Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Research on Multivariate Financial Time-Series Forecasting Considering Endogenous Variables

Yungao Wu[1,2,*], Qixin Zhao[3], Yujia Li[3], Pingan Hao[4]

[1] School of Economics and Management, North China Electric Power University, Beijing, 102206, P. R. China

[2] Department of mathematics and Computer Engineering, Ordos Institute of Technology, Ordos, 017000, P. R. China

[3] Inner Mongolia Power Economic and Technological Research Institute, Inner Mongolia Power Group, China

[4] Transmission Management Office, Bayannaoer Electric Power Bureau, Inner Mongolia, China

*Corresponding Author.

Abstract:

The ensemble learning method is outstanding in the prediction accuracy of composite stock index and foreign exchange time-series. Still, it is slightly worse in the prediction accuracy of multivariate financial time-series when considering endogenous variables. This paper proposes a new method to identify similar data patterns and cluster them into groups by Fuzzy C-Means (FCM). The results of an example prove that the improved FCM clustering algorithm and the improved Elman network can be used to classify and forecast the stock data of Shanghai and Shenzhen A-stock markets.

Keywords*: Cluster analysis, Fuzzy C-Means, Elman Networks, Multivariate time-series prediction model.*

I. INTRODUCTION

With the continuous development of artificial intelligence technology, multiple financial time series has become an important research field. The data and feature selection in the multivariate financial time series will directly determine the effect of the forecasting model. However, in the process of financial time series forecasting, there are many sample data and factors that affect the sample, including the stock market, relevant policies, and interest rates. At the same time, compared with the traditional time series method, the multivariate time series method is better in forecasting accuracy and forecasting effect. Therefore, it is extremely

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

important to analyze the current financial situation and use a variety of time series methods for research.

Time-series is a set of random data collected and sorted according to time intervals. In the same time intervals, according to the sequence of activities, observe and analyze this group of data [1]. Literature [2] pointed out that the essence of time-series is to reflect the changing trend of some variables with time synchronization so that the key of time-series prediction is to dig out this change rule from a series of random data. The essence of time-series theory is to deal with time-series and to predict the changing trend of future time data according to the law obtained by analysis. In the literature [3], the characteristics of time-series are described more accurately: The data in time-series are arranged on the scale of time without being time-related; The data values in time-series have intense randomness, which cannot be deduced directly from historical data; The data in time-series have intrinsic relevance and may present some dynamic laws; There are periodic fluctuations in time-series.

Many scholars have combed and studied the characteristics of financial time-series. Literature [4-5] took the financial time-series of Chinese and American stock markets as an example to study the correlation between American and Chinese stocks. The results showed that the correlation coefficient between stocks in American and Chinese financial markets is stable. Based on the government supervision system, Literature [6] analyzed the prudential and healthy development of finance in City T, and used the time series method to analyze that consumer finance has a positive effect on economic growth. Literature [7] used Copula's multi-dimensional time series theory to analyze the risk management, option pricing, portfolio investment, credit risk, and financial market contagion in economic issues, and revealed the current problems. Literature [8] Based on the background of "Internet + Big Data", they used the improved BEMD algorithm to decompose the interval financial time series at multiple scales and used three interval-type single-term forecasting methods to separate the decomposed series. Forecast, and select the Shanghai Composite Index for analysis of examples, further verifying that the combined forecasting model is better than the single forecasting model. Literature [9] Faced with the strong nonlinear, non-stationary, and high-noise characteristics of financial time series, they explored and analyzed their internal correlations and complementary advantages, and fully compiled the research results applied to financial time series. And prospects, and provide a solid theoretical support for follow-up research. Literature [10] through in-depth analysis of the imbalanced relationship between financial development and income, collected macro data from 1973 to 2017, used multiple regression analysis to verify the relationship between each other, and pointed out that there is a U-shaped relationship It also analyzes and discusses in depth from three aspects: education level, foreign trade dependence, and trade union density.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

The deep learning method can strengthen learning according to the input data characteristics, gradually ignore the non-influencing factors in the process of learning heterogeneous communication, and strengthen the practical communication so that it can obtain better prediction results [11]. The application of deep learning technology in the financial field mainly focused on the market forecast of stocks, futures, securities and other products [12].

In the literature [13], the DBN network structure was built by continuously restricted Boltzmann mechanism, and three different exchange rates were selected to compare and test the predicting effect of this model. Literature [14] took the trend of the closing price of S&P500 on the second day as the research object, and used three different neural networks (MLP, CNN and LSTM) to predict the closing price. The experimental results showed that the CNN neural network structure has the smallest error. Literature [15] used deep learning technology to design a high-frequency trading strategy. Literature [16] compared the application of LSTM network with the prediction of S&P500 index, and selected standard depth network (DNN), logistic regression classifier and random forest (RF) for model comparison. Literature [17] creatively applies the stacked self-encoder to financial prediction. Firstly, the financial time-series was processed by wavelet transform. Then, the LSTM model was optimized by using stacked self-encoder, and the prediction accuracy of this model was dramatically improved. In the Literature [18], they used two different deep neural network structures, namely recurrent neural network (RNN) and extended-short-term memory network (LSTM) to predict the correlation coefficient of stock price. The prediction results show that RNN is highly time-dependent. If the time intervals of sampling points in time-series is extended, RNN is prone to the phenomenon of gradient disappearance. LSTM model has intense time connectivity, and the prediction effect is better when the amount of data is significant. Literature [19] based on the idea of “decomposition-reconstruction-synthesis”.

The above research realizes the application of deep learning method in financial time-series forecasting. Literature [20], constructed model to predict 30 stock indexes. Meanwhile, indicators were selected to evaluate. Literature [21] constructed an improved Deep Belief Network (DBN) to train the unsteady, nonlinear and irregular financial time-series curves. Then, they used the trained model to filter financial data. The decision accuracy of DBN network is high, which reaching 90%. Literature [22] put forward an optimization function for the stationarity test of prediction series. Literature [23] constructed an Alpha model based on LSTM stock forecasting with controlled cost, which can predict the return vector in stock index. This model can control the trading frequency and ensure more stable return.

At present, the deep learning method has achieved good results in financial time-series prediction. However, in practical application, it is necessary to adopt appropriate deep learning methods in view of the specific problems and environment in the financial market, and strive to obtain higher prediction accuracy and practically avoid market risks.

In this paper, an improved FCM-Elman multivariate financial time-series forecasting model is constructed, and the model is applied to cluster analysis of financial time-series. The section 1 introduces the clustering method of multivariate financial time-series. The section 2 introduces multivariate financial time-series clustering based on improved FCM. The section 3 introduces the improved Elman network forecasting model, the algorithm based on FCM-Elman multivariate financial time-series forecasting model and the algorithm flow chart. The section 4 evaluates and analyzes the algorithms of each model to scientifically verify the accuracy and practicalness of the models. In the last section, the summary and prospect are made.

II. CLUSTERING METHOD OF MULTIVARIATE FINANCIAL TIME-SERIES

Multivariate financial time series are predicted by the relationship between the internal yuan:

2.1 Introduction of K-means Clustering Algorithm

The principle of K-means clustering algorithm is relatively simple. Firstly, Select K initial centroids, where K is a parameter specified by the user, that is, the number of expected clusters. Each point is assigned to the nearest centroid, and the set of points assigned to a centroid is a cluster. Then, according to the points assigned to the clusters, the centroid of each cluster is updated. Repeat the assignment and update steps until the cluster does not change, or until the centroid does not change. The formal description of the K-means clustering algorithm is [24].

Step 1: Select K points as the initial centroid.

Step 2: Assign each point to the nearest centroid, forming K clusters.

Step 3: Recalculate the centroid of each cluster.

Step 4: Determine whether the centroid has changed, if it does not change, the algorithm ends, if it changes, go to step 2.

K-means reaches a state where all points will not transfer from one cluster to another so that the centroid will not change [25]. In order to assign points to the nearest centroid, aproximity metric is needed to quantify the concept of “nearest” of the data under consideration. However, for a given data type, there may be a variety of suitable proximity measures. K-means clustering algorithm of Step 4 is described as "Recalculating the centroid of each cluster", because the centroid may change with different data proximity measures and clustering targets.

2.2 Hierarchical Clustering Algorithm

Hierarchical clustering algorithm is also an important type of clustering method. There are two basic methods to generate hierarchical clustering [26]:

(1) Condensed hierarchical clustering method: starting from points as individual clusters, each step merges the two closest clusters;

(2) Split hierarchical clustering method: Start from a cluster containing all points, and split a cluster at each step until only a single point cluster remains.

Compared with the two methods, the agglomerated hierarchical clustering method is more commonly used. This section only introduces the agglomerated hierarchical clustering method. The algorithm is described as:

Step 1: Calculate the proximity matrix;

Step 2: Merge the two closest clusters;

Step 3: Update the proximity matrix to reflect the proximity between the new cluster and the original cluster;

Step 4: Determine whether there is only one cluster left, if yes, the algorithm ends, if not, go to step 2.

The calculation of the proximity of clusters is often related to the definition of cluster types. Many agglomerated hierarchical clustering techniques, such as group averaging, MAX and MIN, are derived from the graph-based view of clusters. The proximity of clusters defined by MAX is the proximity between the two farthest points in different clusters which is described as “the extended edge between two nodes in different node subsets” in terms of graphs. The proximity of clusters defined by MIN is the proximity between two nearest points of different

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

clusters. The cluster proximity defined by average group count is the average pair-by-pair proximity of all point pairs taken from different clusters, which is described as "the average side length between all pairwise nodes in different node subsets" by using the term of graph. The basic cohesive hierarchical clustering method uses the proximity matrix [27].

2.3 Clustering Algorithm based on SNN Density

As we know, standard methods that basis on density and similarity cannot produce ideal clustering results occasionally so that an indirect method of similarity calculation is introduced. The principle of the this method is given below: If two points are in the same point Most of them are similar, even if a direct similarity measure cannot be obtained, they are considered similar. Shared Nearest Neighbor (SNN) of similarity is defined as:

Step 1: Find out the k nearest neighbors of all points;

Step 2: If two points X and Y are not in each other's K nearest neighbor;

Step 3: Discrimination $Similarity(x, y) \neg\ 0$;

Step 4: Else;

Step 5: $Similarity(x, y) \neg$ Number of neighbors shared;

Step 6: End if

Essentially, as extended as two objects are in each other's nearest neighbor list, the number of neighbors they share is the SNN similarity, and any dissimilarity measure or meaningful similarity can be the proximity measure.

Some problems which arise when using direct similarity can be solved by the SNN similarity.

In a first step, it considers the environment of objects by sharing the number of nearest neighbors in order to be able to handle the case where two objects are relatively close but do not extend to the same class. Then, SNN similarity can handle the variable density clustering problem. In low-density regions, objects are separated further than those in high-density regions. However, the SNN similarity between two objects depends only on the number of

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

nearest neighbors shared by the two objects, so the SNN similarity can be automatically scaled according to the density of object points.

Use NN density to measure the extent to which a point is surrounded by similar points (with respect to nearest neighbors) [28]. The following is the clustering algorithm based on SNN density:

Step 1: Calculate SNN similarity map;

Step 2: Use DBSCAN with parameters Eps and MinPts specified by the user;

The number of clusters in the data can be determined automatically by the algorithm. DBSCAN is the most the most basic density-based clustering algorithm, and its algorithm is described as follows:

Step 1: Mark all points as core points, boundary points or noise points;

Step 2: Delete noise points;

Step 3: Give an edge between all core points whose distance is within Eps;

Step 4: Each group of connected core points forms a cluster;

Step 5: Assign each boundary point to a cluster of associated core points.

III. MULTI-FINANCIAL TIME-SERIES CLUSTERING BASED ON IMPROVED FCM

Although the Fuzzy c-means clustering (FCM) algorithm has a fast search speed, local optimality still occurs, and the initial clustering center is randomly selected, which will directly affect the convergence speed and clustering effect. In other words, when the cluster center is close to the local minimum, it is easy to fall into the local optimal situation; when the number of samples is too large, the convergence speed becomes slow; When the sample data has small differences or overlaps, the same effect will appear. Under this status quo, it needs to be improved, as follows:

3.1 The Main Idea of FCM Algorithm

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Fuzzy c-means clustering (FCM) algorithm is an improved algorithm of hard c-means clustering algorithm, which was proposed by Bezdek [29] in 1981. Its idea is to randomly select several clustering centers in euclidean space.

FCM algorithm is a kind of classification number $c(2 \le c \le n)$ Under the given premise, $\{A_1, A_2, ..., A_c\}$ Represents the corresponding c Category, by finding the appropriate similar classification matrix U and cluster center matrix $V = \{v_1, v_2, ..., v_c\}$ To make the accurate function J_m The minimum value is reached to determine the sample space $X = \{x_1, x_2, ..., x_n\}$ The method of the best classification scheme, wherein m_{ik}. It's a sample x_i for classes A_k Membership degree of.

The accurate function of FCM algorithm is

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (m_{ik})^m (d_{ik})^2 \tag{1}$$

v is the cluster center matrix; m $(1 \le m \le \infty)$ is a weighted parameter; d_k stands the distance of I and the $k-th$ cluster center. When the accurate function value reaches the minimum, it is considered that all the sample data have been divided into the best clusters. In order to minimize the accurate function value, the following calculations should also be made.

$$d_{ik} = d(x_i - v_k) = \sqrt{\sum_{j=1}^{m} (x_{ij} - v_{kj})^2} \tag{2}$$

$$m_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \tag{3}$$

$$v_{ik} = \frac{\sum_{k=1}^{n} (m_{ik})^m x_{kj}}{\sum_{k=1}^{n} (m_{ik})^m} \tag{4}$$

Through the above two formulas, the clustering center and membership degree of data are adjusted repeatedly. The sum of membership degrees of all cluster centers is 1, i.e.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

$$\sum_{i=1}^{c} m_{ik} = 1, k = 1, 2, \ldots, n \qquad (5)$$

At this time, the algorithm converges, and the theoretical values of membership degree of various clustering centers and samples for each pattern class are obtained, thus completing the division of fuzzy clustering.

In FCM algorithm, the value of fuzzy index m affects the fuzzy degree and convergence of the algorithm. Therefore, it is extremely important to select a reasonable m value in the process of cluster analysis. At present, many literatures have studied the value of m [30]; Cheung and Chan studied the application background of Chinese character recognition, and gave the best value range of M as [1.25, 1.75]. Pal and Bezdek studied the value range of m from the perspective of clustering validity, and thought that the reasonable value range was [1.5, 2.5], and that m could take the median value of the intervals [1.5, 2.5], that is, m=2, without specific requirements. By studying the existing literature, the fuzzy index m=2 is selected in this algorithm.

In essence, the FCM algorithm is a local optimization algorithm, and the basic cluster center selection process of the algorithm is a random allocation mode, which makes the stability of the FCM algorithm dramatically influenced by the randomness of the initial cluster center selection. Therefore, the results obtained by the FCM algorithm are not necessarily the same every time it is executed. Because the clustering result depends on the initial value of the cluster center to a great extent, if the initial value is not selected properly, the FCM algorithm will converge to the local extreme point, which will lead to the deviation of the state recognition result. Therefore, this study combines genetic algorithm with simulated annealing algorithm, to improve the problems existing in the initial clustering center selection process of FCM algorithm.

3.2 Improved FCM Algorithm

In 1953, Metropolis et al. [31] proposed Simulated Annealing Algorithm (SA), which is a global optimization algorithm, and its idea comes from the heat balance problem in statistical thermodynamics.

Simulated annealing algorithm is divided into the following three processes:

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

(1) Heating process. Due to the increase of thermal motion, the heating process causes the particles to deviate from the original equilibrium position, and when the temperature reaches high enough, the solid is converted into liquid.

(2) Isothermal process. In a closed system with constant temperature and heat exchange with the surrounding environment, the spontaneous change of the system state often reduces the free energy. Until the free energy is minimized, the system reaches equilibrium.

(3) Cooling process. During the cooling process, as the thermal motion of particles weakens and the energy of the system decreases, the crystal structure is finally formed.

Among them, the initial temperature of SA algorithm corresponds to the above process (1), the sampling process using Metropolis method corresponds to the above process (2), and the corresponding control parameter reduction process in SA algorithm corresponds to the above process (3) [32]. The process of obtaining the optimal solution in SA algorithm corresponds to the lowest energy state above. Therefore, the simulated annealing algorithm steps are as follows:

Step 1: Set the initial temperature of simulated annealing algorithm and initialize the iteration times.

Step 2: Generate a new solution from the current solution, which is located in the solution space.

Step 3: Calculate the accurate function difference corresponding to the new solution.

Step 4: Metropolis criterion judges whether to accept the new solution.

Step 5: Determine whether the termination condition is satisfied, if yes, the algorithm ends, if not, go to step 2.

In the simulated annealing algorithm, Metropolis criterion accepts the deteriorating solution with a certain probability. Genetic algorithm (GA) [33], which is derived from the phenomenon of biological evolution in nature, is a global optimal algorithm based on Mendelian genetics and Darwin's biological evolution theory. The GA algorithm is described as follows:

Step 1: Coding.

In the GA algorithm, the genotype string structure data in the genetic space is used to represent the solution data in the solution space, and various points are composed of different combinations of these string structure data.

Step 2: Generate initial population.

Generate n initialized data in a random way. When each structured data is defined as an individual, n structured data form a group. Genetic algorithm starts iteration with n structured data as initial points. Generally, there are two cases that produce the initial solution, one is completely random, which is suitable for the case where there is no prior knowledge of the problem to be solved. The other is not completely random, which is suitable for the case that the problem to be solved has a certain understanding and the prior knowledge can be transformed into initial conditions, and then the sample data is randomly selected from the solution space that meets these requirements.

Step 3: Individual fitness.

Individual fitness is the standard of genetic selection, which is usually evaluated by fitness function which can express the merits and demerits of individuals, and can be defined differently for different problems. The possibility of individuals in the population being passed on to the next generation is determined by fitness function, That is to say, the higher the individual fitness, the more likely it is to be inherited into the next generation population; On the contrary, the lower the individual fitness, the less likely it is to be passed on to the next generation population so that the individual fitness function is the only criterion for genetic selection.

Step 4: Select.

Selection operation refers to selecting excellent individuals from the current population according to the fitness value of each individual and specific rules or methods, and passing them on to future generations. Selection operation embodies Darwin's genetic law of survival of the fittest to the greatest extent, and its purpose is to ensure that the excellent genes in the parent population are inherited to the offspring population, thus practically improving the global convergence. Reduce the possibility of genetic defects.

Step 5: Cross.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

The crossover operation simulates the crossover process of species genes, which means selecting two individuals from a population and exchanging one or more genes of the two individuals according to a greater probability. Cross-operation is a crucial link in genetic operation, which means that individuals with different genes can exchange some genes with each other, to produce a new generation of individuals combined with the excellent individual characteristics of their parents.

Step 6: Mutation.

Variation occurs in the evolution of natural species. The reason for the mutation is that some accidental factors affect the process of individual cell division and replication, which leads to the mutation of individual genes, and the individuals with gene mutation show new individual characteristics. Reasonable mutation is an important part of genetic algorithm. Mutation operation is similar to the mutation process in the genetic evolution of species. The method is to select one of all individuals, and then replace one or more gene values on a single coding string with other alleles with mutation probability so that mutation can produce new individuals.

3.3 SAGA-FCM Algorithm

If SA algorithm with intense local search ability and GA algorithm with good global search ability are combined, the fitness of GA algorithm can be extended appropriately. That is to say, when the temperature of SA algorithm is high, that is, the early stage of GA algorithm, the probability of offspring produced by individuals with similar fitness in GA algorithm is similar; While the temperature of SA algorithm continues to decrease, The stretching effect of GA algorithm is enhanced, and the advantages of outstanding individuals are more obvious. Therefore, the premature phenomenon of GA algorithm can be practically overcome.

Therefore, aiming at the problem that FCM algorithm is very sensitive to the initial value and easily falls into local extremum, this paper improves the traditional FCM algorithm based on the existing research, and puts forward the combination of genetic algorithm and simulated annealing algorithm (SAGA) to improve the problems existing in the process of selecting the initial clustering center of fuzzy C-means clustering (FCM) algorithm so that as to realize the reasonable division of financial time-series. And provide support for further predicting that correspond time-series. SAGA-FCM algorithm steps are as follows:

Step 1: Parameter initialization

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Determine the population individual size n, the maximum iteration times MAXGEN, and the crossover probability P_c, mutation probability P_m, initial temperature T_o, cooling coefficient k, termination temperature T_{end}.

Step 2: Initialize the cluster center, divide the initial population, and calculate the membership degree of each sample and the fitness degree of each sample according to formula (3) f_i, where I = 1, 2, n.

Step 3: Initial iteration times GEN=0.

Step 4: Carry out genetic operations such as crossover and mutation on the initial classification results. According to the operation results and formulas (3) and (5), the membership degree between the new clustering center and each sample and the fitness degree of each sample are calculated $f_i^{'}$. If $f_i^{'} > f_i$ Replace the old individual with the new individual; Otherwise, by probability $P = \exp((f_i - f_i^{'})T)$ Accept new individuals.

Step 5: If GEN<MAXGEN, then GEN=GEN+1, then go to step 4; Otherwise, go to step 6.

Step 6: If $T_i < T_{end}$, the algorithm ends and the optimal solution is output; Otherwise, right T_i Carry out cooling operation $T_{i+1} = kT_i$, go to step 3.

SAGA-FCM algorithm flow is shown in Fig 1.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

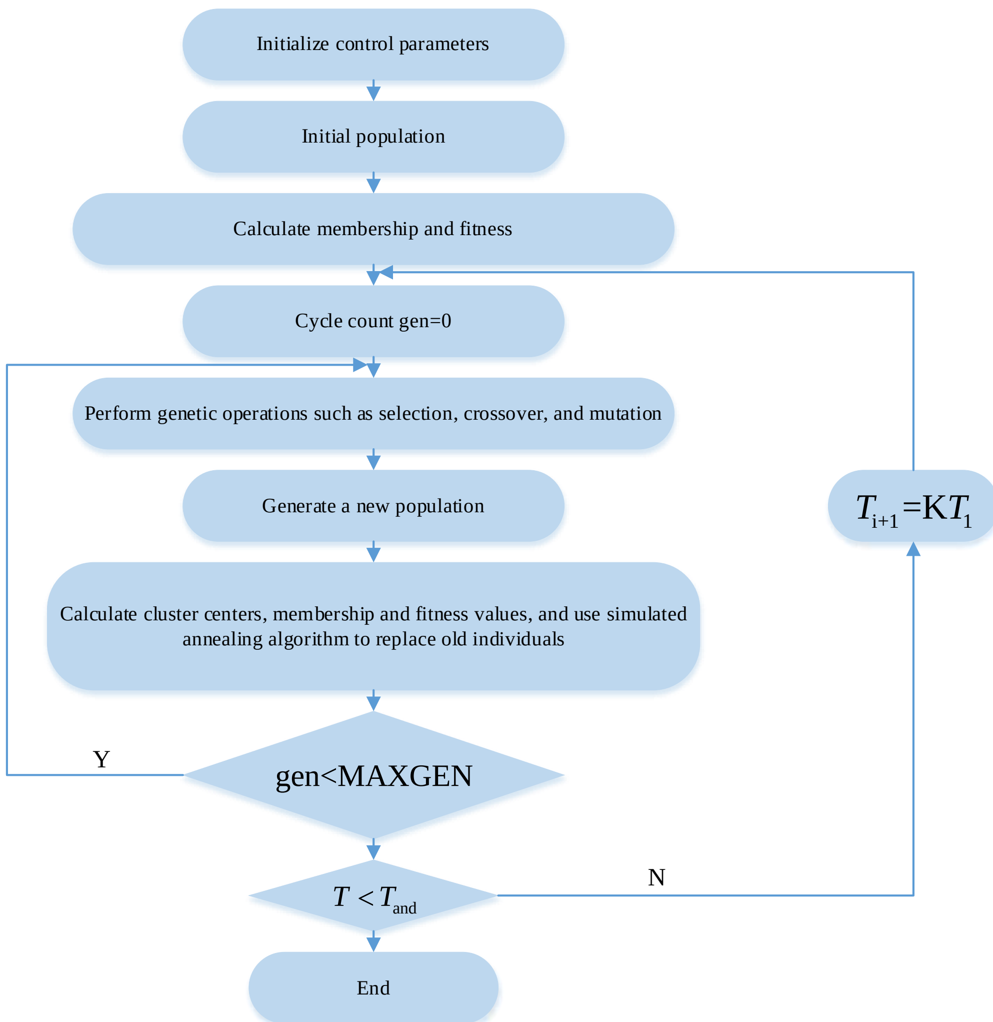

Fig 1: SAGA-FCM algorithm flow chart

IV. OPTIMIZED PREDICTION MODEL OF ELMAN NETWORK

4.1 Elman network

Feedback networks are an important type of neural networks. Due to the addition of between or within feedback connections, which can represent the time delay between input and output and therefore need to be described by dynamic equations, feed forward network implement only nonlinear mappings. It is precisely because of this feedback that the network has a memory function so that it was widely used in sequence analysis, system identification, control and other fields [34-39]. Elman network was originally studied by Elman in 1990

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

according to Jordan network [40]. Later, Pham et al. put forward modified Elman networks [41], which is now regarded as a standard Elman network. Its structure is sh in Fig 2.

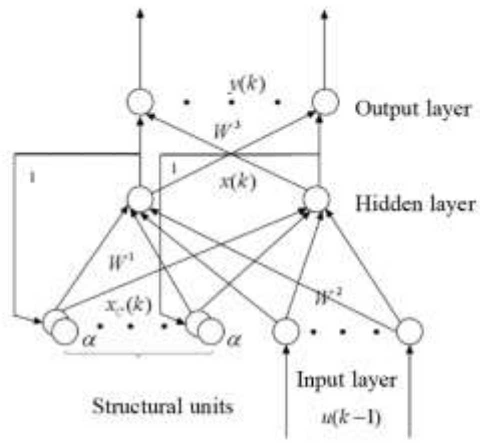

Fig 2: Elman network structure

Therefore, here, the feed forward connection part can make connection weight correction, while the recursive part is fixed, that is, it cannot make learning correction. Specifically, the structural unit k the output of the moment is equal to the hidden layer $k-1$ Output at time plus structural unit $k-1$ of the output value at the moment a Times, that is:

$$x_{C,l}(k) = a \times x_{C,l}(k-1) + x_l(k-1), l = 1, 2, ..., n \tag{6}$$

$x_{C,l}(k)$ and $x_l(k)$ are the outputs of the first structural unit and the first hidden layer unit respectively, a is the self-connected feedback gain factor. When a fixed to zero, this network is a standard Elman network, a If it is not zero, it is a modified Elman network. The mathematical model of the network is:

$$x(k) = f(W^1 x_c(k) + W^1 u(k-1)) \tag{7}$$

$$x_c(k) = a \times x_c(k-1) + x(k-1) \tag{8}$$

$$y_k = g(W^3 x(k)) \tag{9}$$

Most of $f(x)$ is sigmoid functions, i.e.:

$$f(x) = \frac{1}{1+e^{-x}} \tag{10}$$

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Most of $g(x)$ is linear functions, namely:

$$y_k = W^3 x(k) \tag{11}$$

4.2 Learning Algorithm of Elman Network

The error function is defined as:

$$E = \frac{1}{2}(y_d(k) - y(k))^T (y_d(k) - y(k)) \tag{12}$$

E right to connect hidden layer to output layer W^3 Seek partial derivative, get:

$$\frac{\partial E}{\partial w_{ij}^3} = -d_i^0 x_j(k), i = 1,2,...,m; j = 1,2,...,n \tag{13}$$

Make $d_i^0 = (y_{di}(k) - y(k)) g_i(\cdot)$, then

$$\frac{\partial E}{\partial w_{ij}^3} = -d_i^0 x_j(k), i = 1,2,...,m; j = 1,2,...,n \tag{14}$$

E right to connect the input layer to the hidden layer W^2 Seek partial derivative, get:

$$\frac{\partial E}{\partial w_{ij}^2} = \frac{\partial E}{\partial x_j(k)} \frac{\partial x_j(k)}{\partial w_{ij}^2} = \sum_{i=1}^{m} (-d_i^0 w_{ij}^3) f_j^{'}(\cdot) u_q(k-1) \tag{15}$$

Same order $d_i^h = \sum (d_i^0 w_{ij}^3) f_j^{'}(\cdot)$, there are:

$$\frac{\partial E}{\partial w_{ij}^3} = -d_j^h u_q(k-1), j = 1,2,...,n; q = 1,2,...,r \tag{16}$$

Similarly, the right to connect structural units to hidden layers W^1 Seek partial derivative, get:

$$\frac{\partial E}{\partial w_{jl}^1} = \sum_{i=1}^{m} (-d_i^0 w_{ij}^3) \frac{\partial x_j(k)}{\partial w_{jl}^1}, j = 1,2,...,n; l = 1,2,...,n \tag{17}$$

Note that in the above formula:

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

$$\begin{aligned}\frac{\partial x_j(k)}{\partial w_{jl}^1} &= \frac{\partial}{\partial w_{jl}^1}\left(f_j\left(\sum_{i=1}^{n} w_{jl}^1 x_{Cl}(k) + \sum_{i=1}^{r} w_{jl}^2 u_i(k-1)\right)\right) \\ &= f_j^{'}(\cdot)\left(x_{Cl}(k) + \sum w_{jl}^1 \frac{\partial x_{Cl}(k)}{\partial w_{jl}^1}\right)\end{aligned} \tag{18}$$

Substitute (17) to get:

$$\frac{\partial x_j(k)}{\partial w_j^1} = f_j^{'}(\cdot)x_l(k-1) + \alpha\frac{\partial x_j(k-1)}{\partial w_{jl}^1} \tag{19}$$

$\Delta W = -\eta\dfrac{\partial E}{\partial W}$ so that the improved Elman learning algorithm can be derived as follows:

$$\Delta w_{ij}^3 = \eta\delta_i^0 x_j(k), i = 1,2,...,m; j = 1,2,...,n \tag{20}$$

$$\frac{\partial x_j(k)}{\partial w_j^1} = f_j^{'}(\cdot)x_l(k-1) + \alpha\frac{\partial x_j(k-1)}{\partial w_{jl}^1} \tag{21}$$

$$\Delta w_{ij}^3 = \eta\sum_{i=1}^{m}(\delta_i^0 w_{ij}^3)\frac{\partial x_j(k)}{\partial w_{jl}^1}\delta_i^0 x_j(k), i = 1,2,...,m; j = 1,2,...,n; l = 1,2,....,n \tag{22}$$

Among them:

$$\delta_i^0 = (y_{di}(k) - y(k))g_i^{'}(\cdot) \tag{23}$$

$$\delta_j^h = \sum (\delta_i^0 w_{ij}^3) f_i^{'}(\cdot) \tag{24}$$

$\dfrac{\partial x_j(k)}{\partial w_{jl}^1}$ is obtained by recursion of formula (19).

4.3 Improved Elman Network Model

In the Elman network model in the previous section, feedback from nodes at the output layer is not considered, only feedback from nodes at the hidden layer is considered. The two network models we give below both consider the feedback of output nodes. The structural schematic diagrams are sh in Fig 3 and Fig 4 respectively. Among them, OIF Elman network increases the feedback of output nodes, which we call structural unit 2, and puts it in the first layer, which is used as the input of hidden layer nodes together with input unit and structural unit. However, OHFElman network only puts structural unit 2 on the second layer, and takes it as the input of the output layer together with the hidden layer nodes.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

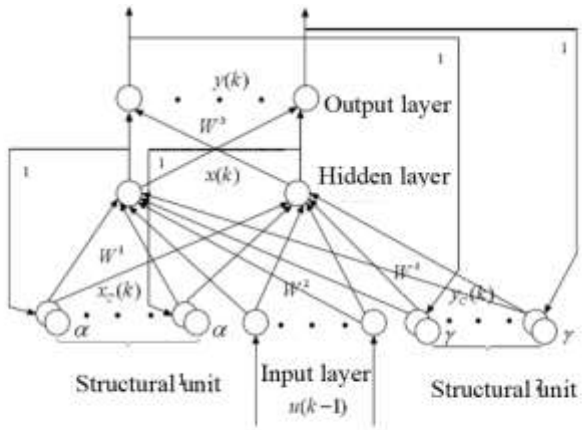

Fig 3: OIF Elman network structure

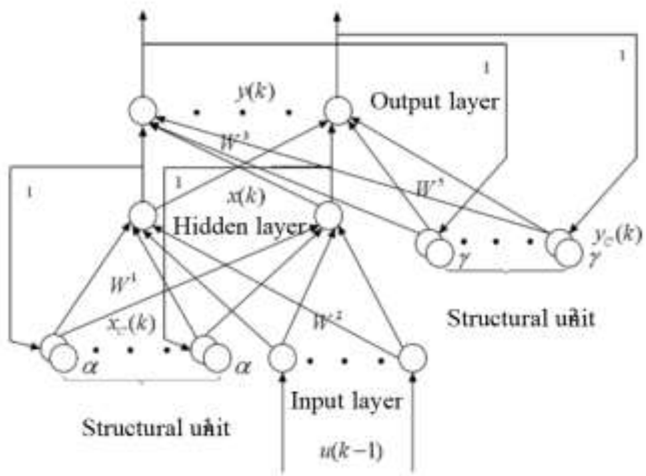

Fig 4: OHF Elman network structure

The mathematical model of OIF Elman network is:

$$x(k) = f(W^1 x_c(k) + W^2 u(k-1) + W^4 y_c(k)) \tag{25}$$

$$x_c(k) = a \times x_c(k-1) + x(k-1) \tag{26}$$

$$y_c(k) = a \times y_c(k-1) + y(k-1) \tag{27}$$

$$y(k) = g(W^3 x(k)) \tag{28}$$

The mathematical model of OHF Elman network is:

$$x(k) = f(W^1 x_c(k) + W^2 u(k-1)) \tag{29}$$

$$x_c(k) = a \times x_c(k-1) + x(k-1) \tag{30}$$

$$y(k) = g(W^3 x(k) + W^5 y_c(k)) \tag{31}$$

$$y_c(k) = a \times y_c(k-1) + y(k-1) \tag{32}$$

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

OIF Elman network only adds structural unit 2 as input in the hidden layer nodes of Elman network so that E Yes W^1, W^2, W^3 There is no change in partial derivativeso that their weight correction formula is the same as Elman network, see formulas (20)-(22) for details. Will E We can get the weight by finding the partial derivative of the connection weight between structural unit 2 and hidden layer and making it equal to zero W^4. The modified formula of:

$$\Delta w_{jl}^4 = \eta \sum_{i=1}^{m} (\delta_i^0 w_{ij}^3) \frac{\partial x_j(k)}{\partial w_{jl}^4}, j = 1,2,...,n; l = 1,2,...,n \tag{33}$$

$$\frac{\partial x_j(k)}{\partial w_{jl}^4} = f_j'(\cdot) y_l(k-1) + \alpha \cdot \frac{\partial x_j(k-1)}{\partial w_{jl}^4} \tag{34}$$

δ_i^0 is given by formula (23).

Similarly, for OHF Elman networks, W^1, W^2, W^3 The weight correction formula of is the same as Elman network, and will be E By taking the partial derivative of the connection weight from structural unit 2 to the output layer and making it equal to zero, we can get the modified formula of the weight:

$$\Delta w_{ij}^5 = \eta \delta_i^0 y_{c,j}(k), i = 1,2,...,m; j = 1,2,...,n \tag{35}$$

δ_i^0 is given by formula (23).

4.4 Multiple Financial Time-series Forecasting Model Based on FCM-Elman

The improved FCM clustering algorithm in Section 2 and Section 3 and the improved Elman network in Section 4.3 are used to predict the financial time-series. The specific steps of the multivariate financial time-series prediction model constructed in this paper are as follows:

(1) Using the improved FCM clustering algorithm to cluster all the sequences in the stock time-series set, and dividing them into big categories $A, B, C, ...$;

(2) According to the need to choose to forecast the stock, as a stock A_1, which be extended to the following categories A Major categories;

(3) The stock A_1 the improved Elman network is used to predict and debug the appropriate network parameters.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

(4) Elman network pairs with the same parameters A Train and forecast other stocks in the major categories.

V. ALGORITHM EVALUATION AND RESULT ANALYSIS

In order to verify the scientificity and effectiveness of the proposed method, a large number of sample data were selected for analysis and discussion, as follows:

5.1 SAGA-FCM Model

5.1.1 Data set

Dataset 1: SCCTS (Synthetic Control Chart Time-series) is a network open test set, which is specially used for the algorithm test of time-series clustering. SCCTS has 600 test which beextended to 6 different categories. Ten groups of data are extracted from the data set (SCCTS). Each group of data contains 60 time-series, which be extended to 6 different categories, with 10 items in each category, and each data is normalized and preprocessed, Make it become a number between 0 and 1. In this paper, the data processing mainly adopts the normalization method as sh in Formula (35), where n is the number of data samples.

$$x_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}, i = 1, 2, ..., N \tag{36}$$

Data set 2: 10 groups of data in data set 4-1 are processed with unequal length so that that the lengths of 60 time-series in each group of data are not completely the same without affecting the clustering results, and each data is normalized and preprocessed.

Data set 3: KDD (1999) data set is used to classify and detect attacks or normal network behaviors. There are 41 features in the data set, of which 34 are continuous and the remaining 7 are classified communication. Data sets are divided into two categories: normal data (no attack) and attack data. Only 11 of the 41 features are used in this paper. Compared with other continuous features, the selected 11 continuous features have more important value. There are 4,898,431 records in the data set, of which 3,925,651(80.1%) records are classified as attacks. The rest of the data is classified as normal. In this paper so that me data are selected from the data set to evaluate the clustering algorithm. A total of 26,829 data are selected, of which 25,620 are classified as attacks.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

5.1.2 Empirical research results

Comparing the clustering results of standard data sets with other three well-kn data clustering algorithms, the three data clustering algorithms are: k-means algorithm, hierarchical clustering method, and clustering algorithm based on SNN density. Table I shows the clustering results of each algorithm, and the evaluation standard is the wrong classification rate. With regard to k-means algorithm and SAGA-FCM algorithm, the relevant classification numbers are provided for each data set. For k-means algorithm, the clustering program is stopped after 10 iterations.

Table I. Comparison of the error rates of the four algorithms

Data set	SAGA-FCM algorithm	K-means	Hierarchical clustering method	SNN density clustering
1	4.37	10.62	11.78	7.85
2	5.89	12.77	14.65	10.74
3	1.03	1.27	1.29	1.23

5.1.3 Result analysis

Although the clustering algorithm based on GASA-FCM proposed in this paper is superior to the clustering results of other three clustering algorithms for a given data set, it is worth noting that when there are clusters overlapping (that is, it is difficult to clearly determine the cluster boundary), the performance of the clustering algorithm based on GASA-FCM may be reduced (that is, the wrong classification rate may increase).In order to solve this problem, it is a good choice to introduce other sensitive factors to determine clear boundaries.

5.2 Evaluation and Analysis of Improved Elman Network Prediction Model Algorithm

5.2.1 Data set

Three significant-scale indexes, SSE Index (SSEC), SZI and S&P 500, are used as experimental data to train and test the model. These three significant-cap indexes have certain representativeness in the financial markets of China and the United States. The reason why the market index is chosen as the experimental data is that the market is the concentrated expression of the fluctuation of the whole financial market. The data is loaded from Flush Finance (http://www.10jqka.com.cn/) website. When selecting the input of the model, it is

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

necessary to determine the time scale of selecting financial time-series. In this paper, the analysis and prediction of financial time-series are mainly aimed at daily so the day is used as the basic time scale, that is, arbitrary time t corresponds to one day. Time-series $X(t)$ Represents the first t Days of including the opening price, the highest price, the lowest price, the closing price, the amount of rise and fall, and the six price components and trading volume. Use the above five components as input and the opening price as output. Both the training set and the test set are normalized in the input process, and the least square error is used to express the prediction accuracy of the network. The absolute average error (AAE) formula (37). See table II for detailed experimental data.

$$AAE = \frac{1}{N}\sum_{i=1}^{N} |y_d(i) - y(i)|, i = 1, 2, ..., N \tag{37}$$

Table II. The experimental data

Stock index		Data selection range	Total sample size	Training sample size	Test sample size
Chinese stock market	Shanghai Stock Exchange Index (SSEC)	1990.12.20-2017.3.10	6412	5412	1000
	Shenzhen Stock Exchange Index (SZI)	1991.4.3-2016.3.16	6126	5126	1000
American stock market	Standard & Poor's Index (S&P500)	2001.3.5-2014.5.1	3295	2295	1000

5.2.2 Empirical research results

Compare the three groups of data with Elman network, OIF Elman network, OHF Elman network and traditional BP neural network respectively. Because there are many test it is inconvenient to put The prediction results of each group of networks in one picture so that compare The prediction accuracy of the networks by the size of least square error, as sh in Table III.

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Table III. Predicted performance results of various networks

Data set	BP	Elman	OIF Elman	OHF Elman
SSEC	0.00445	0.00325	0.00116	0.00213
FALL	0.00469	0.00397	0.00160	0.00231
S&P500	0.00458	0.00371	0.00164	0.00246

5.2.3 Result analysis

It can be clear from the simulation results in Table III that the prediction accuracy of OIF Elman neural network is obviously better than that of traditional BP neural network and Elman network, and compared with OHF Elman network, OIF Elman network has better prediction performance. From the absolute average error table, it can be seen that the order of prediction effect of four stock index prediction models on SSEC, SZI and S&P500 stock indexes is OIF Elman, OHF Elman, Elman and BP network. Fig 5 and fig 6 are the prediction curves of Shanghai Composite Index (SSEC) and Standard & Poor's Index (S&P500) in three data sets by OIF Elman network, respectively.

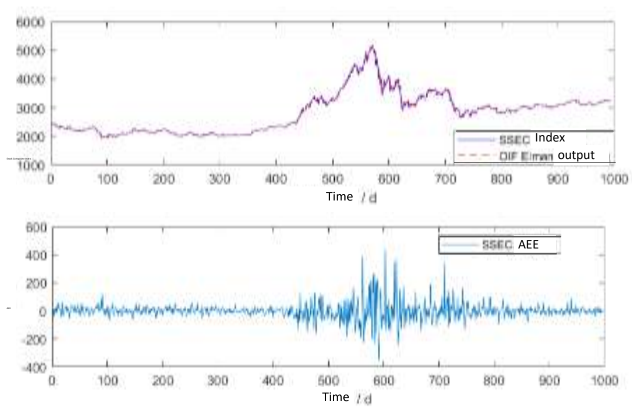

Fig 5: SSEC prediction curve and test residuals

It can be clear from fig 5 that when The prediction time is less than 400 days, the residual error of OIF Elman network to SSEC index is less than 100, and after 400 days, the residual error changes dramatically, which may be due to the margin financing and securities lending in China in 2015. At this time, the abundant funds in China's stock market led to a sharp rise in stocks. As China began to control margin financing and securities lending, the market fell into a dturn. In addition, with the economic return in China at this time, the SSEC index tends to be

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

stable, and the residual value gradually narrows to within 200. The prediction curves of OIF Elman network for SSEC are highly coincident.

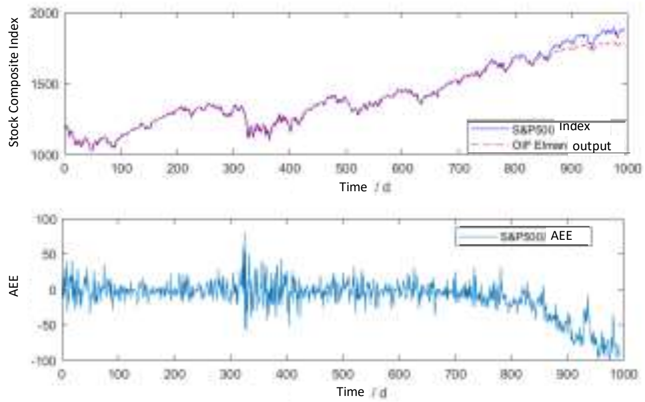

Fig 6: S&P500 prediction curve and test residuals

It can be clear from fig 6 that when The prediction time is within 300 days, the residual error of OIF Elman network to S&P500 index is within 50, but after 300 days, the residual error changes dramatically, and the reason for the sharp decline and shock of US stocks may be related to the European debt crisis in 2011.The subsequent rise of US stocks may be related to the loose monetary policy favorable to investors issued by the Federal Reserve after 2013 and the overall improvement of the US economy so that that the residual value gradually increased to about 100. The prediction curve of OIF Elman network for S&P500 is highly consistent in short-term and medium-term prediction, but there is a big deviation in extended-term prediction.

According to the above experiments, it is feasible and practical to forecast the comprehensive index of stocks by OIF Elman network. The prediction accuracy is higher than Elman network. From the experimental results, we can see that OIF Elman artificial network model has strong applicability in time series prediction, and the prediction accuracy can be obviously improved. Therefore, it is an practical forecasting tool in the financial field, which provides a new and practical way to forecast the stock market and has certain application value. However, the prediction accuracy of the network is dramatically influenced by the overall situation of policy and economy so that the influence of other factors will be added in the following chapters. In order to improve the accuracy of the model in extended-term financial time-series prediction.

5.3 Evaluation and Analysis of FCM-Elman Multivariate Time-series Prediction Model

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

5.3.1 Data set

This example adopts the stock data set of China's Shanghai and Shenzhen A-stock markets. There are currently 3782 stocks in the Shanghai and Shenzhen A-stock markets. Due to the significant number, the first 1000 stocks in the default order of the A-stock markets are taken as the data set. The data is loaded from Flush Finance (http://www.10jqka.com.cn/）. In the section experiment, six price components and trading volume of each stock are selected as Elman network input and opening price as output, which are also normalized. Based on the data from January 2017 to January 2019, the stock price in February 2019 is predicted.

5.3.2 Empirical research results

Using the improved FCM clustering algorithm to cluster, 1000 stocks are classified with 50 categories. Cluster similar stocks in the same category.

All stock time-series in the cluster of Changjiang Media (600757) are predicted by improved Elman network. There are 19 stocks in Changjiang Media's category, and the names and codes of the remaining 18 stocks are: China Sports Industry (600158), Yuye Stocks (600226), Dr. Peng (600804), Zhongchang Data (600242), Mengzhou Stocks (600255), cits joint (600358), Beiba Media (600386), Times Publishing (600555) Changjiang Communication (600345), Jinglun Electronics (600355), Fiberhome Communications (600498), Jichuan Pharmaceutical (600566), Zhongzhu Medical (600568), Baichuan Energy (600681) and Sanan Optoelectronics (600703).

Fig 7 is the prediction curve of stock price of Changjiang Media by OIF Elman network and the residual error of test results.

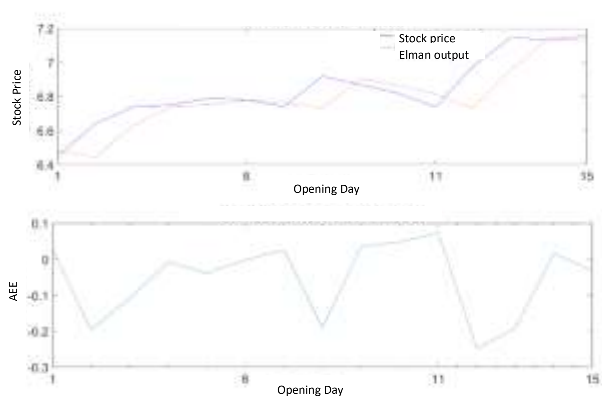

Fig 7: Changjiang Media prediction curve and AEE

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

The least square error of each stock in Changjiang Media Group is brought to test the prediction accuracy of the network. Because of the significant number of stocks, 19 stocks are divided into multiple accuracy ranges, as sh in Table IV.

Table IV. Least squares error of each stock

AEE intervals	Stock name
<0.005	China Sports Industry and Zhongchang Data
0.005-0.007	Changjiang Media, Mengzhou Stocks, Beiba Media, Time Publishing, Dr. Peng, Three Gorges New Materials and Yuye Stocks
0.007-0.009	Cits joint, Angel Yeast, Changjiang Communication, Fiberhome Communication, Kaile Technology
0.009-0.011	Zhongzhu Medical, Jinglun Electronics
> 0.011	Jichuan Pharmaceutical, Baichuan Energy and Sanan Optoelectronics

It can be seen from fig 7 that OIF Elman network has a certain lag in forecasting individual stocks, but the overall trend of stocks is basically the same. The AEE values of most stocks in Table IV are concentrated in the range of 0.005-0.009, and OIF Elman network has a poor performance in the comprehensive stock index, but considering the fluctuation uncertainty of a single stock, it is still a very accurate prediction result.

VI. CONCLUSIONS

Based on the current changes in the financial market and intelligent algorithms, this article has drawn the following and research conclusions, which are as follows:

(1) Several clustering methods commonly used in time-series are proposed, and then the fuzzy C-means (SAGA-FCM) algorithm based on simulated annealing genetic algorithm is applied to the clustering analysis of time-series, and good clustering results are obtained.

(2) The basic Elman network and its two improved networks are introduced. By forecasting the financial time-series. It is proved that OIF Elman network is more suitable for the prediction of financial time-series.

(3) Combining SAGA-FCM clustering algorithm and OIF Elman network, the stock data of Shanghai and Shenzhen A-stock trading markets are classified and predicted, and a significant-scale stock prediction method is realized. By clustering first and then forecasting, similar stocks can be better classified into one class. Training with unified network structure can not only

improve the accuracy of financial time-series prediction, but also expand the prediction range. The accurate prediction results are obtained by experiments.

In addition, combined with the conclusions of this paper, combined with the current development of intelligent algorithms, follow-up research continues to track research, making the prediction results better and closer to the actual situation.

REFERENCES

[1] Dong J, Wang ZH (2015) Overview of time-series representation and classification algorithms. Computer Science 42(3):1-7
[2] Yang HM, Pan ZS, Bai W (2019) Review of time-series prediction methods. Computer Science 46(1):21-28
[3] Wang Z extended (2006) Time-series analysis. Beijing: China Statistics Press
[4] Boginski V, Butenko S, Pardalos PM. (2003) On structural properties of the market graph. Innovations in financial and economic networks 48: 29-35
[5] Chen H, Sun QM. Directed Network Research in Chinese Stock Market. *IEEE International Conference on Emergency Management & Management Sciences, IEEE* 2011: 883-886
[6] Zhang MF, Dai YF. Research on the Strategy of Government Supervision to Promote the Prudent and Healthy Development of Consumer Finance——Based on the Time Series Analysis of T City Samples. Modern Commerce, 2021(27), 162-168.
[7] Yi WD, Huang AH (2021) Summarize on Applications of Copula Theory in Financial Management. *Journal of Chongqing University of Arts and Sciences(Social Sciences Edition)* 09, 17, https://kns.cnki.net/kcms/detail/50.1182.C.20210917.1052.008.html
[8] Ma T, Wang, J Ding SW et al. Combination Forecasting Model of Interval Financial Time Series from Multi-Scale Perspective. Finance and Tax, 2021(9), 59-66
[9] Yan HJ Financial Time Series Prediction: A Literature Review and Prospect. 2021(5), 33-41
[10] Cui GW Research on the Impact of Financial Development on Income Inequality——Based on Macro Time Series Data from 1973 to 2017 in the United States. Operation and Management 2021 (05), 145-150
[11] Yang Q, Wang CW (2019) Research on global stock index prediction based on deep learning LSTM neural network. Statistical Research 36(3):65-77
[12] Ieee fellow (2019) Design and Implementation of Financial Data Analysis System Based on Deep Learning. Ningxia University
[13] Shen F, Chao J, Zhao J (2015) Forecasting exchange rate using deep belief networks and conjugate gradient method. Neurocomputing 167:243-253
[14] Di Persio L, Honchar O. Artificial Neural Networks Architectures for Stock Price Prediction: Comparisons and Applications. *International Journal of Circuits, Systems and Signal Processing* 2016(10): 403-413

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 607-635
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

[15] Arévalo A, Niño J, Hernández G, et al. High-Frequency Trading Strategy Based on Deep Neural Networks. International Conference on Intelligent Computing. Berlin: Springer International Publishing, 2016:424-436
[16] Fischer T, Krauss C (2017) Deep learning with extended short-term memory networks for financial market predictions. European Journal of Operational Research 270(2): 654-669
[17] Wei B, Jun Y, Yu l R, et al. (2018) A deep learning framework for financial time-series using stacked autoencoders and extended-short ter Choi HK. Stock price correlation coefficient prediction with ARIMA-LSTM hybrid model
[18] Bai X , Wang X , Liu X , et al. Explainable Deep Learning for Efficient and Robust Pattern Recognition: A Survey of Recent Developments. Pattern Recognition, 2021(2):108102.
[19] Zhang YA, Yan BB (2020) Deep Learning Hybrid Forecasting Model for Stock Market. Computer Science 47(11), 255-267
[20] Yang Q, Wang CW. (2019) Research on global stock index prediction based on deep learning LSTM neural network. Statistical Research 36(3):65-77
[21].Zeng ZP, Xiao HD, Zhang X P. (2017) Modeling and decision-making of financial time-series data based on DBN. Computer Technology and Development 27(4):1-5
[22] Zhao HK, Wu LK, Zheng L, Zhang X, Liu Q, Chen EH (2019) Dynamic prediction of Internet financial market based on deep neural network structure. Computer Research and Development 56(8):1621-1631
[23] Fu ZJ (2018) Research on Stock Alpha Quantitative Model Based on Deep Learning. Xiamen University
[24] Stoean R, Stoean C, Sandita A (2017) Evolutionary Regressor Selection in ARIMA Model for Stock Price Time-series Forecasting. Smart Innovation, Systems and Technologies, 117-126
[25] Huang J, Breheny P, Ma S, et al. (2010) The mnet method for variable selection. Department of Statistics and Actuarial Science, The University of Lowa
[26] Nonejad N (2017) Forecasting aggregate stock market volatility using financial and macroeconomic predictors: Which models forecast best, when and why? *Journal of Empirical Finance* 42:131-154
[27] Lin YF, Huang CF, Tseng VS. A Novel Methodology for Stock Investment using High Utility Episode Mining and Genetic algorithm. Applied Soft Computing, 2017, 59: 303-315.
[28] Mitchell T, Buchanan B, Dejong G, et al. (2003) Machine Learning. Beijing: China Machine Press
[29] Fischer T, Krauss C (2018) Deep learning with extended short-term memory networks for financial market predictions. *European Journal of Operational Research* 270(2): 654-669
[30] Cavalcante RC, Brasileiro RCso thatuza V L F, et al. (2016) Computational Intelligence and Financial Markets: A Survey and Future Directions. Expert Systems with Applications, , 55: 194-211
[31] Rosenblatt F. Perceptron simulation experiments. Proceedings of the IRE, 1960, 48(3): 301-309.
[32] Zhou KL. (2005) Neural network model and MATLAB simulation program design. Beijing: Tsinghua University Press
[33] Rosenblatt F (1960) Perceptron simulation experiments. Proceedings of the IRE 48(3): 301-309
[34] Widrow B, Hoff ME (1960) Adaptive Switching Circuits. Neurocomputing 4: 126-134
[35] Sun ZJ, Xue L, Xu YM, Wang Z (2012) A summary of deep learning research. Computer Applied Research 29(8):2806-2810

[36] Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proceedings of the national academy of sciences 79(8): 2554-2558
[37] Jordan MI. Serial Order: A Parallel Distributed Processing Approach. Advances in psychology, 1997: 471-495.
[38] Elman J L (1990) Finding structure in time. Cognitive science, 14(2): 179-211
[39] Chen P, Chang L, Chang F, et al. (2013) Reinforced recurrent neural networks for multi-step-ahead flood forecasts. Journal of Hydrology 497:71-79
[40] Hochreiter S, Schmidhuber J (1997) Extended short-term memory. Neural Computation 9(8): 1735-1780
[41] Chung J, Gulcehre C, Cho K, et al. Gated feedback recurrent neural networks. International Conference on Machine Learning. 2015: 2067-2075