

Student's Skills Evaluation Techniques using Data Mining.

Dr. K.Kala

Head of the Department, Department of Computer Science, Nachiappa Swamigal Arts&Science College,
Koviloor-Karaiikudi.

Abstract: Technological Advancement is the mainstay of the Indian economy. Software Development is a primary factor which offers many sources of employment across the state. The major role of Software Company is to provide quality product by Managerial Decisions. To improve the Software Quality and Maintenance, Organization concentrates the Programmer's ability and restricts the Placement with various kinds of audiences such as Written test, Personal Interview, Technical Round, Group Discussion etc. Accordingly, Programmer's has to be formed. So Higher Educational institutes have to amend the student's Performance as their expectations. The student's Programming Skill is analyzed with the help of Decision making and Knowledge Discovery in Data Mining Techniques. In this example, the actual data are collected from Private College, which holds data about academic performance for pupils. This report suggests a method to judge their accomplishments through the utilization of data mining methods and specifies the means to identify their skills and assist them to improve their Knowledge by predicting Training Programs.

Keyword: Decision_Making, Association_Rule, Clustering, Classification, Rule Prediction, Preprocessing

I. Introduction

Due to high competition, it is essential to consider the customer relationship management. Here analyze the massive bulk of client data and classify them based on their behaviors and prediction. Customer relationship management is mainly used in sales forecasting. Data mining offers the technology to examine the mass bulk of data and detect hidden patterns in data to convert raw data into valuable information. It is a powerful new technology with outstanding potential to help institutions focus on the most information in their data warehouse. Data mining is the extraction of required data or info from large databases. The key thoughts are to practice data mining techniques to separate the customer information according to the posterior probability. Here the Data mining concept is employed to perform the classification and prediction of Programming Skill.

Skill Evaluation is important for Institutions, organizations due to some difficulties in linking up with inappropriate selection that may reflect in less number of placement opportunities. When considering the Programming skill, decisions regarding the application for a Student application such as Project experience, Software ability, Logical ethics, etc. tends to utilize the direct scrutiny committee. To examine their programming skills with experience. Before selecting the scholars based on their category, governance has to take various steps such as performance of the firm by finding out the last year academic performance status and Project Experience.

Sometimes with huge data and technology slackness, decisions may be faulty and resulted in project failure defaults. Here data mining offers Clustering algorithm which is based on intelligent information system for providing instruments with additional required information, to reduce the uncertainty of the outcome for enhancing the service quality. The objective of Skill_Set evaluation is to manage the risk involved in providing grant risks. And then that the system creates better decisions quickly. Selection scoring helps to increase the speed and consistency of the project application. The main objective of this paper is to propose Association Rule Mining techniques for analysing Student's Programming Skills and producing an accurate result. Association Rule Mining is one of the data mining techniques which is required to generate associations between the items. Minimum Support and Minimum Confidence can be used to compose the relationships and predict the interesting patterns. Support and confidence measures the interesting patterns. The rule composed of two itemsets

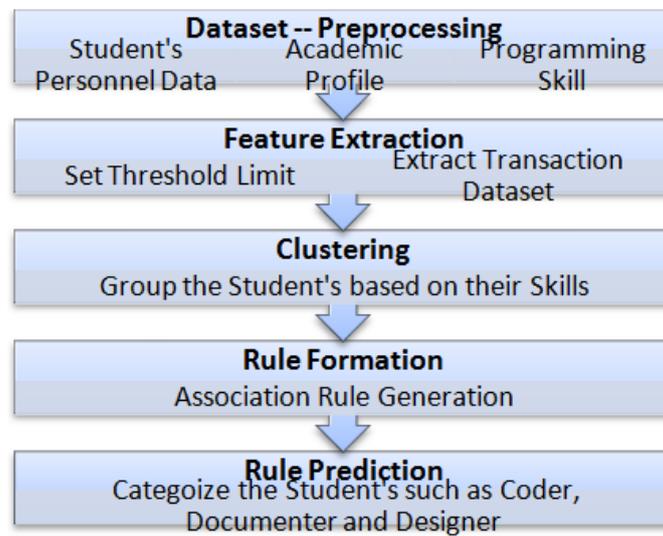
- Antecedent
- Consequent

The above attributes represents the relationship between support, confidence and lift. Association rule mining is the process of finding all the rules with the condition of support and confidence. Initially, Support and confidence is calculated for all the rules and it is compared with the user specified or predefined threshold values. Data mining techniques help the institution to extract patterns like students having extra ordinary features.

Rest of this paper is structured as follows. Research works related to Risk Assessment are discussed in Section2. Detailed explanations of implementing Student's Skill in Programming Side using Prediction Algorithm presented in section 3. Finally section 4 concludes this research work

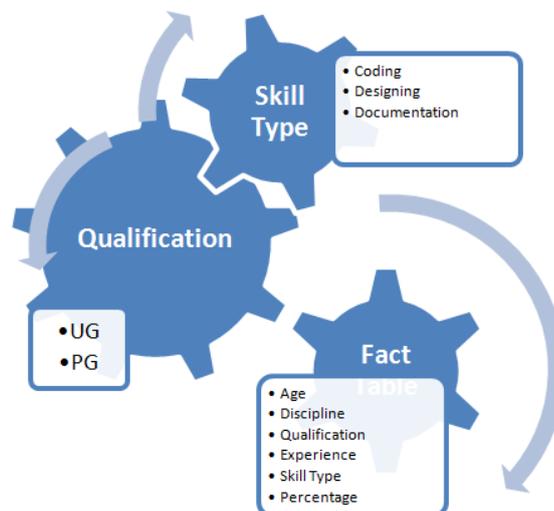
Implementing Proposed Methodology in Students Database

Recognizing the Programming Skill is an important problem in the Higher Educational institutes. The decisions for the best programming skill students should be evaluated properly so that, it may increase the number of placements. ERPCA[] implemented to make the evaluation for Programming skills in an enhanced manner. The overall flow of the proposed work is as follows.



In this Paper, Student details are collected and feature extraction is made on these attributes. Process the data includes the accumulation of details from the Students and converting into a format suited for excavation. The data cleaning process includes filtration by filling the missing values, extraction of duplications, removing outliers and resolving the inconsistencies. Patterns are made for each Skill type like Coding, Designing, Documentation. Final Assessment is done in two levels such as Students Skills are Clustered based on the prediction as Coder, Designer and Documenter.

Here, each student who applies for an interview has to provide their personal, experience and Skill details. These are stored in the dataset for further access. Student details are to be prepared in such a manner suitable for Data Mining. The dataset contains attributes like age, Qualification, Experience, Skill type etc are as follows



Based on the requirements extracted from the placement cell, Valuable attributes are selected using feature Extraction.

Data Preprocessing

The following steps are performed as part of the preprocessing in the dataset. Skill type contains large number of values. So, it is grouped into three categorical segments such as Coder, Designer and Documenter as follows.

- If Logic= TRUE && Syntax= TRUE then Skill_Type= "Coder"
- If Creativity= TRUE && Software Skill = TRUE then Skill_Type= "Coder"
- If Content_Writing= TRUE && Grammar_Skill= TRUE then Skill_Type= "Coder"

Feature Extraction

Dataset contains many attributes which may also contain irrelevant data for the mining task. Though it is possible for domain experts which makes the task difficult and consumes time. Irrelevant attributes leads lots of confusion and increases the data size. Hence the corresponding attribute has to be eliminated to decrease the size. The target of this process is to find a minimum number of attributes and reduce the Computational Complexity. Information gain is calculated for extracting the features.

$$Infogain(A) = Info(D) - Info_A(D)$$

Where A is the attribute investigated.

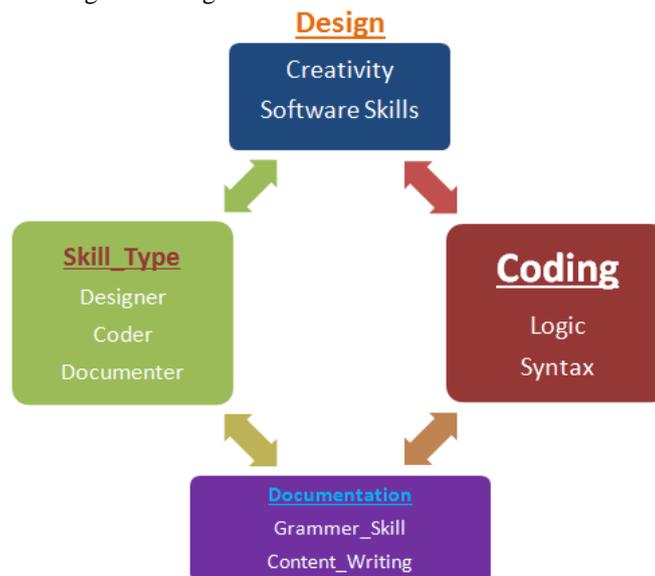
$$Info(D) = - \sum_{i=1}^m p_i \log_2(p_i)$$

Where P_i is the probability (Class t in Dataset D)

In this technique, entropy of random variables has introduced which provides the long term behavior of random process for analyzing the data. Random process behavior is the key factor for developing the coding for information theory. Entropy is the measurement of average uncertainty of data collection. Infogain of an attribute is used to select the best splitting criterion attribute. Highest value Infogain is used to build the decision tree.

Data Clustering

In Computing, Student segmentation allows reduced exposure to best assessment. It allows personalized Services according to their skills and matching Campaigns to Students. A Classifier model is essential in Student segmentation area. Students are classified based on the Skill type that is Students belong to the same category are combined together using Classifier Method.



Rule Formation

Each Software Company has different rule criteria that has to be satisfied by the student to get the Job. Apart from the rules created before, the new rules also can be created in this method. To get a job, Student has to satisfy particular touchstones like Qualification, age, Skill_type, Academic_Percentage, Discipline as shown in Table 1.

Skill_Type	Age	Qualification	Acadmeic_Percentage	Discipline
Coder	25	PG	80	Computer
Designer	22	UG	75	Other
Documenter	25	PG	75	Computer
Designer	23	UG	65	Computer

Rule Prediction

To identify the Student's Programming Skill Type, Threshold value is evaluated with extracted one. Rules are predicted based on their Skills which satisfy the Threshold limit. ie Organization can easily identify the types of Skill and classify it using classifier methods. Then the Skill Students are clustered separately for efficient retrieval.

II. Conclusion

This study examines the factors associated with the assessment of Student with Programming performance. Better Prediction of the training course, obtains its way to reach the highest level of quality in their performance. The proposed model evaluates their programming skills through the utilization of data mining techniques like Association, Classification and Clustering techniques. Hopefully it serves them to better attend to the training process and improve their Programming Skills.

References

- [1]. D.Magdalene Delighta Angeline, "Association Rule Generation For Student Performance Analysis Using Apriori Algorithm", The SIJ Transactions on Computer Science Engineering and its Applications(CSEA), Vol. ,No. 1, March –April 2013.
- [2]. Randa Kh. Hemaïd, Alaa M.EL-Halees, "Improving Teacher Performance Using Data Mining", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.
- [3]. Charanjit Bambrah, Minakshi Bhandari, Nirali Maniar, Prof. Vandana Munde[†], "Mining Association Rules in Student Assessment Data", International journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 3, March 2014.
- [4]. Mrs.Smita Bhanap, Mrs.Rasika Kulkarni, "Student – Teacher Model for Higher Education System", Current Trends in Technology and Science, ISSN: 2279-0535, Volume: Volume-II, Issue: Issue-III.
- [5]. P.Ajith, B.Tejaswi, M.S.S.Sai, "Rule Mining Framework for Students Performance Evaluation", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307. Volume-2. Issue – 6, January 2013.
- [6]. Jayshree Jha¹ and Leena Ragha², "Educational Data Mining Using Improved Apriori Algorithm", International journal of Information and Computation Technology .ISSN 0974-2239 Volume 3, Number 5 (2013).pp. 411- 418.
- [7]. T. Jacobson and K. Roszbach, "Bank lending policy, credit scoring and value-at-risk," Journal of banking & finance, vol. 27, pp. 615-633, 2003.
- [8]. G. Kabir, I. Jahan, M. H. Chisty, and M. A. A. Hasin, "Credit Risk Assessment and Evaluation System for Industrial Project."
- [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. M. Mazid, S. Ali, and K. Tickle, "Improved C4. 5 algorithm for rule based classification," in Proceedings of the 9th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases, 2010, pp. 296-301.