

Research on Learning Prediction Model Based on Optimized Rbf Network

Li Qi

Chongqing Vocational Institute of Engineering 402260, Chongqing, China

Keywords: Rbf network, Learning prediction, Neural network, Sequence, Nonlinearity

Abstract: Learning prediction is an important research content of data mining. It is widely used in many fields to describe the potential relationship between prediction indicators and influencing factors. The classical prediction methods have some difficulties in the application of nonlinear system prediction, while the RBF neural network has better nonlinear characteristics, which is especially suitable for highly nonlinear systems, and opens up a new development space for multi-factor time series prediction. In this paper, the prediction model based on RBF neural network is studied in depth, and the dimensionality reduction reconstruction of network input space is studied in detail. RBF neural network was used for modelling training, and the results were compared with BP network. The training speed of RBF network was significantly faster than that of BP network, with better generalization ability. The simulation experiment showed that it could be effectively applied to multi-factor time series prediction.

1. Introduction

In recent years, time series analysis has developed rapidly in theory and application, and has become an important branch of statistics and data mining. In the objective world, the development of things is the result of many intricate factors. The change and development of one factor in the system must be affected by other factors. This kind of situation reflects in the forecast work, requests to consider its forecast value and the correlation influence factor forecast value relations.

Multivariate statistical analysis is used to analyse the correlation between various variables and the degree of influence of each variable on the research index. The study of time series with certain correlation as a whole, namely multi-variable (factor) time series analysis, can better understand the characteristics of each time series and reduce the complexity of the problem. With the development and application of artificial neural network theory, it is possible to establish the nonlinear mapping model of multi-variable time series with high accuracy. Neural network can realize nonlinear mapping, has good pattern recognition ability and nonlinear prediction ability of arbitrary precision approximation, and can be used to establish nonlinear mapping between various indicators and influencing factors, to reflect the relationship between multi-variables, and has a good prediction effect. In time series data mining, the description, analysis and prediction of multivariate time series with the combination of multivariate statistical analysis and other mining techniques have always been a research topic widely concerned by people.

Data mining is extracted from a large database implied in it, the unknown, and potentially useful information in advance, the extracted knowledge representation for the concept, rule, rule and mode, etc., can help decision makers analyse historical and current data, discover hidden relationships and patterns, then possible to predict the future behavior. The combination of computer technology, statistical methods and various algorithms promotes the rapid development of data mining technology. Before data mining was formally proposed as a research topic, statistical analysis technology has been the main means for people to process data and find rules. Statistics plays an important role in data sample selection, data pre-processing, data mining and knowledge extraction evaluation. Data mining has fully absorbed a variety of statistical analysis technology, combined with its own characteristics, so that it becomes a part of data mining technology.

Multivariate Statistical Analysis, also known as Multivariate, Multivariate, and Multivariate Analysis, is a Statistical discipline that studies the interdependence and inherent Statistical regularity between multiple random variables. Multivariate statistical methods have been widely

used in various fields. At home and abroad, from natural science to social science in many aspects of the actual application has confirmed that it is a very useful data processing method, and has made great achievements, by the general praise.

Radial basis function neural network (RBFNN) is a kind of commonly used three-layer feed forward network, which can be used for both function approximation and pattern classification. Compared with other types of artificial neural networks, RBF networks have the characteristics of physiological basis, simple structure, fast learning speed and excellent approximation performance.

RBF network is a three-layer feed forward network with single hidden layer proposed by Monody and Darken in the late 1980s, in which each layer has a completely different role. The input layer consists of a number of sensing units, which connect the network with the external environment. Different from BP network, there is only one hidden layer in the second layer of RBF network, whose function is to conduct nonlinear transformation from the input space to the hidden layer space. In most cases, the hidden layer space has a higher dimension, which is determined by the need. The third layer is the output layer, which is linear and provides a response to the activation mode applied to the input layer.

The basic radial basis function (RBF) network is a three-layer feed forward network with a single hidden layer, and its topology is shown in figure 1. Since it simulates the neural network structure of local adjustment and mutual coverage in human brain, RBF network is a kind of local approximation network, and its activation function of hidden nodes is a radial basis function, which is a local distribution nonlinear function of symmetry about the center point.

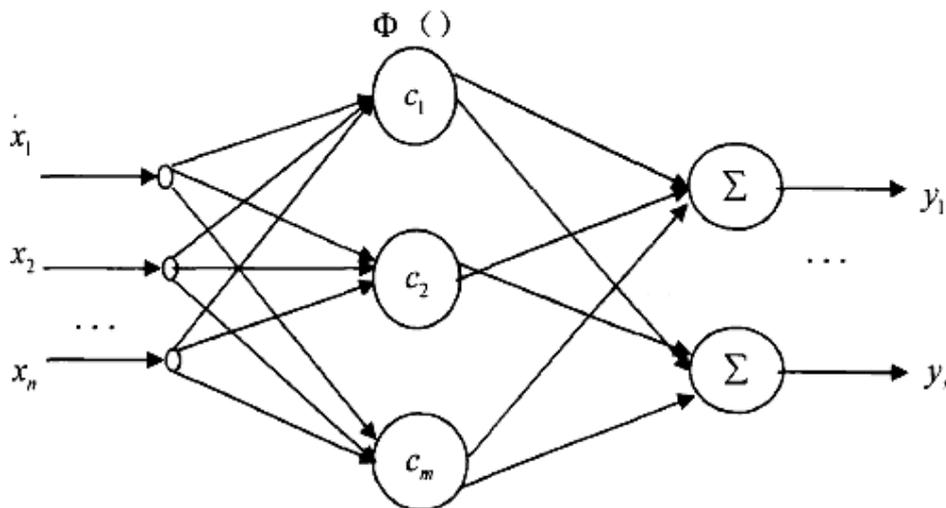


Fig.1 Radial Basis Function Neural Network

RBF network is widely used in many fields, such as function or signal approximation, data or graph compression, pattern recognition, nonlinear system modelling and mechanical fault signal analysis and processing, mainly based on the following functions : (1) any continuous function can be expressed as a weighted sum of a group of radial basis functions, RBF can approximate functions. (2) RBF can be used as a classifier of linear indivisible patterns. The original linear inseparable mode is transformed into a new high-dimensional space after being processed by the hidden layer and becomes linear separable, and the linear classification is realized by the output layer. Choosing the right radial basis function is the key to realize the above functions. The radial basis function is a locally distributed, radial symmetric, non-negative decaying nonlinear function. Its two parameters, the base centre and the base width, will determine the range of significant response to the input. Radial basis functions include Gaussian function, multi-quadratic function, inverse multi-quadratic function and thin plate spline function, etc., all of which have good approximation ability, among which Gaussian function is the most commonly used one.

2. Experimental Procedure

Time series is the basis of prediction and the external manifestation of the behavior characteristics of things under investigation. It is through the analysis of time series that the forecaster excavates the internal regularity of things and makes predictions on the development trend of things by combining the existing or future conditions. In the prediction work, sometimes the behavioural characteristics of a certain aspect of the system are mapped into a time series, that is, a single factor time series. Based on this time series, modelling and prediction are conducted to study the changes and future development trend of a certain behavioural characteristic of the system in isolation.

In order to build the learner model, a large amount of learner information data must be collected and pre-processed. Therefore, all the required historical data in the model need to be analysed and processed in order to fully dig out the key behavior characteristics of learners. This study selected three types of characteristic data: (1) basic characteristics of learners, namely, learner data. The characteristic data set includes 2, 841 records, each record including 21 characteristics. (2) learner behavior characteristics, that is, learner learning behavior characteristics data, the feature data set contains 9 tables, 2, 548, 709 records, 137 behavior characteristics; (3) learners' learning effect characteristics, namely learners' scores, this feature set contains 6 tables, 55 and 957 records.

Data pre-processing is an important and non-negligible link in the early stage of data mining. According to statistics, in the complete data mining process, data pre-processing takes about 60% of the time, and the subsequent mining only accounts for about 10% of the total workload. Data pre-processing mainly includes data cleaning, data integration, data transformation and data specification. Through data pre-processing, this study extracts 9 features, which are learning time span, that is, the time period from the first time the learner logs in the system to the last time he logs out of the system. The total learning time, that is, the learning time accumulated by learners on the network learning platform; Number of learning, that is, the number of times learners log on the learning platform; The average length of stay of each online learning, that is, the average length of each learning period of the learner; Repeat learning rate, that is, the average number of times learners repeat learning for each knowledge point; Discussion and communication, that is, the characteristics of participating in BBS discussion and communication; Learning notes, that is, the characteristics of learners' participation in digital learning notes; The number of received feedback, that is, the number of learners receiving urging information in the learning process; Total score, that is, the learner learning total score.

Data cleaning, that is, processing noise data, irrelevant data, missing values, dirty data, and outliers and so on in the original data. For example: education master's degree information is wrong, correct through SQL language; Delete the information of registered but not learning learners, information of exchange students and rest students, and the information of learners whose examination authority is 0; Delete information of teachers, other users and managers; Delete vacancy value information, etc. The data mining results of the total score decision tree are shown in figure 2.

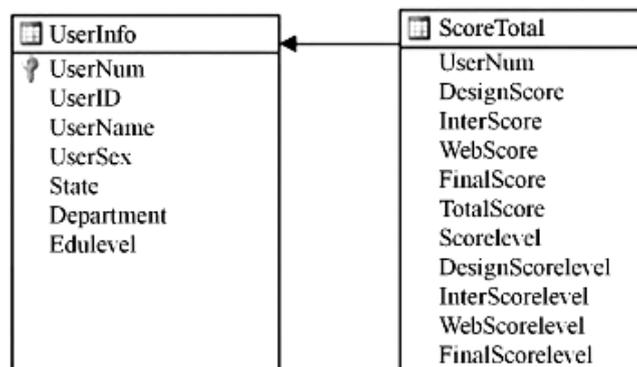


Fig.2 Mining Structure of Decision Tree of Learners' Total Scores

In this approach, the centre of the RBF network is located through self-organizing learning. In a sense, self-organizing learning is to allocate network resources. The purpose of learning is to make the centre of RBF network located in an important area of input space, so that the selected centre can form a specific distribution rule. Common self-organizing learning algorithms include Hard C means and Kinwomen self-organizing mapping algorithm. HCM algorithm is the most commonly used clustering algorithm in the application of RBF at home and abroad. This algorithm is used for clustering with low input space dimension. When the input space dimension is very high and the data distribution in the input space is highly uneven, it is difficult to use HCM algorithm for clustering.

Self-organizing learning algorithm must obtain all possible sample data, dynamic input mode cannot be used for learning algorithm, and the centre of the input data before study number (that is, the RBF network hidden layer unit number) to the determination of human not only that, the width of the Gaussian function is to choose according to different types of data distribution, increase the difficulty to solve the problem.

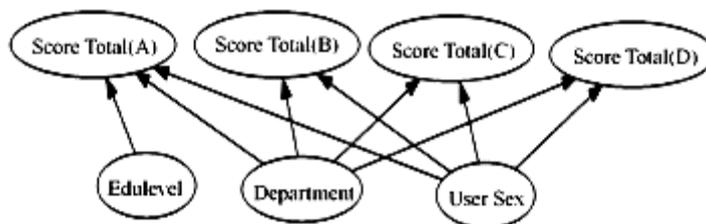


Fig.3 Network Diagram of the Total Score Decision Tree Dependence

Through the above “decision tree” and “dependency network”, the following conclusions can be drawn: the achievement is most closely related to the subject background, the probability of high scores for liberal arts students is high, and the probability of low scores for arts students is high; secondly, there is gender. Girls are more likely to get high marks, while boys are more likely to get low marks. Academic qualification has little influence on the total score, which is only closely related to the total score of A, and there is no significant difference between undergraduates and postgraduates.

3. Results and Discussion

From the perspective of nature, it can be considered that time series system is a way of connection, in which a number of factors with specific attributes constitute a whole with specific functions through specific relations. The definition of the system indicates that there are many interrelated factors in the system, and the change and development of one factor in the system must be affected by other factors. Time series y_t is also affected by various external random factors ($X_1, X_2... X_p$) when it is formed. Ignoring the influence of these factors on the output will lead to a large error, so the model should be changed to a multi-factor time series prediction model considering the influence factors.

Let's take n samples, each of which contains two observations X_1 and X_2 . In the two-dimensional plane determined by the variables X_1 and X_2 , the distribution of n sample points is like a ribbon. It can be seen from FIG. 4 that these n sample points are of great discreteness no matter along the direction of X_1 : axis or the direction of X_2 axis. The degree of discreteness can be expressed by the variance and quantification of the observed quantity respectively. Obviously, if you consider only X_1 and any of them, there is a significant loss of information contained in the raw data.

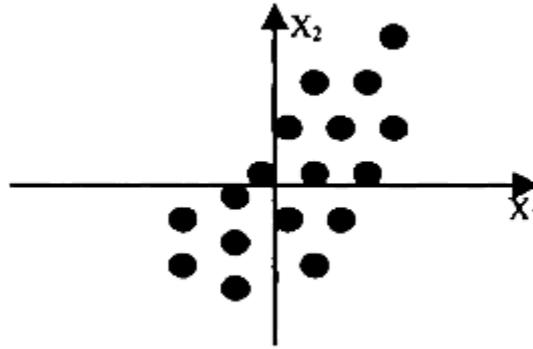


Fig.4 Original Index Distribution

If you rotate the X_1 and X_2 axes counter clockwise by the Angle e , you get the new axes Z_1 and Z_2 , as shown in figure 5. Z_1 and Z_2 are the two new variables.

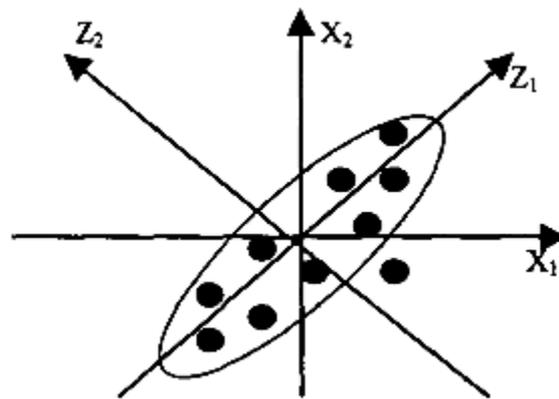


Fig.5 Distributions Of New Indicators

The purpose of rotation is to maximize the dispersion degree of n sample points in the direction of Z_1 axis, that is, the variance of Z_1 is the largest, and the variables represent most of the information of the original data. Therefore, even if Z_2 is not taken into account, the overall situation will not be damaged when the problem is studied. In this way, the information in the original data can be concentrated on the Z_1 axis after rotation and transformation, and the information contained in the data can be concentrated. In addition to being able to condense the information contained in X_1 and X_2 , Z_1 and Z_2 also have irrelevant properties, which avoid the falseness caused by information overlapping when studying complex problems. The variance of n points in a two-dimensional plane is mostly due to the Z_1 axis, and the variance of the Z_2 axis is small. Z_1 and Z_2 are called synthetic variables of the original variables X_1 and X_2 . Since the variance of the variable on the Z_1 axis is the largest, the description of points in the two-dimensional space is described by the comprehensive variable Z_1 , with the minimum loss of information. Therefore, Z_1 is the first principal component and Z_2 is the second principal component. So in our problem, we just have to think about Z_1 ; The direction information, ignore the Z_2 : information and the loss of information is not much. In this way, two-dimensional space becomes one-dimensional space, and Z_1 is a linear combination of X_1 and X_2 , which simplifies the system structure and captures the main contradiction.

Principal component analysis is a method of dimension reduction, the original multiple factors into a few discrete (or independent) a statistical method of comprehensive index, can achieve digital Jane, to reveal the relationship between the variables and the statistical interpretation of purpose, in order to further analyse the nature of the overall and statistical characteristic of data provide some important information.

In the study of multi-factor time series, due to the large number of influencing factors and the correlation between them, the observed data are often overlapped to some extent. When there are

more variables, it is more complicated to study the distribution law of samples in high-dimensional space.

Because the multi-factor time series itself is a dynamic stochastic nonlinear system, various factors interact with each other and influence each other. Therefore, the prediction model is often too complex to determine accurately, or contains fuzziness and chaos. Artificial neural network simulates the basic functions of human brain neurons and the basic structural characteristics of the network, and has the learning ability of distributed memory, parallel processing, association and self-organization. The good nonlinear mapping ability of neural network can avoid the complex parameter estimation process, and at the same time, it can flexibly and conveniently conduct high-precision modelling for the complex unknown systems with multiple factors, which provides a new method for the prediction of nonlinear systems and has been widely used in economic forecasting.

4. Conclusion

Based on RBF network, result analysis and model verification, this study finds that the decision tree method can better predict the network learning behavior and performance, and can better explain some phenomena in the current network learning, such as the total score of male students or art students is prone to crisis, and male students receive more feedback. At the same time, there are some new findings, such as female master's degree students and female arts students do better in academic performance, art students do worse in discussion and communication, female art students do worse in study notes and so on.

Firstly, it is suggested that learners should consciously improve their participation in online learning, increase the learning time span, develop the good habit of recording and reflecting notes, increase the number of repeated learning for complex knowledge, actively participate in discussion and communication, and improve their sense of belonging and immersion in online learning as well as their learning effects. Secondly, it is suggested that learners should consciously improve their learning strategies: complete the learning tasks step by step according to the teaching arrangement of teachers, establish clear learning objectives and make clear learning plans (especially the formulation of learning time and content) before online learning, and complete the learning tasks in strict accordance with the learning plan; Actively regulate their own emotional experience, especially those who have a negative attitude towards online learning, should often participate in discussions, exchanges, reflection notes and other interactions, from which they can get a sense of success and belonging in learning; Timely control, adjustment and evaluation of their own behavior and psychology.

References

- [1] Vachkov G L, Kiyota Y, Komatsu K. Learning of RBF network models for prediction of unmeasured parameters by use of rules extraction algorithm[C]// Fuzzy Information Processing Society, Nafips Meeting of the North American. 2005.
- [2] Ding S, Li X, Su C, et al. An optimizing method of RBF neural network based on genetic algorithm [J]. *Neural Computing & Applications*, 2012, 21(2):333-336.
- [3] An R, Li W J, Han H G, et al. An improved Levenberg-Marquardt algorithm with adaptive learning rate for RBF neural network[C]// Control Conference. 2016.
- [4] Zhen Z, Wang Z, Yong H, et al. Learning method of RBF network based on FCM and ACO[C]// Control & Decision Conference. 2008.
- [5] Min G, Hui P, Dong X P. A hybrid algorithm to optimize RBF network architecture and parameters for nonlinear time series prediction[J]. *Applied Mathematical Modelling*, 2012, 36(7):2911-2919.
- [6] Leung C S, Sum J P. RBF networks under the concurrent fault situation[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2012, 23(7):1148-1155.

- [7] Huang M, Liu X, Li J. Melt index prediction by RBF neural network optimized with an adaptive new ant colony optimization algorithm[J]. Journal of Applied Polymer Science, 2010, 119(5):3093-3100.
- [8] Zhang X M, Ning G L. An Improved RBF Network On-Line Learning Algorithm[C]// International Symposium on Information Science & Engineering. 2009.
- [9] Leung C S, Wan W Y, Feng R. A Regularizer Approach for RBF Networks Under the Concurrent Weight Failure Situation[J]. IEEE Transactions on Neural Networks & Learning Systems, 2016, 28(6):1360-1372.