Stock Price Prediction Based on CPP-GAM

Yong Li¹, Mingrui Bai¹* and Jianguo Liu²
¹Computer Science and Engineering, NorthWest Normal University, China
²Shanghai University of Finance and Economics, Institute of Fin.tech, China

*Corresponding author email: baimr_yk@163.com

Abstract. Based on the generalized additive model, we propose a CPP-GAM algorithm which transforms the non-linear problem into a linear one. We apply this algorithm to predict the closing price of international and domestic stocks. We train the history data of stocks through back-fitting algorithm. In order to make the effect of prediction better, we get trend lines based on the method of changing point prediction, the regressive algorithm of OLS, and the Fourier series. Through a large number of empirical data analysis, we found the predictive accuracy of CPP-GAM algorithm is 89%, which is 15% higher than that of RBN, SVM, SSA-SVM and so on.

1. Introduction
Stock is an important part of the financial market, so the analysis and prediction of stock price has important theoretical significance and practical value. According to market efficiency theory [1], the stock market of U.S. is a semi-efficient, the current price of stock is derived from all public information. For example, national policies, markets, and the choices of investors have all affecting the stock market, which leading to the stock price is complex and changeable. So, Many scholars not only use dimension reduction techniques [2] to extract features, but also use multiple machine learning models [3],[4] to predict the stock price, which can provide investors with investment advice to better help them maintain and increase their assets in the capital market.

In recent years, machine learning has been widely used for prediction. Machine learning is a data analysis method that does not rely on rule design. It is an important branch of computer science and artificial intelligence. It is used to discover the laws which are hidden in data by iterative learning. Statistical model is the stylized expression of the variables in the form of mathematical equations, which is an important branch of mathematics and statistics that is used to find the correlation between variables and depends on parametric estimation. Previous studies have shown [5] that the optimal parameter makes statistical models perform better than machine learning.

Generalized Additive Model (GAM)[6] is a flexible statistical model, which was proposed by Trevor Hastie and Tibshirani in 1990. The principle of GAM is similar to the regression method, but the difference lies in the sum effect instead of the individual predictors. In the GAM model, the independent variable is linear with the smoothing function, while the dependent variable and the smoothing function are non-linear. Therefore, GAM is a model which based on smoothing function, and the dependent variable is not limited to the linear assumption of the independent variable, the model is more in line with the non-linear characteristics of financial markets. GAM uses back-fitting algorithm to fit the function, and performs spline smoothing on the predicted variable by constraint function to minimize the error term. GAM and other models are shown in Table 1.

The GAM model is flexible, so the function of the GAM model can be determined according to actual problems. Because the predictive accuracy of the GAM model is greatly affected by parameters, we
proposes CPP-GAM algorithm which based on the GAM model, we use the Change Point Prediction method and Ordinary Least Square (OLS) regression algorithm to find the better bias parameter, which will make the predictive accuracy of stock price better. According to the "Long Tail Distribution and Pareto's Law "[7], only 20% of the stocks occupy 80% of the market, which has a market-oriented role. Therefore, we select more than 600 stocks in the New York Stock Market (about 20% of the total number of stocks) to analyze and predict based on the CPP-GAM model. The main contribution result is shown below:

(1) Through a large number of empirical data analysis, it is found that the predictive accuracy of CPP-GAM algorithm is 89% based on AAPL stock, which is 15% higher than that of RBN, SVM, SSA-SVM and so on.

2. Related Work

Machine learning technology has the ability to identify stock trends from massive data, and can capture the inherent dynamic trends of stock price, such as Support Vector Machine (SVM)[8], the hybrid model of Self Organizing Map and Support Vector Machine(Self Organizing Map-Support Vector Machine, SOM-SVM)[3], the hybrid model of Self Organizing Map and Fuzzy Support Vector Machine (Self Organizing Map-Fuzzy Support Vector Machine, SOM-fSVM)][3], Radial Basis Function (RBF)[3], and Singular Spectrum Analysis-Support Vector Machine (SSA-SVM)[4], etc. Support Vector Machine is based on the Vapnik-Chervonenkis[9] dimensional theory and the principle of structural risk minimization, and seek the best way between the complexity of the model and the learning ability based on the limited sample information. With the introduction of the insensitive loss function, the Support Vector Machine, which not only deal with classification problems, but also be able to handle regression problems. By constructing a kernel function, the non-linear problem can be mapped to the high-dimensional linear problem, and it plays a decisive role in the predictive performance of SVM. The disadvantage of the SVM is that in order to improve the accuracy of the model, the number of samples will increase, which will lead to the increase of fuzzy rules.

The hybrid model of self-organizing map and support vector machine is a time series model, which has been widely used in stock market[10]. For each small data set, a set of fuzzy rules is extracted by using the fuzzy support vector machine model, and the data is predicted based on the fuzzy rule set. The hybrid model of singular spectrum analysis and support vector machine is a method to analyze the problem of non-linear and non-stationary time series. Time series with different characteristics can be derived by analyzing singular values of different information, and the singular values are usually used to extract noise from time series. The stock price sequence can be decomposed into trend, market fluctuation and noise by using singular spectrum analysis, and then can be introduced into support vector machine for the prediction of stock price. Previous studies have found that[4], the method with the highest hit rate of prediction analysis is the hybrid model of singular spectrum analysis and support vector machine. For the stock data of AAPL, the predictive accuracy of stock price by the hybrid model of the singular spectrum analysis and support vector machine is 69.9%[4]. Radial basis neural network[11] model has also been used for the prediction of stock. Previous studies have shown[12] that the RBF model has a lower accuracy than other models[4].

In addition, Li Z R][13],[14] et al. has analyzed the temporal and spatial characteristics with volatility as the carrier in the financial market adjustment period, and studied the asset pricing model[13], the fluctuation characteristics and the SSE index pricing model[14] which based on the diffusion of market information. Behavioral finance believes that[15], because investors have different "information cognition abilities" for different industries and individual stocks, the information can only gradually spread in the market, which leads to biased asset pricing. In the 21st century, attention is considered as a scarce cognitive resource. Existing literatures[16] have studied the collective attention model of investors to predict the trend of stock market. Some Studies show that[17] the investor's cognitive ability of different stocks in different industries will also have different impacts on stock price, it is helpful to understand the response mode that investor's to industry information by constructing an asset pricing model which slowly spreads within industries due to the limited attention of investors, but different indicators need to be used to measure the investors' attention, such as:
company size, transaction volume, turnover rate et al, the results obtained by different indicators will be different, and how to measure attention uniformly remains to be further studied.

### Table 1. Model Comparison.

| Advantage | Disadvantage | Main Function | Key Technology | Research Hotspots |
|-----------|--------------|---------------|----------------|-------------------|
| GAM       | 1. Without being limited to the linear hypothesis, study each function separately. | Using non-parametric functions to detect non-monotonic and nonlinear relationships between variables. | 1. Finding the optimal parameter and the right function to achieve the best prediction effect. | 1. Determining the optimal decomposition function. 2. Finding optimal parameter. |
| SVM       | 1. Solving high-dimensional problems by optimizing penalty variables and kernel functions to achieve linear inseparability of sample data. | Based on the structural risk minimization theory, so that the learner can be globally optimized. | Selecting the appropriate kernel function according to the actual data model. | 1. Constructing a better multi-class classifier based on SVM. |
| RBF       | 1. Non-linear fitting for large-scale samples. 2. Good generalization ability and fast learning convergence speed. | It is suitable for prediction and classification of various nonlinear problems. | Finding the right activation function. | It’s pre-processing and coding are still to be discussed. |
| SOM       | 1. SOM uses an analytical method to avoid the complex iterative process of secondary planning. | The kernel functions and kernel parameters involved in the model have a great influence on the prediction accuracy. | Clustering the data can get the best sample data. 2. Choosing the right kernel function and kernel parameters. | How to find better parameters and kernel functions. |
| SVM - FSVM| 1. It’s suitable for a large number of sample. 2. The complexity of the algorithm is reduced, and the execution time of the algorithm is improved. | Determine the appropriate and unified fuzzy membership function. | Extracting fuzzy rules from support vectors with fSVM. | 1. Avoiding a surge in the number of fuzzy rules. 2. Researching the unified fuzzy membership function. |
| SSA-SVM   | 1. Fast calculation speed and high accuracy, suitable for short-term prediction methods. | Using SSA to decompose stock prices into trends, market fluctuations, and noise with different economic characteristics. | Breaking down the stock price into different items. | 1. How to find the optimal singular value and construct the optimal solution. |

Through the above analysis, it is found that these models have their own shortcomings, but the GAM model can avoid many problems. Compared with other models, the GAM model is more applicable. However, the parameter of the GAM model will affect the result of prediction. In order to avoid this problem, we propose the CPP-GAM algorithm which based on the generalized additive model. Lots of experimental results show that the predictive accuracy of CPP-GAM algorithm is 15% higher than other models.

### 3. Introduction to Models and Algorithm

#### 3.1. Generalized Additive Model

The mathematical expression of the generalized additive model (GAM)[6][18] is:

\[
Y_i = \alpha + \sum_{j=1}^{N} f_j(X_{ij}) + \epsilon
\]  

(1)

Where, \( \alpha \) is a constant term, \( \epsilon \) is an error term, \( f_j(X_{ij}) \) is an unknown non-parametric smoothing function, and \( X_{ij} \) is a predictive variable.

#### 3.2. Back-fitting Algorithm

In statistics, the back-fitting algorithm[18] is a simple iterative process, which is mainly used for model
fitting[19]. In most cases, the back-fitting algorithm is similar to Gauss Seidel algorithm[20] to solve linear equations. For equation (1), due to the high flexibility of $f_i(X_i)$, GAM does not have a unique solution, so the value of $\alpha$ cannot be determined. However, $\alpha$ can be obtained by a constraint function, which is as follows in equation (2):

$$\sum_{i=1}^{N} f_i(X_i) = -\alpha_i$$ (2)

For the GAM expression shown in equation (1), for a given $x_i, y_i$, if its error term is close to 0, then a spline smoothing[21] algorithm is needed, and $\alpha = -1/N \sum_{i=1}^{N} y_i$. The specific process of the back-fitting algorithm and the smoothing algorithm are shown below:

**Algorithm 1. Back-fitting ($i, t, s, \delta, y, \alpha(t), k, m, T$ )**

1. Initialize the constraint of $e = 0$
2. for $i \in N$ do
3. $\alpha = -1/N \sum_{i=1}^{N} y_i$
4. end for
5. Quantifying the transition point
6. for each $t \in T$ do
7. $a_j(t) = \begin{cases} \frac{1}{t}, & k \geq s_l \\ 0, & \text{otherwise} \end{cases}$
8. end for
9. Finding the smooth function $f(t)$ over time
10. for each $t \in T$ do
11. $A(t) = (k + ZIP(*a(t)))*\delta$
12. $B(t) = t - (m + ZIP(*a(t)))*y_i$
13. $g(t) = C(t) + \exp(A(t) * B(t))$
14. for $i=1$ to $N$ do
15. $Z(t) = a_n \cos((2\pi n t)/T) + b_n \sin((2\pi n t)/T)$
16. end for
17. $f(t) = g(t) + Z(t) + \alpha$
18. end for
19. for $i=1$ to $N$ do
20. $f_i = f_0 + S_1 f_0 + S_2 f_1 + ... + S_{i-1} f_{i-1} + \alpha$
21. return $f_i$

**Algorithm 2. Smoothing ($i, j, f, y_i$ )**

1. Translation of historical data
2. for $i=1$ to $N$ do
3. $\bar{y}_i = \{(i + 1) * (1/10) \}^{1/(b)} * \bar{y}_i$
4. end for
5. The result is processed by compression and expansion
6. for $i=1$ to $N$ do
7. $f = \sum_{i=1}^{N} (f_i)^{j}$
8. $\bar{y}_i = (\{(i + 1) * (1/10) \}^{1/(b)} * \bar{y}_i$
9. end for
10. The matrix coefficient of the smoothing function
11. for $i=1$ to $N$ do
12. for $j=1$ to $N+1$ do
13. $S_j = \bar{y}_j^{2j}$
14. $S_j = (y_j)^{j} (y_j)^{j}$
15. $F = f \times (\bar{y}_j)^{j}$
16. $(s_k, l - s_k) / (s_k - s_k) = (\bar{y}_j)$
17. end for
18. end for
19. for $i=1$ to $N$ do
20. $\tilde{f}_i(\tilde{y}_i) = \tilde{f}_i^{j}$
21. $\tilde{y}_i = f \times (y_i)^{j} (y_i)^{j}$
22. $f = \sum_{i=1}^{N} f_i(y_i) = f_0 + f_1 y_1 + f_2 y_1^2 + f_3 y_1^3 + ... + f_i y_i$
23. end for

### 3.3. Change-Point Predictive Analysis

This paper proposes the Change-Point Predictive Analysis(CPP) which is defined as follows: Given a set $m_i = \{0.001, 0.005, 0.01, 0.015, 0.02, \ldots\}, i=1,2,\ldots,N$, the values in the set $m_i$ are increase in step of 0.005, the set controls the overfitting and underfitting of the GAM, and $Y_i$ represents the predictive accuracy of the stock price. The relationship between the $m_i$ and the $Y_i$ can be described as: $Y_i \cap m_i$. The realization process of the CPP is shown in algorithm 3:
Algorithm 3. CPP \((i, N, f_i, y_i)\)

| Input: | 7. for \(i=1\) to \(N\) do  
|--------|-----------------------------------------------|
| The value of the prediction function at point \(i\): \(f_i\)  
| variable parameter: \(i\)  
| The true value at point \(t\): \(y_j\)  
| Iterations: \(N\) | 8. if \(\frac{1}{N} \sum_{i=1}^{N} \frac{y - f_i}{y_i} \leq 1\)  
| Output: Optimal bias parameter: \(m\)  
| Best stock price predicted accuracy: \(P\) | 9. \(\text{count} += 1\)  
| 1. Generate bias parameter set \(m\) in steps of 0.05  
| 2. for \(i=1\) to \(N\) do | 10. return \(\text{count}, \ P/\text{count}\)  
| 3. \(m_i = m_{i+1} + 0.05\) | 11. end for  
| 4. return \(m\) | 12. Fitting the \(m\) and the stock prediction accuracy  
| 5. end for | 13. for \(i=1\) to \(N\) do | 14. \(P_i = 1 - np.\exp(3.431*(-m_i))\)  
| | 15. end for | 16. return max \(P_i, m_i\) | 17. end for

4. Model Composition

In order to more easily analyze the equation (1), the predictive function \(f_i(X_i)\) is decomposed into \(g(t)\), \(Z(t)\) and \(\epsilon\), then the equation (1) can be decomposed into the following form:

\[
Y_i = \alpha + \sum_{i=1}^{N} (g(t) + Z(t)) + \epsilon
\]  

\[ (3) \]

Where, \(g(t)\) represents the trend function, which is used to analyze non-periodic changes in the time series; \(Z(t)\) represents the periodic changes; \(\epsilon\) is an error term, which represents the error effect that the model does not consider.

4.1. Trend Function

The increase of the population in city[22] or the viruses in ecosystem[23] is a complex system, which is also affected by many factors, it is similar to the increase or decline of the stock price in financial market, which will reach saturation after experiencing non-linear growth. This type of growth is similar to the logistic growth model[24]. Therefore, the definition of the trend function is shown in equation (4):

\[
g(t) = \frac{C}{1 + \exp(-k(t - m))}
\]  

\[ (4) \]

Where, \(C\) is the saturation value, \(k\) is the growth rate, and \(m\) is the bias parameter. For stock prediction, the saturation value \(C\) will change with time, and the growth rate will also change with some external factors. Set several turning points, that is, \(S_t, t = 1, 2, ..., N\), and construct a vector \(a(t) \in \{0, 1\}^{S_t}\) :

\[
a(t) = \begin{cases} 
1, \quad t \in S_t \\
0, \text{otherwise}
\end{cases}
\]  

\[ (5) \]

In the process of fitting the logistic curve[25], the expression of the growth rate at time \(t\) is set as: \(k + a(t)^T \delta_i\). In order to make the function continuous, it is need to set the bias adjustment quantity \(\gamma_i\) at the turning point \(t\) when the growth rate changes, the bias adjustment quantity is as shown in equation (6):

\[
r_i = (S_t - m) - \sum_{i \in t} r_i \left( k + \sum_{j \in t} \delta_j \right)
\]  

\[ (6) \]

get the trend function \(g(t)\):

\[
g(t) = \frac{C}{1 + \exp(-A(t) \ast B(t))}
\]  

\[ (7) \]
among them:

\[ A(t) = (k + a(t)\hat{T})\delta_t \]  \hspace{1cm} (8)

\[ B(t) = t - (m + a(t)\hat{T})\gamma_t \]  \hspace{1cm} (9)

Where, \( k \) is the growth rate, \( \delta_t \) is the growth rate adjustment value at the turning point \( t \), \( \gamma_t \) is the bias adjustment quantity at the turning point \( t \), and \( m \) is the bias parameter.

### 4.2. Periodic Function

In order to fit the noise generated by the periodicity, we use the fourier series\[26\] to construct flexible periodicity rule, we set \( P \) as the period length of time series, as shown in equation (10):

\[ Z(t) = \sum_{n=1}^{N} \sum_{t=1}^{N} (a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right)) \]  \hspace{1cm} (10)

In order to fit the periodicity, we need some parameters, that is, \( \beta = [a_1, b_1, ..., a_N, b_N] \), and it is achieved by constructing a seasonal vector matrix for each \( t \) in history and the future. Supposing that \( N = 10 \), The periodicity of each year is as follows:

\[ \begin{vmatrix} \cos(\frac{2\pi(1)t}{365}) \\ \sin(\frac{2\pi(1)t}{365}) \end{vmatrix} \]

The annual periodic estimate is: \( Z(t) = x(t)\beta \)

### 5. Model Data Introduction and GAM Model Fitting

#### 5.1. Data

Based on the Quandl WIKI financial database\[27\], we obtained more than 3,000 U.S. stocks. According to the "Long Tail Distribution and Pareto's Law"\[7\], this paper selects more than 600 representative stocks for research. Due to the stock data of several years ago is no longer has reference significance to the current stock price trend, so the past 10 years are selected for research, and we select four stocks, they are SP500 index, AAPL, TSAL and BIDU, among them, the AAPL is used as an example for analysis. First, we get the change of closing price. The formula for calculating the increase in closing price is as follows:

\[ y(t) = \frac{V(t) - V(t-1)}{V(t-1)} \]  \hspace{1cm} (11)

Where, \( y(t) \) represents the increase in closing price at \( t \) time, and \( V(t) \) represents the closing price of the stock at \( t \) time. The change in the closing price of AAPL stock is shown in Figure 1.

It can be seen from Figure 1 that the increase in the closing price of AAPL stock is between -20% and 20%, it can be considered as a relatively stable development situation. Therefore, models can be used to predict and analyze the future closing prices of stock. GAM has the ability to solve the problem of highly non-linear and non-monotonic, this paper use GAM to predict the stock market.
5.2. Model Fitting

According to the assumption and analysis of non-periodic functions and periodic functions in Section 4, we use GAM model to train the historical data of AAPL, BIDU, SP500 and TSLA. Taking 2000 as the starting calculation time, we found that the 2005, 2009, and 2015 are the rising times of the Chinese market; the 2000, 2007, 2011, 2015, and 2018 are the falling times of the US market, so this paper chooses the data of 5 years to train model, that is, the stock price from 2014 to 2016 is used as training data, and the stock price from 2017 to 2018 is used as validation data. Within the range of error, the predictive results are shown in Figure 2. The red mark indicates the start of the verification.

Figure 1. The change of closing price of AAPL.

Figure 2. The prediction of stock price based on GAM.
It can be seen from Fig 2 that the trend of different stocks price based on GAM model. Among them, the predictive accuracy of AAPL data based on GAM model can only reach 63.86%, while the predictive accuracy of AAPL data based on machine learning can reach 69.9%[3], so GAM model needs to be optimized to improve the predictive accuracy of stock price.

6. Model Optimization

Due to the low predictive accuracy of stock price based on the GAM in Figure 2, the CPP-GAM algorithm based on GAM is proposed. The implementation process of the CPP-GAM algorithm is shown in Section 3.2. The core idea of the CPP-GAM algorithm is to use the method of change point prediction to find the optimal bias parameter, which controls overfitting and underfitting of model. The specific definition is shown in section 3.3. The result of OLS algorithm on the stock price of AAPL are shown in Figure 3.

![Figure 3. OLS fitting of AAPL stock.](image)

Obviously, the distribution form of the observation points in Fig3 is similar to the negative exponential distribution curve. In this paper, ordinary least squares regression is used to obtain the relationship between $Y_i$ and $m_i$, as follows:

$$ y = 1 - np.exp(3.431*(-m)) $$

The analysis finds when the $m_i$ is close to 1.0, the predictive accuracy of stock price may be maximum. Therefore, if the parameter $m_i = 1.0$, the predictive result of different stocks are shown in Fig 4, in which the predictive accuracy of AAPL stock is 89%. The predictive accuracy of stock price based on multiple models are shown in Table 2.

![AAPL](image)

![BIDU](image)

(a) AAPL

(b) BIDU
7. Experimental Comparison

In this paper, the five models with the best performance in literatures are selected for comparison and analysis with CPP-GAM model. In order to show the experimental results, we selected four representative stock samples: AAPL, IBM, SP500 index and Shanghai Stock Exchange Composite Index. Different models use the same test data. As shown in Table 2, the CPP-GAM model has a higher predictive accuracy of stock price, it is at least 15% higher for AAPL stock than other models.

| Code/Models | CPP-GAM | RBN  | SOM-1SVM | SOM-SVM | SVM  | SSA-SVM |
|-------------|---------|------|----------|---------|------|---------|
| AAPL        | 89.1    | 45.73| 53.27    | 52.76   | 49.67| 69.92   |
| IBM         | 88.66   | 43.72| 50.75    | 44.22   | 47.35| 66.75   |
| SP500       | 80.72   | 51.76| 53.27    | 52.76   | 48.75| 69.57   |
| SSE         | 83.52   | 49.25| 51.37    | 55.96   | 52.49| 67.97   |

The histogram of the predictive accuracy of the stock price based on different models are shown in Fig. 5. It can be clearly seen that the predictive accuracy of the stock price based on CPP-GAM is better than other models.

8. Conclusion

Based on the generalized additive model, this paper proposes a CPP-GAM model which finding the optimal parameter to minimize the error. The CPP-GAM model includes: the predictive function in GAM is decomposed into three factors, which are the trend function, the periodic function and the error term; The back-fitting algorithm is used to fit the trend function; The constraint function is used...
to smooth the predictive variables so that the error term is close to 0; The optimal parameter is found by the method of change point prediction and OLS algorithm, and the trend of stock price with the minimum spline smoothing error is obtained. Through a large number of empirical data analysis, we found the predictive accuracy of AAPL stock based on CPP-GAM algorithm is 89%, which is 15% higher than that of RBN, SVM, SSA-SVM and so on.

The advantages of stock price prediction based on the CPP-GAM model includes: 1) The predictive results are not limited to the linear assumption of the variables based on CPP-GAM model; 2) The different predictors of the CPP-GAM model can be defined for different research problems. However, Because the CPP-GAM model is based on the sum of various predictors, the predictive results may be different by selecting different predictors.

In future research, the CPP-GAM model will be decomposed into different predictive functions according to the characteristics of the financial market, for example, the holiday function will be used to study the impact of holidays on the stock market[28]. Or, combining the GAM and deep neural network[29][30] model to predict the stock price.

Acknowledgments
This research was partially supported by the grants from the Natural Science Foundation of China (No.71764025, 61863032, 61662070); the Research Project on Educational Science Planning of Gansu, China(Grant No. GZ[2018]GHBBKZ021, GZ[2018]GHBBKW007); the Scientific Research Foundation of the Higher Education Department of Gansu, China(Grant No.2018A-001).

References
[1] Mockus, Raudys J, Aistis, 2010, Expert System with Application, On the Efficient-Market Hypothesis and stock exchange game model, vol. 13, no. 8, pp. 5673-5681.
[2] Reddy G T, Reddy M P K, Lakshmana K, et al, 2020, IEEE Access, Analysis of dimensionality reduction techniques on big data, vol. 8, pp. 54776-54788.
[3] Nguyen D H, Le M T, 2014, (IJCISIS) International Journal of Computer Science and Information Security, A two-stage architecture for stock price forecasting by combining SOM and fuzzy-SVM, vol. 12, no. 8.
[4] Wen F H, Xiao J H, He Z F et al, 2014, Procedia Computer Science, Stock Price Prediction Based on SSA and SVM, vol. 31, pp. 625-631.
[5] Tang L W, Lu Y Y, 2018, System Engineering-Theory & Practice, The optimization of GM (2,1) model based on parameter estimation of grade difference format, vol. 38, no. 2, pp. 502-508.
[6] Hastie T J, Tibshirani R J, 1990, London: Chapman Hall, Generalized Additive Models.
[7] Reed W J, 2003, Physica A: Statistical Mechanics and its Applications, The Pareto law of incomes–An explanation and an extension, vol. 319, pp. 469–486.
[8] Lin Y L, Guo H X, Hu J L, 2013, in Proc (IJCNN) International Joint Conference on Neural Networks, An SVM-based Approach for Stock Market Trend Prediction, USA, , pp. 236-242.
[9] Yaser S, Abu M. 1989, Neural Computation, The Vapnik-Chervonenkis Dimension: Information versus Complexity in Learning, vol 1, no. 3, pp. 312-317.
[10] Juan G R, Juan C, 2017, Soft Comput, Burguillo:Topology-based analysis of self-organizing maps for time series prediction, vol. 21, no. 6, pp. 1601-1618.
[11] Surechev S, Sotirov S, 2014, IEEE Intelligent Systems, Modeling the Process of Face Recognition with Pattern Neural Network Using a Generalized Net, vol. 2, pp. 693-702.
[12] Hanafizadeh P, Hashemi A, 2014, (IJBIR)International Journal of Business Innovation and Research, The Effect of Behavioral Factors on Stock Price Prediction using Generalized Regression and Backpropagation Neural Networks Models, vol. 5, no. 4, pp. 44-57.
[13] Li Z R, Qiao Z R, Wang S R et al, 2018, Systems Engineering-Theory & Practice, An information diffusion-based model of stock market cycle analysis: International experience and domestic application, vol. 38, no. 10, pp. 2449-2465.
[14] Li Z R, Zu L, 2015, Systems Engineering-Theory & Practice, Study on the pricing mechanism and fluctuation features for Shanghai Composite Index using information diffusion-based model, vol. 35, no. 6, pp. 1416-1424.
[15] Hong H, Stein J C, 1999, Journal of Finance, *A unified theory of under-reaction, momentum trading, and overreaction in asset markets*, vol 54, no. 6, pp. 2143–2184.

[16] Liu J G, Yang Z H, Li S N, Yu C R, 2018, Physica A: Statistical Mechanics and its Applications, *A generative model for the Collective attention of Chinese stock market investors*, vol 512, pp. 1175–1182.

[17] Xiang C, Lu J, 2018, Systems Engineering-Theory & Practice, *Investor limited attention, industrial information diffusion and stock pricing*, vol 38, no. 4, pp. 817-835.

[18] Hastie T, Tibshirani R, Friedman J, 2010, Germany: Springer Series in Statistics, *The Element of Statistical Learning: Data Mining, Inference, and Prediction*, pp. 257-261,127-129.

[19] Friedman J, WSuetzle W, 1981, Journal of the American Statistical Association, *Projection pursuit regression*, vol. 376, no. 76, pp. 817-823.

[20] Kahan W, 1958, Canada: University of Toronto, *Gauss-Seidel Methods of Solving Large System of Linear Equations*.

[21] Buja A, Hastie T J, Tibshirani R J, 1989, The Annals of Statistical, *Linear smoother and additive models*, vol. 17, no. 2, pp. 453-510.

[22] Andreis S D, Ricci P E, 2005, Mathimatical and Computer Modelling, *Modelling population growth via laguerre-type exponentials*, vol. 42, no. 13, pp. 1421-1428.

[23] Abbiana R A, Neil B T, Evan C H, 2017, Mathematics and Computers in Simulation, *The logistic growth model as an approximating model for viral load measurements of influenza Avirus*, vol. 133, pp. 206-222.

[24] Swishchuk A, Wu J H, 2003, Germany: Springer, *Logistic Growth Models*, pp. 175-185.

[25] Caballini F, 1993, College Mathematics Journal, *Fitting a logistic curve to data*, vol. 24, no. 3, pp. 247-253.

[26] Tang K T, 2007, Germany: Springer, *Fourier Series*.

[27] Nasdaq. (2018, Dec, 4). *The world’s most powerful data lives on Quandl*. [Online]. Available: https://www.quandl.com.

[28] Andrey K, 2018, Journal Asset Management, *Holiday Effect on Stock Price Reactions the Analyst Recommendation Revisions*, vol. 19, pp. 507-521.

[29] Reddy T, RM S P, Parimala M, et al, 2020, Computer Communications, *A deep neural networks based model for uninterrupted marine environment monitoring*, vol. 157, pp. 64-75.

[30] Yu P F, Yan X S, 2020, Neural Computing and Applications, *Stock Price Prediction based on deep Networks*, vol. 32, no. 6, pp. 1609-1628.