Renewable energy company stock dynamics forecast in the period of sustainable development based on Fractal-FOA-LSTM

Guofeng Ma*, Yingjie Wang, and Junhong Yang
Zhengzhou Railway vocational & technical college, 450000 Zhengzhou, China.

Abstract. In stock trend forecasting system, feature selection and model building are two major factors that affect prediction performance. In order to improve the accuracy of prediction and the stability of the model, a stock trend prediction model of Fractal-FOA-LSTM is proposed. Firstly, the features are selected by using the FOA (fruit fly algorithm) combined with the fractal dimension to reduce the redundancy of the features, and the selected indexes are used as the system input. And proposing a double input LSTM (long-short term memory) network prediction model and optimizing its parameters, it can select the best parameters for different data automatically. This paper test on 4 sets of UCI database and Shanghai Composite Index and proved the feature selection method is effective, through the empirical analysis of the Shanghai Composite Index and S&P500, and compared the results with SVM and PNN, verified the feasibility and superiority of the stock trend forecasting system base on fractal-FOA-LSTM.

1 Introduction

A large number of experimental studies show that the stock market is nonlinear, non-stationary and highly noisy[9,6]. In prediction system, feature selection and model building are two major factors that affect prediction performance. In the study of feature selection, literature[8] proposed the stock prediction model of CFS and fractal feature selection algorithm respectively, which proved the validity of the fractal feature selection, but it was not combined with the fruit fly algorithm in the swarm intelligence algorithm.

In the study of prediction model, literature[12] constructed the PCA-ANN stock price prediction model, but the prediction training speed is slow and easy to fall into the local optimal, and the selection of the number of hidden layer nodes greatly affects the prediction accuracy. The literature[5] combined the Boruta algorithm feature selection with the PSO optimization SVM to classify the return rate, and the literature[4] used SVM combines the feature selection algorithm to predict the stock trend. Deep learning is a new field in the study of machine learning, Chen K and others verify the validity of LSTM in China's stock market prediction[3], and Akita R uses the basic analysis and LSTM model to predict the stock market[2]. The parameter of the prediction models has a great influence on the

* Corresponding author: 13096743@qq.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
prediction effect, and the best parameters is different for different input data. The LSTM prediction model involves more parameters, and the choice of parameters has a greater impact on the prediction effect.

In order to improve the accuracy of stock trend prediction, combined with the above analysis, the fast optimization ability of FOA is combined with fractal dimension, and the parameters of double input LSTM are optimized by the global optimization ability of the FOA, and a prediction model combine the deep learning and fruit fly algorithm is designed. Through the simulation and empirical analysis, proved the prediction system enables to select better feature sets and model parameters for different data. Due to the improvement of feature selection and model construction, the accuracy of prediction has been greatly improved.

2 The Prediction Model

2.1 The LSTM

The deep learning was put forward by Hinton and others in 2006. It is a new field in machine learning. It uses different training mechanisms from traditional artificial neural networks to overcome the problem of traditional neural network[7,1]. Deep learning has been widely used in various fields and achieved good results[10,14]. Recurrent neural network is an effective network for analyzing time series. The LSTM (long and short term memory) network effectively solves the problem of gradient disappearance or gradient explosion of ordinary recurrent neural network. Fig.1 is structure diagram of LSTM, which describes the state of memory cell at time \( t \) and the recursive connection at time \( t-1, t+1 \).

![Fig.1. Structure diagram of LSTM](image)

As shown in the Fig.1, the enlarged part is the structure of a single memory cell. \( x_t \) is the input at time \( t \), and the \( W_{ij} \) is a weight matrix from \( i \) to \( j \), \( b \) is bias, \( \sigma \) and \( \text{tanh} \) respectively is sigmoid function and hyperbolic tangent function. The memory cell state vector, the input vector and hidden layer vector propagates along the blue line, the black line and the yellow line respectively. Then illustrating the structure combine with the formula.

At the time \( t-1 \), there are two statuses: last memory unit status \( s_{t-1} \) and the last hidden status \( h_{t-1} \), and the information that is not important in reverse transmission is forgot by forget gate through...
\[ f_t = \sigma(W_{sf}x_t + W_{hf}h_{t-1} + b_f) \]  

where \( \sigma \) is sigmoid function, \( W_{sf} \) is weight matrix from input \( x \) to the forget gate \( f_t \), \( x_t \) is the input at time \( t \), \( W_{hf} \) is weight matrix from the last hidden status \( h_{t-1} \) to the forget gate \( f_t \), \( b_f \) is the linear bias of forget gate.

To be updated information \( \tilde{s}_t \) is determined by the input gate \( i_t \), which is added to the memory cell at time \( t \) and output from a fully connected tanh layer according \( x_t \) and \( h_{t-1} \), the formula is:

\[ i_t = \sigma(W_{si}x_t + W_{hi}h_{t-1} + b_i) \]  

\[ \tilde{s}_t = \tanh(W_{sx}x_t + W_{hx}h_{t-1} + b_s) \]  

The current state is combining the past memory with new memory, the past memory is the last memory unit status \( s_{t-1} \) and the new memory is the To be updated information \( \tilde{s}_t \), the formula is as follows:

\[ s_t = i_t \ast \tilde{s}_t + f_t \ast s_{t-1} \]  

The output \( o_t \) is determined by current input status \( x_t \) and the last hidden status \( h_{t-1} \), and the current hidden status \( h_t \) is determined by \( o_t \), the \( o_t \) and \( h_t \) is calculated by:

\[ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \]  

\[ h_t = o_t \ast \tanh(s_t) \]

Each layer of neurons of LSTM has three "gates," so that some of the errors can propagate through the "gate" to the next without depending on the current neuron, which makes no matter how far the gradient propagates will not disappear, so the convergence is good.

### 2.2 Feature Selection Based on Fractal Dimension and FOA

The traditional FOA is aimed at the optimization of continuous space, and the flight mode is unrestricted in the optimization process. However, it is not suitable for the optimization of discrete space. Therefore, this paper improve the FOA and combine with fractal dimension for feature selection, which improves the efficiency of feature selection. The following steps are mainly designed:

- **Flight mode**

The feature selection is a discrete problem, this paper redefined a flight mode for this problem. Assuming the current position of fruit fly \( i \) is \( x_i \), the distance flying forward
is $r = \left[ R \times \text{rand} \right]$, and $\text{rand} = [-1, 1]$, the $R$ is the group radius, and $\text{rand}$ is the maximum distance of a flight, then the position coordinates update to $(x_i + r) \% N$, $N$ is the total number of features, $\%$ is the remainder of the operation.

- Search stopping condition
  The number of features to be selected is the upper bound of fractal dimension. Drosophila does not have to reach every feature node in the process of optimization, as long as the number of features to be searched reaches the fractal dimension of the upper bound, that is, the number of features, the search process can be stopped to get a subset of features.

- The fitness
  The connection between FOA and feature selection is based on the fractal dimension through the fitness function. The fitness function is to evaluate the quality of the subset of features, which determines the next round of fruit flies search. The closer the fractal dimension of feature subset is to the fractal dimension of the original dataset, the better the subsets of features are. So the fitness of this paper is $f = \left( \sum_{k=1}^{m} \text{frac}(k) - D \right) / m$, $D$ is the fractal dimension of the original data, $\text{frac}(k)$ is the fractal dimension of the feature, and the process of searching is to find the minimum value of $f$.

- Variation operation
  In order to prevent the flies from getting into the local optimum in the optimization process, the introduction of mutation operation in this paper. Using non-uniform mutation, according to the mutation probability $p_0$ to determine whether mutation, the all mutation individual variation of each dimension according the formula is as follows:

$$x_i = x_i \times (\text{rand} \times (1 - t / T))^b$$

where $t$ is number of the current iterations, $T$ is the maximum number of iterations, $b$ is the parameters of degree of non-uniformity of variation, which general value is from 2 to 5, the smaller value, the greater the range of variation, this paper takes 2 to make the magnitude of variation larger.

Through the above analysis, the steps of the feature selection algorithm based on fractal dimension and FOA combination of are as follows:
- Step1. Find the original sequence fractal dimension $D$, determine the number of features is $m = \lceil D \rceil$.
- Step2. Give the population size $\text{SizePop}$, the maximum number of iterations $\text{MaxGen}$, randomly initialize the departure feature position, that is, the initial population of drosophila position $x_i(t)$;
- Step3. According to the flight direction and distance to determine the flight, select the property, if the number of choices is less than $m$, then continue to choose;
- Step4. The fitness value is calculated according to the fitness evaluation function $f$;
- Step5. Find the best taste of drosophila, its best taste concentration $\text{bestSmell}$ corresponding to the best subset of the fractal dimension;
- Step6. Save the best flavor concentration $\text{bestSmell}$ and corresponding subset encoding $x_i(t)$;
\begin{equation}
\text{Smellbest} = \text{bestSmell} \\
x_i(t) = x(\text{bestindex}) \tag{8}
\end{equation}

- Step7. According to the mutation probability $p_0$ to determine whether the variation, non-uniform variation is adopted, and if the fitness of the subset corresponding to the mutation is better, and the subset mutation is determined as the optimal subset;
- Step8. Save the best flavor concentration and corresponding subset encoding, fruit fly flew to this location visually;
- Step9. Repeat steps 2-8, when the number of iterations reaches the maximum number of iterations or the optimal solution does not change within a certain number of generations, the process is over.

2.3 Improved double-input LSTM model

The Keras is the most convenient Python library for constructing LSTM, which complied with the fully connected layer (Dense layer), active layer, LSTM layer, Dropout layer and optimizer module. The traditional LSTM time series prediction is an input and an output, the method of feature processing will make the model lost some important information during the prediction process. So a two input and one output LSTM prediction model can be constructed, which can input another indicator like price information to prevent information loss. We called this structure as double-input LSTM prediction model, that is to add an additional indicator in the LSTM layer, take the LSTM layer output and the indicator together as the input of the back of the Dense layer, the structure as follows:

![Fig.2. The structure of double-input LSTM model](image_url)

The parameters of double-input LSTM prediction model include weight regular terms, epoch and batch and so on. Weight regularization is an important method to prevent over fitting, a epoch means that all samples of the training set are involved in a round of training, taking batch as the preface, the size of batch determines the gradient value and the frequency of weight update. These parameter are selected by people. In order to analyze the influence of each parameter on prediction model, take the data from 2005.01.01 to 2010.07.19 of Shanghai composite index to predict the stock trend, take 20 numbers dropout_W and dropout_U increased by 0.02 from 0.02 to 0.4 respectively, take 20 numbers epoch increased by 100 from 100 to 2000, and take 20 numbers batch size increased by 1 from 1 to 20, when take one parameter, the others unchanged, the result of the prediction accuracy under each parameter as shown in Fig.3.
As shown in Fig. 3, epoch has the greatest impact on the accuracy of stock trend prediction. Dropout_W, dropout_U and batch also have some effect on the prediction accuracy. Therefore, in order to get effective parameters, optimizing the parameters by FOA to get the best parameters, the following steps are designed:

• Flight mode
Since the batch and epoch are integers, the flight mode is designed to be \((x_i + r)\%N\), the current position of fruit fly \(i\) is \(x_i\), the distance flying forward is \(r = \left\lfloor R \times \text{rand} \right\rfloor\), and \(\text{rand} = [-1,1]\), the R is cardinal number of the parameter. In this article, the value of batch should be between 1-20, so the N is 20; the value of epoch is at 100-4000, the flight radius is not too large, the N is 100 and then multiply10; the values of the dropout_W and dropout_U are generally between 0-0.5, thus N is 5, and then multiply 0.1.

• Search stopping condition
When the number of iterations is reached, or when the best flavor concentration Smellbest<0.1, jump out of the loop iteration and search stop.

• The fitness
Take the error of LSTM prediction S as the evaluation function. The process of fruit fly search is to find the minimum value of S.

Based on the above analysis, the use FOA to optimize the double-input LSTM parameters as follows:

- Step1. Randomly initialize the starting position of each parameter \(x_{ij}(t)\), where

  \[ j = 1, 2, 3, 4 \]

  and is four parameters.

- Step2. According to the flight direction and distance to determine the flight, calculate the concentration of flavor and the determine the parameters.

- Step3. Calculate the fitness value S according to fitness evaluation function

- Step4. Find fruit flies with best taste concentration, and the best value.;

- Step5. Preserve the best flavor concentration \(\text{bestSmell}\) and the best location for each column \(x_{ij}(t)\) through the following equations:
Smellbest = bestSmell
\[ x_j(t) = x_{bestindex}(t) \]  \hspace{1cm} (9)

- Step6. Update the best and corresponding locations, drosophila groups fly to the site with vision.

### 3 Feature selection analysis

In order to verify the validity of the feature selection algorithm based on fractal dimension and FOA, this paper test on the tree, breast-cancer, waveform and segmentation data sets of UCI. The basic information of the data set and the number of to be selected features are shown in Table1. The features selected by the method of proved feature selection and the classification results are shown in Table2.

**Table1.** The description of data set

| data set     | samples | features | Fractal dimension | Need to select the number of features |
|--------------|---------|----------|-------------------|--------------------------------------|
| tree         | 700     | 18       | 2.4215            | 3                                    |
| breast-cancer| 683     | 9        | 2.4731            | 3                                    |
| waveform     | 5000    | 21       | 6.3273            | 7                                    |
| segmentation | 2100    | 19       | 3.7825            | 4                                    |

**Table2.** The feature selection and classification results of data set

| data set     | features selected | fractal dimension of Subset | classification accuracy before feature selection | classification accuracy after feature selection |
|--------------|-------------------|-----------------------------|-----------------------------------------------|-----------------------------------------------|
| tree         | 4,6,17            | 2.3427                      | 80.67%                                        | 81.78%                                        |
| breast-cancer| 1,5,7             | 1.8851                      | 90.45%                                        | 92.21%                                        |
| waveform     | 5,6,10,12,13,14,17| 4.2738                      | 75.63%                                        | 73.22%                                        |
| segmentation | 1,2,16,17         | 3.2162                      | 86.15%                                        | 94.88%                                        |

In Table1, the fractal dimension of each data set is calculated and the number of features to be selected is determined. In Table2, the population size is initialized to Sizepop = 20, the maximum number of iterations is initialized to Maxgen = 50, through the comparison of the results, the dimension of data set after feature selection is greatly reduced, the classification accuracy is only one data set decreased, and the other three are improved, especially the segmentation data set with obvious fractal characteristics, the classification accuracy is greatly improved after the feature selection.

Literature[8] compared different feature selection algorithms for SSCI, and it take the 20 technical indexes of SSCI in 2000 -2008 to select features. On this basis, taking the same data and use the fractal combined with FOA proposed by this paper to predict the stock trend, and the results as shown in Table3.

As shown in Table3, the fractal combined with FOA method select index VR, MTM, RSI, MFI, though the number of features is not the least, the classification accuracy is highest. The UCI database analysis and the empirical analysis of SSCI prove the fractal combined with FOA feature selection method is effective.
Table 3. The feature selection and classification results of 2000-2008 SSCI

| feature selection algorithm | selected features | classification accuracy |
|-----------------------------|-------------------|-------------------------|
| Cfs                         | MFI               | 61.984%                 |
| OneR                        | VR, PDI, RSI, ROC | 64.463%                 |
| Information Gain            | MTM, MFI          | 52.479%                 |
| Fractal-FOA                 | VR, MTM, RSI, MFI | 64.625%                 |
| None                        | All features      | 62.810%                 |

4 Empirical analysis

In order to verify the effectiveness of the prediction model proposed in this paper, SVM, PNN, LSTM and this paper are used to empirical analysis on the Shanghai Composite Index (SSCI) and the S&P 500 index (S&P500) from 2005.01.01 to 2015.05.01. The prediction results of each model are as shown in Table 4.

Table 4. The results comparison of prediction model

| Predict model                | SSCI  | S&P500 |
|------------------------------|-------|--------|
| SVM                          | 62    | 55     |
| SVM+selected features        | 65.07 | 57.67  |
| PNN                          | 60.33 | 56     |
| PNN+selected features        | 62.5  | 58     |
| LSTM                         | 62.17 | 55.46  |
| Fractal-FOA-LSTM             | 66.26 | 59.17  |

As shown in Table 4, the accuracy of all models using selected features is higher than using all features with each model, and prove the validity of feature selection method with fractal and FOA proposed in this paper. The highest prediction accuracy is this paper method, and it respectively higher than LSTM model 4.09% and 3.71% for SSCI and S&P500, and prove the validity and accuracy of this paper prediction model in stock trend prediction. Although in the SSCI prediction, this paper algorithm is 1.19% higher than the second good SVM combined with the feature selection model, in the S&P500 prediction, this paper algorithm is 1.17% higher than the second better PNN combined with the feature selection model, the slight improvement in prediction performance is also important for the exploration and research of stock trend prediction.

5 Conclusion

Stock market is a dynamic and nonlinear complex system. There are a lot of research on stock trend forecasting. The features selection and the model construction are the two factors that affect the accuracy of the forecast. There are many parameters involved in the construction of long and short time memory network model. The selection of parameters has a great influence on the accuracy of the prediction, and the optimal parameters for the model are different for the different input data.

Therefore, in order to improve the prediction accuracy and the model stability, this paper combines the fractal dimension and the fruit fly algorithm to select the best features from characteristics technical, and uses the fruit fly algorithm to optimize the parameters of the double-input LSTM model, and the optimal parameters can be selected automatically for
different data to get a suitable prediction model. This paper test on 4 sets of UCI database and Shanghai Composite Index proved the feature selection method is effective, through the empirical analysis of the Shanghai Composite Index and S&P500 verified the feasibility of this model, and through the experiments compared with SVM and PNN verified the superiority of fractal-FOA-LSTM model.

References
1. C. Affonso, A. L. D. Rossi, F. H. A. Vieira, Expert Systems with Applications, 85 (2017)
2. R. Akita, A. Yoshihara, T. Matsubara, International Conference on Computer and Information Science, 1 (2016)
3. K. Chen, Y. Zhou, F. Dai, IEEE International Conference on Big Data, 2823 (2015)
4. H Grigoryan, International Conference on Applied Mathematics, Modelling and Statistics Application, 5 (2017).
5. H S Guo, B Y Gao, H J Lu., Sensor and Micro System, 3 (2018)
6. S. Lahmiri, Expert Systems with Applications, 55, 268 (2016)
7. Y. Lecun, Y. Bengio, G. Hinton, “ Deep Learning,” Nature, vol.521, no.7553, pp.436-444, 2015.
8. L. P. Ni, Z. W. Ni, Y. Z. Gao, Expert Systems with Applications, 38 (5), 559 (2011)
9. J. Patel, S. Shah, P. Thakkar, Expert Systems with Applications, 42(1), 259 (2015)
10. S. Sun, B. Zhang, L. Xie, Neurocomputing, 8811996 (2017)
11. X. Sun, Z. Wu, Z. Huang, China University of Science and Technology Press, 33 (2003)
12. J. Wang, J. Wang, Neurocomputing, 156, 68 (2015)
13. X. Wu, Q. Li, Firepower and Command & Control, 38(4), 17 (20130
14. S. Yu, Y. Wu, W. Li, Neurocomputing, 4, 438 (2017)