IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 09-12
www.iosrjournals.org

Measuring Student Performance with Data Mining

Archana Rajan Nair[1], Anusha Jagannathachari[2]
[1](KJ Somaiya Institute Of Management Studies & Research, Mumbai University, India)
[2](KJ Somaiya Institute Of Management Studies & Research, Mumbai University, India)

Abstract: *This paper focuses on how data mining can be applied to analyze educational data. With data mining students academic progress can be tracked also the prediction of final grade and results can be done. Based on the result of data mining measures can also be taken to improve student's performance. In an academic year lot of educational data is produced also every teacher maintains an individual record. All such data is collected at a central storage so that the teachers and the authorities can access the data and analyze it whenever required. In the proposed system all such data is collected at a unified data hub and then mining is done .Based on the results of mining proper action can be taken.*

Keywords: *Algorithms, Educational Data Mining, Random Forest technique*

I. Introduction

Data mining techniques are applied in many fields. Educational data mining is related to data mining in the educational field. Educational data mining have been applied to predict students academic performance based on their current and previous performance. Database of colleges and universities collect a lot of information about students every year and this information keeps increasing every year. Right from student's admission his personal data, previous educational details, every semester's mark including assignments, practical, viva, presentations marks all such data is stored in the database. There is no action taken to gain knowledge from this database. With the use of data mining techniques, patterns, trends in educational field can be obtained. Based on the result of data mining decision can be taken to improve students' performance.

II. Objectives

- To predict student performance
- To improve and manage the college or university to be more efficient in decision making on basis of correct results.
- To use data mining techniques to study performance of the students.

III. Related Work

Amjad Abu Saa made use of different data mining techniques to predict results of students from a dataset and found that student performance is majorly dependent on their academic efforts though there are many other factors influencing academic performance [2].

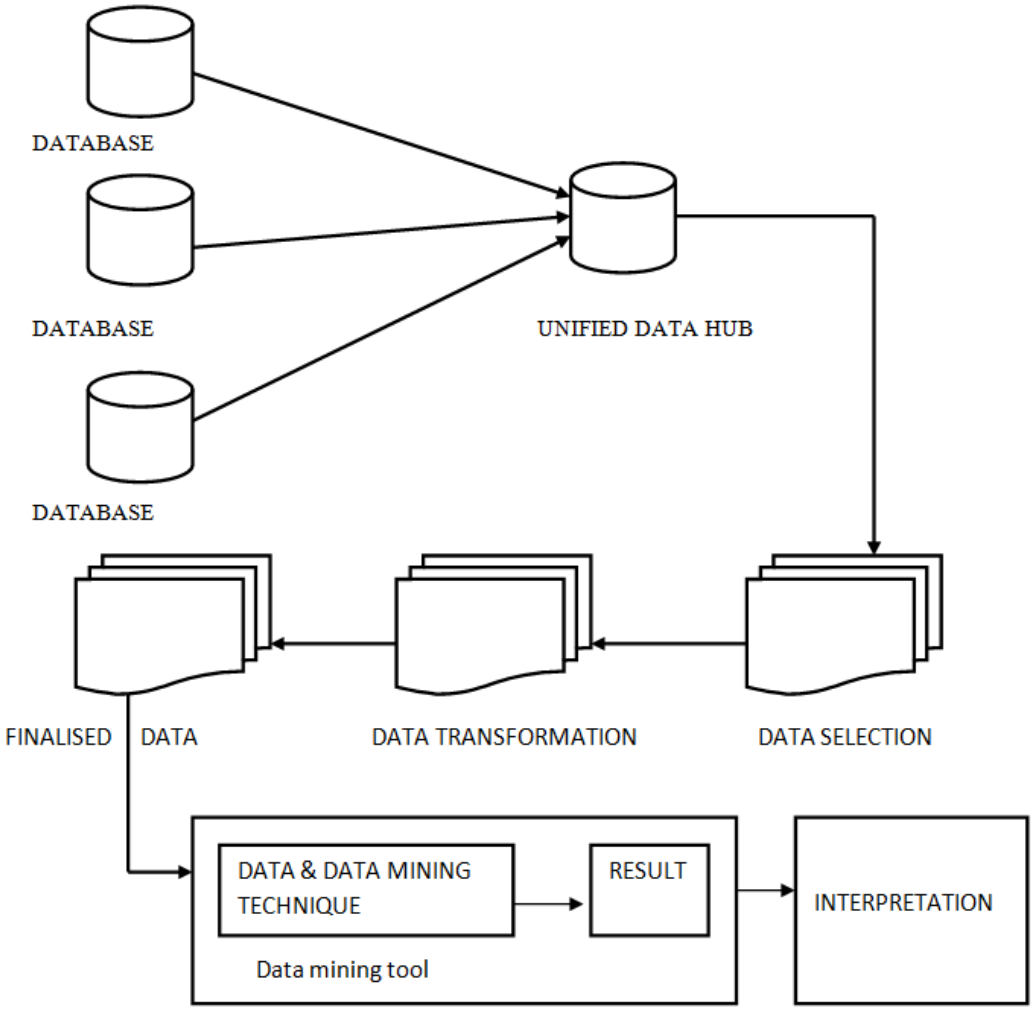
Suchita Borkar, K. Rajeswari did a study on a group of MCA students from Pimpri college of Engineering based on the percentage of assignment, attendance, unit test percentage and found association based on the input using Weka data mining tool[4].

Fadhilah Ahmad, Nur Hafieza Ismail, and Azwa Abdul Aziz did a study on a group of students' GPA (Grade Point Average). Different parameters and associated values were decided and three algorithms were used to predict results and compare these algorithms. They found out that the algorithms used can achieve the good accuracy under some circumstances [5].

IV. Design Overview

Data collected or generated at each end is stored in its respective database. All databases are connected using a unified data hub. Once this is done data is extracted depending on the requirements or the data on which knowledge discovery has to be done. The range of attributes can be decided if required. The selected data is then given as input to the data mining tool. Data mining techniques are applied and results are interpreted.

V. Diagramatic Representation Of System



VI. Working

In this paper, we have some dataset. To predict student's academic performance their previous results are considered, his social behavior is considered. Ratings are done by the lecturer on observing students class behavior. All these data are considered as input. The whole process of prediction of the result is divided into some steps as follows.

1. Collecting Data

Previous semester or previous degree's marks are considered. Current semester's assignment marks, unit test marks are considered. Along with this, teacher will rate students on their behavior, communication skills and class participation.

2. Pre Processing Data

Previous semester or previous degree's marks are considered. Current semester's assignment marks, unit test marks are considered. Along with this, the teacher will rate students on their behavior, communication skills, and class participation.

3. Data Mining And Techniques Applied

To identify patterns and trends in the large datasets data mining is used. In educational data mining patterns and trends in educational field can be identified. Input datasets in educational data mining can be student grades, attendance marks, and personal details. To find patterns and trends in data mining, data mining tool can be used. Data mining tools will have many algorithms that can be applied to large datasets to find meaningful relationships between them. Data mining tool used in this paper is Rapid Miner tool. Rapid Miner tool allows working on large datasets and applying different algorithms and data mining techniques to easily analyze the input. In this paper, we will use one data mining techniques – Random Forest technique. Random Forest technique helps to do classification of data. More number of data, more accurate the result produced by it.

4. Interpretation

Results of this technique are then interpreted and the pattern is judged.

VII. Input Data Set

For analysis and prediction of the result, we have used dummy data set. The following table shows the set of historical data to find trends. With the use of historical data, the analysis and prediction can be simpler. Data considered are results of previous education, assignment marks, unit test marks, behavior, class participation and Grade on basis of marks of end term. Based on this record set we will do prediction of results.

RESULTS	ASSIGNMENT	UNIT TEST	BEHAVIOUR	CLASS PARTICIPATION	Grade
77.33	69.33	82.00	59.33	82.67	A
83.33	55.33	66.67	76.67	58.67	A
78.67	70.00	75.33	75.33	56.00	A
57.33	58.00	84.00	85.33	67.33	A
90.00	79.33	73.33	76.00	73.33	O
86.67	60.67	72.67	77.33	95.00	A
67.33	69.33	62.00	59.33	52.67	B
63.33	55.33	66.67	66.67	58.67	B
92.00	79.33	73.33	76.00	73.33	O
80.67	60.67	72.67	77.33	95.00	A
67.33	69.33	62.00	59.33	52.67	B
63.33	55.33	66.67	66.67	58.67	B
88.00	79.33	73.33	76.00	73.33	O
67.67	67.67	72.67	75.33	95.00	A
62.33	69.33	62.00	59.33	52.67	B
63.33	55.33	66.67	66.67	58.67	B
57.33	69.33	82.00	29.33	42.67	C
50.33	55.33	66.67	76.67	58.67	C
51.67	70.00	75.33	75.33	56.00	C

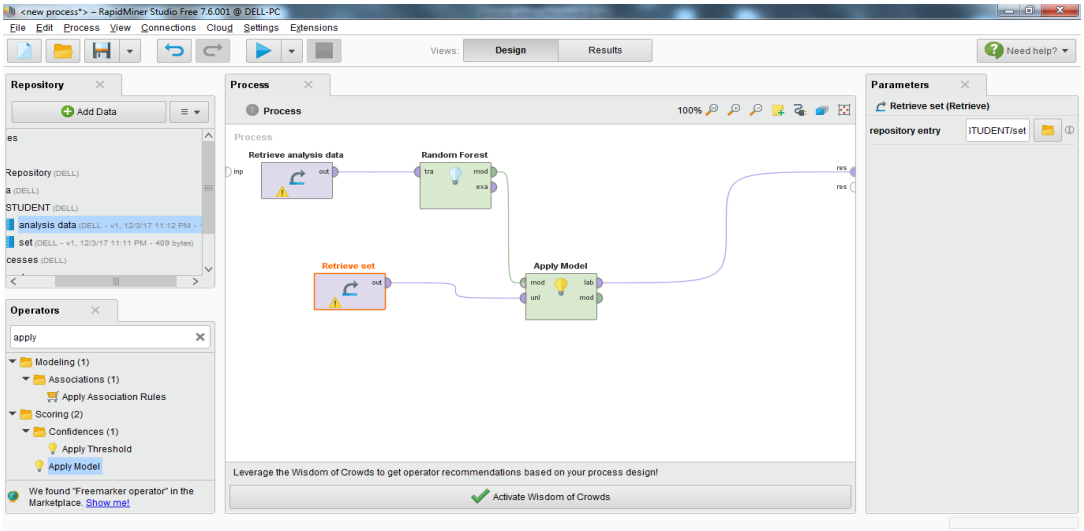
Table: 1 –Historical Data

The following table shows the set of data whose results have to be predicted. Before semester exam, prediction of the result is done based on the historical data. This dataset also has results of previous education, assignment marks, unit test marks, and behavior and class participation

RESULTS	ASSIGNMENT	UNIT TEST	BEHAVIOUR	CLASS PARTICIPATION
94	83	87	89	70
75	48	29	60	54
35	56	38	47	67
70	79	73	76	73
56	57	54	58	67
85	68	89	70	66
41	67	40	45	54

Table: 2 –Data set to predict result

VIII. Results



Screen Shot: 1 –Rapid Miner tool

Results of random forest show the prediction of result and show the confidence level of each grade. Higher the confidence of grade more chance of scoring that grade.

Prediction	Confidence(0)	Confidence(A)	Confidence(B)	Confidence(C)
O	0.883	0.117	0	0
A	0.883	0.470	0.385	0.062
C	0	0.053	0.465	0.482
A	0.383	0.400	0.217	0
A	0	0.371	0.350	0.279
A	0.383	0.821	0.070	0.026
C	0	0.221	0.333	0.446

Table: 3- Result of Random Forest Technique

IX. Conclusion

Prediction of result can be calculated with ease and measures can be taken to improve students' performance. Prediction of result helps teachers to focus more on students with poor performance. With early prediction, the score of students can be improved which will be very helpful for students as well as teachers. This study also helps to identify students who need special attention and take appropriate decision for next semester too.

References

[1]. Poonam Gupta, Pooja Jadhav , Bhagyashree Kadam , Amruta Kedari *"*International Journal of Innovative Research in Computer and Communication Engineering*"* Vol. 4, Issue 3, March 2016, pg.no: 3040-3047.

[2]. Amjad Abu Saa *"*Educational Data Mining & Students' Performance Prediction*"* International Journal of Advanced Computer Science and Applications, *Vol. 7, No. 5, 2016* ,pg.no: 212-220.

[3]. Zahyah Alharbi, James Cornford, Liam Dolder and Beatriz De La Iglesia, "Using Data Mining Techniques to Predict Students at Risk of Poor Performance", *SAI Computing Conference*, pg.no:523-531

[4]. Suchita Borkar1, K. Rajeswari2 *"* Predicting Students Academic Performance Using Education Data Mining*"* International Journal of Computer Science and Mobile Computing, Vol. 2, Issue. 7, July 2013, pg.no:273 – 279

[5]. Fadhilah Ahmad, Nur Hafieza Ismail and Azwa Abdul Aziz *"* The Prediction of Students' Academic Performance Using Classification Data Mining Techniques*"* Applied Mathematical Sciences, Vol. 9, 2015, pg.no:6415 - 6426