

Here for stock trading, p_{win} is the probability of upward movements for the stock derived from LSTM. p_{loss} is the probability of downward movements for the stock derived from LSTM. b is the increase rate and c is the absolute value of decrease rate. b and c are the statistical price change based on training period data. We assume no leverage but Kelly Criterion often wagers too much [27]. So some small modifications should be taken:

$$\tilde{f} = min(max(\beta * f, 0), 1), \tag{6}$$

where β is the shrinkage coefficient and f stands for the position ratio. Here, β represents the ability to take the risk, as stated in [28].

At each time, the transaction amount is determined by the current remaining cash multiplying the position ratio which is

TABLE I
BASIC PARAMETERS IN KELLY CRITERION

Name	Description
f	the proportion of the money for buying the stock in the remaining capital at a time
p_{win}	the probability of upward movements of the stock
c	the ratio of losses to capital under the circumstance of lose
p_{loss}	the probability of downward movements of the stock
b	the ratio of profits (excluding capital) to capital under the circumstance of win

determined by the Kelly formula. If the stock continues to rise or fall, we will continue to hold it until the stock is predicted to move in the opposite direction.

IV. COMPARISONS AND RESULTS

A. Experimental Settings

1) Data Descriptions: CSI300 constituents have the characteristics of high liquidity, active trading, and good market representation. The constituents of CSI300 in granularity of five minutes dating from January to June in 2018 are chosen as the study example. Considering the high transaction costs, high liquidity is what we prefer. So those stocks whose standard deviation of close price less than one are eliminated from the stock pool. Suspended stocks are also eliminated. The total number of stocks in our stock pool is 165. The training samples include 4 months of stock data, with 19200 minutes in total. The following two weeks are trading samples, with 2400 minutes in total.

2) Parameter Settings: We trained all LSTM-individual models on a computer with 1 CPU (Intel i7-8700K 3.70GHz). The maximum training duration of each model is 1000 epochs, and early stopping patience is 10. γ which is mentioned in III-A is set to 0.001.

3) Benchmark Models: In order to verify the superiority of our proposed method, we compare the results of our methods with many other counterparts. The comparisons are conducted from the following three aspects, prediction, stock ranking, and portfolio management. To keep the fairness of comparisons, comparing objects for the same task is set to the same environment.

For the prediction model, pseudo-random (randomly choose K stocks to buy), random forest, and convolution neural network (CNN) are adopted. For fair comparisons, CNN is set as the same layer number and neuron number as LSTM. For the stock ranking task, we set several ranges of K, and compare the dynamic K-top with traditional fixed K-top. For the portfolio management task, we compare the Kelly Criterion with the equally weighted portfolio and minimum variance portfolio. CSI300 index return and average market return are also set as baselines.

B. Evaluation Criteria

The circumstances of stocks classified into category '1' and category '2' are recorded as RFs. To evaluate the performance of stock prediction, we adopt three metrics $AccRF$ (precision rate), $RecRF$ (recall rate), and $CerRF$ (critical error rate), as stated in [18]:

$$AccRF = \frac{NAcc}{NRF_p}, \tag{7}$$

$$RecRF = \frac{NAcc}{NRF}, \tag{8}$$

$$CerRF = \frac{NCer}{NRF_p}, \tag{9}$$

where NRF_p denotes the number of predicted RFs. $NAcc$ denotes the number of RFs that are classified correctly. $NCer$ stands for the number of RFs that are classified into antitype. NRF represents the number of predicted RFs that are truly in category '1' or '2'.

Furthermore, to verify the superiority of our method's returns, we also calculate the daily return, accumulative return, maximum drawdown, winning rate and so on for evaluation disregarding the transaction costs.

C. Empirical Results

1) Comparisons with Other Prediction Models: In Table II, we compare the proposed LSTM model with pseudo-random, random forest, and CNN when K is set to 5. It is shown that the LSTM model has the best performance in terms of AccRF, RecRF, and CerRF. AccRF and CerRF are related to the profitability while Rec is associated with the missed opportunity. Therefore, AccRF and CerRF are more important than RecRF in the stock trading process. With the same number of layers and neurons as LSTM, CNN outperforms pseudo-random and random forest but performs less well than LSTM in terms of prediction accuracy. The results can prove that LSTM has a strong ability to forecast financial time series precisely even if the task of predicting the stock movements is of high complexity.

TABLE II
PREDICTION RESULTS OF DIFFERENT MODELS

Model	AccRF	RecRF	CerRF
Pseudo-random	26.3%	45.5%	31.5%
Random forest	35.0%	51.7%	32.6%
CNN	35.2%	54.4%	29.6%
LSTM	**41.2**%	**58.9**%	**28.7**%

2) Comparisons with Fixed K-Top: To make the comparisons fair and valid, the dynamic K-top and fixed K-top both use LSTM as the prediction model and equally weighted portfolio as portfolio management. We take several ranges of dynamic K-top and the corresponding values of fixed K as examples.

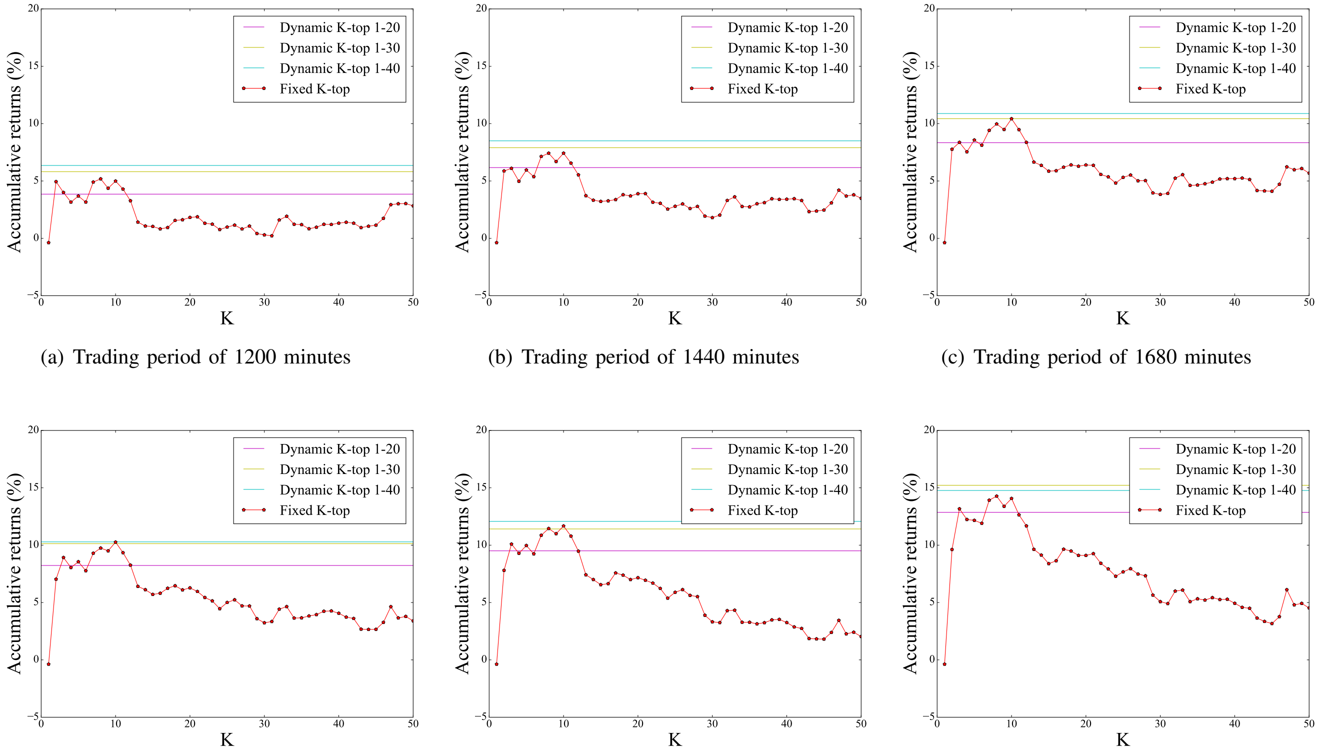
TABLE III
RESULTS OF DYNAMIC K-TOP AND FIXED K-TOP

Evaluation Metrics	**Different ranges of dynamic K**			
	1-10	**1-20**	**1-30**	**1-40**
Accumulative return	10.14%	12.86%	**15.22**%	14.77%
Winning rate	51.70%	**53.67**%	52.17%	52.49%
Daily return	0.81%	0.89%	1.08%	**1.12**%
Maximum drawdown	2.61%	2.46%	1.96%	**1.78%**
Evaluation Metrics	**Different values of K**			
	10	**20**	**30**	**40**
Accumulative return	14.07%	9.10%	5.07%	4.93%
Winning rate	51.73%	53.60%	53.60%	53.60%
Daily return	1.11%	0.67%	0.34%	0.42%
Maximum drawdown	2.15%	4.35%	2.55%	2.99%

In Table III, detailed comparisons between dynamic K-top and fixed K-top are listed. Fig. 4 illustrates the relationship with accumulative return and different values of fixed K. Besides, we draw the accumulative return of dynamic K-top for vivid comparisons. It can be seen from the figure that different values of K have different returns after a trading period. This means that it is possible to gain more profit by changing the K value. There is always a best K with regard to the time in terms of return. We draw six different trading periods for comparisons. During the whole trading period, the accumulative return is affected by the K value constantly. The volatility of return will be much larger when considering the return each time instead of a long period. If the number of stocks that makes up a portfolio is not very large, then much risk is taken.

As K increases, the accumulative return becomes lower with small volatility. This agrees with the above argument that as the number of chosen stock increases, the return will become more stable and less volatile. In these figures, dynamic K-top ranging from 1 to 40 and ranging from 1 to 30 always perform better than the best fixed K value of different trading periods in terms of return. When the trading period is 2400 minutes, dynamic K-top 1-30 exceeds the best fixed K-top with the return rate of 0.95%. In addition, the return of dynamic K-top 1-40 is 0.5% over the best fixed K-top. A wider range means that there are more stocks to choose from the stock pool. More stocks can better represent the trend of the whole stocks. Extensive experiments have shown that by employing dynamic K-top method, it is possible to obtain an excess return in comparison with almost all the fixed K values.

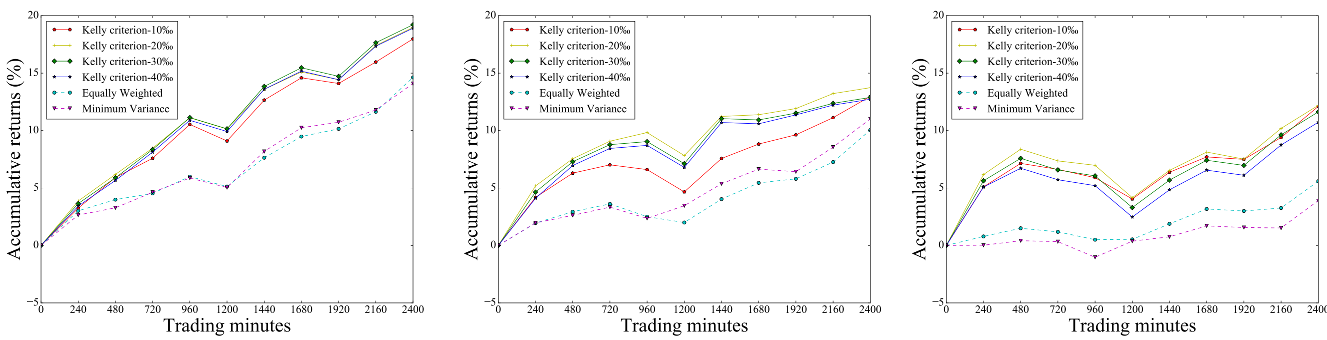
3) Comparisons with Other Portfolio Methods: For portfolio management comparison, all the different portfolio optimization methods take LSTM as prediction and choose $K \in \{10, 20, 30\}$ and shrinkage coefficient $\in \{10‰, 20‰, 30‰, 40‰\}$. The detailed results are shown in Fig. 5. As shown in these figures, the accumulative returns of LSTM with Kelly Criterion with all coefficients are higher than that of LSTM with equally weighted portfolio and minimum variance portfolio. When K is 30 and the trading period is 2400 minutes, the return of Kelly Criterion with shrinkage coefficient 10‰is 6.15% over the equally weighted portfolio and 8.2% over the minimum variance portfolio. Different shrinkage coefficient doesn't matter too much as they almost follow the same trend during the trading period. A lower shrinkage coefficient means the weaker ability to shoulder much risk. In these three figures, the lowest shrinkage coefficient brings the lowest accumulative return. It can also be seen in the figures that the traditional portfolio shows very low volatility. It agrees with the rule that lower volatility means a lower return. Compared with



(a) Trading period of 1200 minutes (b) Trading period of 1440 minutes (c) Trading period of 1680 minutes

(d) Trading period of 1920 minutes (e) Trading period of 2160 minutes (f) Trading period of 2400 minutes

Fig. 4. Comparisons between dynamic K-top with fixed K-top in terms of accumulative returns (%) in different trading periods.



(a) k=10 (b) k=20 (c) k=30

Fig. 5. Comparisons between Kelly Criterion and other portfolio optimization methods in different fixed k values.

the equally weighted portfolio method and minimum variance portfolio method, it is shown that the Kelly Criterion can maximize the profit given an appropriate level of risk exposure.

4) Comparisons with Other Formulated Benchmark Models: We further formulate other baselines to prove the outstanding performance of our method. In Fig. 6, average market return and CSI 300 return are set as our baselines for comparisons. K is set to 10 and the shrinkage coefficient is 10‰. LSTM-DK represents the proposed LSTM-based model with the Kelly Criterion and dynamic K-top. With the Kelly Criterion and dynamic K-top method, LSTM-DK shows a very big improvement. It greatly outperforms the average market return and CSI300 return.

In Table IV, detailed comparisons are made between the proposed model and many similar counterparts in terms of winning rate, accumulative return, maximum increase per operation, daily return, maximum drawdown, monthly alpha, monthly Sharpe ratio, and monthly Sortino ratio. K is set to 30 and the shrinkage coefficient is 10‰.

According to the majority of evaluation criteria, LSTM performs better than CNN because LSTM has better forecasting ability than CNN in financial series prediction. According to

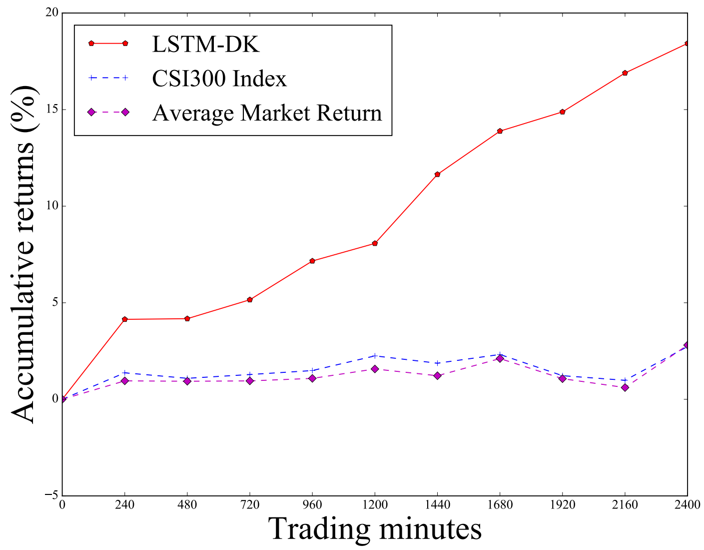


Fig. 6. Comparisons between the proposed model with market performances in terms of accumulative returns (%).

TABLE IV
RESULTS OF THE PROPOSED MODEL AND OTHER FORMULATED BENCHMARK MODELS

Evaluation Metrics	fixed K-top equally weighted CNN	LSTM	fixed K-top Kelly Criterion CNN	LSTM	dynamic K-top Kelly Criterion CNN	LSTM
Accumulative return	3.11%	5.07%	3.64%	7.03%	1.0%	**7.35%**
Winning rate	42.86%	54.70%	43.19%	55.52%	43.38%	**55.95%**
Maximum increase per operation	0.46%	0.59%	0.91%	2.04%	0.84%	**2.04%**
Daily return	0.28%	0.45%	0.34%	**0.63%**	0.32%	0.44%
Maximum drawdown	**1.49%**	2.55%	5.35%	4.66%	5.35%	4.70%
Alpha	0.02	0.10	0.05	0.15	0.02	**0.16**
Sharpe ratio	-0.03	1.32	0.21	1.81	-1.12	**2.17**
Sortino raio	-0.04	1.91	0.28	2.53	-1.64	**2.96**

most of the evaluation criteria, LSTM with Kelly Criterion is better than LSTM. Moreover, CNN with Kelly Criterion is better than CNN. The reason is that the Kelly Criterion can maximum the limiting exponential growth rate of wealth.

As for return analysis, the LSTM-based trading strategy with the dynamic K-top and Kelly Criterion yields an increase of 44.79% than the vanilla LSTM-based trading strategy in terms of accumulative return. This means that dynamic K-top with the Kelly Criterion can strengthen the ability to earn more money.

As for risk analysis, the aims of investors are to maximize the investments' expected return given an appropriate level of risk exposure. The Sharpe ratio and Sortino ratio are applied to measure the return of an investment compared to its risk. From Table IV, it is shown that LSTM-DK has a higher value of monthly Sharpe ratio and monthly Sortino ratio compared with other formulated models. Depending on the risk exposure investors can take, Kelly Criterion can be further tuned and the shortcoming of high volatility can be further reduced.

Considering the high transaction fees, the net return of LSTM with the dynamic K-top and Kelly Criterion is 4.40%. Higher net returns indicate better-performing investments. It provides solid evidence for the effectiveness of the proposed LSTM-based quantitative trading system with the dynamic K-top and Kelly Criterion.

V. CONCLUSION

In this paper, we use the dynamic K-top stock selection method and Kelly-Criterion-based portfolio management strategy to improve the performance of an LSTM-based quantitative trading system. We combine market timing, stock selection, and portfolio management to improve the quantitative trading system to a large extent. From prediction, ranking, and portfolio management aspects, extensive comparisons and analyses have been conducted to validate the effectiveness and reasonableness of our methods in an all-around way. The results show that the proposed model is superior to similar counterparts and greatly outperforms the CSI300 index and the market average. Compared with the straight forward LSTM-based trading strategy, the LSTM-based trading strategy with the dynamic K-top and Kelly Criterion can achieve an increase of 44.97% over ten days in terms of accumulative return.

Further work will be carried out in the following aspects. We will test our proposed method on more datasets to explore deeply into the effect of Kelly Criterion and dynamic K-top. Then, more optimization methods can be added to the task of stock prediction.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61774082 and Grant 61604068 and in part by the Fundamental Research Funds for the Central Universities under Grant 021014380065. (Corresponding authors: Jun Lin; Zhongfeng Wang.)

REFERENCES

[1] Werner FM De Bondt and Richard Thaler. Does the stock market overreact? *The Journal of finance*, 40(3):793–805, 1985.

[2] Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993.

[3] Lawrence Blume, David Easley, and Maureen O'hara. Market statistics and technical analysis: The role of volume. *The Journal of Finance*, 49(1):153–181, 1994.

[4] Sima Siami-Namini and Akbar Siami Namin. Forecasting economics and financial time series: Arima vs. lstm. *arXiv preprint arXiv:1803.06386*, 2018.

[5] Francis EH Tay and Lijuan Cao. Application of support vector machines in financial time series forecasting. *omega*, 29(4):309–317, 2001.

[6] Sima Siami-Namini, Neda Tavakoli, and Akbar Siami Namin. A comparison of arima and lstm in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1394–1401. IEEE, 2018.

[7] Eduardo A Gerlein, Martin McGinnity, Ammar Belatreche, and Sonya Coleman. Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications*, 54:193–207, 2016.

[8] David MQ Nelson, Adriano CM Pereira, and Renato A de Oliveira. Stock market's price movement prediction with lstm neural networks. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 1419–1426. IEEE, 2017.

[9] Siyuan Liu, Guangzhong Liao, and Yifan Ding. Stock transaction prediction modeling and analysis based on lstm. In *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 2787–2790. IEEE, 2018.

[10] Xiaodan Liang, Zhaodi Ge, Liling Sun, Maowei He, and Hanning Chen. Lstm with wavelet transform based data preprocessing for stock price prediction. *Mathematical Problems in Engineering*, 2019, 2019.

[11] Jian Cao, Zhi Li, and Jian Li. Financial time series forecasting model based on ceemdan and lstm. *Physica A: Statistical Mechanics and its Applications*, 519:127–139, 2019.

[12] Fengmei Yang, Zhiwen Chen, Jingjing Li, and Ling Tang. A novel hybrid stock selection method with stock prediction. *Applied Soft Computing*, 80:820–831, 2019.

[13] Lean Yu, Lunchao Hu, and Ling Tang. Stock selection with a novel sigmoid-based mixed discrete-continuous differential evolution algorithm. *IEEE Transactions on Knowledge and Data Engineering*, 28(7):1891–1904, 2016.

[14] Fischer Black and Robert Litterman. Global portfolio optimization. *Financial analysts journal*, 48(5):28–43, 1992.

[15] Andre F Perold. Large-scale portfolio optimization. *Management science*, 30(10):1143–1160, 1984.

[16] Yuan Song. *Stock trend prediction: Based on machine learning methods*. PhD thesis, UCLA, 2018.

[17] Svetlana Borovkova and Ioannis Tsiamas. An ensemble of lstm neural networks for high-frequency stock market classification. *Journal of Forecasting*, 2019.

[18] Siyu Yao, Linkai Luo, and Hong Peng. High-frequency stock trend forecast using lstm model. In *2018 13th International Conference on Computer Science & Education (ICCSE)*, pages 1–4. IEEE, 2018.

[19] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.

[20] Satit Yodmun and Wichai Witayakiattilerd. Stock selection into portfolio by fuzzy quantitative analysis and fuzzy multicriteria decision making. *Advances in Operations Research*, 2016, 2016.

[21] Mohaddeseh Hajjami and Gholam R Amin. Modelling stock selection using ordered weighted averaging operator. *International Journal of Intelligent Systems*, 33(11):2283–2292, 2018.

[22] Ratchata Peachavanish. Stock selection and trading based on cluster analysis of trend and momentum indicators. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, pages 317–321, 2016.

[23] XingYu Fu, JinHong Du, YiFeng Guo, MingWen Liu, Tao Dong, and XiuWen Duan. A machine learning framework for stock selection. *arXiv preprint arXiv:1806.01743*, 2018.

[24] Chien-Feng Huang. A hybrid stock selection model using genetic algorithms and support vector regression. *Applied Soft Computing*, 12(2):807–818, 2012.

[25] Edward W Piotrowski and Malgorzata Schroeder. Kelly criterion revisited: optimal bets. *The European Physical Journal B*, 57(2):201–203, 2007.

[26] Oskar Markusson and Emil Ohlsson. Application of the kelly criterion on a self-financing trading portfolio-an empirical study on the swedish stock market from 2005-2015. 2017.

[27] Leonard C MacLean, Edward O Thorp, and William T Ziemba. Good and bad properties of the kelly criterion. *Risk*, 20(2):1, 2010.

[28] Rose D Baker and Ian G McHale. Optimal betting under parameter uncertainty: improving the kelly criterion. *Decision Analysis*, 10(3):189–199, 2013.

[29] Yarin Gal and Zoubin Ghahramani. A theoretically grounded application of dropout in recurrent neural networks. In *Advances in neural information processing systems*, pages 1019–1027, 2016.