ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Data Mining With Neural Networks to Predict Students Academic Achievements

[1]Richa Shambhulal Agrawal, [2]Mitula H. Pandya
[1]Alpha College of Engineering and Technology, Khatraj, Gujarat, India
[2]Dept. of Computer Engineering, Alpha College of Engg. and Technology, Khatraj, Gujarat, India

Abstract
In this paper we propose to apply data mining techniques to Improve Student's Overall Performance in Exams. The focus is on the educational Data Mining and Classification Techniques along with Attribute Evaluator Techniques. The classification algorithms Naïve Bayes, J48, Random forest, Multi-Layer Perceptron, are grouped with analysis was experimented using WEKA tool. The Evaluators are Chi Squared Attribute Eval, Filtered Attribute Eval, Gain Ratio Attribute Eval, Info Gain Attribute Eval, Relief Attribute Eval, One R Attribute Eval and Symmetrical Uncert Attribute Eval methods are used. Combination of such techniques helps us to predict Student's performance accurately and quickly.

Keywords
Attribute Evaluators, Classifiers, Data Mining, Neural Network, WEKA Tool

I. Introduction
Recent years have shown a growing interest and concern in many countries about problem of college students, their achievements, and the factors related to their education. At present, a lot of research [2] has been covered on understanding the factors that affect the performance of students (ICAI) at different educational levels (ICAI, The Institute of Chartered Accountants of India, Ahmedabad Branch) using the large amount of information that current computers can store in databases.
The same, using varied methods of data mining, has also been done. Several combinations of Evaluators and Algorithms to improve the efficiency and output have been carried out.
Educational data mining is an emerging trend, concerned with developing techniques for exploring, and analyzing the huge data that come from the educational context. EDM is constituted mainly as a liaison between the humongous amount of data from the data mining association and several educational issues in learning and student education.
Recently, Educational data mining and various methods of data mining in this thesis like Classification and evaluation system, are more successful in solving many of the student's problems, as a reason of the high rate of industrial tools like WEKA and Data Mining Algorithms like J48, Naïve Bayes, Multi-Layer Perceptron etc. The aim of educational institutions and other coaching facilities is providing quality education to students. A fool proof method for achieving the quality education is to improve and re-improve the education quality by constantly taking feedbacks from students and predicting their results based on their past and present data.

II. Background and Related Work
Data mining is also called the "Knowledge Discovery in Databases" process.

A. Modules
- Data Gathering
- Pre-Processing
- Data mining
- Interpretation

Neural Networks are statistical modelling tools. They are also non-linear models. Complex relationships can be formed and modelled. Relationships can be modelled between the input and output in order to find patterns in the data. Using Neural Networks, complex relationships can be modelled also. Different data warehousing firms are obtaining information from data sets and utilizing the same to create big data base management systems.

B. WEKA
In the flow of work, WEKA (Waikato Environment for Knowledge Analysis), a popular machine learning software is used. It is written in Java and developed by University of Waikato. The University of Waikato is located in New Zealand. The visualization tools, classification algorithms and attribute evaluators, of the WEKA tool, together, make data analysis easier and also give graphical user interfaces. The tool helps in providing easy access functionality.

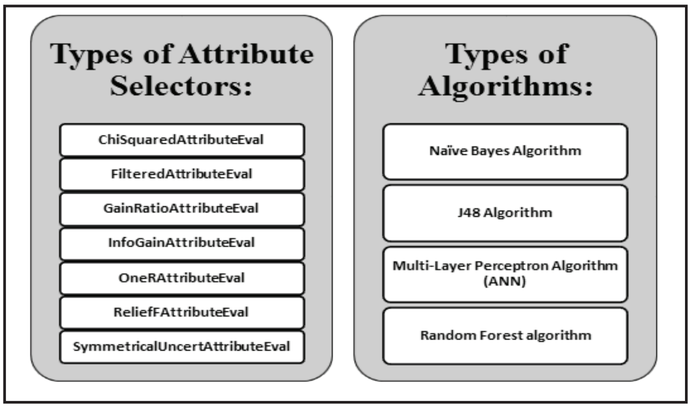

Fig. 1: Combination of Attribute Evaluators and Algorithms Used

C. Data Set
It is written in Java framework. The use of WEKA on almost any platform and the application of algorithms can be direct on a dataset. The data set can be fed in .arff file or as .csv file or fed in java framework coding also.
Here, we have used about 46 attributes and 100 instances for each attribute. We also used the version WEKA 3.6 for our project.

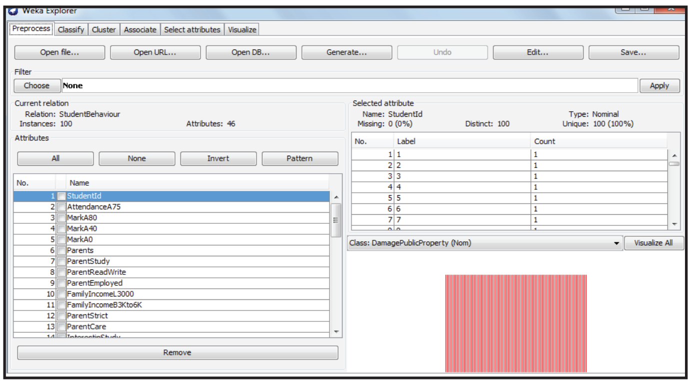

Fig. 2: Data Set Fed in the WEKA Tool

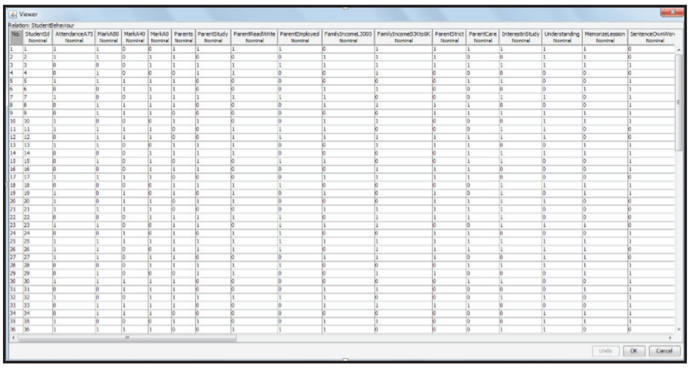

Fig. 3: View of the Data Set in WEKA

III. Methodology

The implementation is solely in the WEKA Tool, version 3.6. The user has defined a step by step process to support the same. The algorithm is as follows:

1. The user selects a dataset, preprocessed as input into the WEKA tool.
2. The Preprocess tab shows the number of attributes, number of instances, name of all the class attributes etc.
3. Following classification algorithms are run on the selected data set:
   - Naïve Bayes Algorithm
   - J48 Algorithm
   - Multi- Layer Perceptron Algorithm
   - Random Forest algorithm
4. The user then goes to the "SELECT ATTRIBUTES" Tab. This tab contains the Attribute Evaluator and the Search Method Types.
5. For each Attribute Evaluator, with ranker's Method as the search method, we run the project for the training set.
6. The Ranker's search method is supported by following Attribute Evaluators:
   - ChiSquaredAttributeEval
   - FilteredAttributeEval
   - GainRatioAttributeEval
   - InfoGainAttributeEval
   - OneRAttributeEval
   - ReliefFAttributeEval
   - SymmetricalUncertAttributeEval
7. The above mentioned Attribute Evaluators Provide a complete count of all attributes by ranks ranging from highest to lowest.
8. The user selects a threshold value of ranks at each Attribute Evaluator.
9. The Attributes lower than the threshold rank, are removed from the Preprocess Tab.
10. The following classification algorithms are run on the removed attributes:
    - Naïve Bayes Algorithm
    - J48 Algorithm
    - Multi-Layer Perceptron Algorithm
    - Random Forest algorithm
11. The results of 10) and 3) are compared and the outputs are analyzed
12. Further according to the requirement, Accuracy, Total Accuracy, Random Accuracy and different Errors can be compared in order to find which combination of Attribute Evaluator and Classifier gives the best accuracy and classifier.

ChiSquaredAttributeEval	Output of Naïve Bayes before use of Evaluator	Output of Naïve Bayes after use of Evaluator
Correctly Classified instances	80	84
Incorrectly Classified Instances	20	16
Kappa Statistic	0.5981	0.6795
Mean Absolute Error	0.2429	0.1947
Relative Absolute Error	49.2169%	38.9799%
Weighted Avg. TP Rate	0.8	0.84
Weighted Avg. FP Rate	0.196	0.161
Weighted Avg. Precision	0.804	0.84
Weighted Avg. Recall	0.8	0.84
F-Measure	0.801	0.84
ROC Area	0.881	0.934

Fig. 4: First Level Output of the Algorithm Discussed

IV. Experimental Results

WEKA contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization.

WEKA 3.6 version is used in the stated work. Mainly, two features of WEKA are used as stated:

Attribute Selection

Classification Algorithms

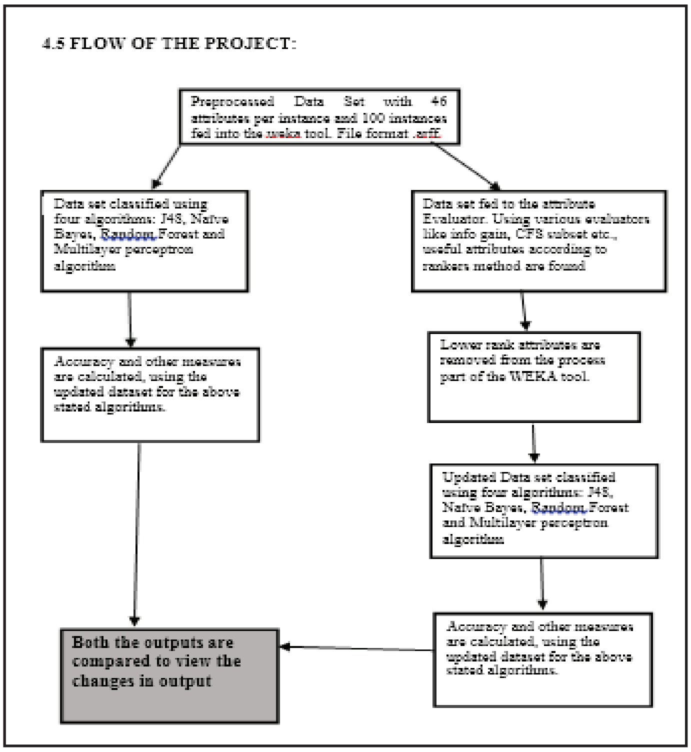

Fig. 5: Flow of the Project

 ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Firstly, Naïve Bayes algorithm in combination of all evaluators and Naïve Bayes Algorithm simply, was conducted. Further, a combination of all the algorithms with all the evaluators was carried out.

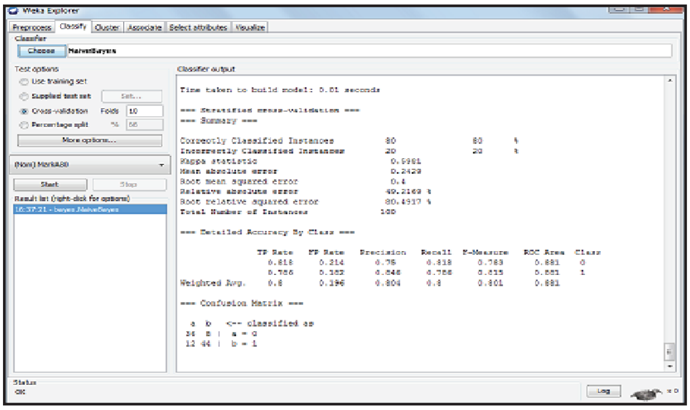

Fig. 6: Simple Naïve Bayes Classifier is Run on the Provided Data Set

After running simple algorithms, various evaluators were run on the data set, and based on the threshold values, the attributes were discarded.

	attributes removed	threshold value
ChisquaredAttributeEval	9,33,14,15,35,34,5,31,16,32,27,29,23,46,24	attributes < 1.0, removed
filtered attribute Eval	40,24,9,33,14,15,35,34,5,31,16,32,27,29,23,46	attributes < 0.01, removed
GainRatioAttributeEval	40,24,9,33,14,15,35,34,5,31,16,32,27,29,23,46	attributes < 0.01, removed
InfoGainAttributeEval	40,24,9,33,14,15,35,34,5,31,16,32,27,29,23,46	attributes < 0.01, removed
OneRAttributeEval	34,16,23,46,35,27,29,31,32,25,15,14,43,21,19,33,22,24,30,38,37,40,9,5,39,1	attributes <57, removed
ReliefAttributeEval	32,25,27,33,21,12,14,43,39,9,42,15,19,31,16,45,46,1,30,22,29,24	attributes < 0.1, removed
SymmetricalUncertAttributeEval	40,24,9,33,14,15,35,34,5,31,16,32,27,29,23,46	attributes <0.01 removed

Fig. 7: Threshold Value Calculations and Attribute Ids Which are Removed

	J48	Naïve Bayes		Random Forest	Multi Layer Perceptron
Correctly Classified instances	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Incorrectly Classified Instances	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Mean Absolute Error	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	Multi Layer Perceptron
Root mean Squared Error	J48	OneRAttributeEval		ReliefAttributeEval	Multi Layer Perceptron
Relative Absolute Error	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ReliefAttributeEval
Weighted Avg. TP Rate	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Weighted Avg. FP Rate	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Weighted Avg. Precision	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Weighted Avg. Recall	OneRAttributeEval/ ReliefAttributeEval	ChisquaredAttributeEval	Naive Bayes	ChisquaredAttributeEval	OneRAttributeEval
F-Measure	OneRAttributeEval/ ReliefAttributeEval	ChisquaredAttributeEval		ChisquaredAttributeEval	OneRAttributeEval
Total Accuracy	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Precision	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Sensitivity	J48	Naive Bayes Algorithm	OneRAttributeEval	ReliefAttributeEval	ChisquaredAttributeEval
Specificity	J48	filtered attribute Eval	SymmetricalUncertAttributeEval	All except SymmetricalUncertAttributeEval	ChisquaredAttributeEval / Multi Layer Perceptron

Fig. 8: Overall Analysis of Fifferent Algorithms and Evaluators With Respect to the Properties

Algorithm	Best Considered Evaluator
J48	J48
Naïve Bayes	filtered attribute Eval
	SymmetricalUncertAttributeEval
Random Forest	ReliefAttributeEval
Multi Layer Perceptron	ChisquaredAttributeEval

Fig. 9: Analysis of Algorithms and the Best fit Evaluator for Each.

After the complete analysis was conducted, graphs for total accuracy for all four algorithms, Naïve Bayes algorithm, J48 Algorithm, Random Forest Algorithm and Multi-Layer Perceptron Algorithm was done.

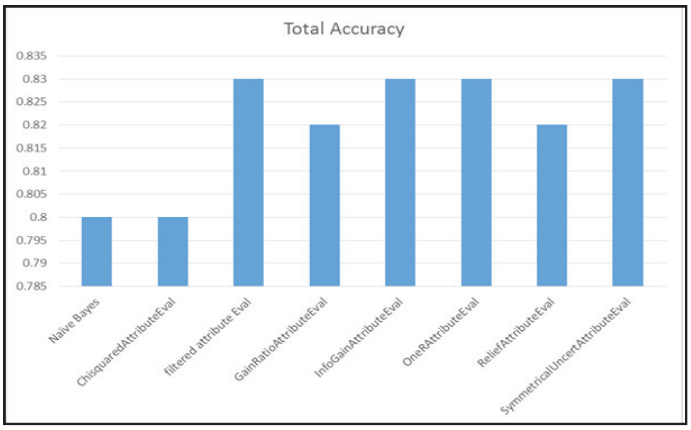

Fig. 10: Total Accuracy Graph of Naïve Bayes Algorithm

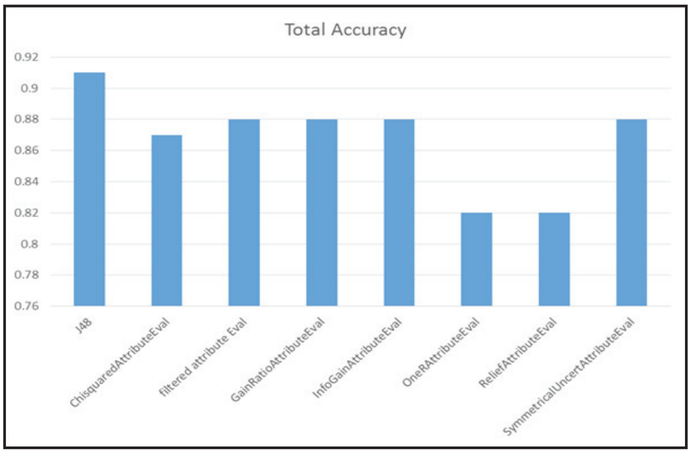

Fig. 11: Total Accuracy Graph of J48 Algorithm

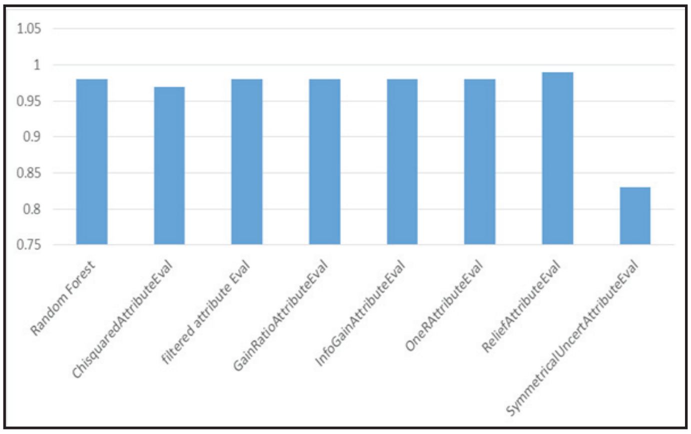

Fig. 12: Total Accuracy Graph of Random Forest Algorithm

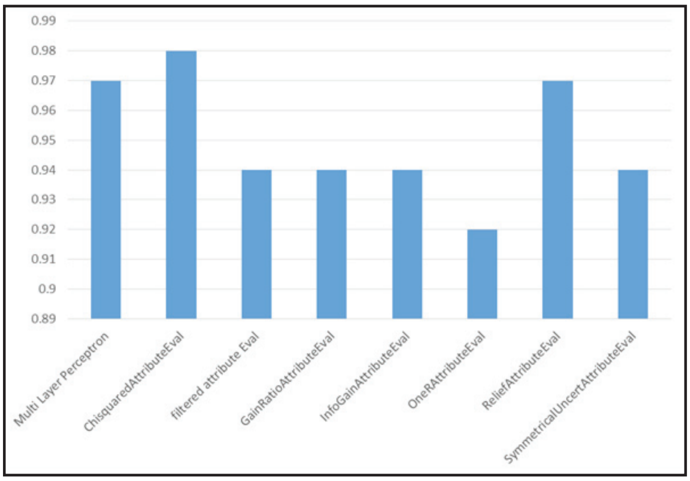

Fig. 13: Total Accuracy Graph of Multi-Layer Perceptron Algorithm

IV. Conclusion and Future Work

In the stated work, a combination of four different algorithms (Naïve Bayes, J48, Random Forest and Multi-Layer Perceptron) and seven different Attribute Evaluators based on the rankers' method have been used. A comparison of output attributes like Accuracy, Mean Absolute Error, Precision, Sensitivity, Specificity, Precision, Recall etc.

It was found that for J48 algorithm, no Attribute Evaluator Method was working as the best output was obtained, apart from the J48 Algorithm itself. For Naïve Bayes Algorithm, FilteredAttributeEval and SymmetricUncertAttributeEval were found to be giving better outputs for the factors stated before. For Random Forest Algorithm the best output was given by ReliefAttributeEval and finally for Multi-Layer Perceptron Algorithm, ChiSquaredAttributeEval worked the best.

An alternative means by which the academic performance of undergraduate students can be predicted; and comparison of this study with the use of machine learning technique based on several metrics shows that, the proposed model generalizes well and performs much better.

We hope to extend this study to encompass the integration of mechanisms for the construction of a more robust prediction model that can interact well with the students' portal for real-time exploration in order to strengthen the institution's information system.

Also, related to the study provided in this document, further more combinations of different algorithms with different other evaluators can be studied upon and other methods of evaluators can also be worked upon.

References

[1] Richa Agrawal, Mitula Pandya,"Survey of Papers for Data Mining with Neural Networks to Predict the Student's Academic Achievements", International Journal of Computer Science Trends and Technology (IJCST) – Vol. 3, Issue 5, Sep-Oct 2015.

[2] R. S. Naik B,"Using Neural Network to Predict MBA Student Success", College Student Journal, 2004.

[3] Alan Cheah Kah Hoe, Mohd Sharifuddin Ahmad, Tan Chin Hooi, Mohana Shanmugam, Saraswathy Shamini Gunasekaran, Zaihisma Che Cob, Ammuthavali Ramasamy College of Information and Technology, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN 43009 Kajang, Selangor, Malaysia. Grant Scheme no.: J510050264. Analyzing Students records to Identify Patterns of Student's performance.

[4] O. M. N. Alfan E,"Undergraduate Students' Performance: a case of University Malaya," Emerald Quality Assurance in Education, Vol. 13, No. 4, 2005.

[5] Alan Cheah Kah Hoe, Mohd Sharifuddin Ahmad, Tan Chin Hooi, Mohana Shanmugam, Saraswathy Shamini Gunasekaran, Zaihisma Che Cob, Ammuthavali Ramasamy College of Information and Technology, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN 43009 Kajang, Selangor, Malaysia. Grant Scheme no.: J510050264. Analyzing Students records to Identify Patterns of Student's performance.

[6] R. D. Ibrahim Z, "Predicting Students' Academic Performance: Comparing Artificial Neural Network, Decision Tree and Linear Regression, 2007," In 21st Annual SAS Malaysia Forum, 5th September 2007, Shangri-La Hotel, Kuala Lumpur, 2007.

[7] I. Z. Rusli N M, Janor R M "Predicting Students' Academic Achievement: Comparison between Logistic Regression, Artificial Neural Network, and Neuro-Fuzzy," In IEEE, 2008.

[8] N. M. M. Kadirgama K, Sani M S M, M.M.Rahman, M M, Rejab M R M, M.Y.Taib M Y, Sulaiman A, Ibrahim A , Bakar R A "Development of Genetic Algorithms and (MPNN) model to Study The Student.

[9] Pauziah Mohd Arsad, Norlida Buniyamin, Faculty of Electrical Engineering, Universiti Teknologi Mara, 40450 Shah Alam Malaysia, IEEE 2013, Prediction of Engineering Students' Academic Performance Using Artificial Neural Network and Linear Regression: A comparison.

[10] Pauziah Mohd Arsad, Norlida Buniyamin, Faculty of Electrical Engineering, Universiti Teknologi Mara, 40450 Shah Alam Malaysia, IEEE 2014. Neural Network and Linear Regression Methods for Prediction of Students' Academic Achievement.

[11] [Online] Available: http://ordnet.princeton.edu/

[12] [Online] Available: http://ieee.org/

[13] STANDUP lexical database API: [Online] Available: http://inf.abdn.ac.uk/research/standup/downloads/resources/lexicon/