Student Performance Prediction System

PROF.MUSILEK
Undergraduate Student
Rajarshi Shahu college of Engineering
Maharashtra, INDIA- 411033
Email: sohankamble8@gmail.com

Dr.JULIE
Undergraduate Student
Rajarshi Shahu college of Engineering
Maharashtra, INDIA- 411033
Email: saurabhkodam@gmail.com

.Dr.HENRY
Undergraduate Student
Rajarshi Shahu college of Engineering
Maharashtra, INDIA- 411033
Email: adeshmaske@outlook.com

***Abstract*—Now, a days every university has their own management system to manage the student's records. Education can be utilized as a tool to face the problems, over many hurdles in life.The Knowledge obtained from the education helps to get the opportunities.Extracting the information from the student records data mining techniques used. Discovers the unique pattern of students annual progress and behavior provides a data mining tools and techniques to analyze and visualize. Selecting the required area build Future.Students bunk the class and wondering here and there and tell the staff I'm in hospital and there and there, By using the gps sensors it helps to staff where is the student actually.**

Keywords- Analysis of Student Performance,Decision Tree(Fuzzi logic),Naive Bayes,Educational Data Mining.

I. INTRODUCTION

Student is the main role in the education system. Education, Sports, Behavior plays an important roles for achieving the goal's of the student. There are the more no of student failed in a particular subject.For example,according to a 2016 report by job skills credential company Aspire minds,nearly 80% of engineering graduate in India are not employable. Most of them are forced to take up job in non-engineering field or remain unemployed. There is a perennial problem about the quality of (engineering) students. Unemployment ratio for the age group of 18-19 years is 12.9%, while for above 30 years it is 1.4%, which points out that there are very few opportunities for fresh entrance.

For the overcome this entrance the students have to improve the qualities.The Student Performance Prediction System(SPPS) helps to improve the qualities and skills of the students.As usual every staff stores the record of the each an every students in a file format or in the document format on computer, by analyzing these records the SPPS system helps to improve which skills and qualities mostly improve in students. However,the increasing amount of students and their data become hard to store, analyze and visualize, but by using Data mining tools and techniques it is easy.Student Performance Prediction System is the system to predict the performance of the student in a particular field.It analyze the records of the student and give the predicted output.

And for the attendance management gps and heart beat sensors are Used .These Sensors sence particular students Heartbeat and locate the position.

A. RELATED WORK

The Significant feature of Student Performance Prediction System(SPPS) it predict student performance in different area. The Student Performance Prediction System, takes the input as the students records analyze it, perform the classification techniques and predict the result according to input records.Naive Bayes classifier classify the records by using classification techniques.Ms.Tisvi Devasia,Ms.Vinushree T P, Mr.Vinayak Hegde used the Naive Bayes Algorithm and scanning the student records calculated the probability of each value, by applying the formula after that they multiplied the probability which calculated by P and finally all the values compared and classified the attribute values to one of the predefined set of class [5].

Chew Li Sa,Dayang Hanani bt.Abang Ibrahim, Emmy Dahliana Hossain, Mohammad bin Hossain used the three systems to perform student performance these systems are Faculty Support System(FSS),Student Performance Analyzer(SPA)and the Intelligent mining and decision support system [1].

B. PROPOSED SYSTEM

In the proposed system we have used first added the student records in the database in the form of key value pair and we have put search methodology in it to find the student interested or strong in which field.The classification is don by using the classification technique Naive Bayes,Decision tree,KNN.The Student Performance Prediction is included into the proposed system to make sure the objectives are achieved. Furthermore, the generation of reports in portable Document format and illustration display such as charts in PDF makes student performance easier.

From these features found in proposed system, all the user requirements would be fulfilled. The users requirements collected from lecturers, Sport teacher,culture head,social activity head college during th system analysis phase are follows:

i. Able to help to lecturers to automatically predict students performance System analysis and performance.

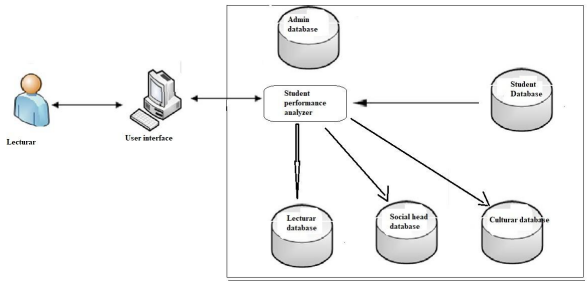

Fig. 1. Evaluation of LOLP

Attribute	Description (Domain)
sex	student's sex (binary: female or male)
age	student's age (numeric: from 15 to 22)
school	student's school (binary: *Gabriel Pereira* or *Mousinho da Silveira*)
address	student's home address type (binary: urban or rural)
Pstatus	parent's cohabitation status (binary: living together or apart)
Medu	mother's education (numeric: from 0 to 4[a])
Mjob	mother's job (nominal[b])
Fedu	father's education (numeric: from 0 to 4[a])
Fjob	father's job (nominal[b])
guardian	student's guardian (nominal: mother, father or other)
famsize	family size (binary: ≤ 3 or > 3)
famrel	quality of family relationships (numeric: from 1 – very bad to 5 – excellent)
reason	reason to choose this school (nominal: close to home, school reputation, course preference or other)
traveltime	home to school travel time (numeric: 1 – < 15 min., 2 – 15 to 30 min., 3 – 30 min. to 1 hour or 4 – > 1 hour).
studytime	weekly study time (numeric: 1 – < 2 hours, 2 – 2 to 5 hours, 3 – 5 to 10 hours or 4 – > 10 hours)
failures	number of past class failures (numeric: n if $1 \leq n < 3$, else 4)
schoolsup	extra educational school support (binary: yes or no)
famsup	family educational support (binary: yes or no)
activities	extra-curricular activities (binary: yes or no)
paidclass	extra paid classes (binary: yes or no)
internet	Internet access at home (binary: yes or no)
nursery	attended nursery school (binary: yes or no)
higher	wants to take higher education (binary: yes or no)
romantic	with a romantic relationship (binary: yes or no)
freetime	free time after school (numeric: from 1 – very low to 5 – very high)
goout	going out with friends (numeric: from 1 – very low to 5 – very high)
Walc	weekend alcohol consumption (numeric: from 1 – very low to 5 – very high)
Dalc	workday alcohol consumption (numeric: from 1 – very low to 5 – very high)
health	current health status (numeric: from 1 – very bad to 5 – very good)
absences	number of school absences (numeric: from 0 to 93)
G1	first period grade (numeric: from 0 to 20)
G2	second period grade (numeric: from 0 to 20)
G3	final grade (numeric: from 0 to 20)

Fig. 3. The details of a student which forms the data-set

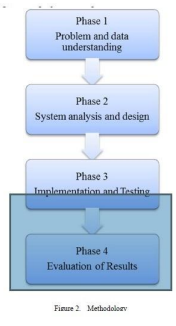

Fig. 2. Methodology

ii. Able to keep track and retrive student performance in particular course and field.
iii. Able to view the factors that affects the student prediction result.
iv. Able to generate Student Report.

The proposed system architecture is designed as shown in below:

C. MOTIVATION

The data-set considered consists of 395 tuples and 34 attributes [1]. Each tuple represents the attribute values of a student or it provides the details of the student in terms of academic performance and social behavior.

D. CLASSIFICATION ALGORITHM

I. Na¨ıve Bayes: Na¨ıve Bayes Algorithm: Na¨ıve Bayes Algorithm are easy to implement. The classification rate of Na¨ıve Bayes is very high. It predicts accurate results for most of the classification and predication problems. Na¨ıve Bayes algorithm needs large training dataset. Na¨ıve Bayes algorithm's precision rate decreases if dataset is small. Results can be good only if data is large.

II. K-Nearest Neighbor Algorithm: K-NN algorithm has one distinct advantage as classes do not have to be separated linearly. There is no cost in the learning process. It is strong enough to handle noisy data. K-NN has many limitations. If the dataset is large, it can be very time- consuming. K-NN is sensitive to noise. The performance of K-NN algorithm depends on the dimensions used.

III. A decision tree is a simple structure where each nonterminal node represents a test or decision on the considered data item. A decision tree can be used to classify an instance by starting at the root of the tree and moving through it until a leaf node
Build Tree (Node t, Training Database D, Split Selection Method S)
1. Apply S to D to find Splitting criterion
2. If (t is not a leaf node)
3. Create Children nodes of t
4. Partition D into children Partitions
5. Recurse on each partition
End if

II. PREDICTION OF CLASSIFIER FOR STUDENT BEHAVIOR

Prediction of student behavior is made by the classifier on the test data-set after once the model is build in training phase. Prediction is made by the in which, 34 attributes

are given as an input and prediction is made for any one remaining attribute.
The accuracy of two different type is calculated separately. steps involved are

1) Load the data-set

2) Calculating mean and standard deviation: The mean is the central tendency of the data, and it is used as the middle of our normal distribution when calculating the probabilities. The standard deviation is calculated for each attribute for a class value. The standard deviation describes the variation of spread of data which is used to characterize the expected spread of each attribute in our normal distribution when calculating probabilities.

3) Separation of data: In the proposed system, data-set is separated by class values.

4) Summarize the data: The naive bayes model consists the information about the summary of the data in the training data-set. This summary is used to make the predictions when test data-set is given as an input. The summary obtained from the classifier are the mean and standard deviation values for each attribute. These values will be used to calculate the probability of an attribute belonging to a particular class.

5) Making predictions: In this phase, prediction is made based on the summaries obtained from the training data. Predictions are made by calculating the normal probability density function which uses mean and standard deviation for the attribute from the training data.

III. Conclusion

In conclusion,the project 'Student Performance Prediction System' it describes the overall performance of the student ,in academic, sports ,practicals,culture, social as well as extra curricular activities.This project help to select the field in future as well as the student weak in which area how to improve his knowledge in that area.

Acknowledgment

The completion of any interdisciplinary project depends upon co-operation, co-ordination and combined efforts of several sources of knowledge. We are grateful to Prof. J.P.Kshirsagar for her even willingness to give us valuable advice and direction whenever we approached her with a problem. We are thankful to her providing immense guidance for this project. We are also thankful to Dr. Seema Kedar head of the department of computer engineering and all staff member for their immense co-operation and motivation of completing our project

References

[1] Chew Li Sa,Dayang Hanani bt.Abang Ibrahim, Emmy Dahliana Hossain, Mohammad bin Hossain , *"Student Performance Analysis System (SPAS)"*.

[2] Sagardeep Roy,Anchal Garg , *"Analysing Performance of Students by Using Data Mining Techniques"*.

[3] Suhas S Athani,Sharath A Kodli, Mayur N Banavasi, P. G. Sunitha Hiremath, *"Student Academic Performance and Social Behavior Predictor using Data Mining Techniques"*.

[4] V. Shanmugarajeshwari,R. Lawrance , *"Analysis of Students' Performance Evaluation using Classification Techniques)"*.

[5] Ms.Tismy Devasia1 ,Ms.Vinushree T P2, Mr.Vinayak Hegde3, *"Prediction of Students Performance using Educational Data Mining"*.