International Journal of Computer and Communication Engineering, Vol. 1, No. 1, May 2012

Land Use Efficiency Appraisal Based on Cadastral Spatial Data Mining

Huang Hai

***Abstract*—Land use efficiency is a quite important index for land management. And there is much information hidden in the land information which has great potential for land management. A data mining method based on fuzzy neural network for land use data mining is presented in this paper. The fuzzy neural network based on fuzzy logic and neural network and its study algorithm are introduced. Considering its features as fast speed, steady studying course, global dynamic optimization, the fuzzy neural network is applied to establish the model for data mining. The urban cadastral data mining for land use efficiency appraisal is used as an example to proves it is feasible.**

***Index Terms*—Land use efficiency, data mining, fuzzy neural network, cadastral data.**

I. Introduction

With the urban cadastral investigation and updating in the past decades, huge data have been gathered and stored in the land management bureau of government. However, the knowledge and information hidden in these data have not been discovered and put to use yet, thus leading to “explosive data but scarce knowledge”. This has caused great data waste of data gathered with much human and financial resource.

Data mining means extracting data patterns and features, data relations, and other implicit knowledge which users are interested in. It is a new field of multi-disciplinary and multi-technology. In recent years, the major data mining methods include data analysis methods, statistical methods, inductive learning method, clustering and classification methods, detective data analysis methods, rough set method, cloud theory, detection methods of data features and trends, digital image analysis and pattern recognition, visualization methods and so on. These methods are often integrated, and sometimes it is necessary to search for new methods for spatial data mining.

Neural network, a nonlinear self-adaptive dynamic system, can extract the internal features of information and simulate any complex nonlinear relationship, so it is very suitable for solving the problems of data mining. However, the existing neural network model can only deal with precise data, and in reality we often be faced with some imprecise or fuzzy information in data mining. Therefore the integration of fuzzy theory and neural network has gradually attracted the researchers’ attention and it has showed extensive prospects in the research and application of this field.

Manuscript received February 27, 2012; revised April 28, 2012.

Huang Hai is with the School of River and Ocean, ChongQing JiaoTong University, ChongQing, China (e-mail: lottery98@163.com, tel: 18983250429).

II. Fuzzy Neural Network and Its Algorithms

Artificial neural network (ANN) simulates the functions of neurons, with the capabilities of distributed storage and parallel processing for information. Thus it can compute collectively and learn self-adaptively. Similar to the human brain, fuzzy theory can express the experience of operators as rules and convert it into computer algorithms. Fuzzy neural network embeds fuzzy logic into neural networks, integrating the strong knowledge expressing ability of fuzzy systems and powerful learning ability of neural networks. Thus it is better than conventional neural network in learning time, training steps and precision. As the existing fuzzy neural network has some technical difficulties in optimization of fuzzy membership functions, fuzzy logic reasoning, optimized fuzzy computing and anti-fuzzy functions designing, we contrive a fuzzy neural network model with fast learning algorithms and fuzzy reasoning ability.

A. Fuzzy Neural Network and Its Structure

Comprising some fuzzy neurons orienting toward control and decision-making, fuzzy neural network is a combined system of fuzzy logic and neural network. These fuzzy neurons are defined as implementing fuzzy computing, fuzzy reasoning, fuzzy computing and anti-fuzzy computing. Fuzzy neurons adopted in the fuzzy neural network make it be properly defined from the initial training of the fuzzy rules, which makes the network more fault-tolerant and the system more stable. In the conventional fuzzy neural networks, stationary and partial optimization computing methods such as minimum and maximum operations are often used for fuzzy computing. But in compensative fuzzy neural network, dynamic and global optimization methods are adopted. And in the learning algorithms, compensative fuzzy computing is also dynamically optimized, which makes the network optimized. The network not only can adjust the input and output fuzzy membership functions, but also can dynamically optimize the fuzzy reasoning with the aid of compensative logic algorithms. Its parameters, which can be preset by a heuristic algorithm to enhance the training speed, have definite physical meanings.

Each compensative fuzzy neural network has five layers: input layer, fuzzy layer, fuzzy reasoning layer, compensative computing layer and anti-fuzzy layer. Each neural node of the first layer is directly connected to each input vector. Each second layer’s node represents a fuzzy linguistic variable, which can calculate the membership degree of every branch of the input vector belonging to the corresponding fuzzy set of each linguistic variable. Each neural node of the third layer represents a fuzzy rule. It can match the fuzzy rules and calculate the fitness degree of each rule. The fourth layer's neurons carry out the compensative fuzzy computations. And the fifth layer's neurons carry out the anti-fuzzy computations to obtain the exact values of the network's outputs. The network structure is shown in Fig. 1.

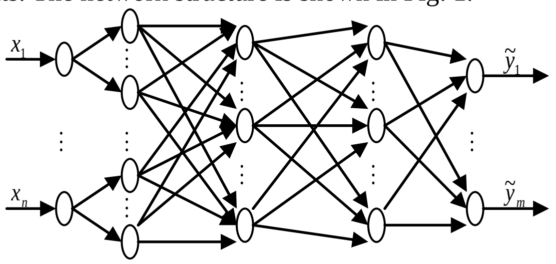

Fig. 1. Structural chart of compensated neural network

B. Training Algorithms

In the training process of a compensative fuzzy neural network, iteration may be divided into two processes: the forward compensative fuzzy reasoning process and the backward error propagation process. At the beginning of training, the initial values should be preset for the network's parameters including the center and width of fuzzy membership functions for input and output, together with the compensation degree. Then the compensative fuzzy reasoning and error's counter spreading process will be iterated until the result is satisfying.

Suppose a compensative fuzzy neural network with *n* inputs, *m* outputs and *s* fuzzy reasoning rules and take *P* sample signals (x^p, y^p) to train the network. The fuzzy membership functions of input and output are Gaussian as follows.

$$\mu_{A_i^k}(x_i) = \exp[-(\frac{x_i - a_i^k}{\sigma_i^k})^2] \qquad (1)$$

where a_i^k represents the center and σ_i^k represents the width of the input of membership function.

$$\mu_{B_j^K}(y_i) = \exp[-(\frac{y_j - b_j^k}{\omega_j^k})^2] \qquad (2)$$

where b_j^k represents the center and ω_j^k represents the width of the output of membership function.

In practical use, the sum of width for the input and output variables' membership functions is generally equal to the value range of the variables. The division of the fuzzy sector (i.e. nodes number of the concealed layers), the center and width of the membership functions are often determined according to the distribution features of the variables. As the training precision and time are directly influenced by the nodes number of the concealed layers, these two factors should also be considered when we determine it.

The fuzzy reasoning rule *k* among the *R* ones is described as follows:

R_K : If (x_1 is A_1^K and x_2 is A_2^k and ··· and x_n is A_n^k), then (y_1 is B_1^k and y_2 is B_2^k and ··· and y_m is B_m^k).

The fuzzy set *A* in the range $U=U_1\times U_2\times\cdots\times U_n$ produces the fuzzy set *B* in the range $V=V_1\times V_2\times\cdots\times V_m$ after it is transformed by the fuzzy reasoning rule k in the third layer. After the reasoning operation, the compensative operation and the single-value fuzzy operation, its degree of membership is obtained as follows.

$$\mu_B k = [\prod_{i=1}^{n} \mu_{Ai} k(x_i)]^{1-r_k+r_k/n} \qquad (3)$$

where the compensation degree $r_k \in [0，1]$.

The output value is obtained after the anti-fuzzy operation of the fifth layer.

$$\tilde{y}_j = \frac{\sum_{k=1}^{s} b_j^k \omega_j^k z_k}{\sum_{k=1}^{s} \omega_j^k z_k} \qquad (4)$$

where $z_k = [\prod_{i=1}^{n} \mu_{Ai} k(x_i)]^{1-r_k+r_k/n}$ (5)

The target function is

$$E = \frac{1}{2}\sum_{j=1}^{m}(\tilde{y}_j^p - y_j^p)^2 . \qquad (6)$$

The gradient descent law which can adjust the pace dynamically is taken to train the compensation degree in the compensative computing, the center and width of the input / output membership functions. The corresponding iteration formulas are listed as follows.

$$b_j^k(t+1) = b_j^k(t) - \eta(t)\frac{\partial E}{\partial b_j^k(t)} \qquad (7)$$

$$\omega_j^k(t+1) = \omega_j^k(t) - \eta(t)\frac{\partial E}{\partial \omega_j^k(t)} \qquad (8)$$

$$a_i^k(t+1) = a_i^k(t) - \eta(t)\frac{\partial E}{\partial a_i^k(t)} \qquad (9)$$

$$\sigma_i^k(t+1) = \sigma_i^k(t) - \eta(t)\frac{\partial E}{\partial \sigma_i^k(t)} \qquad (10)$$

III. Land Use Efficiency Data Mining Based on Fuzzy Neural Network

As a type of spatial data, cadastral data record such basic information of land parcels as area, ownership, location and use and so on, which can reflect the land use status of a district. So it is very important for the administrative organs such as land management and urban planning. Analysis of the existing cadastral data is helpful for the prediction of the future land use and the scientific land management. Because cadastral data is of huge volume, either manual analysis or traditional statistical methods are difficult to discover the information implied in these data. So data mining can be effective in dealing with such a problem. This paper attempts to use fuzzy neural network for the analysis of cadastral data in order to find the implied knowledge and relationship.

Compensative fuzzy neural network can express the uncertainty of entity in fuzzy means, and condense the knowledge implied in the data to the weight between the nodes after they are trained with many spatial data samples. Cadastral data mining based on compensative fuzzy neural network mainly includes the following procedures.

1) Data choosing and pre-processing. Extract the necessary data from the cadastral database and check its completeness and uniformity. The "noise" data should be canceled and the lost data should be filled with statistical methods. Then store the processed data into a new database.
2) Construct the data mining model. The training and non-linear reflection of the network are completed by adopting the compensative fuzzy network structure and improved algorithms mentioned above. The prepared cadastral data are used as the input and output for the training.
3) Select the input columns and predictable columns. The input columns, strongly related to the predictable columns, are used to train the mining model. "Construction area" and "land use type" are chosen as the input columns for the example. The predictable columns are the expected result for the data mining model. The "building capacity rate" is chosen as the predictable column for the example.
4) Test and evaluate the model. Verify the testing data and if the given requirement of the total errors is satisfied, the training can be ended and the model can be put to use. If not, the training data should be adjusted until satisfaction.
5) Appraise the urban land use efficiency. The tested model is put to use for urban land use efficiency of the whole city.

According to the "maximum membership" principle, the "building capacity rate" appraisal results for commercial, residential and industrial land are general, high and low respectively after the data mining form the urban cadastral database of ChongQing city with the compensative fuzzy neural network. This shows the land use of this region is at a low efficiency and there is still much potential. So in the future land use, attention should be focused on enhancing the land efficiency through measures such as land transformation and old city reform and so on.

IV. Conclusions

Spatial data mining is a hot field in data mining for its special features related to space. Urban land use has spatial features, so methods for spatial data mining can also be used for cadastral data mining. Fuzzy neural network, which can deal with complex data and prediction process that other algorithms can not accomplish, has become a focus in recent years in many fields. Data mining can extract such information and knowledge as data classification, spatial evolution and prediction and so on, and in the huge cadastral data find the implied information which is helpful for our urban construction and management. Research of this paper shows we can use neural network for land use efficiency data mining from urban cadastral database. Spatial data mining from cadastral database can also help us get much information about urban land use, so the fuzzy neural network method has a nice prospect in data mining from cadastral database.

Acknowledgement

The research of this paper is supported by Education Committee of Chongqing Municipality, grant No. KJ110406.

References

[1] H. Lu, R. Setiono, and H. Liu, "Effective Data Mining Using Neural Network," *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, June, pp. 957-961, 1996.
[2] H. J. Milller and Han Jia-wei, "Geographic data mining and knowledge discovery, London: Taylor and Francis," 2001, pp. 3-32
[3] D. Li, S. Wang, and W. Shi, "Spatial Data Mining and Knowledge Discovery," *Geomatics and Information Science of Wuhan University*, vol. 26, pp. 491-499, June, 2001.
[4] H. Han and L. Zhang, "Application of Neural Network in Data Mining," *Journal of Systems Engineering*, vol. 23, April, 2004, pp: 88-90.
[5] J. Feng and R. Ma, "Stock price forecast based on data mining of neural networks," *Journal of Computer Applications*, vol. 29, pp. 155-156, June, 2009.
[6] S. Bengio and Y. Bengio "Taking on the curse of dimensionality in joint distribution using neural networks," *IEEE Transactions on Neural Networks*, vol. 11, March, 2000, pp. 550-557.
[7] S. Wu and Tommy WSC. "Clustering of the self organizing map using a clustering validity index based on inter-cluster and infra-cluster density," *Pattern Recognition*, Jan, 2004, pp. 175-188.
[8] C. Li. "Knowledge Discovery in Database," *Journal of Northwest University: Natural Science Edition*, 1999, vol. 29, no. 1, pp. 114-119.
[9] D. Qu, "Design and Simulation of the Improved Performance Index Function with Enhanced Generalization Ability for the ANN's Identification Model," *Journal of Naval Aeronautical Engineering Institute*, 2007, vol. 22, no. 1, pp. 109-113.
[10] H. Han and L. Zhang, "Application of Neural Network in Data Mining ," *Journal of Systems Engineering*, vol. 23, no. 4, pp. 88-90, 2004.