Can the Futures Market be Predicted-Perspective based on AutoGluon

YangChun Xiong[1], ZiXuan Pan[2]

Department of Finance and Accounting

Guangzhou Huashang Vocational College, Guangzhou, China

BaiFu Chen[3]

School of Economics and Trade

Guangdong University of Finance, Guangzhou, China

Abstract—This paper discusses how to raise efficiency of predicting the Chinese futures market correlation coefficient. First, the predicted periods are divided by major events and the predictabilities between different periods are compared at the same time. Second, on this basis, an automatic machine learning framework, AutoGluon is applied to compare the predictive ability between different deep learning models such as LSTM and GRU. Results demonstrate that: (1) Compared by LSTM and GRU, AutoGluon can indeed raise efficiency of predicting. (2) The changes of prediction error between different periods can explain the influence of major events happened in futures market. (3) Although the predictive ability of many models decline over time, the performance of XGBoost is relatively stable, which can provide useful tools for market participants.

Keywords—AutoGluon; LSTM; GRU; Chinese futures market

I. INTRODUCTION

Financial time series forecasting methods include econometric method represented by ARIMA model and GARCH model, and deep learning methods represented by LSTM and GRU model. Futures market forecasting can also adopt the same method. The traditional econometric methods based on linear function hypothesis and model driven show good applicability in dealing with small amount of calculation and low-dimensional data. However, with the explosive growth of data volume in the era of big data, econometric method has gradually exposed its weaknesses.

The LSTM model was originally designed for natural language processing tasks. Now it has attracted more and more attentions in time series forecasting tasks. A large number of in-depth studies have been carried out on the use of LSTM model to predict future prices. The results showed that deep learning method has obvious advantages in forecasting accuracy compared with econometric methods. However, the deep learning models are too dependent on model structure and parameter adjustment, which makes it is difficult to deploy rapidly in different situations. In recent years, the automatic machine learning method represented by AutoGluon framework has performed excellently in various tasks relying on bagging and stacking strategy, and has attracted more and more attentions due to its ease of use.

Whether deep learning model or automatic machine learning method, the premise for predicting the future is that the data obeys the assumption of independent and identical distribution. A large number of studies have proved that it is difficult for financial market participants to get rid of psychological effects such as greed and fear, resulting in the fact that historical data have a certain impact on future data, which leads to the fact that financial time series data are not completely independent before and after. At the same time, the impact of major events may change the expectations of market participants, making it difficult for financial time series data in different periods to maintain the same distribution assumption. These problems make many models highly fitting historical data in the training process have poor generalization ability in the testing process. Models that perfectly fit existing data cannot guarantee the same prediction accuracy in the future.

In short, the assumption that financial time series data obey the independent and identical distribution fundamentally challenges the basis of machine learning model to predict the future. The specific manifestation is the difference between training error and testing error, but the changes of difference may indicate the changes of market risk.

The innovation of this paper is that according to the time point of major public events (China's supply-side reform and the new corona epidemic), the futures index time series data are segmented and the correlation coefficients between varieties are calculated. By analyzing the difference between the training error and the testing error of the futures index correlation coefficient, new ideas are provided for the futures market prediction.

The structure of this paper is as follows: the second section is the literature review; the third section introduces the models used in this paper; the fourth section is the empirical analysis process; and the fifth section is the summary and enlightenment.

II. LITERATURE REVIEW

There are many studies on the use of econometric models to predict financial markets. Li Hongquan [1] used interval measurement method to study the crude oil price prediction. Zhang Y J, Yao T, He L Y [2] compared the abilities of different GARCH models to predict the crude oil market. Li Hongquan and Zhou Liang [3] used CoVaR, cross-sectional VaR, absorption ratio, Granger causality index and information spillover index to measure systemic financial risk, and examined the predictive ability of five indicators on macroeconomy in detail. Hong Yongmiao, Wang Shouyang [4] pointed out that the econometric methods focus on the relationships between the economic variables to reveal the inherent nature of economic operation, but due to the highly simplified and abstract mathematical model, many other

factors in reality may be not taken into account, which often results in model misdesign.

With the rapid development of artificial intelligence technology in the context of big data, machine learning, deep learning and text analysis have been widely used in the research of financial market prediction. Chen Y, He K, Tso G K F.[5] used deep learning model to predict the crude oil prices, R.A.de Oliveira, D.M.Q.Nelson, A.C.M.Pereira. The author in [6] studied the application of LSTM model in stock market forecasting. Mu Nianguo, Yao Honggang [7] proposed a prediction model of recurrent neural network based on attention mechanism, and found that the prediction effect of gated recurrent network was improved after adding attention mechanism. In addition, a large number of literatures focus on improving the prediction ability of deep learning model in stock and commodity markets [8]-[16]. The common point of the above research is using machine learning models to predict future prices directly, and the researches focused on improving the accuracy and speed of model prediction. Ensembles that combine predictions from multiple models have long been known to outperform individual models. Wang Y, Liu L, Wu C.[17] studied the effect of using time-varying parameter models to predict the crude oil prices. Sun Fuxiong et al. [18] took the Chinese listed companies as the research object, and put forward the combination model of stock suspension prediction. The empirical analysis results showed that the combination model prediction has achieved high accuracy. Zhou Hao et al. [19] proposed an improved crude oil price combination forecasting model, therefore proposed a dynamic particle swarm optimization algorithm. The experimental results showed that the predictions of combined model can greatly reduce the computational complexity and improve the prediction accuracy. Nick Erickson, Jonas Mueller et al. [20] proposed the AutoGluon framework based on automatic machine learning, which greatly simplifies the preliminary work such as feature engineering and parameter debugging of traditional machine learning models, and performs well in the prediction task of structured data. To our knowledge, there is no precedent to apply AutoGluon to the prediction of financial time series. In general, the research on the prediction of financial time series using single or combined models of econometrics and deep learning has been quite sufficient. But the research on the prediction performance of automatic machine learning is not sufficient enough, and the financial time series data do not obey the assumption of independent and identical distribution is always an unavoidable matter. In this paper, the AutoGluon framework is used for predicting the financial time series for the first time. By analyzing the difference between the training error and the testing error of the correlation coefficients of the futures index, the influence of major public events on the futures market is studied to explore the method of predicting the risk of China's futures market and provide reference for researchers.

III. MODELS DESCRIPTION

A. AutoGluon based on Automatic Machine Learning

In the past decades, many powerful machine learning models have emerged. But how to integrate these models is faced with many obstacles, such as model selection, model integration, super-parameter adjustment, feature engineering, and data preprocessing. Automatic machine learning(AutoML) provides a possible solution through the combination of model selection algorithm and super-parameter optimization strategy. As a representative of AutoML, AutoGluon arranges and trains different models hierarchically, which saves training time and reduces overfitting by bagging and stacking strategy. It has long been found that the combination of multiple models can achieve better performance than single models. The popular AutoML uses bagging and stacking strategy to improve prediction ability and reduce variance. Specifically, several 'base' models are trained separately at each layer, then the outputs of each model are aggregated as features to be transmitted to the next layer for further training (stack) to achieve performance beyond the 'base' models. As a typical AutoML, AutoGluon embodies these ideas in Fig. 1.
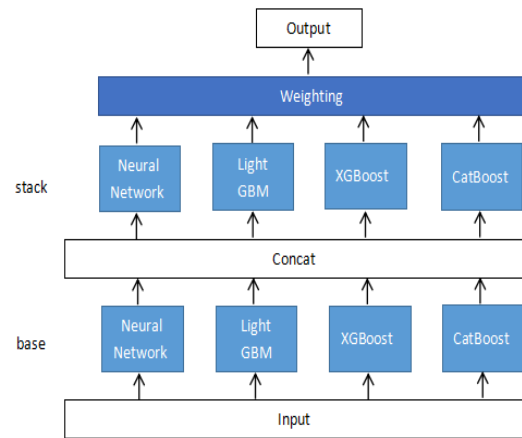

Fig. 1. AutoGluon's Multi-layer Stacking Strategy.

This paper select four representative machine learning algorithms to generate the basic model of AutoGluon, including artificial neural network, LightGBM algorithm, XGBoost algorithm and CatBoost algorithm. After the data were input into the model, different samples are formed by random repeated sampling, then the bagging strategy is applied to each layer to train the basic model on different samples by using four algorithms. At the same time, the stacking strategy is used to train the basic model on the same original data sample of each layer. Finally, all the scalars of each model output are connected to obtain a vector, and then a linear combination is made to obtain the final output of the model. The key codes are given in Fig. 2.

```
1.hyperparameters={'GBM':gbm_options,'NN':nn_options,'XGB':xgb_options,'CAT':cbm_options,}

2.time_limit = 1000 num_trials = 5 search_strategy = 'auto'

3.hyperparameter_tune_kwargs={'num_trials':num_trials,'scheduler':'local','searcher':search_strategy,}predictor=TabularPredictor(label=label).fit(train_data,time_limit=time_limit,num_stack_levels=1,num_bag_folds=3)
```

Fig. 2. Key Codes of AutoGluon.

B. LSTM and GRU

As branches of the recurrent neural network, LSTM and GRU models can solve the problem of gradient disappearance, and are often used for time series prediction. In order to compare with AutoGluon, Keras platform is applied to build LSTM and GRU models, and the key codes are given in Fig. 3.

```
1.Key codes of LSTM

model=Sequential()
model.add(LSTM(units=4,activation='tanh',recurrent_activation='hard_sigmoid',input_shape =
(15 , 1)))model.add(Dense (units =1, activation
= 'linear'))

model.compile(loss='mean_absolute_error',optimizer = 'rmsprop')

history=model.fit(x_train,y_train,batch_size=1,
epochs =30, shuffle = True )

2.Key codes of GRU

model = Sequential()

model.add(GRU(units=4,return_sequences=False,activation='tanh',recurrent_activation='hard_sigmoid', input_shape =(15 , 1)))

model.add(Dense(units=1,activation ='linear'))

model.compile(loss='mean_squared_error',optimizer ='rmsprop')

history=model.fit(x_train,y_train, batch_size
=1,epochs =30)
```

Fig. 3.   Key Codes of LSTM and GRU.

C. Model Assessment Index

There are many indexes to evaluate the fitting ability of machine learning model. This paper use mean square error (MSE) and mean absolute error (MAE) as model evaluation indexes. These can be calculated using the following formulas.

$$\begin{cases} MSE = \dfrac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2 \\ \\ MAE = \dfrac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}| \end{cases} \qquad (1)$$

where n is the total number of samples, is the actual value and is the predicted value.

Compared with MAE, MSE gives greater weight to outliers, so it is not as stable as MAE. For the fixed learning rate, the effective convergence of MSE is better than that of MAE, so MSE and MAE are used to evaluate the performance of the models.

IV. DATA DESCRIPTION AND EMPIRICAL ANALYSIS

A. Variables and Data

It is common to select extra-price indicators as explanatory variables to forecast future price. However, the available time of extra-price indicators often lags behind the price itself, which leads to the fact that the hindsight predictable phenomenon cannot be realized in real time. Moreover, the reflexivity between some extra-price indicators and prices is hard to be falsified. For example, oil prices influence oil production and vice versa.

Therefore, this paper study the risk measurement of futures market by establishing the correlation coefficient time series between the futures index of rebar, iron ore and coke. The explained variable is the current value of the correlation coefficient of China's futures market price index, and the explanatory variable is the historical value of the correlation coefficient. The specific algorithm is to use the corr function of math module in python to calculate the correlation coefficient based on the daily closing price of futures index, and the number of cycles is 100.

This paper adopts the black industry index of South China Futures released by the tushare data community, which includes rebar, hot coil, iron ore, coke, coking coal, wire rod, manganese silicon and ferrosilicon. However, due to the different listing dates of each variety, the historical transactions of wire rod, manganese silicon, ferrosilicon, hot coil and coking coal are not active and the market influence is small. Considering the above factors, this paper only analyzes the futures price index of rebar, iron ore and coke for 2001 trading days from October 21, 2013 to December 31, 2021.

It can be seen from Table I that the original data (from 21 October 2013 to 31 December 2021) is divided into six intervals according to the approximate time points of China's supply-side reform and the new coronavirus epidemic. Then the correlation coefficients are calculated in each intervals. Finally, six training sets and six testing sets are generated, which are divided as follows.

TABLE I.   DATASET PARTITION

Interval number	training sets date	testing sets date	reference
1	20131021-20141114	20141117-20151211	Before supply-side reform
2	20131021-20151211	20151214-20180205	In supply-side reform
3	20151214-20161108	20161109-20170927	After supply-side reform
4	20180223-20190212	20190213-20200123	Before new coronavirus
5	20180223-20200123	20200203-20211231	In new coronavirus
6	20200203-20210113	20210114-20211231	After new coronavirus

Then the training set and testing set are input into the model respectively. Finally, the results are compared and analyzed. The specific process is given in Fig. 4.

B. Empirical Analysis

For market participants, when a good fitting model of historical data (training set) can predict future data (testing set) within a certain error range, the risk is low. On the other hand the risk rises when the prediction error increases. Based on this, this paper proposes two hypotheses: 1) When the market is

influenced by external events, the risk will increase characterized by greater prediction error between the training set and the testing set. 2) When the market gradually adapts to the influence of external events, the risk will be reduced which is characterized by the decrease of prediction error between the training set and the testing set. If these two assumptions hold, it can be estimated that the market risk level by observing the change of the prediction error between the training set and the testing set, and then replace the model when the original model is obviously unable to adapt to market changes.
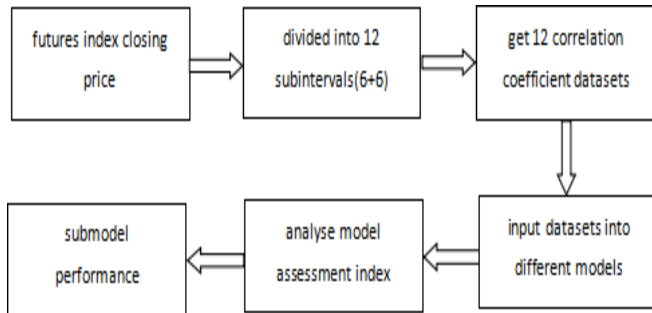

Fig. 4.   Data Processing Procedure.

After inputting the dataset into different models, the output results are as follows.

where Ratio of error=Testing MSE / Training MSE, and the smaller the MSE index is, the better the fitting degree of the model to the dataset is.

It can be seen from Table II that AutoGluon framework has obvious advantages in fitting degree compared with LSTM model and GRU model in each training set, but it is completely backward in the testing set. Especially in interval 5, the error ratio of AutoGluon framework is as high as 24.22, which is far higher than that of other models. If the corresponding MSE index is carefully observed, it can be found that the MSE of the training set is only 0.01, and the MSE of the testing set is 0.35, which indicates that the AutoGluon framework has a certain overfitting phenomenon and leads to poor generalization ability of the model. Therefore, when measuring the prediction accuracy of the model, the MSE value of the model in a single interval cannot be used as the sole criterion, but the performance of the model on the training set and the testing set should be compared. However, even if the model performs well in both training set and testing set, it cannot guarantee that the model will have the same stable performance in the future.

Taking interval 2 (supply side reform) and interval 5 (new corona epidemic) as reference points, from interval 1 to interval 3 and from interval 4 to interval 6, it can be seen that the occurrence of two major events increases the error ratio of each model. This phenomenon confirms the first assumption mentioned above. One possible explanation is that the occurrence of major events leads to the increase of market risk, which is manifested as the decrease of model prediction ability. By comparing interval 3 and interval 4, the impact of old major events on the market gradually decreases as the error ratio decreases. This phenomenon confirms the second assumption mentioned above. But as major new events occur, the error ratio expands again. Although the error ratio has fluctuation,

but if the interval 1, 3, 4, 6 is divided into a group and the interval 2, 5 is divided into a group, the overall error ratio increases gradually. This shows that as time goes by and major events influence the market, the overall forecasting ability of the model is declining. The following Table III MAE index descriptive statistics also reflects the same characteristics.

TABLE II.   DESCRIPTIVE STATISTICS OF MSE INDEX OF EACH MODEL

Interval number	1	2	3	4	5	6
LSTM						
Training MSE	0.15740	0.08780	0.08670	0.12440	0.11230	0.10930
Testing MSE	0.10410	0.13340	0.20610	0.16120	0.38290	0.41890
Ratio of error	0.66137	1.51936	2.37716	1.29582	3.40962	3.83257
GRU						
Training MSE	0.16062	0.08905	0.08879	0.11950	0.10798	0.10410
Testing MSE	0.10126	0.13755	0.21996	0.15596	0.35992	0.38120
Ratio of error	0.63043	1.54464	2.47731	1.30510	3.33321	3.66186
AutoGluon						
Training MSE	0.03370	0.03310	0.02400	0.03420	0.01480	0.04050
Testing MSE	0.24950	0.17170	0.19630	0.20220	0.35860	0.40490
Ratio of error	7.40356	5.18731	8.17917	5.91228	24.22973	9.99753

TABLE III.   DESCRIPTIVE STATISTICS OF MAE INDEX OF EACH MODEL

Interval number	1	2	3	4	5	6
LSTM						
Training MAE	0.30630	0.21950	0.18480	0.27670	0.25220	0.24100
Testing MAE	0.24460	0.25990	0.31830	0.31240	0.42840	0.49410
Ratio of error	0.79856	1.18405	1.72240	1.12902	1.69865	2.05021
GRU						
Training MAE	0.30367	0.21882	0.18551	0.26809	0.24083	0.23651
Testing MAE	0.24080	0.26103	0.32663	0.30724	0.42986	0.47335
Ratio of error	0.79297	1.19290	1.76071	1.14603	1.78491	2.00140
AutoGluon						
Training MAE	0.30367	0.21882	0.18551	0.26809	0.24083	0.23651
Testing MAE	0.24080	0.26103	0.32663	0.30724	0.42986	0.47335
Ratio of error	0.79297	1.19290	1.76071	1.14603	1.78491	2.00140

This paper use artificial neural network, LightGBM algorithm, XGBoost algorithm and CatBoost algorithm to generate AutoGluon framework sub-model for prediction. The dataset is divided into six intervals, and AutoGluon framework generates more than 30 sub-models in each interval. For simplification, this paper selects the interval before and after the outbreak of the new coronavirus (interval 5), and studies the top 10 performance sub-models in the training set and the testing set, respectively. The evaluation index is MAE as follows. It should be noted that the research conclusions of other intervals and MSE are basically consistent with this.

TABLE IV.   COMPARISON OF AUTOGLUON SUBMODEL (INTERVAL 5)

Model ranking	Training set	MAE	Testing set	MAE
1	WeightedEnsemble _L3	0.084 51	XGBoost_BAG_L 1/T1	0.365 06
2	LightGBM_BAG_ L2/T4	0.089 70	XGBoost_BAG_L 1/T4	0.365 29
3	LightGBM_BAG_ L2/T1	0.091 08	XGBoost_BAG_L 1/T2	0.366 55
4	LightGBM_BAG_ L2/T3	0.091 37	XGBoost_BAG_L 1/T3	0.374 48
5	LightGBM_BAG_ L2/T2	0.091 96	CatBoost_BAG_L1 /T0	0.374 94
6	CatBoost_BAG_L2 /T2	0.092 03	XGBoost_BAG_L 1/T0	0.376 67
7	CatBoost_BAG_L2 /T0	0.092 18	LightGBM_BAG_ L1/T2	0.380 39
8	CatBoost_BAG_L2 /T1	0.092 74	LightGBM_BAG_ L1/T1	0.381 74
9	WeightedEnsemble _L2	0.093 06	LightGBM_BAG_ L1/T0	0.384 50
10	LightGBM_BAG_ L2/T0	0.094 74	LightGBM_BAG_ L1/T4	0.385 54

where 'L' represents the number of stacking layers, 'T' represents the parameter search times and 'BAG' represents the use of bagging strategy.

In the training set of six intervals, the prediction accuracy of the model is the highest, whether measured by MAE or MSE. However, in the testing set of each interval, Weighted Ensemble _ L3 performs quite backward, which may be due to overfitting in the training process of weighted combination model, which also shows that the generalization ability of weighted combination model is weak. In the testing set, it is found that several models trained by XGBoost algorithm perform well. It is also worth noting that the same model performs poorly in the training set and does not enter the top ten.

Based on the characteristics of the AutoGluon framework, the more stacking layers and parameter search times, the more complex the sub-model trained will be. However, more complex models do not necessarily achieve better prediction results, which is quite obvious on several sub-models generated by XGBoost algorithm in Table IV testing set.

In general, from the perspective of model prediction accuracy, AutoGluon framework is generally better than LSTM model and GRU model, especially in the training set. There is no significant difference between LSTM model and GRU model. It is particularly noteworthy that the testing/ training error ratio of AutoGluon framework is much larger than that of LSTM model and GRU model, which indicates that AutoGluon framework has certain over-fitting phenomenon.But this does not affect the conclusion that AutoGluon framework has stronger overall prediction performance. At the same time, it is noteworthy that the testing/training error ratios of LSTM model and GRU model are smaller than those of AutoGluon framework, indicating that the prediction performance of LSTM model and GRU model is more stable.

V. CONCLUSION

In this paper, it is found that: (1) the impact of major events increases the difficulty of futures market prediction. At the same time, with the passage of time, it is more difficult to accurately predict the market through a single model, which is verified by comparing the change of interval error ratio before and after the event. (2) Although over-fitting phenomenon exist, the prediction accuracy of AutoGluon framework which consumes more resources is generally better than LSTM model and GRU model, but the overall performance difference between LSTM model and GRU model is trivial. (3) It may be meaningful to compare model performance only on specific datasets or tasks. By consuming more resources to train more complex weighted combination models, it is not certain to achieve better prediction results in specific tasks, while simple models are not necessarily inferior to complex models. The diversity of specific tasks and the ease of use of AutoGluon framework will make AutoGluon framework based on automatic machine learning have greater advantages over traditional machine learning methods in the future.

Based on these findings, such following suggestions are put forward: (1) In addition to MSE or MAE, ratio of error may be more suitable to measure the model prediction ability. (2) In order to improve the performance of time series prediction task model, XGBoost algorithm is worth being studied in the future.

REFERENCES

[1]   LI Hong-quan. Application of interval econometrics for crude oil price forecasting[J]. Computer Engineering and Applications, 2010, 46(6): 23-25.

[2]   Zhang Y J, Yao T, He L Y, et al. Volatility forecasting of crude oil market: Can the regime switching GARCH model beat the single-regime GARCH models[J]. International Review of Economics Finance, 2019, 59: 302-31.

[3]   LI Hongquan, Liang ZHOU.Systemic Financial Risk Measurement and Its Economic Forecasting Ability[J].China Journal of Econometrics, 2021,1(4): 892-903.

[4]   Hong Yongmiao,Wang Shouyang.How Is Big Data Changing Economic Research Paradigms?[J].Management World.2021,37(10).

[5]   Chen Y, He K, Tso G K F. Forecasting crude oil prices: A deep learning based model[J]. Procedia Computer Science, 2017, 122: 300–307.

[6]   R.A.de Oliveira,D.M.Q.Nelson,A.C.M.Pereira. Stock markets price movement prediction with lstm neural networks[C]. //In 2017 International Joint Conference on Neural Networks (IJCNN), pages 1419–1426, 2017.

[7]   MU Nianguo, YAO Honggang.Application of attention mechanism based recurrent neural network in financial time series[J].Modern Electronics Technique，2021，44（14）: 1-5.

[8] Zhang Xu,Xue Lei,Li Dun-Yu.Stock Analysis System Based on Financial TimeSeries and Public Opinion Analysis[J]. Industrial Control Computer,2018,31(04).

[9] CHEN Jia, LIU Dongxue, WU Dashuo. Stock Index Forecasting Method Based on Feature Selection and LSTM Model[J]. Computer Engineering and Applications, 2019, 55(6): 108-112.

[10] PENG Yan, LIU Yuhong, ZHANG Rongfen. Modeling and Analysis of Stock Price Forecast Based on LSTM[J]. Computer Engineering and Applications, 2019, 55(11): 209-212.

[11] WANG Yan, GUO Yuankai. Application of Improved XGBoost Model in Stock Forecasting[J]. Computer Engineering and Applications, 2019, 55(20): 202-207.

[12] SONG G，ZHANG Y F，BAO F X，et al．Stock prediction model based on particle swarm optimization LSTM［J］．Journal of Beijing University of Aeronautics and Astronautics，2019，45(12）: 2533-2542.

[13] GE Dao-Hui，LI Hong-Sheng，ZHANG Liang,LIU Ru-Yi,SHEN Pei-Yi, MIAO Qi-Guang. Survey of Lightweight Neural Network[J],Journal of Software, 2020, 31(9):2627-2653.

[14] XU Haoran, XU Bo, XU Kewen. Analysis on Application of Machine Learning in Stock Forecasting[J]. Computer Engineering and Applications, 2020, 56(12): 19-24.

[15] DANG Jianwu, CONG Xiaoqing. Research on Hybrid Stock Index Forecasting Model Based on CNN and GRU[J]. Computer Engineering and Applications, 2021, 57(16): 167-174.

[16] ZHAO Hongrui, XUE Lei. Research on Stock Forecasting Based on LSTM-CNN-CBAM Model[J]. Computer Engineering and Applications, 2021, 57(3): 203-207.

[17] Wang Y, Liu L, Wu C. Forecasting the real prices of crude oil using forecast combinations over time-varying parameter models[J]. Energy Economics, 2017, 66: 337–348.

[18] SUN Fuxiong, LIU Guangming, ZENG Zixuan, PENG Mengqi. Research on Stock Suspension Prediction Based on Combination Model[J]. Computer Engineering and Applications, 2020, 56(18): 272-278.

[19] ZHOUHao, ZHANGYifei,WANG Zhen,WANG Jue,WANG Shouyang. An improved research on optimization strategy of crude oil price forecast combination[J].System Engineering-Theory &Practice,2021,41 (10): 2660-2668.

[20] Nick Erickson, Jonas Mueller, Alexander Shirkov, Hang Zhang,PedroLarroy,MuLi,AlexanderSmola.AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data[C]. //7th ICML Workshop on Automated Machine Learning (2020) Vienna.