



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 11, November 2018

An overview of Machine Learning Techniques for Student Performance Analysis

Nilesh Machale, Ashitosh Dhaware, Amruta Khandare, Prof. S. R. Vispute

U.G. Students, Department of Computer Engineering, Pimpri-Chinchwad College of Engineering, Pune,
Maharashtra, India

ABSTRACT:The main objective of the paper is to have review of the Machine Learning Techniques for developing a system to support student for their jobs and further academics. The system will help student to provide advice in Easy format in the form of prediction. The Objective of this work is to analyze student performance for suitable job prediction and higher study guidance using machine learning techniques. Predictive analytical techniques helps user to predict future for new instances using trained module. Survey shows that Naïve Bayes classifiers and Regression analysis better performs as compared to other classifiers. System will provide platform to develop predictive model to analyze student performance using parameters – Attendance, unit test, progressive assessment Semester marks and overall performance. This process is heavy by manual for teachers to classify or predict the result accordingly student performance. An educational institute should have a prior knowledge of registered students to predict their performance in future academics. This helps them to identify bright students and also gives them an opportunity to pay attention and improve those students who would probably get lower grades. As a solution, work focuses to develop a system which can predict the performance of students from their previous performances using concepts of data mining techniques.

KEYWORDS:-Career zone, Prediction and Data Mining, Performance Analyzer, Student Performance, Classification.

I. INTRODUCTION

As the world is trending into new technologies and implementations it's a necessity goal to trend up in Education to boot. Every year, instructional institutes and MNC corporations admit students for courses and Jobs severally, with varied advantage scores in entrance examinations. Analyzing the past performance of scholars we are able to predict the long run scope of scholars. this may all right be achieved mistreatment the ideas of knowledge mining. For this purpose, we've got analyzed the info of scholars listed in third and final year of engineering. This knowledge was obtained from the knowledge provided by the admitted students to the institute. It includes their full name, scores teachers, semesters of Engineering, scores in entrance examinations, class and admission sort. wetend to then applied the Naïve Bayes, KNN and regression formula. This method can offer steerage to students for his or her jobs and admission per their teachers. This method aims at reducing the efforts of the scholar and school it creates the one style of awareness in fourth year student to induce their Jobs in future and admission for educational activity in next school year.

II. LITERATURE REVIEW

A. Existing system

Upon review of similar works done by existing author's, a background study is done to perform analysis of student performance. Here we have selected three existing systems as study of these systems helped us to propose our system. Automatic Student Performance Analysis and Monitoring[1] at el SnehalKekane. In this survey to understand and analyze the existing system and the algorithms that are used in it and to propose a system that will analyze student performance and will guide them by displaying the areas where they need improvement, in order to contribute to a student's overall development by generating a score card for the same in 2016. Computer assisted Decision Support for student Admission Based on Their Predicted Academics Performance. American Journal of



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 11, November 2018

Pharmaceutical Education 2017. Admissions committees of health sciences health sciences schools spend a significant amount of time and effort screening applications to assess candidates' readiness for a professional program. Competition for admission is very high when applicants are more than available slots.

Predicting student's performance using ID3 and C4.5 classification algorithm. International Journal of Data Mining & Knowledge Management Process (IJDKP). An educational institution needs to have an approximate prior knowledge of enrolled students to predict their performance in future academics. This helps them to identify promising students and also provides them an opportunity to pay attention to and improve those who would probably get lower grades. Student placement and skill ranking predictors for programming classes using Class attitude, psychological, scale, and code metrics. Ishizu et al. Research and Practice in Technology Enhanced Learning (2018). Abstract In some situations, it is necessary to measure personal programming skills. For example, often students must be divided according to skill level and motivation to learn or companies recruiting employees must rank candidates by evaluating programming skills through programming tests, programming contests, etc. A Review on Predicting Student's Performance using Data Mining Techniques. Amirah Mohamed Shahiria, Wahidah Husaina, Nur'aini Abdul Rashida,

In this System, we tend to see the key issue within the Student community area unit at the stage of post-graduation and need to pick their career. It's chiefly because of lack of data within the space that they require to settle on. Cognitive content is that the initial issue they face that blocks them from continued to the proper destination. Finally they opt for some course and establishment indiscriminately when sacrificing their own dream. As a result of this our country loses varied many alternative potential students in various areas.

B. Algorithms Used in Existing Systems

The prediction methods used for student performance In educational data mining method, predictive modeling is usually used in predicting student performance. In order to build the predictive modeling, there are several tasks used, which are classification, regression and categorization. The most popular task to predict students performance is classification. There are several algorithms under classification task that have been applied to predict students performance. Among the algorithms used are Decision tree, Artificial Neural Networks, Naive Bayes, K-Nearest Neighbor and Support Vector Machine. Next, the specific application of data mining techniques grouped by algorithms in predicting student performance will be described in the next section.

1. Decision Tree: Decision Tree is one of a popular technique for prediction. Most of researchers have used this technique because of its simplicity and comprehensibility to uncover small or large data structure and predict the value [10]. Romero et al. (2008) said that the decision tree models are easily understood because of their reasoning process and can be directly converted into set of IF-THEN rules [10].

2.

3. Naive Bayes: Naive Bayes algorithm is also an option for researchers to make a prediction. The objective of all these papers is to find the most effective prediction technique in predicting student's performance by making comparisons [11]. Their research showed that Naive Bayes has used all of attributes contained in the data. Then, it analyzed each one of them to show the importance and independency of each attributes.

3. K-Nearest Neighbor: All papers studied in this research showed that K-Nearest Neighbor gave the best performance with the good accuracy. According to Bigdoli et al. (2003), K-Nearest Neighbor method had taken less time to identify the students performance as a slow learner, average learner, good learner and excellent learner [12]. K-Nearest Neighbor gives a good accuracy in estimating the detailed pattern for learner's progression in tertiary education.

4. Simplecart: It is a non-parametric decision tree learning technique that produces either classification or regression trees [15], depending on whether the dependent variable is categorical or numeric, respectively. It is used for

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 11, November 2018

implementing minimal cost complexity pruning.

5.NBTree: The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier[15].

r. No	Algorithm Used	Accuracy
1.	Naive Baye's	94%
2.	Decision Tree	90%
3.	KNN	84%
4.	SimpleCart	82%
5.	NB tree	86%

Table. 1. Algorithm Accuracy Table

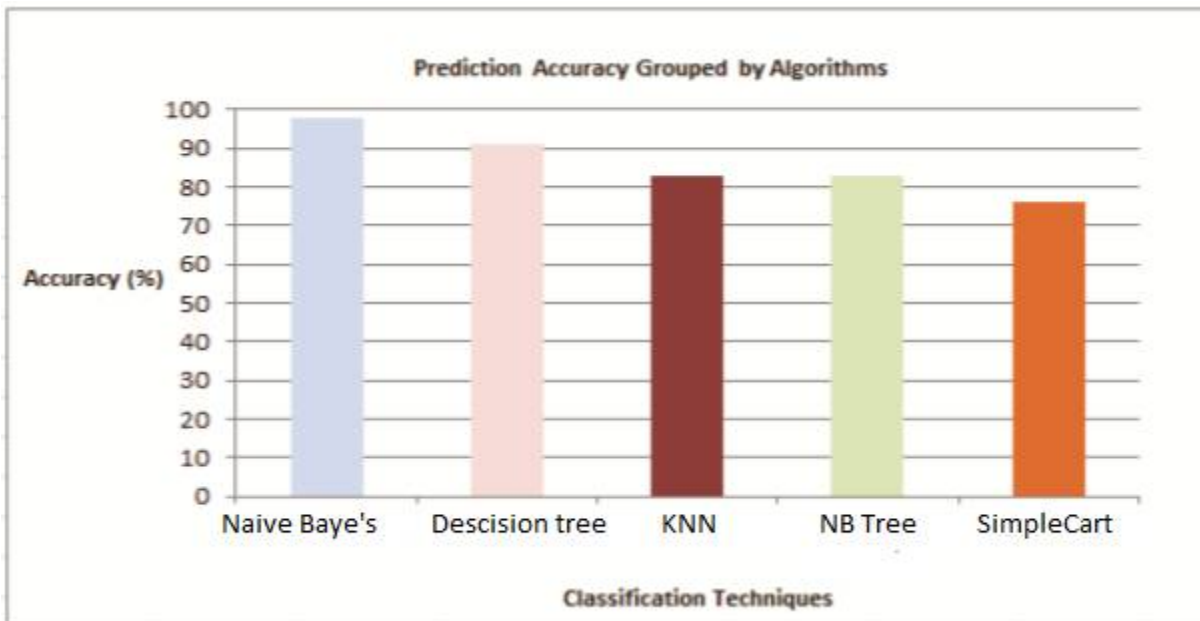


Figure 1. Prediction Accuracy Graph for Existing System

III. MATHEMATICAL MODEL

Mathematical Model for Classification techniques:

$$S2 = \{s, e, X, Y, F\}$$

Where,

s = Initial State: Input student performance data set without classification

e = End State: Classified dataset

X = Input to the system. Here it is training and testing data set in any suitable file format such as XLS,

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 11, November 2018

CSV, ARFF, class attribute with defined class
Y = Output. Classified dataset as per defined class F = Function used in the program.

Algorithm:

1. Naive Baye's algorithm

- Step 1: Convert the data set into frequency table by taking attributes for X.
- Step 2: The dataset is validated for X..
- Step 3: By using Naive Baye's equation it declares the result by validating attributes for X.
- Step 4: It jumps for another attribute for another class.

IV. PROPOSED SYSTEM

This system can give idea to students for his or her jobs and admission as per their academics. This technique aims at reducing the efforts of the student and faculty to create the awareness in 4th year student to urge their Jobs and admission in next year. Figure 2 shows block diagram for proposed system work.

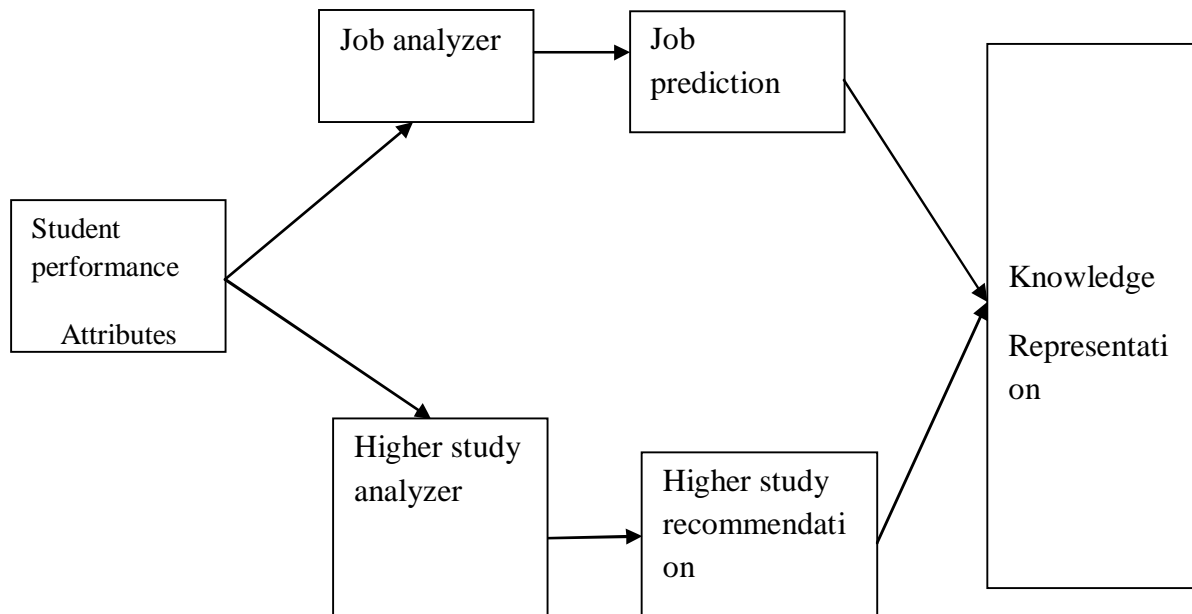


Figure 2. System Block Diagram



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 11, November 2018

V. CONCLUSION

The main objective of the system is to develop a system using machine learning technique to support and guide students for their jobs and admission, The main objective of this work is to analyse student performance and as per the basis of the students performance we can predict the job opportunity for a student as well as can provide the further options for admissions for pursuing higher studies

REFERENCES

1. SnehalKekane, DipikaKhairmar, RohiniPatil, Prof. S. R. Vispute, Prof. N. Gawande Automatic Student Performance Analysis and Monitoring, on IJIRCCE Vol. 4, Issue 1, January 2016
2. HRSPCA: Hybrid Recommender System for Predicting College Admission, ISDA International Conference on Intelligent Systems Design and Applications 2012
3. Thai-Nghe, Andre Busche, and Lars Schmidt-Thieme Information “Improving Academic Performance Prediction by Dealing with Class Imbalance Nguyen “Systems and Machine Learning Lab University of Hildesheim.
4. PERFORMANCE Paulo Cortez and Alice Silva Dep. Information Systems/Algoritmi R&D Centre UniversityofMinho“UsingDataMiningtoPredictSecondarySchoolStudent”.
5. Techniques S. B. Kotsiantis, C. J. Pierrakeas, and P. E. Pintelas Educational Software Development Laboratory, Department of Mathematics University of Patras, Greece “Preventing Student Dropout in Distance Learning Using MachineLearning”.
6. B. O. Ehigbor, and T. N. Akinlosotu, “Parents’ Occupation as Correlate of Students’ Career Aspiration in Public Secondary Schools in Ekpoma Metropolis,” Int. Journal of Arts and Humanities, 2016, vol. 5 (3), pp. 197 –212.
7. F. Olayinka, “Choosing librarianship as a career: a study of public Secondary school students in Ekiti State,” Int. Journal of Afri. &Afri. American Studies, 2008, vol. 7 (2), pp. 1 –8.
8. N. S. Ezeonu, “Career choice: A basic issue in primary and secondary school level,” Arabian Journal of Bus. andMgt Review, 2012, vol. 2, pp. 18 -28.
9. M. D. Eremie, “Comparative Analysis of Factors Influencing Career Choices among Senior Secondary School Students in Rivers State, Nigeria,” Arabian Journal of Bus. andMgt Review (OMAN Chapter), 2014, vol. 4 (4), pp. 20 -25.
10. Amirah Mohamed Shahiria,*, WahidahHusaina , Nur’aini AbdulRashidaA Review on Predicting Student’s Performance usingData Mining Techniques
11. C. Romero, S. Ventura, Educational data mining: A review of the state of the art, Trans. Sys. Man Cyber Part C 40 (6) (2010) 601–618.doi:10.1109/TSMCC.2010.2053532.
12. E. Osmanbegovic, M. Sulji c, Data mining approach for predicting student performance, Economic Review 10(1)
13. B. M. Bidgoli, D. Kashy, G. Kortemeyer, W. Punch, Predicting student performance: An application of data mining methods with the educational web-based system lon-capa, in: Proceedings of ASEE/IEEE frontiers in education conference, 2003.
14. Automated System for Soil Health Card Generation and Analysis to Satisfy Local Agriculture Needs: An Overview.Artikelkarprof.sushma vispute.2017