Application on Development of Modern Service Industry in China Based on Artificial Intelligent Model Optimized by TANSAFOA

Tian Wang*, Jianbang Lin*

1Management School, South China Business College Guangdong University of Foreign Studies, Guangzhou, China
2Business School, Nanfang College of Sun Yat-sen University, Guangzhou, China
*Corresponding author’s e-mail: bangnfxysu@163.com

Abstract. Modern service industry is one of the industries that widely applies artificial intelligent (AI) in the era of big data. The development of modern service industry plays a very important role in industrial transformation. The emergence of AI technology improves the effectiveness of data mining, and more and more scholars suggest a series of optimization methods to lift the prediction ability. In this study, tangent fruit fly optimization algorithm with step adjust (TANSAFOA) is proposed and it is used to optimize multivariate adaptive regression splines (MARS) and back propagation neural network (BPNN) to construct a prediction model of business performance. The result shows TANSAFOA can effectively optimize the prediction model and the BPNN model optimized by TANSAFOA has higher prediction performance than MARS. TANSAFOA BPNN-model is the most appropriate prediction model for modern service industry in China.

1. Introduction
In big data era, the modern information technology of artificial intelligent (AI) has been applied in all walks of life and modern service industry is one of them. In 2017, the 13th Five-year Plan for Modern Service Industry Science and Technology Innovation pointed out that the industrial structure is transforming from industry to service industry in China under the situation of increasingly fierce international competition and the development of modern service industry plays a very important role in industrial transformation.

In recent years, the proportion of China's service industry has grown rapidly. In 2019, the added value of the service sector accounts for 53.9% of the GDP, and it is estimated that the added value of service industry will increase to 60% of GDP by 2025 [1]. Despite the rapid development of China's modern service industry, if the performance and resources of enterprises are not well predicted and allocated, enterprises are prone to loss.

The emergence of AI technology improves the effectiveness of data mining, and more and more scholars suggest a series of optimization methods to lift the prediction ability. In recent years, swarm intelligence optimization algorithms have become popular methods which are widely discussed in various fields. As mentioned above, this study suggests a novel improved swarm intelligence optimization algorithm, i.e. tangent fruit fly optimization algorithm with step adjust (TANSAFOA), to correct the problems of basic FOA. TANSAFOA is used to optimize both back propagation neural network (BPNN) and multivariate adaptive regression splines (MARS), which are widely applied to...
assess the business performance. To achieve the goals of this study, first, the important indicators of performance prediction of modern service industry are selected. And then, the prediction ability of TANSAFOA-model is compared to FOA-model and non-optimized prediction model. Finally, the effective performance prediction model of modern service industry is confirmed in China.

2. Method

2.1. Fruit fly optimization algorithm (FOA)
One of the most popular swarm intelligence optimization algorithms is FOA which is proposed by pan in 2011 [2]. The following steps are introductions of the FOA based on the process of fruit fly’s searching for food [3]:

1) Giving the the random direction and distance to initial random fruit fly groups.

\[
X_i = \text{Init}X_{\text{axis}} + \text{Random Value} \tag{1}
\]

\[
Y_i = \text{Init}Y_{\text{axis}} + \text{Random Value} \tag{2}
\]

2) Calculating the distance (Disti) of every initial fly group and the smell concentration value (Si), which is the reciprocal of distance.

\[
\text{Dist}_i = \sqrt{X_i^2 + Y_i^2} \tag{3}
\]

\[
S_i = \frac{1}{\text{Dist}_i} \tag{4}
\]

3) Transforming the smell concentration value (Si) into the smell value (Smelli) by using a smell function (Fitness function).

\[
\text{Smell}_i = \text{Function}(S_i) \tag{5}
\]

4) Finding out the fruit fly at the smallest Smell, of all groups. Keeping the best smell concentration values and its x, y coordinates of that location.

\[
[\text{best Smell best Index}] = \min(\text{Smell}_i) \tag{6}
\]

\[
X_{\text{axis}} = X(\text{best Index}) \tag{7}
\]

\[
Y_{\text{axis}} = Y(\text{best Index}) \tag{8}
\]

\[
\text{Smell best} = \text{best Smell} \tag{9}
\]

There are some advantages of basic FOA, such as easy, simple, accurate and applicable and many scholars have widely applied FOA on various fields. For example, Pan (2011) and Pan et al. (2018) used FOA on financial field [3~4], Pan and Shi (2017) forecasted short-term traffic [5], Zhong et al. (2017) built electricity price model [6], Han et al. (2018) constructed the prediction model for China agricultural output value [7], Wu et al. (2019) solved the engineering optimization problems [8].

2.2. Tangent fruit fly optimization algorithm with step adjust (TANSAFOA)
Though FOA has widely adopted on different fields, most scholars consider that FOA have the problems of non-negative smell concentration value and local extreme value. It is unable to solve the optimization problem involving negative domain due to the non-negative smell concentration value [9]. Moreover, FOA is often prone to a local extreme because the smell concentration value is usually small [10]. So, a novel TANSAFOA is suggested in this study to solve the limitations of FOA.

Tangent function exists the range value of real numbers, and it improves the problem of non-negative smell concentration value of FOA by tangent transformation. On the other hand, the appearance of extreme is influenced by searching step size and the searching step size is adjusted as following. It can prevent from the problem of easily falling into local extreme value by increasing and
decreasing searching step. Therefore, the adjustment details of TANSAFOA are described as followings.

First, TANSAFOA begins to randomize the fruit fly groups and add the adjustment factor \( t \) to change the searching space.

\[
X_i = \text{Init}_X\_axis + (\text{Random Value}) \times t
\]

\[
Y_i = \text{Init}_Y\_axis + (\text{Random Value}) \times t
\]

Next, figuring out the distance of every initial fruit fly group (see eq. (3)-(4)). And then, it is transformed by tangent function to form a new smell concentration value \( S_i \) from FOA.

\[
S_i = \tan(Dist_i)
\]

Finally, according to the different prediction models which are used in this study, TANSAFOA optimizes parameters of different models (see section C) and follows the eq. (6)-(7) to find the best prediction result.

2.3. Prediction models

To achieve the aim of this study, AI models (i.e. MARS and BPNN) are used to predict the performance of modern service industry.

MARS is a nonlinear and non-parametric regression model first proposed by Friedman in 1991 [11]. The MARS splits input data into different intervals, and then builds its own linear regression model. Furthermore, a regression equation is individually built to fit the different region data in the independent variable space. Therefore, there are different linear regression lines separated by several knots and finally are combined as MARS model. So, it can also suitable for financial data with high-dimensional problems. The MARS model is shown as below:

\[
\hat{y} = a_0 + \sum_{m=1}^{M} a_m + B_m(x)
\]

where \( \hat{y} \) is the dependent variable, \( a_0 \) is the constant, \( B_m(x) \) is the \( m \)th basis function, \( a_m \) is the coefficient of the \( m \)th basis function and \( M \) is the number of basis functions.

Minspan and Endsapn of MARS are used to split the input data into separate regions. During the modeling process, Minspan is the minimum step size which is set to reduce the local variance, and Endsapn is used to set the both sides of the smallest place node distance. This study suggests TANSAFOA to find these two parameters which can lift the global search ability and gain better prediction performance.

The BPNN algorithm for training multilayer neural network is proposed by Rumelhart et al. [12]. There are multilayer neurons such as input layer, hidden layer and output layer. \([X_1, X_2, \ldots, X_n]\) is the input of the input layer, \([\Sigma, f]\) are both an adding function and an activation function in the hidden layer, and \(Y\) is the output. \(w_{ij}\) and \(w_j\) are the weight values of neural network. The neuron numbers of input layer are the numbers of input variable, and the neuron numbers of output layer are the numbers of output variable. The hidden layer can be one or more layers to solve the complex non-linear relationship between variables.

The weights and thresholds of BPNN are the key factors which influence the prediction performance. In this study, the prediction performance can be improved by finding the best weights and thresholds of BPNN through TANSAFOA. In this study, there are 8 or 9 inputs (see Table I.), one hidden layer with 6 neurons, and 1 output.

3. Empirical study

3.1. Data and variables

According to the guidance in 13th Five-year Plan for Modern Service Industry Science and Technology Innovation, one of the applications of modern service industry is culture and travel. The qualified listed companies whose business including culture and travel in Wind database from 2016 to
2018 are selected as research samples in this study. Meanwhile, financial indicators which are suggested to affect business performance by literature are selected from the database as the candidate list of input variables. The 3 different indexes, return on equity (ROE), return on assets (ROA), earnings per share (EPS) [13], which are widely used to evaluate business performance are further selected as the output variables of prediction model in this study. Thus, the effectiveness of TANSAFOA can be examined in different indexes.

Before constructing the business performance prediction model of modern service industry, the selected sample data is cleaned and normalized first to reduce the inference of different dimensions. Next, the influential variables are found by using multiple regression from candidate list of input variables in this study. With variables selection, it not only effectively simplifies the complexity of model but reaches the good prediction accuracy. There are respectively 8 or 9 influential variables in different 3 models (see Table 1).

| Variable                                    | P-value |      |      | Variable                                    | P-value |      |      |
|---------------------------------------------|---------|------|------|---------------------------------------------|---------|------|------|
| Net profit margin on sales                  | 0.001   | 0.000| 0.000| Cash ratio                                  | 0.034   |      |      |
| Operating cost ratio                        | 0.000   | 0.003| 0.000| Current assets ratio                        | 0.000   |      |      |
| Cash recovery rate for total assets         | 0.008   | 0.000| 0.020| Operating cash flow ratio                   | 0.024   |      |      |
| Current assets turnover                     | 0.034   | 0.039|      | Operating revenue per share                 | 0.000   |      |      |
| Equity ratio                                | -       | 0.000| 0.000| Administrative expenses ratio               | -       | 0.029|      |
| EBIT per share                              | -       | 0.000| 0.000| Current ratio                               | -       | 0.009|      |
| Book value per share                        | -       | 0.010| 0.004| Net cash flows from operating activities ratio| -       | 0.013|      |
| Total Assets Turnover                       | 0.000   |      |      | Tangible assets ratio                       | -       | 0.001|      |

Note: “-” means the variable is not included in the model.

3.2. Accuracy of prediction model

Different models by MARS and BPNN are constructed to predict the business performance of modern service industry. Meanwhile, to prove the optimization effectiveness of TANSAFOA, the prediction models are further optimized by FOA and TANSAFOA. The population size and iteration number of FOA and TANSAFOA are both set to 5 and 1000 in this study.

| Model | Index | Optimization | Training | Testing | Model | Index | Optimization | Training | Testing |
|-------|-------|--------------|----------|---------|-------|-------|--------------|----------|---------|
| EPS   | Non-optimized | FOA          | 0.351   | 0.626   | EPS   | Non-optimized | FOA          | 0.936   | 1.062   |
| MARS  | TANSAFOA     |              | 0.323   | 0.596   | TANSAFOA |              | 0.381   | 0.471   |
| ROE   | Non-optimized | FOA          | 0.190   | 0.466   | ROE   | Non-optimized | FOA          | 0.337   | 0.459   |
|       | TANSAFOA     |              | 0.137   | 0.440   | TANSAFOA |              | 0.277   | 0.308   |
| ROA   | Non-optimized | FOA          | 0.293   | 0.656   | ROA   | Non-optimized | FOA          | 1.523   | 1.382   |
|       | TANSAFOA     |              | 0.235   | 0.604   | TANSAFOA |              | 0.499   | 0.488   |

| Model | Index | Optimization | Training | Testing | Model | Index | Optimization | Training | Testing |
|-------|-------|--------------|----------|---------|-------|-------|--------------|----------|---------|
| EPS   | Non-optimized | FOA          | 0.354   | 0.615   | EPS   | Non-optimized | FOA          | 0.311   | 0.510   |
| MARS  | TANSAFOA     |              | 0.232   | 0.596   | TANSAFOA |              | 0.381   | 0.471   |
| ROE   | Non-optimized | FOA          | 0.077   | 0.683   | TANSAFOA |              | 0.340   | 0.329   |
|       | TANSAFOA     |              | 0.137   | 0.440   | TANSAFOA |              | 0.277   | 0.308   |
| ROA   | Non-optimized | FOA          | 0.325   | 0.606   | TANSAFOA |              | 1.365   | 0.501   |

| Model | Index | Optimization | Training | Testing | Model | Index | Optimization | Training | Testing |
|-------|-------|--------------|----------|---------|-------|-------|--------------|----------|---------|
| EPS   | Non-optimized | BPNN        | 0.331   | 0.626   | EPS   | Non-optimized | BPNN    | 0.936   | 1.062   |
|       | TANSAFOA     |              | 0.323   | 0.596   | TANSAFOA |              | 0.381   | 0.471   |
| MARS  | TANSAFOA     |              | 0.190   | 0.466   | TANSAFOA |              | 0.337   | 0.459   |
| ROE   | Non-optimized | BPNN        | 0.077   | 0.683   | TANSAFOA |              | 0.340   | 0.329   |
|       | TANSAFOA     |              | 0.077   | 0.683   | TANSAFOA |              | 0.340   | 0.329   |
| ROA   | Non-optimized | BPNN        | 0.293   | 0.656   | TANSAFOA |              | 1.523   | 1.382   |
|       | TANSAFOA     |              | 0.235   | 0.604   | TANSAFOA |              | 0.499   | 0.488   |

There are 3 training models and 3 testing models of each output variable in this study and the prediction performance of 3 prediction models are compared by RMSE (see Table 2). In the Table 2, the RMSE of training results of TANSAFOA-model are the least besides FOA-MARS-ROE and FOA-BPNN-EPS. In addition, it shows all of the RMSE of testing results of TANSAFOA-model are superior to non-optimized model and FOA-model (see gray cells of table). To avoid overfitting of
model, the results of out of sample data are most cared by scholars in modeling process. Thus, TANSAFOA has the best prediction accuracy in this study.

After proving TANSAFOA can effectively improve the performance of optimization, the most appropriate prediction model is further suggested for modern service industry in China. The performance of MARS model optimized by TANSAFOA is compared to the one of BPNN model optimized by TANSAFOA again and it shows that the prediction of business performance of modern service industry by BPNN is better than MARS.

3.3. Convergence of TANSAFOA

Convergence speed can also be an alternative to show the effectiveness of an optimization method. The convergence speed refers to the time required (i.e. iteration numbers) to approach the optimal output value. Fig. 3 displays the iteration charts for 3 indexes of MARS and BPNN optimized by TANSAFOA and FOA. In the charts, the solid line is the 1000 iteration process which RMSE of the TANSAFOA model declines, and the dotted line shows the situation of FOA model. In the Fig. 3, all charts show that the solid line have the lower RMSE and most solid line descend more steeply and find the optimal output value early. The result shows, no matter MARS or BPNN model, TANSAFOA proposed in this has better convergence effectiveness than FOA.

4. Conclusions

In this paper, TANSAFOA is proposed to correct the basic FOA. In TANSAFOA, the smell value is transformed by tangent function to find the optimal parameters of prediction model. The AI model optimized by TANSAFOA has the best prediction performance and it proves that TANSAFOA successfully improves the limitations of the basic FOA. Furthermore, TANSAFOA BPNN-model is the most appropriate prediction model for modern service industry in China. What’s more, one of the purposes in this study is to explore the influential variables to business performance of modern service industry in China. The results show that “Net profit margin on sales”, “Operating cost ratio” and “Cash recovery rate for total assets” are the common variables when predicting the 3 indexes. Therefore, the managers of modern service industry should use them well to maintain the good
development of the company. On the other hand, the different business performance evaluation indexes have their own influential variables, and then input variables can be selected according to different prediction projects for modern service industry in China. To sum up, this study proposes a novel TANSAFOA and the most appropriate prediction model and influential factors to the business performance prediction model of modern service industry in China have also been determined.

Acknowledgment
This research was funded by Characteristic Innovation Projects of Universities in Guangdong (2018WTSCX226), Discipline Co-construction Project of Guangdong Province Philosophy and Social Science Planning (GD18XGL33) and Program for Innovative Research Team of Guangdong University of Foreign Studies South China Business College (Team number: 2018TD05).

REFERENCES
[1] National Development and Reform Commission. (2017) Service industry innovation development outline (2017-2025). http://www.gov.cn/xinwen/2017-06/21/content_5204377.htm.
[2] Pan, W.T. (2011) A new evolutionary computation approach: fruit fly optimization algorithm. In 2011 Conference of Digital Technology and Innovation Management. 382-391.
[3] Pan, W.T. (2011) Using fruit fly optimization algorithm optimized general regression neural network to construct the operating performance of enterprises model. Journal of Taiyuan University of Technology (Social Sciences Edition). 29(4): 1-5.
[4] Pan, W.T., Han, S.Z., Yang, H.L., Chen, X.Y. (2018) Prediction of mutual fund net value based on data mining model. Cluster Computing. 3: 1-6.
[5] Pan, Y., Shi, Y. (2017) A Grey Neural Network Model Optimized by Fruit Fly Optimization Algorithm for Short-term Traffic Forecasting. Engineering Letters. 25(2): 198-204.
[6] Zhong, Z.C., Pan, W.T., Luo, S.H., Yang, T.T. (2017) Acceptable if building electricity price trend change to some biology trends. Journal of Intelligent & Fuzzy System. 33(5): 3069-3077.
[7] Han, S.Z., Pan, W.T., Zhou, Y.Y., Liu, Z.L. (2018) Construct the prediction model for China agricultural output value based on the optimization neural network of fruit fly optimization algorithm. Future Generation Computer Systems. 86: 663-669.
[8] Wu, L., Liu, Q., Tian, X., Zhang, J., Xiao, W.S. (2018) A new improved fruit fly optimization algorithm IAFOA and its application to solve engineering optimization problems. Knowledge-Based Systems. 144: 153-173.
[9] Niu, J., Zhong, W., Liang, Y., Luo, N., Qian, F. (2015) Fruit fly optimization algorithm based on differential evolution and its application on gasification process operation optimization. Knowledge-Based Systems. 88: 253-263.
[10] Zhang, Y., Cui, G., Wu, J., Pan, W.T., He, Q. (2016) A novel multi-scale cooperative mutation fruit fly optimization algorithm. Knowledge-Based Systems. 114: 24-35.
[11] Friedman, J.H. (1991) Multivariate Adaptive Regression Splines. The Annals of Statistics. 19(1): 1-67.
[12] Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986) Learning Representations by Back-Propagating Errors. Nature. 323: 533–536.
[13] Téllez, D., González, M., Guzmán, A., Trujillo, M.A. (2017) What do you say and how do you say it: Information disclosure in Latin American firms. Universidad EAFIT.