STUDENT IMPROVEMENT SYSTEM USING NOVEL ALGORITHM

Vikas Singh Rajput[1], Dr. Rohit Miri[2] , Mr S. R. Tandan[3]

[1]M. TECH (CSE) Scholar, [2]Associate Prof. [3]Associate Prof.

[1]Department of CSE

[1]Dr. C.V. Raman University Bilaspur, India.

ABSTRACT: This paper describes a project which aims at the implementation of an online student attendance management system which can act as an efficient means in maintaining proper correct and updated records. With the help of the various tools and techniques in data base, data mining and web technologies. Currently the amount huge of data stored in educational database these database contain the useful information for predict of students performance. we can implement an attendance management system. Attendance can be accessed anytime and personalized reports can be generated for each student. 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 their parents also provide them an opportunity to pay attention to and improve those who would probably get lower attendance. This system which can predict the attendance of students from their previous performances using concepts of data mining techniques under decision tree based novel algorithm. It tracks all the attendance details of a student from the day one to the end of his course which can be used for all reporting purpose, tracking of attendance, completed semester's years, coming semester year curriculum details.

1. INTRODUCTION

People have always been concerned about attendance management in any Institute. Since attendance plays a very important role in the studying of the Institute. People have used paper and pen to take the attendance. This method is very tedious, time consuming and not interactive. For over decades people have been dreaming about a better way to manage and view attendance, which is easily accessible and usable. The goal of student performance improvement system (SPIS) by Attendance is to bring computer technologies into attendance management. This would make the attendance of student for managing attendance more efficient and easier. SPIS is highly dynamic and easy to access in nature; any person can access it any time from the centralized database system and generate required report.

In this project database is used to keep a record of attendance and also use data mining to generate attendance management report for individual students. The ability of updating attendance live makes the system more robust to check ones attendance.

To provide interaction with a teacher, we can use a database server and web interface to mark the attendance of students for various lectures.

In this paper at the end of month total attendance of a student are calculated and find out those students name who are short listed and then generate automatic message to their parents for take care of their children.

The object under consideration in the project is to identify areas in the Education System which can be compiled together with the help of data mining techniques to determine trends which can be used to improve student performance.

IndexTerms: decision tree, classification, prune.

2. RELATED RESEARCH

Data mining is an exploratory data analysis.

[1] This paper gave us insights into how the process of data mining can be used to gain valuable insights in the educational sector. Predicting the performance of a student is a great concern to the higher education managements.

[2] This paper helped us in judging the criteria which are useful and important to determine their usefulness in using them in our data collection process.

[3] The following paper gave us insights into the various techniques and algorithms which are effective in classifying and clustering our data into useful format for our use.

[4] This paper also helped us in gaining more knowledge and familiarity with the whole process of result prediction from data mining and generating results for our processing.

According to [1] Srishti Taneja et al. designed an approach for Implementation of Novel Algorithm (SPruning Algorithm). She described that how the decision tree work and how it is useful for tree generation. For classification of data she used an approach by which accuracy of data is improved.

According to [2] Dr. Neeraj Bhargava et al. (2013) had given a review of basic concepts of data mining i.e. he told about decision tree, characteristics of tree, approaches used, and also described about pruning concept in tree by which accuracy of data will improve. He designed a novel approach of decision tree for huge amount of data. He discussed that data mining is removal of noisy data and knowledge can be obtained from it.

According to [3] W.Nor Haizan W. Mohamed et al. (2012) discussed about Reduced Error Pruning technique in decision tree algorithms. He discussed that Classification trees are most accepted and well-organized technique used in data mining. Various tree algorithms are used to create decision trees, some of them are complex and some are simple it depends upon size of data. Complex trees are difficult to recognize.

According to [4] Goyal Anshul et al. (2012) compared 2 decision tree algorithms, he presents that classification is used in every field of life. Classification utilized to categorize each entry in a dataset into already defined classes or groups. He carried out a survey to make a performance evaluation of Naïve Bayes and j48 classification algorithm. Naive Bayes algorithm is based on chance and j48 algorithm is based on choice. A tree is generated using J48 algorithm and the tree generated is improved as compared to other algorithms. J48 is a simple classifier method to construct a decision tree, well-organized consequence has been taken from dataset of a bank using weka tool. Naive Bayesian classifier also present superior outcome. The experiments consequences shown are for accuracy of classification for doing analysis of cost. The consequence in the research on datasets also shows that the efficiency and accuracy of j48 and Naive bayes algorithm is superior.

According to [5]   Abeer Badr El Din Ahmed, Ibrahim Sayed Elaraby, discussed about the basic concepts of data mining techniques and algorithm how it works and there results.

According to [6] Prachi s. Bhokare and Dr . Rekha Sharma, had given the basic idea of frequent item how to find and how it works

3. IMPLEMENTED SYSTEM

With help of good user interface and proper database management we can make attendance management easier to view and manage. Our System maintains a centralized database for the attendance of all students. As the data is readily available at any time, reports can be easily generated in an organized form using data mining techniques on the available data.

- **Admin control**: The proposed system is going to be controlled by admin so no one can update/change the level of attendance. The system follows a hierarchal flow of data hence the process seems more organized and easy to use.

- **Notification System**: In the proposed system in case the student's overall attendance drops below a certain percentage an message will be sent to the student reporting his/her current attendance levels in each subject. This helps in managing their attendance better.

- **Digital working:** Since the whole process is managed by the computer there will be a minimal instance of error generation during attendance calculation and report creation.

- **Report Generation:** This system can be used to generate reports as per month wise, semester wise and as well as year wise.

- **Messaging system**: This system can be used to send their attendance result to their parents.

4. DATA MINING TECHNIQUES

Prediction of the outcome of the student in the academic scenario requires the analysis of various different parameters and criteria. The model which is to be created after the analysis requires the inclusion of a variety of variables involving the academic attendance factors. Due to current project constraints we have decided to work only with the academic variables.

Through the extensive research and study of various literatures and discussions with academic professionals a few key factors were identified as the most important in the analysis process. **Table I** lists out the various attributes shortlisted for the study to be conducted.

These factors were categorized as input variables for the data set to be considered for the data mining model. On the other hand the output variables or the rank values represent the possible outcomes of the criteria after the decision tree algorithm (i.e. ID3 or C4.5) has been implemented on it.

For this experiment before the actual calculations could be finalized, data from our college was collected. The data set consisted of students from all branches from all the semesters over the span of annual academic sessions.

Table 1

Sr.no	Attributes for attendance		
	Criteria	**Variables**	**Remark**
1	Yearly Attendance	YA	>75% = good <75% &>60% = poor <60% = very poor
2	Half yearly attendance	HYA	>75% = good <75% &>60% = poor <60% = very poor
3	Monthly attendance	MA	>75% = good <75% &>60% = poor <60% = very poor
4	Average attendance	AA	>75% = good <75% &>60% = poor <60% = very poor

5. Methodology

Many algorithms are already designed for generation of classification tree. Different algorithms have different efficiency or accuracy issue. Accuracy of algorithm differs from application to application; it may be possible that an algorithm is accurate in one but not much accurate in another area. Analysis of various algorithm was done by which a proposal came into mind that a new algorithm design will improve performance. In this it not only find the name of short listed candidate name but also sends automatic message to their parents. For this we use the concept of decision tree. Which take nlogn space as compare to n^2.

Novel approach is based on :

1 Classification of files.

2 Enhanced Pruning in searching method.

3 Tree generation.

4 message generation.

6. Building blocks

1 Tree generation

This is an important step from where our algorithm take data and organize it in hierarchical manner as a tree.

2 Tree pruning

After the first step find out and prune the name of those student who are below the attendance level.

3 Name finding

Find the name of those students which are frequent item set in each month and give them specially notification about their attendance level.

4 Messaging

Send message to their parents who are below the attendance level.

7. Result

Using the below formula to calculate our sample mean:

$$\overline{x} = \frac{\sum x}{N}$$

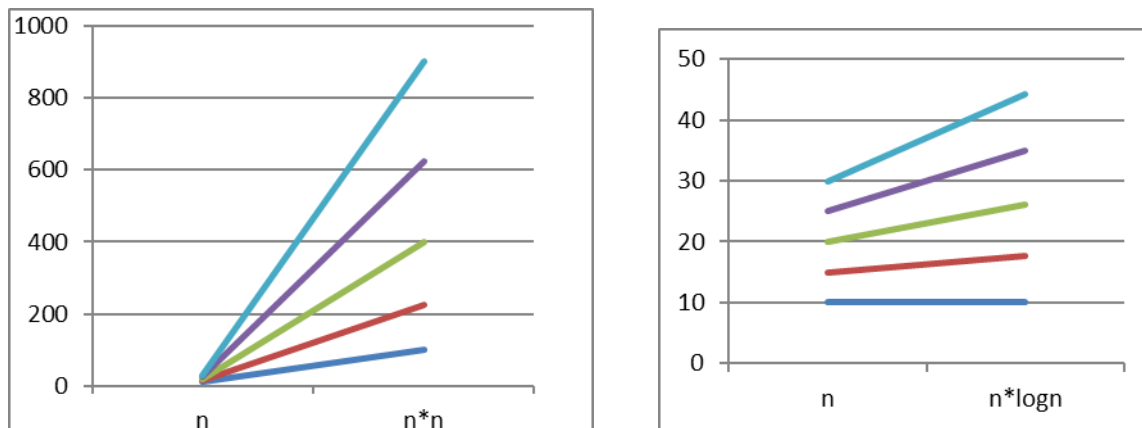| Sr. no.	Value of x	Remark
1	YA>=75% YA>=60% && =<45 YA<=45%	Good Poor Very Poor
2	HYA>=75% HYA>=60% && =<45 HYA<=45%	Good Poor Very Poor
3	MA>=75% MA>=60% && =<45 MA<=45%	Good Poor Very Poor
4	AA>=75% AA>=60% && =<45 AA<=45%	Good Poor Very Poor

8. Comparison of n with n^2 and nlogn

Fig comparison with n, n^2 and nlogn

9. Conclusion and future work:

This system work on manually take attendance each and every day, generate total attendance and then the system sends the message to the short listed candidate. The space and time complexity taken by the system is reduced from n^2 to nlogn.

Now in future it is implemented using biometric device which helps to generate attendance automatically and then generate the attendance report and sends message automatically.

REFERENCES:

[1] Implementation of Novel Algorithm (SPruning Algorithm) by Srishti Taneja.

[2] Dr. Neeraj Bhargava, "Decision Tree Analysis on J48 algorithm for Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering, June 2013.

[3] W. Nor Haizan W. Mohamed , " A Comparative Study of Reduced Error Pruning Method in Decision Tree Algorithms", IEEE International Conference on Control System,Computing and Engineering, pp.23 - 25 Nov. 2012.

[4] Goyal Anshul, "Performance Comparison of Naïve Bayes and J48 Classification Algorithms", International Journal of Applied Engineering Research.

[5] Abeer Badr El Din Ahmed, Ibrahim Sayed Elaraby Data Mining: A prediction for Student's Performance Using Classification Method.

[6] Prachi s. Bhokare and Dr . Rekha Sharma A Novel Algorithm PDA (Parallel And Distributed Apriori) for Frequent Pattern Mining
.

[7] A Novel Algorithm for Mining Rare-Utility Itemsets in a Multi-Database Environment by Guo-Cheng Lana, Tzung-Pei Hongb,c, and Vincent S. Tsenga.

[8] Vikas Singh Rajput, Dr. Rohit Miri, Student improvement system using novel Algorithm.

[9] Du Zhang & Jeffrey J.P. Tsai, Advances in Machine Learning Applications in Computer Engineering(2000)

[10] W. Nor Haizan W. Mohamed , " A Comparative Study of Reduced Error Pruning Method in Decision Tree Algorithms", IEEE International Conference on Control System,Computing and Engineering, pp.23 - 25 Nov. 2012.

[11] Goyal Anshul, "Performance Comparison of Naïve Bayes and J48 Classification Algorithms", International Journal of Applied Engineering Research.

[12] http://ieeexplore.ieee.org/document/4378362/

[13] [A+10] Ayesha S., Mustafa T., Raza Sattar A., Khan M. I. - Data Mining Model for Higher Education Systems. European Journal of Scientific Research, Vol.43, No.1, pp.24-29, 2010.

[14] [A+14] Abdulsalam S. O., Adewole K. S., Akintola A. G., Hambali M. A. – Data Mining in Market Basket Transaction: An Association Rule Mining Approach, International Journal of Applied Information Systems, Vol. 7, No. 10, pp. 15-20, 2014.

[15] [A+15] Abdulsalam S. O., Babatunde A. N., Hambali M. A., Babatunde R. S. - Comparative Analysis of Decision Tree Algorithms for Predicting Undergraduate Students' Performance in Computer Programming. Journal of Advances in Scientific Research & Its Application (JASRA), 2, Pg. 79 – 92, 2015.

[16] [BP11] Bharadwaj B.K., Pal S. - Data Mining A Prediction for Performance Improvement Using Classification. International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.

[17] [FPS96] Fayadd U., Piatesky-Shapiro G, Smyth P. - From Data Mining to Knowledge Discovery in Databases. AAAI Press / The MIT Press, Massachusetts Institute of Technology. ISBN 0–26256097–6, 1996.

[18] [G+07] Galit B. Z., Hershkovitz A., Mintz R., Nachmias R. - Examining Online Learning Processes Based on Log Files Analysis: A Case Study. Research, Reflection and Innovations in Integrating ICT in Education, 2007.

[19] [KBP11] Kumar S., Bharadwaj B., Pal S. - Data Mining Applications: A Comparative Study for Predicting Students Performance. International Journal of Innovative Technology and Creative Engineering (ISSN:2045-711)VOL. 1 No. 12, December, 2011.

[20] [RSN06] Al-Radaideh Q. A., Al-Shawakfa E. W., Al-Najjar M. I. - Mining Student Data Using Decision Trees. International Arab Conference on Information Technology (ACIT'2006), Yarmouk University, Jordan, 2006.