International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.6, No.10, October 2017 DOI:10.15693/ijaist/2017.v6i5.29-37

A Review on Learner's Performance Forecast Model in Data Mining

[1]N. Muneeslakshmi, [2]P. Sundari,
[2]Assistant Professor, [1]Research Scholar,
[1, 2,] Department of Computer Science
[1, 2,] Government Arts College (Autonomous) Coimbatore-18, Tamil Nadu

Abstract: In recent days, the performance of learners is important to evaluate. The individual student's level has great importance to process and maintain the whole organization. Accordingly, performance management is needed at each learner's level and at business level to implement a system in order to measure the overall performance. In data mining applications the knowledge discovery of interest in Human Resources Management (HRM) is applicable. To extract the knowledge significant data mining classification techniques were used. This review is presented here to monitor the importance and significance of performance forecast that helps to coin out the learners from a group to monitor the overall outcome.

Keywords---Prediction, Data Mining, Feature Selection Techniques, Classification, Students' Performance Prediction.

I.INTRODUCTION

In the updating world the database management plays a major role. In this research the learner's performance is evaluated using the classification techniques. This section discusses the brief introduction about the survey. Data mining is sometimes called as Knowledge Discovery in Databases (KDD). Generally data mining is a combination of machine learning, statistics and visualization. It is also used to extract the information. Educational data mining is wide area that provides machine learning and statistical information. Schools, colleges as well as universities have necessity to judge the academic performance of students by grades or external and internal marks. The general prediction model is show in the figure 1.

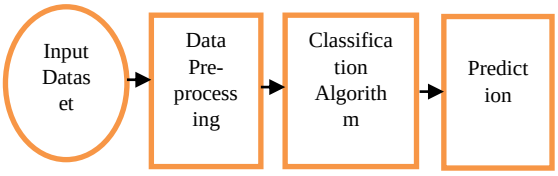

Figure 1 Basic Prediction model

Learner's retention is one of the indication of organizational performance and the management enrolment. To process a database it is necessary to follow the steps such as classification, clustering and association. For any applications the data is to be iterative i.e., the data transferred from the source to destination is always varied with respect to the iteration.

In any school/ colleges, the students' performance is to be maintained and evaluated carefully in order to maintain the effective path between them. It helps to make a reward point or focus them into any extracurricular activities. The student's database is maintained based on his name, age, gender, qualification, professional training, and extracurricular activities and so on.

An overview of the data mining and future directions are carried out by the Kurgan and Musilek [14], it focusses on Knowledge Discovery and Data Mining process model. Based on the classification techniques in data mining several methods processes in Neural Network by Phyu (2009) [19], Rough Set Theory, Decision Tree and Bayesian theory and Fuzzy logic. In some cases the Decision tree is the most popular classification techniques. The rules which is selected from decision tree is used for future prediction.

The rest of this article is organized as follows, in section 2 the Importance and significance of performance forecast is discussed. In section 3, the detail survey is framed based on several prediction models. Finally, the research is concluded with some suggestions to improve the performance factors, in section 4.

II.IMPORTANCE OF PERFORMANCE PREDICTION

In recent days, the expansion in the number of individuals entering college has added to a more significant fluctuation out of sight of those initial programs. Two principle explanations behind this are to deliver an appropriate first year educational programs and to find out whether a learners need any extra help or not. Along these lines, all together that any vital strides can be taken to enhance a learner's execution, a definitive objective would be the capacity to anticipate the future prediction.

III.LITERATURE SURVEY

The evolutionary algorithms for data mining and knowledge discovery are described by Freitas [9]. The detail description about the individual representation, fitness function and genetic operations were adapted here. The use of genetic programming is represented with detail steps of processing. Some of the properties of the extraction also mentioned for some applications.

The students' performance based on decision tree classifier is evaluated by Yadav [21]. Classification and Regression Trees (CART), ID3 (Iterative Dichotomiser and their extension C4.5 algorithm were discussed. The student dataset from VBS Purvanchal University is considered and processed. A staff Performance Appraisal is framed by utilizing Fuzzy Evaluation by Arbaiy and Suradi [3]. An application of Multi-factorial Evaluation Model is proposed in terms of performance appraisal system by Yee and Chen [22]. This work contributes and concentrates only on Quantity of Working Output, Quality of Working Output, Punctuality and Effectiveness of Working Output. The dataset were taken from Malaysia named Information and Communication Technology based company. The staff performance were evaluated and indicated the weightage for each of these aspects need to be define in the system.

Some challenges and issues based on the employees are examined by Yusof [23]. Moon, C., Lee, J., Jeong, C., Lee, J., Park, S., & Lim, S presented a Systematic performance appraisal and ranking of candidates for promotion and maintain the statistics [16]. Here the framework is used to evaluate from the unique performance evaluation criteria. The framework allows some criteria such as quantitative scores and qualitative characteristics of each candidate, it ensures the fairness, objectivity and transparency. Based on the membership function and fuzzy set operations the ranking is applied.

The examination carried out by Agodini is to inspect the relationship of parental support, free learning and accomplishment [1]. A national specimen of tenth review 15,362 undergraduate

from the national instructive longitudinal investigation 2002 was utilized. The discoveries demonstrated that parental support had positive and noteworthy association with scholarly accomplishment. Results additionally uncovered that parental control was contrarily connected with accomplishment measures. It was additionally discovered that parental activities related with definitive child rearing style were emphatically related with scholarly accomplishment.

The analysis of parental contribution in their children's scholarly inspiration in provincial territories at essential level by Ghazi, S. R., Shahzada, G., & Ullah [10]. The examination was led on a specimen of 250 undergraduate from Bannu in Pakistan. Information were gathered through organized meeting from the outcomes. Demonstrated that parents‟ support, discourse of significance of instruction and instructive issues had coordinate positive impact on accomplishment inspiration. The discoveries likewise uncovered that the majority of the guardians were not very much aware of their part for their children's training.

Muola explored the connection between scholarly accomplishment inspiration and home condition among standard eight understudies and their folks.. The example contained of 235 Kenyan understudies and their folks. Between the ages run 13 to 17 years from six urban and rustic grade schools of Machakos area. Two surveys, the straightforward profile and home condition poll, were utilized to assemble information. Results demonstrated that among all the factors parental support was the main factor that was not altogether (r = 0.03) identified with scholarly accomplishment inspiration.

Ghazi, S. R., Shahzada, G., & Ullah investigated the relationship between students' personality traits and their academic achievement in Khyber Pakhtunkhwa, Pakistan [11]. Theoretical framework of this study is based on Big Five Personality Trait Theory (Cattell's and Eysenck's 1973). The objective of the study was to find out different personality traits of secondary school students and to identify the relationship between personality traits and their overall academic achievement. Results of the study revealed that no significant relationship was found between the students' personality traits and their academic achievement.

Dutt, A., Aghabozrgi, S., Ismail, M. A. B., & Mahroeian analyzed the application of data mining methods in the educational sector is an interesting phenomenon [7]. It sets to uncover the previously hidden data to meaningful information that could be used for both strategic as well as learning gains. In this author have detailed the various disparate entities that are widely spread across in the educational forum. However, collectively they have not been addressed and this paper serves to bridge this gap. We would continue to pursue our research in clustering algorithms as applied to educational context and will also be working towards generating a unified clustering approach such that it could easily be applied to any educational institutional data set without any much overhead.

Okafor and Anaduaka stated that "The Nigerian School Children and Mathematics Phobia: How the Mathematics Teacher Can Help" [18]. It considered the importance of mathematics and why it is of utmost necessity that all Nigerian children acquire

Mathematical knowledge. It noted with dismay the general students' negative attitude towards the subject and their consequent poor performance in it especially in the West African Senior School Certificate Examination (WASSCE). It then looked

at the attributes of a mathematics teacher that can bring about a change in the teaching and learning of mathematics and in the attitude of students towards the subject. It finally considered best practices for mathematics instruction, such practices that can stimulate and sustain students' interest in mathematics learning.

The recent literature in the area of telecom customer churn mainly with two perspectives, i.e. technique being applied to telecom churn prediction and the publication year, it is framed by Madan, M., & Chopra, M [15]. Also the aim is to help the researchers to gain insight into the recent trends in this area, which will guide them in finding the possible reasons for the churn and consequently reducing them and helping the telecom sector to reduce their financial loss.

Akagah F.A. (2013) carried out a study to know the influence of study habits on academic performance of Junior high school students in the Gomoa west district. Findings revealed that despite other dimensions of study habits thus; examination, homework and assignment, reading and note-taking, concentration, only time management explains the bulk of the variables that predicts students' study habits. It was found that male students used examination, homework and assignment, reading and note-taking, concentration and time management related study habits more than their female counterparts. Older students develop better study habits than their younger counterparts and hence, perform better academically than younger students. Study habits were found to be significantly related to all socio-economic variables (such as sex, age, parents' level of education, fathers' and mothers' sector of education) of a student [2].

Mutodi and Ngirande carried out an investigation related as "The Exploring Mathematics Anxiety: Mathematics Students' Experiences" [17]. The purpose of their research was to explore students 'mathematics anxiety levels at a selected tertiary institution in South Africa. The results also show high levels of mathematics anxiety among female students. The t-test showed that the mean difference between mathematics anxiety and gender is significant. Based on the findings of this study, it is worth noting that mathematics anxiety is one psychological factor that affects students' achievement and their general practices.

In pre-processing step highlight determination calculations are utilized to distinguish highlights that will influence the forecast procedure the most. It can be seen many components like statistic factors, financial variables, family factors, and so on can be in charge of understudies from dropping out from that point courses. Subsequent to planning information for mining grouping calculations are connected and by examination of choice tree and enlistment rules we get the forecast display that is tried on test information can discover helpful learning. Result got from such models can help instructors and administration to recognize the issue zones and reasons that influence dropout the most. Creator presents investigation of informational collection utilizing information mining calculations. After investigation the result will be the central point that influence understudy dropping out of the open courses the most (dropout rae). Before applying order calculations some component determination calculations are likewise utilized to get refined forecast comes about. Such examination and forecast data will help school administration and instructors to improve essential changes for bestowing training. Mining of valuable learning should be possible by utilizing numerous

other mining methods like affiliation, grouping. Apparatus utilized for include choice and mining is weka.

There are a few restrictions of forecast display which envelops following focuses:-

- The social occasion and examination of authentic information, present and future embroils a ton of time and cost. Thusly, administrators need to adjust the cost of forecast with its advantages. Most little associations don't include in estimating on account of the high cost.
- Forecasting can just guestimate the future occasions. It can't make affirmation that these occasions will happen later on. Long haul expectations will be less honest when contrasted with here and now estimate.
- Forecasting depends on specific traditions. On the off chance that these traditions aren't right, the estimating will not be right. Estimating is totally rely upon chronicled occasions. In any case, history may not rehash itself consistently.
- Forecasting requires appropriate judgment and aptitudes with respect to chiefs. Estimates may go incorrect because of devilish judgment and aptitudes with respect to a portion of the directors.

Table-1 exhibit how forecast over informational index functions in various stages.

S.No.	Author Name	Variables Descriptions	Algorithm Used	Merits	De Merits
1.	Quadri, M. M., & Kalyankar, N. V. (2010) [20]	Gender, Attendance Parented Parent income Scholarship First child Working Dropout.	Weka J48 Classifier	It characterize the students 'motivation in terms of the other attributes that are automatically generated from the log data.	Calculations get very complex, particularly if many values are uncertain.
2.	Kabakchieva, D. (2013) [12]	Bad Average Good Very good Excellent Weighted Average.	Naïve Bayes classifier, 10-fold cross validation.	University sample data reveal that the prediction rates are not remarkable (Vary between 52-67 %).	Need to estimate more variables then only it is easy to predict the process.
3.	Elbadrawy, A., Studham, S., & Karypis, G. (2014) [8]	Assignment quiz num-attempts course-level cumGPA	Multi-regression	Accurate ability to identify outliers, or anomalies.	Some incomplete data and falsely concluding results in causation.

		cumGrade n-viewed-mater n-viewed-mater-1 n-viewed-mater-2 n-viewed-mater-4 n-viewed-mater-7 n-other-accesses student-bias Course-bias.	model		
4.	Kakadiya, S., Moodbidri, K., & Nair, S. (2015) [13]	Uniqueness_of_Idea Usability Price and need Location.	Decision-tree algorithm	This method may be made more efficient by identifying the accurate attributes pertaining to Start-ups.	Very long and time consuming.
5	Desai et al., (2016) [5]	Only Marks	Classifier based on K-means algorithm	This study works to identify those Students which need special attention to reduce failing ratios and taking appropriate action at right time.	Less accuracy, need to add detailed profile, Curriculum, extra educational courses.
6	Dey et al., (2017) [6]	Aman, Aus and Boro rice	Modified Nonlinear Regression (MNR)	Accurate set of weather statistics	Need to increases the statistics to improve the prediction
7	Bansal et al., (2017) [4]	Only 5 Attributes	Improved k-mean clustering algorithm.	Cluster quality is improved.	Less accuracy level.

IV.CONCLUSION

In this paper, the arrangement is utilized in learner's concept to anticipate the division on the preface of past data. As there are a few methodologies that utilized for information grouping or classification. Some researchers focused on Data like gathering activity, class test, class and marks were gathered from the understudies' past data, to anticipate the execution at the highest point of the semester. This investigation can encourage the students and the professors to support all types to perform well. This examination makes a difference to spot out those understudies who require uncommon consideration, limit the disappointment proportion and to make satisfactory move for upcoming semester examination.

- Future work incorporates applying information preparing strategies for different data set with extra credits.
- The proposed concept will be in the form of modified K-means classification strategy.
- It must help to improve the classification accuracy among students learning concepts.

References

[1]. Agodini, R., Harris, B., Atkins-Burnett, S., Heaviside, S., Novak, T., & Murphy, R. (2009). Achievement Effects of Four

[2]. Early Elementary School Math Curricula: Findings from First Graders in 39 Schools. NCEE 2009-4052. National Center for Education Evaluation and Regional Assistance.

[3]. Akagah, F. A. Influence of Study Habits on Academic Performance of Junior High School Students in the Gomoa West District of Ghana.

[4]. Arbaiy, N., & Suradi, Z. (2007, September). Staff performance appraisal using fuzzy evaluation. In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 195-203). Springer US.

[5]. Bansal, A., Sharma, M., & Goel, S. (2017). Improved k-mean clustering algorithm for prediction analysis using classification technique in data mining. International Journal of Computer Applications (0975–8887) Volume, 157, 33-40.

[6]. Desai, A., Shah, N., & Dhodi, M. (2016, August). Student profiling to improve teaching and learning: A data mining approach. In Data Science and Engineering (ICDSE), 2016 International Conference on (pp. 1-6). IEEE.

[7]. Dey, U. K., Masud, A. H., & Uddin, M. N. (2017, February). Rice yield prediction model using data mining. In Electrical, Computer and Communication Engineering (ECCE), International Conference on (pp. 321-326). IEEE.

[8]. Dutt, A., Aghabozrgi, S., Ismail, M. A. B., & Mahroeian, H. (2015). Clustering algorithms applied in educational data mining. International Journal of Information and Electronics Engineering, 5(2), 112.

[9]. Elbadrawy, A., Studham, S., & Karypis, G.(2014). Personalized multi-regression models for predicting students performance in course activities. UMN CS, 14-011.

[10]. Freitas, A. A. (2003). A survey of evolutionary algorithms for data mining and knowledge discovery. In Advances in evolutionary

computing (pp. 819-845). Springer Berlin Heidelberg.

[11]. Ghazi, S. R., Ali, R., Shahzad, S., & Hukamdad,H.(2010).Parental involvement in children academic motivation. Asian Social Science, 6(4), 93.

[12]. Ghazi, S. R., Shahzada, G., & Ullah, S. (2013). Relationship between students' personality traits and their academic achievement in khyber pakhtunkhwa, pakistan. Journal of Educational and Social Research, 3(2), 437.

[13]. Kabakchieva, D. (2013). Predicting student performance by using data mining methods for classification. Cybernetics and information technologies, 13(1), 61-72.

[14]. Kakadiya, S., Moodbidri, K., & Nair, S. (2015). Analyzing Start-up Success Possibility Using Data Mining.

[15]. Kurgan, L. A., & Musilek, P. (2006). A survey of Knowledge Discovery and DataMining process models. The Knowledge Engineering Review, 21(1), 1-24.

[16]. Madan, M., & Chopra, M. (2013). Data mining: a mode to reform today's higher learning institutions through performance indicators.

[17]. Moon, C., Lee, J., Jeong, C., Lee, J., Park, S., & Lim, S. (2007, October). An implementation case for the performance appraisal and promotion ranking. In 2007 IEEE International Conference on Systems, Man and Cybernetics (pp. 661-667). IEEE.

[18]. Mutodi, P., & Ngirande, H. (2014). Exploring mathematics anxiety: Mathematics students' experiences. Mediterranean Journal of Social Sciences, 5(1), 283.

[19]. Okafor, C. F., & Anaduaka, U. S. (2013). Nigerian school children and mathematics phobia: How the mathematics teacher can help. American Journal of Educational Research, 1(7), 247-251.

[20]. Phyu, T. N. (2009, March). Survey of classification techniques in data mining. In Proceedings of the International MultiConference of Engineers and Computer Scientists (Vol. 1, pp. 18-20).

[21]. Quadri, M. M., & Kalyankar, N. V. (2010). Drop out feature of student data for academic performance using decision tree techniques. Global Journal of Computer Science and Technology.

[22]. Yadav, S. K., Bharadwaj, B., & Pal, S. (2012). Data mining applications: A comparative study for predicting student's performance. arXiv preprint arXiv:1202.4815.

[23]. Yee, C. C., & Chen, Y. Y. (2009). Performance appraisal system using multifactorial evaluation model.

[24]. Proceedings of World Academy of Science: Engineering & Technology, 41, 231-235.

[25]. Yusof, A. A. (2003). Performance Appraisal: Issues, Challenges & Prospects. Prentice Hall.

Author Profile

P. Sundari received the B.E. degree in Computer Science Engineering in 1992, from Government College of Technology College, Coimbatore India. She received M.C.A degree in 1998 from Bharathiar University, Coimbatore, India and completed MPhil degree in 2009 from Kongunadu Arts& Science College in Coimbatore. Currently, she is working as an Assistant Professor in Department of Computer

Science at Government Arts College (Autonomous) Coimbatore. India.

N. Munees Lakshmi received the B.C.A degree in Computer Science Engineering in 2014 from Bharathiar University, valparai, India and M.SC. Degree in Information Technology in 2016 from Michael job Arts& Science College, Sulur, India. Currently she pursuing as an Mphil Research Scholar in Government Arts College (Autonomous) Coimbatore, India.