



Recommended Framework for Identifying the Best Project Proposal using Risk Evaluation and Prediction

¹Dr. K. Kavitha

¹ Assistant Professor

¹ Department of Computer Science

¹ Mother Teresa Women's University

¹ Kodaikanal.

Abstract:-

Data Mining plays a major role in Knowledge Discovery using various domains. Decision Making and Clustering is an important concept to group the related information and make the better decisions accordingly. Risk Assessment and Evaluation is a significant task of organizations those provide project grants to Contractors. The success of project and Tender Quotation Amount depends largely on Co-Ordinators ability to determine the risk effectively. This paper focuses the key problem in the corresponding organizations such as good project proposal and bad project proposal. Risk Assessment and Prediction method is applied to evaluate the risk for granting project proposals. This method allows the users for finding the risk percentage to determine whether projects can be sanctioned to a Contractor or not. The proposed idea is tested by extracted data from organizations regarding projects approval and the results proven that it provides better accuracy.

Keywords: - Association Rule, Risk Evaluation, Principal Component analysis, Multidimensional Schema, Risk Percentage

1. INTRODUCTION

Data Mining provides many technologies to detect the hidden patterns and predict the interestingness rules. In Business field, project allotments are a

tedious process to decide. Every organizations is having its own criteria for each proposal. Based on the Criteria, Project Proposals has to be identified. But it is very difficult to identify the Contractors. Tender is Quoted and advertized the project details by organizations. Accordingly, Informations are collected along with project experience if any previous and description. The key problem in this process is to extract the information and classification of datas based on criteria. Here datas has to be classified based on the Criteria. To satisfy this problem, Data Mining offers Classification, Rule Generation and Prediction methods to perform the required task. Risk Evaluation in project grant is essential for organizations because inappropriate decisions may produce great losses. To limit this kind of risk, Contractor's has to be screened based on Project Experience, Quoted Amount, Financial Background etc. Organisation has to decide and grant the projects which satisfy their criteria's. Clustering algorithm provides some additional tools to reduce the outcome uncertainty for enhancing the quality. The objective of grant Sanction is to evaluate the risk so that the process makes better decision making. It helps to improve the speed and consistency of the project application.

Related Work

This section deals with the work related to risk evaluation and association rules.

Usman.et.al proposed a method to select a subset of informative dimension and fact identifiers from initial candidate sets. Knowledge Discovered from standard approach for mining original data[1]. Pears.et.al presented a methodology to incorporate semi-automated extraction methods. Binary tree of hierarchical clusters are constructed for each node. This paper presented new method to generates rank for nominal attributes and generate candidate multidimensional schema [2]. Liu.et.al recommended a system to generate association rules for personal financial schemas. The data cubes generated based on financial information and multidimensional rules are generated [3]. Chiang.et.al proposed a model to mine association rules of customer value. Ward's method initiated to partition the online shopping market into three markets. Here supervised learning is employed to create association rules [4]. Herawan.et.al presented an approach for regular and maximal association rules from transactional datasets based on soft set theory. Here Transactional datas are transformed into Boolean values based on information system [5].

2. EXISTING METHODOLOGIES AND ISSUES

(i) Discovery Diverse Association Rule [DDAR]

DDAR framework exposes knowledge discovery process by evaluating statistical scheme for finding interesting rules. Hierarchical Clustering strategy is implemented in the dataset for partitioning the dataset into smaller fragments. Hence Principal Component analysis is presented to rank the nominal attributes for generating multidimensional schema.

Limitations:

DDAR takes more time to construct multidimensional data cube and also categorization of discovered rules are not so efficient. Risk Evaluation process is very

difficult to evaluate the risk percentage for categorized attributes.

(ii) DataCubes Association Rules[DCAR]

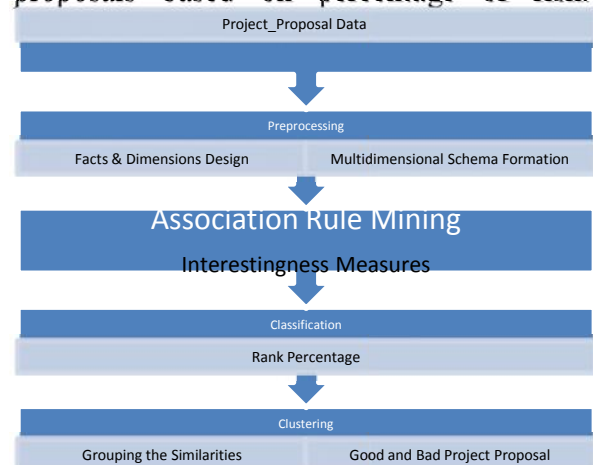
DCAR framework presented an overview of multidimensional schema for discovering association rules. Here PCA and information gain is used to evaluate the dimensions and rank the numerical attributes. Highest rank dimensions and facts are determined in the multidimensional schema using Association rules.

Limitations

Categorizing Numeric and Nominal Ranking are very tedious and takes more time to classify the highest, Medium and low levels based on ranks. Multidimensional schema formations is availed by experts not others so it is not user friendly.

3. PROPOSED METHODOLOGY FOR RISK EVALUATION AND PREDICTION

The following figure shows the main steps involved in proposed framework. The real time project request proposal application dataset is collect and used to diverse the association rules. Initially, dataset is processed to attain the high quality. Initial step removes the irrelevant datas in the dataset. The proposed framework does not include PCA strategy for nominal rankings. Here it applies Clustering strategy to categorize the project proposals based on percentage of risk.



4. PROPOSED ALGORITHM

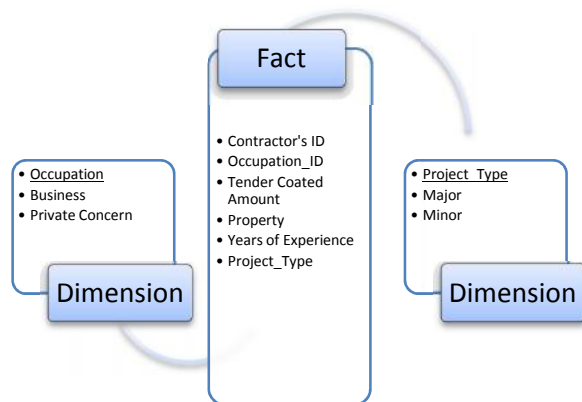
Input: Project_Proposal_Application

Output: Predicted Rules for Project Grant

Steps:

1. Set the Criteria for each Project Type
2. Extract required attributes in the Attributes based on the Criteria
3. Grouping the Attributes based on Similar Values using Clustering
4. Generate Multidimensional Schema Formation
5. Evaluate Ranks for each transaction
6. Categorize the dataset based on High and Low Ranks.
7. Set Threshold Value
8. Implementing Association Rule Mining and Evaluate Support and confidence
9. Identifying Interestingness Measures
10. Classification of Good and Bad Project Proposal based on Mean Values.

5. MULTIDIMENSIONAL SCHEMA FORMAT



Multidimensional Schema represented on the form of attributes as Dimension and Fact. Here Multidimensional Fact Shares the dimension table. Physical Structure of dataset is formed based on SQL queries. It creates essential tables for fact and dimension to define the relationship. The number of dimension can be grouped along hierarchical Structures. Each hierarchy has some levels with set of members that

represents the Subset of dimension. The above Multidimensional Schema contains all the dimensions and facts in project proposal datas. It is used to construct the information transaction set.

6. THE INTERESTINGNESS MEASURES

Rules are evaluated with the help of probability measures. It captures the usefulness of a rule. Parameters support, Confidence, Life and weight are measured to evaluate the probability.

CONCLUSION

Risk Assessment is an important task in the project industry. This paper proposed a framework for risk evaluation where bulk amount of data are engendered and risk assessment is done based on data mining technique. Rule prediction is performed for evaluating the project types. Risk Assessment is examined by using Primary and Secondary levels. Each Risk levels are evaluated based on specific criteria and attributes defined by Co-ordinators. Threshold value is set so that the proposed project_proposal range is below the limit is rejected and the rest are considered for project grant. Good and Bad Project Proposals are evaluated based on Risk Percentage.

REFERENCES

- [1] M.Usman, R.Pears and A.Fong, "Discovering diverse Association Rules from Multidimensional Schema," 2013
- [2] R.Pears, M.Usman and A.Fong, "Data guided approach to generate MultiODimensional schema for targeted Knowledge discovery", 2012.
- [3] G.Liu, H.Jiang, R.Geng and H.Li, "Application of Multidimensional association rules in personal financial services" in Computer design and Applications (ICCD), 2010 International Conference ,pp V5-500-503.

[4] 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

[5] T.Herawan and M.M.Deris, “A soft set approach for association rule mining Knowledge based Systems”, vol 24, pp 186-195,2011.