International Journal of u- and e- Service, Science and Technology
Vol.9, No. 10 (2016), pp.123-130
http://dx.doi.org/10.14257/ijunesst.2016.9.10.11

An Agricultural Land Contractual Management Transfer Prediction Model Based on Analytic Hierarchy Process and Logistic Regression

Hai-ying Cao[1,2] and Ling Wu[1]

[1]*Northeast Agricultural University, Harbin 150030, China*
[2]*Harbin Finance University, Harbin 150030, China*
[1]*caohaiying0813,* [2]*88wuling @163.com*

Abstract

In the field of application prediction, there are two kinds of common methods to establish the prediction model: prediction model established by artificial data analysis relying on expert experience, and prediction model achieved by statistical model exploiting data analysis. However, the prediction accuracy of the model based on expert is restricted by the experiences, while the model based on statistical analysis is limited by the quality and scale of the training data. In view of the advantages and disadvantages of these two kinds of models, this paper presents a prediction model by integrating Analytic Hierarchy Process and Logistic Regression. The proposed prediction model uses the Analytic Hierarchy Process, which are based on the training data and expert experience to obtain the rank of predominant factor in a specific domain, and exploits the logistical regression model to learn the weights of each influencing factor. Finally, the linear combination of the two models is used to obtain the prediction model. Further, we take agricultural land contractual management transfer prediction as an example to test the proposed hybrid prediction model.

Keywords*: Prediction Model, Analytic Hierarchy Process, Logistic Regression, Agricultural Land Contractual Management Transfer Prediction*

1. Introduction

We consider the prediction problem in the application field. When predicting the problem of an application domain, prediction model established by artificial data analysis relying on expert experience and prediction model achieved by statistical model exploiting data analysis are two types of main model. For the first method, prediction model established by artificial data analysis mainly relies on expert experience, however, the prediction accuracy of the model based on statistical model analysis is limited by the quality and scale of the training data.

Considering the advantages and disadvantages of these two kinds of models, a prediction model based on Analytic Hierarchy Process(AHP)[1] and Logistic Regression(LR)[2-3] is proposed. The model combines the Analytic Hierarchy Process and the Logistic Regression model to obtain a hybrid comprehensive prediction model. Firstly, the model uses the Analytic Hierarchy Process, based on the training data and expert experience to obtain the ranking model of important factors in the specific problems in application domain, then, based on the training data, the prediction model uses the Logistical Regression Model to learn the weight of each influencing factor.

Further, in order to verify the performance of the proposed prediction model, this paper takes the problem of agricultural land contractual management transfer as an example to analyze and verify the proposed model.

ISSN: 2005-4246 IJUNESST

The remainder of this paper is organized as follows. In the second section, a prediction model based on Analytic Hierarchy Process and Logistical Regression Model is proposed, In the third section, the application field of the agricultural land contractual management transfer is studied, specifically, and the important characteristics of the influencing factors are given. In the fourth section, we report the experimental results and performance comparisons. And in the last section we conclude our study.

2. Prediction Model based on Analytic Hierarchy Process and Logistic Regression

2.1. Prediction Model

The samples X in domain D can be represented as a feature vector as follow:

$$X=\{x_1,x_2,...,x_n\} \tag{1}$$

where x_i is the value of the i-th feature, and n is the number of feature.

Let w denotes the weight vector of each feature vector, and suppose we have m methods which can be used to obtain the weight vector w of X. For each method, let α_j as the trust degree for the j-th method, then we define a framework of prediction model as follow:

$$H(X) = f(\alpha_1 w_1^T X + ... + \alpha_m w_m^T X) \tag{2}$$

where, H (x) represents the prediction function, and f is a probability function.

This paper uses logistical regression function as the probability function in Eq.2 to integrate Analytic Hierarchy Process and Logistical Regression Model. So we have two models to decide the weight vector w: Analytic Hierarchy Process and Logistical Model. Using these two models, the prediction model can be described as:

$$H(X)= f((1\text{-}\alpha) * w_{AHP}^T + \alpha * w_{LR}^T)X \tag{3}$$

where w_{AHP} is the weight vector learned by Analytic Hierarchy Process, while w_{LR} is the weight vector learned by Logistical Regression Model. α is used to set the weight between Analytic Hierarchy Process and Logistical Regression Model.

We choose Logistical Regression Model as the prediction model f, then the final prediction model can be defined as follow:

$$h_{w,\alpha}(X) = \frac{1}{1 + e^{-((1\text{-}\alpha)*w_{AHP}^T+\alpha*w_{LR}^T)X}} \tag{4}$$

The rest of this paper will introduce how to get w_{AHP} and w_{LR}..

2.2. Analytic Hierarchy Process

Analytic Hierarchy Process is a hierarchical weighted decision analysis method proposed by Thomas L. Saaty. Based on the deeply analyzing the nature of complicated deciding problems, influential factors and internal relation, people can use less quantitative information to make decision process more mathematical. Thus decision problems can be resolved by exploiting more simple deciding method for complicated decision-problems with multiple goals, multiple criteria or non-structural characteristics. Using AHP method, decision-makers can resolve complicated problems into different component factor by combining different level factors according to the correlation and subordinate relations to make a multi-level hierarchical structure. Then we can invite experts, scholars and authorities to analyze the relatively important factors in the same level.

Using math methods, we can construct the judgment matrix. By solving the judgment matrix the maximum characteristic root and characteristic vector method, people can also get the relative weight of each factor. By calculating the rank on the lowest level, people can get the layer of the ranking about influential factors. Finally professors are able to make plan and decisions, and select the measure of solving problems according to ranking result. AHP method is mostly used to solve some issues like complicated structure, many decisions, multiple criteria and the problem hard to quantified.

The steps of AHP can be described as follow.

2.2.1. Constructing Recursiving Hierarchy Model : Firstly, the issues we want to estimate are divided into different levels according to its contained factors. Then we can arrange different levels according to the form of the highest level, the middle level and the lowest level. At last, a multi-level structure can be formed. Elements in the same level are independent by each other. They not only control some elements in next level but also are controlled by the upper level elements.

2.2.2. Constructing Judgment Matrix: The definition of judgment matrix is compared and analyzed the pairwise importance of features. When we can make sure recursive hierarchy model, then we also can confirm the subordinate relations between up and down level of features. The same level of index pairwise will be compared when constructing judgment matrix. When we are constructing the judgment matrix, we will use "1-9" scale table. On the same level of n features, we can compare judgment matrix A={aij} by using pairs.

2.2.3. Calculating Feature Weight: According to the judgment matrix, we can calculate the weight on each level, the calculation steps are shown as follows. Firstly, we will calculate the elements of estimation matrix for each line product M_i:

$$M_i = \prod_{j=1}^{n} \mathrm{b}_{ij}, \quad i=1, 2, 3, \ldots, n \tag{5}$$

Then, the n-th root of M_i can be computed as:

$$\overline{\overline{W}}_i = \sqrt[n]{M_i}, \quad i=1, 2, 3, \ldots, n. \tag{6}$$

where n is the matrix order.

Then, the vector $W=(W_1, W_2,...,W_n)$ will be normalized by using Eq(7):

$$\overline{W}_i = \overline{\overline{W}}_i / \sum_{i=1}^{n} \overline{\overline{W}}_i \tag{7}$$

where W_i is the desires of the index weight.

2.2.4. Calculating the Maximum Feature Value in Judgment Matrix The maximum feature value in judgment matrix can be computed as follows:

$$\lambda_{\max} = \sum_{i=1}^{n} \frac{(A \cdot W)_i}{n \cdot W_i} \tag{8}$$

$$A \cdot W = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_3 \end{bmatrix} \tag{9}$$

$$(A \cdot W)_i = a_{i1} \cdot W_1 + a_{i2} \cdot W_2 + \ldots + a_{in} \cdot W_n \tag{10}$$

2.2.5. Consistency Check: In order to evaluate the effectiveness of the hierarchical ranking, we must also test the consistency of the judgment matrix. The consistency is an evaluation feature of the score is reasonable or not. Because the judgment matrix is predicted quantitatively by the experiences of experts, it is impossible to achieve complete consistency. To address this assues, T.L.Saaty put forward the concept of random consistency ratio, denoted as C•R. And when the R•C < 0.1, the consistency is satisfied which means the comparison matrix of the judgment results can be accepted. R•C's formula is:

$$C \bullet R = C \bullet I / R \bullet I \tag{11}$$

R•I is the average random consistency feature, which is related to the order n of the judgment matrix.

First, calculate the consistency feature C•I：

$$C \bullet I = (\lambda max - n) / (n-1) \tag{12}$$

$$\lambda = \frac{1}{n} \sum_{i=1}^{n} \frac{(A \cdot W)_i}{W_i} \tag{13}$$

where A is known to judge the matrix, N is used to judge the matrix order, W_i is the relative weight column vector.

Secondly, the consistency feature of the same order matrix is R•I.

Finally, calculate the consistency ratio C•R. C•R=C•I/R•I, when C•R=0, A has a complete consistency. When C•R＜0.1, A is with satisfactory consistency. When C•R≥0.1, A has a very satisfactory consistency, it should be adjusted or absent. By assigning and adjusting the weight of each feature, it can show the tendency and flexibility of the evaluation feature system.

The same order matrix average consistency feature R•I as shown in Table 1.

Table 1. The Mean Random Consistency Feature

n	1	2	3	4	5	6	7	8	9
R_1	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

2.3. Prediction based on Logistical Regression

The logistic regression will learn a set of weights for each feature. When a new train case X arrives, LR finds the list of features, and sums the weights associated with those features in the current X. In mathematical terms, we can sign $w^T x + b$. In this notation, w is a vector of weights, and x is a feature vector. The notation $w^T x$ simply means taking the sum of weights associated with each of these features. We then convert this sum of weights into a probability by using the logistic function[18].

$$h_{w,b}(x) = \frac{1}{1 + e^{-w^T x}} \tag{11}$$

This equation converts a number between -1 and +1 to a probability of [0,1].

In this paper, we define the cost function as a logarithmic loss function:

$$cost(h_{w,b,\gamma,\delta}(x,\theta_m,\theta_q),y) = -y\log(h_{w,b,\gamma,\delta}(x,\theta_m,\theta_q)) - (1-y)\log(1-h_{w,b,\gamma,\delta}(x,\theta_m,\theta_q)) \tag{12}$$

Therefore, the loss function is expressed as follows:

$$L(w,b) = -\frac{1}{n}\sum_{i=1}^{n} cost(h^{(i)}_{w,b,\gamma,\delta}(x^{(i)},\theta^{(i)}_{m},\theta_{q}), y^{(i)}) \tag{13}$$

The gradient descent method is used to minimize theloss function L(w, b) to evaluate parameter w and b. Theupdate for each sample (x(i), y(i)) can be shown as follows:

$$w = w + \alpha*(1-\delta)*(y^{(i)} - h_{w,b,\gamma,\delta}(x^{(i)},\theta^{(i)}_{m},\theta_{q}))*x^{(i)} \tag{14}$$

$$b = w + \alpha*(1-\delta)*(y^{(i)} - h_{w,b,\gamma,\delta}(x^{(i)},\theta^{(i)}_{m},\theta_{q}))*b \tag{15}$$

3. Features for Agricultural Land Contractual Management Transfer Prediction

Based on national policies and scholars research situation, this paper selects the affected the economic development of the circulation of the right to the contracted management of rural land, system arrangement, social factors, individual factors and factors of agricultural production, those five main factors. And the factors that affect the agricultural land contractual management transfer are analyzed by selecting several specific indicators [4-7], construct index system as Table 2.

Table 2. Influencing Factors in Agricultural Land Contractual Management Transfer

Category	Features	Meaning
Economic factor	1) Per capita GDP(pGDP)	Ratio of gross domestic product to total population in a country or region
	2)Two or three industrial output value accounted for GDP ratio(iovc)	Of a country or region, the third industry, the second industry output value and the ratio of gross domestic product
	3)Proportion of agricultural output value(paov)	The proportion of agricultural output to GDP
	4)Direct cost index of agricultural land contractual management transfer(dci)	Reflected in the process of rural land transfer, the direct access to land revenue or land transfer in the transfer of land to pay the land rent
Social factor	5)Two or three industry practitioners ratio(ipr)	Reflect the proportion of non farm population in a region
	6)Per capita expenditure on science and education(ese)	Reflect an area of science and education investment situation
	7)Rural Engel coefficient(rec)	Per capita food expenditure in rural areas accounted for the proportion of total cost of living expenses
	8)Normative index of agricultural land contractual management transfer(ni)	Signed a written contract and the circulation of the total area of the transfer of the total area of land circulation
	9)Fiscal expenditure on Agriculture(fea)	Reflect the government's investment in agriculture
Agricultural production factor	10)Cultivated land area per capita(cla)	Reflect the quantity of cultivated land in a certain area
	11)Per capita household production expenditure(hpe)	Reflect the peasant household's investment to agriculture
	12)Economic crop area accounts for the proportion of total farmland area(ecaa)	To reflect the situation of cultivated land use in a certain area
Individual	13)Age(age)	Actual age of farm households

farmer factor[]	14)Sex(sex)	-
	15)Education level(el)	Reflect the ability of farmers to accept new things
	16)Household labor ratio(hlr)	Reflect the family burden of farmers

4. Experiment and Result Analysis

4.1. Experimental Data

In order to further understand the effects of above features in the agricultural land contractual management transfer, we collect 480 related valid questionnaires through making a survey to 4 counties of X province.

Based on data availability and other considerations, the survey indicators mainly affect land transfer to the main indicators of individual farmers, and take into account the economic, social and agricultural production targets.

4.2. Experiment Settings

According to the 480 copies of the questionnaire, we obtained 480 training data. We choose 200 as training data, and the remaining 280 as the test data. In the experiment, we perform 2-fold cross validation to train the parameters.

We set the parameters α in Eq.4 0.5 in the experiment. After the training phase, the final weight is shown in Table 3.

Table 3. Farmland Conversion Factors Affecting the Weight of the Index System

Feature	Weight	Feature	Weight
pGDP	0.4674	fea	0.0173
iovc	0.0954	cla	0.4531
paov	0.2767	hpe	0.3201
dci	0.1605	ecaa	0.2265
ipr	0.2167	age	0.1605
ese	0.6413	sex	0.0954
rec	0.1142	el	0.2767
ni	0.0105	hlr	0.4674

4.3. Experimental Results and Analysis

The experimental results are shown in Table 4.

Table 4. Experimental Results

Feature	Weight	Wald	Probability	Exp(B)
age	.376	5.413	.020	1.456
Education level	-.527	6.283	.012	.590
Family Population size	-.279	2.359	.125	.757
Household labor ratio	-.001	.000	.996	.999
Concurrent industry number	-.392	3.544	.030	.676
Number of farm machinery	-.590	4.427	.035	.554
Cultivated land area per capita	.025	.047	.828	1.025
Number of medical insurance	.006	.001	.970	1.006
Number of Endowment insurance	.126	1.262	.261	1.135
Land policy	.245	.634	.426	.783
Family Non farm income	.435	20.072	.000	1.545

According to the results predicted by the proposed prediction model based on Hierarchy Analysis and Logistic Regression, we can conclude that significant influence factors of farmers land transfer have five points, and significantly from strong to weak are family non farm income, education level, age, concurrent industry number, number of farm machinery. Other significant factors that was not significant, but in the level of significance beyond the scope of 0.05.

5. Conclusion

In this paper, we proposed a new prediction model to combine Analytic Hierarchy Process(AHP) and Logistic Regression(LR) to predict the problem in the field of agricultural land contractual management transfer. This model exploits the advantages and disadvantages of the model based on expert analysis and statistical methods. The weights of influence factors are learned by AHP and LR respectively, and integrated by a probability function to predict the probability of each factors. Evaluated on the test dataset, the experimental results show that the proposed model can predict the all kinds of influence factors in agricultural land contractual management transfer.

Acknowledgment

This work is supported by the Humanities and Social Sciences of Ministry of Education Planning Fund(12YJAZH157) and National Natural Foundation of China-Fund for the Management Science Department(71341049).

References

[1] T. L. Saaty, “How to make a decision: the analytic hierarchy process”, European journal of operational research, vol. 48, no. 1, **(1990)**, pp. 9-26.
[2] J. Goodman and W. Yih, “Online Discriminative Spam Filter Training”, CEAS, **(2006)**, pp. 1-4.
[3] D. W. Hosmer Jr. and S. Lemeshow, “Applied logistic regression”, John Wiley & Sons, **(2004)**.
[4] G. D. Jaynes, “Economic theory and land tenure”, HPB a. MR Rosenzweig, **(1984)**.
[5] K. Deininger and G. Feder, “Land institutions and land markets”, Handbook of agricultural economics, vol. 1, **(2001)**, pp. 288-331.
[6] W. Jiehua and W. Jieyu, “On the benefit game and institutional equilibrium in the process of rural land circulation”, Social Sciences Review, vol. 24, no. 5, **(2009)**, pp. 45-47
[7] L. Qiyu, “Analysis on the game of the interests of rural land transfer based on urban and rural integrated planning”, Agricultural Economy, vol. 12, **(2011)**, pp. 69-71

Author

Hai-ying Cao, Born in August 1979, Candidate Dr. Lecturer. Her research interests include agricultural economic theory and policy, human resource management, land financial and tax research.

Ling Wu, Born in 1970, PhD degree. Professor. Doctoral tutor. Her research interests include Agricultural economic theory and policy, Marxist basic theory research.

