Teacher's Performance Analyzer

[1]Ms. Bhosale Ankita, [2]Ms. Sonar Sukanya, [3]Ms. Khalate Madhavi, [4]Ms. Bhosale Bhagyashali, [5]Prof. Kumbhar Hemant

[1,2,3,4]Student, [5]Professor, [1,2,3,4,5]Department of Computer Engineering, Shivnagar Vidya Prasarak Mandal's COE, Malegoan(bk), Maharashtra, India.

Abstract : **Any Educational system is built on two important processes, teaching and learning. If the question is asked which process among these two is most important, obvious answer is learning. Learning must happen satisfactorily. But how can we neglect the teaching process? For the betterment of educational systems, a lot research is done focuses on students' performance instead of teacher's performance in educational mining. Teaching is the most important process to complete learning process satisfactorily. To improve teaching process teacher's performance is very important. So in this paper we are focusing on discovering knowledge that predicts teacher's performance, which is one of the ways to achieve high quality education. Higher educational system uses data mining applications most widely for understanding and solving the problems faced by students. The analysis shows that the teacher's success mainly depends on the interest of the students in the course.This paper presents an efficient model for prediction of teachers' performance in higher institutions specifically engineering streams of learning using data mining techniques. Questionnaire is the most common tool to evaluate performance of the teacher. In this paper, classifier model is built by applying learning algorithms, such as C4.5,ANN(Artificial Neural Network),Naive bayes Classifier. In addition, an analysis is done for the variable importance for each classifier model.**

Keywords: Data mining, ANN, C4.5, Naïve Bayes classifier.

I. INTRODUCTION

Recent years there has been a wide demand in evaluating the teacher's performance and their effectiveness in teaching and quality of teaching. The main goal of the educational institutes is to improve the quality of education. Data mining techniques are popularly used in higher education, and because of its use in educational institutes such as better allocating resources, predicting student performance, academic planning and intervention transfer prediction, improving the effectiveness of alumni development, a new field has emerged Educational data mining (EDM) that works on the data related to the education field.[2]. The higher education institutions has potential knowledge such as academic performance of students, administrative accounts, potential knowledge of the faculty, demographic details of the students and many other information in a hidden form. The technique behind the extraction of the hidden knowledge is Knowledge Discovery process. Recently Data mining is widely used on educational dataset. Educational Data mining (EDM) has become a very useful research area [6]. Backbones of any educational system are teachers and the students. Teacher does the job of teaching and student learns it. Teacher is the one of the important person through which student gets knowledge. The student is required to recommend additional activities, teaching materials and tasks that would favour and improve his/her learning. Professors would have the feedback, possibilities to classify students into groups based on their need for guidance and monitoring, to find the most made mistakes, find the effective actions. Hence, its' very necessary to improve the quality of teaching[7].A variety of "qualitative and quantitative"

1

measures have been taken to solve the problem but it keeps it away from achieving the quality[1].To evaluate the teacher's performance most widely used tool is surveying student's responses about the course and its' teacher through a questionnaire. Student's opinions about the courses and teachers are collected and stored in databases waiting to be discovered and used for managerial purposes Results and attendance of the student play a vital role in evaluation .Feedback is taken normally from the students with regular attendance. Additionally, variables related to teacher and course characteristics are also included in the study. The information of the faculty, attendance of the student, questionnaire and the results will be the set of input. Using this information analysis is done based on the three algorithms as c4.5, ANN, Naïve bayes classifier. Comparative study of these algorithms will be done with respect to accuracy, precision and specificity. Many of the researchers prefer to apply the single technique from the mentioned above. These algorithms are used to build a classifier model, which will help in analyzing whether the teacher teaches satisfactorily or unsatisfactorily.

II. METHODOLOGY

2.1 Modules of the System

Select Dataset: Selecting a dataset actually includes giving the dataset as an input to the algorithm for processing.

C4.5 Processing: C4.5 processing includes the processing the given input dataset according to the defined algorithm of C4.5 data mining.

Naïve Bayesian Processing: Processing the dataset according to the training dataset and predict the class labels for the new data.

ANN Processing: Processing the dataset by calculating the error function until the correct result is obtained.

2.2 Attribute Selection Measure

The information gain measure is used to select the tested attribute at each node in the tree. Attribute with highest information gain is chosen as test attribute for the current Node. This attribute minimizes the information needed to classify the samples in resulting partition and reflect the least "impurity" in these partitions.

Let S be set consisting of data sample. Suppose the class label attribute has m Distinct values defining m distinct class C_i (for i =1... m).

Let S_i be the number of Sample of S in class C_i. The expected information needed to classify a given sample is given by equation

$$I(S1, S2, \cdots, Sm) = -\sum^{m}_{i=1} Pi \log 2(Pi).$$

Where Pi is probability that an arbitrary sample belongs to classify Ci and estimated by S_i/S. Note that a log function to base 2 is used since the information in encoded in bits.

2.3 Entropy

Entropy is minimum number of bits of information needed to encode the classification of arbitrary members of S. Lets attribute A have v distinct value a_1,............., a_v. Attribute A can be used to Partition S into v number of subsets, S_1, S_2,........, S_v , where S_j contains those samples in S that have value a_j of A. If A were selected as the test attribute, then these subset would corresponds to the branches grown from the node contains the set S. Let S_{ij} be the number of class C_i, in a subset by S_j. The entropy or expected information based on partitioning into subset by A, is given by equation

$$E(A) = \sum^{v}_{j=1} (S_{1j} + S_{2j} + \cdots + S_{mj} / S) * I(S_{ij} + \cdots + S_{mj}).$$

The first term acts as the weight of the jth subset and is the number of samples in subset and is divided by the total number of sample in S. The smaller the entropy value, the greater is the purity of subset partitions as shown in

$$I(S_1, S_2, \cdots, S_m) = -\sum^{m}_{i=1} Pi \log 2(Pi).$$

Where P_i is the probability that a sample in S_j belongs to class C_i.

2.4 Information Gain

It is simply the expected reduction in entropy caused by partitioning the examples according to the attribute. The following equation gives Information gain, Gain(S, A) of an attribute A and relative collection of examples S,

$$\text{Gain }(A) = I\,(S_1, S_2, \cdot\cdot\cdot, S_m) - E\,(A)$$

Gain (A) is the expected reduction in entropy caused by knowing the Value of attribute A. This algorithm computes the information gain of each attribute in the dataset. With highest information gain is chosen as the test attribute for a given set.

III. NAIVE BAYESIAN ALGORITHM

1. Let D be training tuples.

    X = (x1,x2,….,xn)

    n measurements of attributes (a1,a2,….an).

2. Suppose there are m classes (C1,C2,.., Cm)

    X is a given tuple, classifier predicts

    X Є Class having highest posterior probability

    X Є Ci iff

    P(Ci / X) >P(Cj/X)

    Where, j≠1

    The class Ci for which P(Ci/X) is maximized

    posterior Hypothesis.

$$P(C_i/X) = \frac{P\left(\frac{X}{C_i}\right)P(C_i)}{P(X)}$$

3. As, P(X) is constant for all classes only

    P(X/Ci)*P(Ci) needs to be maximized.

4. $P(X/C_i) = \prod_{k=1}^{n} = P(X_k / C_i)$

    Xk refers the value of Attribute Ak

5. Predicted class label Ci for which

    P(X/Ci)*P(Ci) is maximum.

IV. C4.5 ALGORITHM

1. All the samples in the list belongs' to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.

2. None of the feature provides any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class.

3. Instance of previously- unseen class encountered C4.5 creates a decision node higher up the tree using the expected value of the class.

4. For each attribute a, find the normalize information ratio from splitting on a.

Information Gain:

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

$$Info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j)$$

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

Gain Ratio:

$$SplitInfo_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \times \log_2\left(\frac{|D_j|}{|D|}\right)$$

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)}$$

1. Let a be the attribute with the highest normalize information gain.
2. Create the decision node that split on a.
3. Recursively do this on the sub-list obtained by splitting on and add those nodes as children of node.

II. ANN (ARTIFICIAL NEURAL NETWORK) ALGORITHM

1. Initialize the weight in the network (often randomly)

2. Repeat

    *For each example e in the training set do

i. O = neural-net-output (network, e);

Forward pass

ii. T=teacher output for e.

iii. Calculate error (T-O) at the output unit.

iv. Compute delta_wi for all weights from hidden layer to output layer;

Backward pass

v. Compute delta_wi for all weights from input layer to hidden layer;

Backward pass continue

vi. Update the weight in the network

*end

3. Until all examples classified correctly or stopping criterions satisfy.

4. Return (network).

VI. EXPERIMENTAL RESULTS

As discussed earlier, in this paper, we are using three classification algorithms that are Naïve Bayesian algorithm, C4.5 and ANN. In C4.5 algorithm, Gain ratio is used as an impurity measure. In this, we are considering the following questions BASIC: Q3 and Q6, from MODERATE: Q1 and from ADVANCE: Q2 and the Gain Ratio is calculated from the training data set. The question having the lowest Gain Ratio is considered to be the root of the decision tree, and is the sector in which the teacher has to improve and depending on that suggestions are provided to that particular teacher. We expect result as given below:

Result For C4.5

	BASIC: Q3	BASIC: Q6	MODERATE: Q1	ADVANCE: Q2
Gain Ratio	1	0.505	0.638	0.375

From the above table, it can be seen that the Gain Ratio of ADVANCE: Q2 is the lowest among the four. Hence, it is considered to be the weakness of that particular teacher and hence, suggestions are given to improve the efforts in ADVANCE category.

In Naïve Bayesian classification, the Posterior probability is calculated based on the prior probabilities of the class and the tuples and the conditional probability of X given class. For the questions from BASIC: Q3 and Q6, from MODERATE: Q1 and from ADVANCE: Q2, the approximate results are as follows:

If we consider, four different tuples such as,

Tuple X1= (BASIC: Q3=A, Q6=A, MODEARTE: Q1=B, ADVANCE: Q2=A).

Tuple X2= (BASIC: Q3=C, Q6=B, MODEARTE: Q1=C, ADVANCE: Q2=D).

Tuple X3= (BASIC: Q3=B, Q6=B, MODEARTE: Q1=C, ADVANCE: Q2=D).

Tuple X4= (BASIC: Q3=C, Q6=A, MODEARTE: Q1=B, ADVANCE: Q2=D).

Result for Naïve Bayes:

	Posterior Probability of YES	Posterior Probability of NO	Satisfactory or Not?
X1	0.0858	0.00047	YES
X2	0.00003	0.066	NO
X3	0.0081	0.0143	NO
X4	0.0107	0.0216	NO

So, from the above table, we can conclude that the result that is Satisfactory or not depends on the posterior probability of the class value.

Also Appreciation results may be drawn for teachers doing very satisfactory teaching.

VII. CONCLUSION

Thus, the teacher's performance can be evaluated by using the data mining Classification techniques to improve their performance, which will also improve the performance of the students. The Classification techniques can be applied by taking into consideration attributes such as Questionnaires, Answers, Attendance of students, Results of students, etc. Performance of classification algorithm used in building a model does not necessarily indicate that, the algorithm that uses least time is not always the best model to use. Among the three methods, that are Naïve Bayesian classification, C4.5 and Artificial Neural Network, C4.5 provides best Accuracy. It means that C4.5 Decision tree is the best suitable algorithm for analyzing teacher's performance when compared to other algorithms in this work.

REFERENCES

[1] MUSTAFA AGAOGLU ,"Predicting Teacher Performance Using Data Mining Techniques in Higher Education," VOL. 4, 13 May 2016.

[2] K. Sumathi,S. Kannan,K. Nagarajan "Data Mining: Analysis of student database using Classification Techniques",Vol. 141 – No.8, May 2016.

[3] Surjeet Kumar Yadav, Saurabh Pal, " Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification," Vol. 2, No. 2, 51-56, 2012.

[4] V. Ramesh, P. Parkavi, K.RAMAR., "Predicting Student Performance: A Statistical and Data Mining Approach," Volume 63– No.8, February 2013.

[5] Amirah Mohamed Shahiri, "The Third Information Systems International Conference. A Review on Predicting Student's Performance using Data Mining Techniques," Amirah Mohamed Shahiri et al. / Procedia Computer Science 72 (2015) 414 – 422.

[6] R. Sumitha1, E.S. Vinothkumar, "Prediction of Students Outcome Using Data Mining Techniques," Volume-2, Issue-6,June 2016.

[7] Edin Osmanbegoviæ, Mirza Suljiæ , "DATA MINING APPROACH FOR PREDICTING STUDENT PERFORMANCE," Vol. X, Issue 1, May2012.

[8] MardikyanS., and Badur B, "Analyzing teaching Performance of Teachers Using Data Mining techniques," Vol. 10, No. 2, pp 245 – 257.

[9] B. K. Baradwaj and S. Pal, "Mining educational data to analyze students' performance," vol. 2, 2011.

