Anticipation of Best Project Proposals based on Risk Assessment using Clustering Technique

Dr. K.Kavitha
Assistant Professor, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, India.

Abstract – Data Mining techniques has made major improvements in the field of knowledge discovery using various domains. Clustering is an important technique which is used to group the related elements without advance knowledge. Risk Assessment is an important task of project allotment through E-Tender, as Success and Failure of the Project depends largely on Project Co-Ordinator ability to evaluate the risk effectively. Risk evaluation process is used to identify the good and bad Project Proposals. Risk evaluation and Prediction is done based on Clustering and decision making approaches. This method allows the user to generate the risk percentage and determine whether project can be granted to a Contractor or not. This paper concentrates mainly the concept of Multidimensional data clustering for Risk evaluation and Prediction.

Index Terms – Association Rule, Risk Evaluation, Prediction, Clustering.

This paper is presented at the International Conference on Information Technology on 25th September (2015) conducted by Department of computer science and applications, Thiruthangal Nadar College, Selavayal, Chennai, Tamilnadu, India.

1. INTRODUCTION

Due to high competition, Business field is necessary to consider the Contractor's relationship management of the enterprise. This management analyses the large volume of data and classify them based on Contractor's experience, Earnings and Prediction. Data Mining offers many technologies to analyze and detect the hidden pattern and also convert raw data into useful information. The key idea for utilizing data mining techniques is to classify the Contractor's according to the posterior ability. Here data mining offers Classification and prediction methods to perform the task.

Project grant Risk evaluation decisions are important for government or private organizations due to high risk associate with inappropriate grant decisions that may reflect in huge amount of losses. When considering the grant rate decisions regarding the application for a project proposal grant, such as project construction, the lender tends to use the direct scrutiny a committee.

To limit the default risk, Screening the Contractor's financial background and previous project experience. Before granting the projects, organization has to take various steps such as performance of the firm by determining the last year financial status and Contractor's Previous Project Experience.

Sometimes with huge data and technology slackness, decisions may be wrong and resulted in project failure defaults. Here data mining offers Clustering algorithm which is based on intelligent information system for providing tools with additional required information, to reduce the uncertainty of the outcome for enhancing the service quality.

Project grant is defined as a Statistical method to predict the Probability which helps to determine whether project should be granted or not to a Contractor. Project Grant can also be defined as a systematic method for evaluating grant risk that provides Consistent analysis for affecting the risk level. The objective of grant scoring is to manage the risk involved in providing grant risks. So that the system makes better decisions quickly. Grant scoring helps to increase the speed and consistency of the project application.

Rest of this paper is structured as follows. Research works related to Risk Assessment are discussed in Section2. Detailed explanations of implementing project grants using Prediction Algorithm presented in section 3.Experimental results are reported in Section4 to prove the accuracy and finally section 5 concludes this research work.

2. RELATED WORK

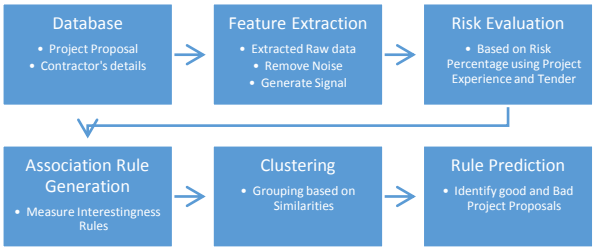
Credit Risk evaluation is an interesting management problem in financial structure. Francesca.et.al proposed hazard model for predicting loan population which involves different probability of risk factors. Probability is modeled into two groups such as good and bad borrowers [1].

Zakrzewska.et.al proposed a technique allows building of different rules for different type of customers; each applicant is assigned to similar group of concerned part [2].Bhasin.et.al presented a model to extract efficient information from existing data and enables the system to make better decisions [3].

2.1. Implementation of Risk Evaluation & Prediction Algorithm for Project Grant

Risk Assessment is one of the existing problems in the project proposal grant. The decision for project sanction to a Contractor should be evaluated properly so that it may not lead project failures. Here ERPCA method aids the organization sector to make the evaluation for project grant in

an enhanced manner. The overall flow of this implementation is shown in figure.



Here, each Customer who needs project has to provide their personal details, financial backgrounds, project details etc to the organization. These details are captured in the database for further processing. The dataset contains attributes like age, qualification, year_of_experience, property, Tender_Amount, Asset value etc. The details of the Contractors are collected and then segmented based on past experience and Tender_amount. Then the valuable attributes are selected using feature selection concept.

2.1.1. Feature Extraction

The main aim of this stage is to find the minimum attribute set for understanding the data easier and reduce the complexity. Irrelevant data can be removed to decrease the dataset size. Here extraction is done by calculated information gain.

It measures the uncertainty of collected data when the outcome is not known. Information gain is used to select the best criterion attribute based on decision tree.

2.1.2. Rule Formation

Every organization has different criteria that has to be satisfied by the contractor to avail the project grant. To receive the project grant, Contractor has to satisfy particular constraints like maximum experience, Income, Property as shown in table. Rule list are framed by the organization as shown in the figure. The following algorithm is used to evaluate and predict the risk percentage using sample dataset and specific constraints.

Algorithm: Risk Assessment

Input: User_Infm, Constraints

Begin

Initialize Threshold Value

If (User.Project_Experience >= r.Project_ Experience) && User.Project_Type=”Major” && User.Tender_Quotation >=

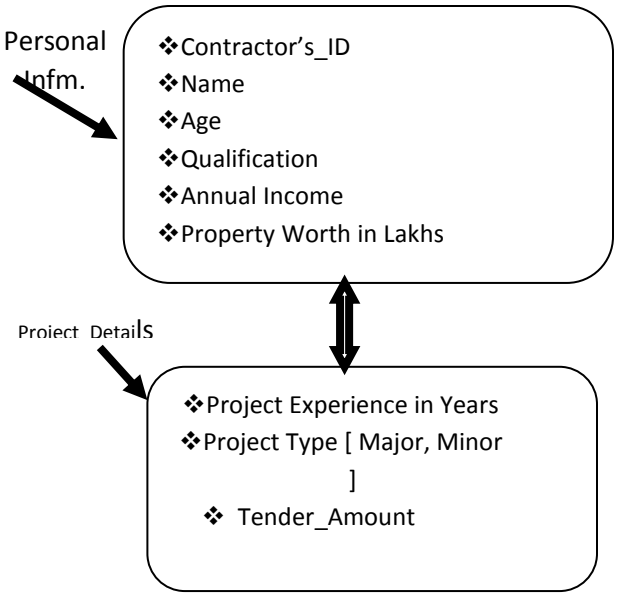
r.tender_Quotation then

Count++

RiskCount/N * 100 // No. of transactions

If (Risk <=Threshold) Return True

Else Return False



2.1.3. Risk Evaluation

Project Co-Ordinator should be capable of risk measurement.Risk evaluation is done by measuring certain attributes. Risks are divided into two categories Primary and Secondary. Primary Risk considers three attributes such as Project_Experience, Project_Type and Tender_Quotation. Secondary Risk considers Property, Age and Qualification. Based on the Predicted Rules, Specific values are assigned for the attributes for project grants and then convert it into binary formation.

2.1.4. Clustering

Clustering method groups the set of objects and checks whether there is some relationship between objects. Associative Clustering Algorithm used to mine Clusters from high dimensional numerical databases. In this algorithm, two vectors can be taken into consideration. By using Risk Percentage, values can be categorized as Good and Bad vectors are

2.1.5. Rule Prediction

For project grant, threshold value of 50% of risk is specified. Risk is predicted based on the risk percentage (ie) Co-Ordinator can decide whether to grant the project or not. If the percentage of risk is greater than 50%, the application is rejected else project is granted. Project_Grant and Project_Reject list are classified using the specified threshold

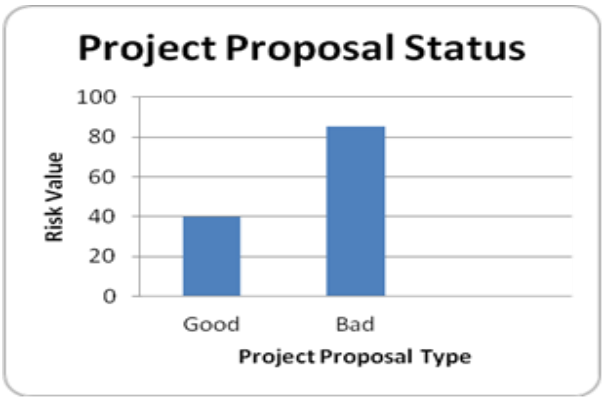
value. Then the approved project contractors and rejected contractors are clustered separately for efficient retrieval.

3. EXPERIMENTAL RESULT

To evaluate the effectiveness of the risk evaluation method, performance evaluation is carried out based on percentage of risks. Experimental datasets are generated with 100 applicants with their personal and experience details. It consists of 15 attributes.

The applicant are segmented based on Project type. Next, Feature selection is performed to extract the attributes which is used to eliminate redundancy and identify irrelevant features. Among 15 attribute, 4 attributes are selected by Feature Extraction process. Rules are predicted after feature extraction process. Organization proposes different rules for applicant based on project type. Then Risk Assessment is evaluated based on Primary and Secondary attributes. Percentage of Risk is calculated based on these attributes. Based on % of Risk, Contractors are classified as High, Medium and Low as shown in Figure.

Finally, Projects are also separated as good and bad project proposal considered to grant the project based on % of Risk Value as shown in figure. Here, Bad Proposals are eliminated and Good project proposals are considered for further proceedings.



4. CONCLUSION

Risk Assessment is an important task in engineering field. This paper proposed implementations of Project Grant approval using Risk Evaluation and Prediction. Contractors are extracted for feature selection and prediction of rules for each project type. Risk Assessment is performed by using Primary and Secondary level of Risk. Clustering technology s also used to classify the risk levels as High, Medium and Low based on the obtained risk values. The granted and Rejected Project Proposals are considered as Good and Bad proposal correspondingly. Experimental Results predicts the accuracy and also expose that this method produces accuracy and consumes less work.

REFERENCES

[1] G.Francesca, "A Discrete-Time Hazard Model for Loans: Some evidence from Italian Banking system", American Journal of Applied Sciences, vol. 9, p. 1337, 2012.
[2] D.Zakrzewska, "On integrating unsupervised and supervised classification for credit risk evaluation", Information Technology and Control, Vol. 36, pp. 98-102, 2007.
[3] M.L.Bhasin, "Data Mining: A Competitive tool in the Banking and Retail Industries" Banking and Finance, vol 588,2006
[4] M.Usman, R.Pears and A.Fong, "Discovering diverse Association Rules from Multidimensional Schema,",2013
[5] R.Pears,M.Usman and A.Fong, "Data guided approach to generate Multi0Dimensional schema for targeted Knowledge discovery", 2012.
[6] G.Liu, H.Jiang, R.Geng and H.Li, "Application of Multidimensional association rules in personal financial services" in Computer design and Applications(ICCDA), 2010 International Conference ,pp V5-500-503
[7] W.Y.Chiang "To mine association rules of Customer values via data mining procedure with improved model:An empirical case study'Expert Systems with Applications, Vol 38, pp 1716-1722,2011
[8] T.Herawan and M.M.Deris, "A soft set approach for association rule mining Knowledge based Systems", vol 24, pp 186-195,2011.
[9] B. Bodla and R. Verma, "Credit Risk Management Framework at Banks in India," *ICFAI Journal of Bank Management, Feb2009,* vol. 8, pp. 47-72, 2009.
[10] R. Raghavan, "Risk Management in Banks," *CHARTERED ACCOUNTANT-NEW DELHI-*, vol. 51, pp. 841-851, 2003.
[11] M. A. Karaolis, J. A. Moutiris, D. Hadjipanayi, and C. S. Pattichis, "Assessment of the risk factors of coronary heart events based on data mining with decision trees," *Information Technology in Biomedicine, IEEE Transactions on,* vol. 14, pp. 559-566, 2010.
[12] M. Anbarasi, E. Anupriya, and N. Iyengar, "Enhanced prediction of heart disease with feature subset selection using genetic algorithm," *International Journal of Engineering Science and Technology,* vol. 2, pp. 5370-5376, 2010.
[13] M. Du, S. M. Wang, and G. Gong, "Research on decision tree algorithm based on information entropy," *Advanced Materials Research,* vol. 267, pp. 732-737, 2011.
[14] X. Liu and X. Zhu, "Study on the Evaluation System of Individual Credit Risk in commercial banks based on data mining," in *Communication Systems, Networks and Applications (ICCSNA), 2010 Second International Conference on*, 2010, pp. 308-311.
[15] B. Azhagusundari and A. S. Thanamani, "Feature selection based on information gain," *International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN,* pp. 2278-3075.