

A Comparative Study of Predicting Student's Performance by use of Data Mining Techniques

Aysha Ashraf^{a*}, Sajid Anwer^b, Muhammad Gufran Khan^c

^{a,b}National University of Computer & Emerging Sciences, Department of Computer Sciences, Pakistan

^cNational University of Computer & Emerging Sciences, Department of Electrical Engineering, Pakistan

^aEmail: f159002@nu.edu.pk

^bEmail: sajid.anwer@nu.edu.pk

^cEmail: m.gufran@nu.edu.pk

Abstract

Educational systems need innovative ways to improve quality of education to achieve the best results and decrease the failure rate. Educational Data Mining (EDM) has boomed in the educational systems recently as it enables to analyze and predict student performance so that measures can be taken in advance. Due to lack of prediction accuracy, improper attribute analysis, and insufficient datasets, the educational systems are facing difficulties and challenges exist to effectively benefit from EDM. In order to improve the prediction process, a thorough study of literature and selection of the best prediction technique is very important. The main objective of this paper is to present a comparative study of various recently used data mining techniques, classification algorithms, their impact on datasets as well as the prediction attribute's result in a clear and concise way. The paper also identifies the best attributes that will help in predicting the student performance in an efficient way.

Keywords: Data mining; Predicative models; Classification; Decision tree; Performance prediction.

1. Introduction

Improvement of student performance and enhancement of quality of education is of utmost importance for all educational institutes. To provide quality education to learners, deep analysis of previous records of the learners can play a vital role.

* Corresponding author.

EDM involves analysis and improvement in the predication methods of student performance. Based on prediction results, if the student needs are fulfilled timely then the overall result and performance will increase year by year. For the purpose of performance analysis and prediction, important attributes and previous records of students are gathered. Subsequently, various data mining techniques and classification algorithms are applied to get deeper insights and predictions.

The purpose of EDM is to reduce the failure rate, improve the educational system and analyze the attributes, which are of key importance and consider the student success and performance. Moreover, it enables us to develop useful predictive models for the performance prediction. It does not only help to immediately take steps for betterment of at risk students but also provides information and insights for the next year planning of education process. In recent years, various data mining techniques and classification algorithms have been used such as Naïve Bayes, Decision tree, neural networks, outlier's detections and advanced statistical techniques. These techniques are applied on the student data in order to get information, to help in decision support systems, and pattern extracting etc. Commonly the student's academic performance is measured by previous CGPA but there are various other important attributes that affect the overall performance of the student. Recently various empirical and statistical based researches have been conducted on student's dataset. Kabakchieva and his colleagues [1] used the Bayes and decision classifiers to predict the performance of students on pre university data. Some other proposed techniques in the literature are based on neural networks, statistical methods and ID3 algorithms [2-7].

The problem of accurate student performance prediction is still a challenging task due to various issues and many other factors are involved in it. The main issues in the performance predication methods are inefficiency and use of improper attributes or variables. The objective of this survey paper is to conduct a comparative study and provide the best methods for performance prediction after analyzing the recent studies. It presents a summary of previous literature from 2011 to 2017 that includes decision tree [ID3, c4.5 and CART], Bayesian network classifier, naïve Bayes classifiers, MLP, NB and J48 algorithm, logistic regression, neural network, clustering/classification, association rule and NBtree classification algorithms [7-13]. The remaining paper is organized as follows. Section 2 presents the previous studies and background in the form of literature review. Section 3 presents the basic research methodology. The comparative study of educational data mining techniques and their impact based on student attributes is given in section 4. In section 5, comparative study of techniques and their results based on various algorithms of data mining are presented. Finally, the last section discuss and presents the conclusions and future work.

2. Literature Review

There is a wide research available on the academic performance prediction using data mining techniques and machine learning that have been carried out for the development of new world.

The comparison on different data mining techniques with neural network, Bayesian classifier, and by the help of decision trees is done by Osmanbegović and his colleagues in [2, 3] work. The neural network has solved the problem of classification. It provides various pattern accuracy methods and recognized process as well as

approximate function for prediction was compared with the help of Bayesian classifiers. The work is based on datasets get from Tuzla students and staffs of economics, for the year is 2010 to 2011, for this purpose their dataset has been used. The weka software is used; different algorithms provide different results; average is the best solution to consider the result. For this purpose, MLP, NB and J48 produce good results. This paper is not considering demographics data and distinctive attributes of students. It needs more experiments for valuable and accurate results. Marquez-Vera and his colleagues and Sergi Rovira and his colleagues [4, 5] proposed a method, used real data of six hundred and seventy students of Zacatecas, school located in Mexico and data of Barcelona university. They used the machine learning, white box method of classification, decision tree algorithms and rules of induction. Three hypothesis-based experiments conducted to find failure ratio and the dropout student's rate in school. They used the feature selection method, out of seventy-seven attributes, only fifteen have been considered, which are considered best in education systems. The dimensional modeling and statistical techniques are also implemented in this work. Weka tool is used, and results are not represented in a graphical form for better understanding. The data imbalanced problem is solved efficiently. In order to predict the student's performance based on pre-university and personal characteristics. Kabakchieva and his colleagues in [1] used the datasets to analyze data by data mining algorithms, two rule for learner. In which two main Bayes classifiers, a decision tree classifier, and nearest neighbor classifier have been used. Weka is used for implementations, the data of 10330 students and 20 parameters are considered. Weka classification filters are applied on datasets by an algorithm, in which J48 and JRip are reliable and provides excellent results. The kNN classifier and Bayes classifiers are not efficient. Another method proposed by Ahmed and his colleagues [6] used to predict the performance of students, classifications methods, clustering, artificial intelligence, neural networks, regression, associate rules, generic and the decision tree included, ID3 methods have been used. Some attribute were collected for the prediction of results of students. For these purposes data sets of 1547 record used to predict the performance. Weka is used for decision tree implementation. This method does not check the attributes like attendance, mood effect, and environmental factors. Another proposed model is based on longitudinal data derived from Gwinnett County Public Schools data, those students who entered in class 8 assessment in Math's and science subjects are done by Tamhane and his colleagues [7] work. For missing values, means imputation is used, logistics regression, decision trees and naïve Bayes are the techniques for implementation. They use weka and SPSS modeler and several demographics factors are considered. From all of these techniques logistic regression provides the best results. There was an issue in missing values filling method. It produces the noise in predictions and one classifier is not enough for entire data of students. The grouping should be feasible option for yielding the optimal risk prediction of student performance. The behavioral data and enrollment related factors should also considered. Arora and his colleagues [17] used the neural network approach to predict the marks of students. A function radial basis is used, to map the inputs by Gaussian method. The information related to grades data of students of the years 2010 to 2011 and 2011 to 2012 totally based on the marks obtained in last past years are helpful to predict the present marks. The record consists of more than thousand student's data and it covers two hundred and fifty subjects. The prediction parameters are not sufficient to predict the performance for next years. No proper algorithm and practical implementations is described. To find out the master's student performance, the method helps the students to divide into clusters with their academic performance. Ktona and his colleagues [18] presented data mining techniques. The extraction of classification rules and the clustering have been implemented by use of c4.5 and k-

means algorithm. By this implementation, the students can choose the area of interest in the best subject, as this platform is not too much efficient. Lime survey had been used and the number of participant was very less about 277 in this survey. Due to these flaws, it does not provide successful way to find the performance of the students. Another better approach had been used. The NBtree classification algorithm is used for predicting the student's performance by Christian and his colleagues work [19]. The datasets consist of academic, education, admission data as well as personal data during their studies. Weka toolkit for data mining is used. The paper is used for building a model of classification for checking the performance of students; gender attribute, GPA, credit, and test score as an important attribute are also used. It is more efficient to use the datasets of higher education institutions to get the model for performance of the students. It is better to use artificial intelligence techniques for more efficient results. Grivokostopoulou and his colleagues in [20] used a method to check the student's learning mechanism as well as semantic rules; it helps in guessing the performances of students. For the improvement of education, delivering qualities and the learning activities, semantic rules and ontologies are also used. Various artificial intelligence methods are used for adaptive learning. An efficient approach of decision tree is used to answer the fail or pass percentages in advance for the artificial intelligence course; c4.5 and cart have been used. Weka is used for experiments. The above-described approach does not provide a larger scale evaluation for understanding the performance of a system. The authors work is not analyzing the performance of gender-related rules, as well as the mistakes happened in examination tests. The missing data is not properly manage in this approach. A simple model of linear regression is used to predict the cumulative GPA proposed by Wang and his colleagues [21] work. It is a new method to automatically infer the study and social behaviors, using passive sensing from smartphones. It used the analysis of behaviors as input to sketch the model. The students of undergraduate at Dartmouth College has been utilized as data sets. The behavioral slope and breakpoints are used for the representation of patterns, least absolute shrink and selecting operator act as model of prediction. This is an efficient approach, because no work had been done by the use of passive sensor data and time series analysis from smart-phones as a predictor. The longitude measure of living, style and behaviors of students are used for prediction of the performances. To predict the course based data from Washington State University is available for the course. Carter and his colleagues [22] proposed a normalized programming state model based on empirical study, a formula is calculated. This model is unreachable in novice programming environments. It is explanation of how a learner online social behavior produces an impact on capabilities of prediction of the normalized programming state model. The work of Carter and his colleagues is specified for programming skill and learning; other studies are ignored for prediction of performances. According to Mgala and his colleagues [23] proposed study based on developing a computer-based prediction tool. A model was built based on a large dataset collected from grade-8 students; they participated in Kenya certificate of primary education exam. The record consisted of 2426 students. The mean imputation is used to treat the missing values. For preprocessing steps, filters are used in feature selection technique, machine-learning techniques, and data mining algorithms have been used for building models of prediction. The logistic regression, multilayer perceptron, sequential minimal optimization algorithm, Bayesian network classifiers, naive Bayes classifier, random forest classifier and J48 algorithm are used. The datasets has been taken from the rural environment is not feasible to provide efficient results for urban education systems as the attributes are taken according to rural factors. This method is not providing the accurate results in some cases as in urban-based student analysis of college or university students. Student personal information and family relevant

information of expenditures are considered very helpful in finding the performance in advance as discussed in [8]. The main purpose of this term paper is to predict performance of learners/students, by comparative study of previous work in form of survey. In which data mining algorithms, classifiers, artificial intelligence techniques and statistical based methods are discussed. This paper is summarization of different techniques used for performance prediction. From 2012 to present studies are discussed for student's performance attributes, as well as it describes the best predicting techniques which are not clearly mention in previous surveys. This paper is also focusing on those important attributes, which affects the performance of student's in academic results.

3. Methods

The purpose of this paper is to find out the mostly used attributes which are used for prediction of performance and determine which algorithm and parameters are best to improve the prediction mechanism in educational system.

Educational Data mining: In order to predict the performance of students, brief description of various data mining algorithms is necessary; this section will discuss the basis of data mining algorithms with their impact on various attributes of students. The flow diagram of prediction model is described in Figure 1.

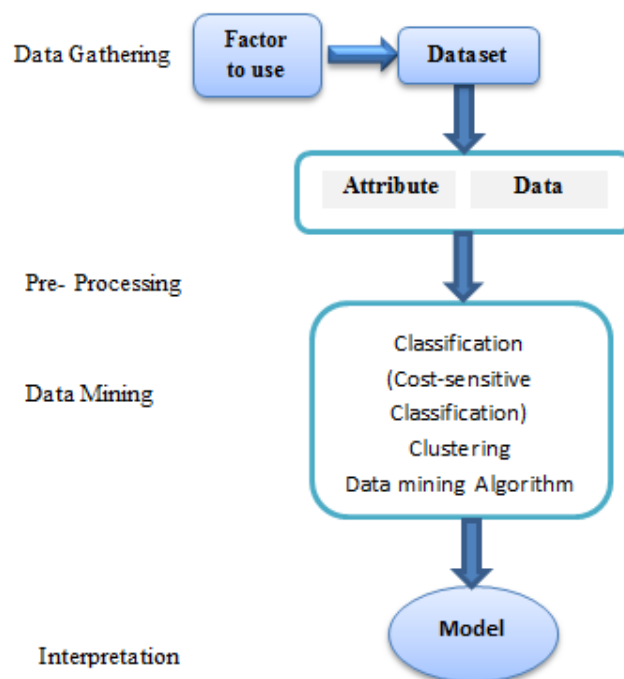


Figure 1: Process for Predication [24]

Evaluation Parameters: These parameters are based on two perspectives, one is based on the dimensions, on which the data mining algorithms are applied for results of prediction. The second is based on how we evaluate which algorithm is better for predicting the performance.

- **Student related variables:** The following evaluation criteria is considered in state of art as described in

below Table 1. In which we use the student demographics information, previous class results, social information, extracurricular activities and other factors are considered.

Table 1: Major evaluation dimensions based on student data

Criteria	Details
Student demographic information	Age, gender, region, residence, guardian info
Previous Results	Cleared certificates, scholarships and Results
Grades	Recent all Assessments results, Quizzes, Final exam, CGPA, Attendance,
Social Network Details	Interaction with social media websites
Extra-Curricular Activities	Games partitions, sports, hobbies
Psychometric Factor	Behavior, absence, Remarks[7, 9-15]

- **Algorithms based parameter:** The following are the algorithms based parameter:
 - a. Accuracy: describes the correctness of value
 - b. Probability Threshold: Presents the True Positive and True Negative rates.
 - c. Execution Time: Time of running the algorithm on dataset.
 - d. Precision: $(\text{number of true positive}) / (\text{number of true positive} + \text{False positives})$
 - e. Recall: $[(\text{number of true positive}) / (\text{number of true positive} / \text{number of false negatives})]$
 - f. Number/Size of Rules
 - g. ROC Area: (Receiver operating characteristics) used alternative to accuracy [12].
 - h. F-measures: $2 * [(precision * recall) / (precision + recall)]$
 - i. Geometric Mean

4. Data mining algorithm and their impact with student's attributes

The following described algorithms and classifications techniques are analyzed for predictions:

A. *Decision Tree:* Decision tree is a method, which is used to predict the performance of students in literature work. It is simple flow chart structure, it consists of internal nodes and leaf nodes, used to put the values and check the values/attributes. Its classifier consists of two steps; one is preparing phase. Other is prune phase. The pruning phase helps to reduce the data and fit it in the decision tree. For decision tree, three algorithms are mostly used that are ID3, C4.5 and ADT [25].

B. *Iterative Dichotomies 3 (ID3):* The algorithm consists of two phases, build and pruning phase. Hunts algorithm is the base of the ID3. In order to split the attributes and dimensions, it uses the information gain. ID3 does not mostly provide accurate results, when data is not properly preprocessed. The noise is mostly present in data. This method uses categorical attributes for building a tree.

C. *C4.5:* For building, a decision tree mostly used the continuous, categorical attributes and various

dimensions. It splits the values on some threshold percentage or value. In tree structure, children are created on above and below threshold values. To remove the extra branches, it uses pessimistic type prune. The accuracy can be efficiently improve by using this method.

D. *Alternative Decision Tree (ADT)*: It consists of various predictions and decision tree nodes. It is different from ID3 and C4.5. The tree is not differentiated into parent and child nodes. Treat the whole tree for interpretation. By using Join nodes sets, it is evaluated [6, 24].

Table 2: Prediction parameter results with Decision Tree

Methods	Attributes	Accurate Results (%)	Papers
Decision Tree	Final Exam	85	Bunkar, and his colleagues (2012) [16]
	CGPA, demographic, Background info, and Scholarship information's	91	Jishan, and his colleagues (2015) [17]
	(Internal) exam, CGPA, and Extra activities	66	Mayilvaganan and his colleagues (2014) [18]
	Exam (internal), CGPA, and extra activities	65	Ramesh and his colleagues (2013) [19]
	Exam (internal), demographic, and extra activities	90	Elakia and his colleagues (2014) [20]
	Psychological factors, extra activities, knowledge Skills	88	Mishra and his colleagues (2014) [21]

E. *Naive Bayes algorithm (NB)*: It is simple algorithm for the prediction of performance based on classification method by using the probability theory; this method makes the problem very simple by use of assumptions. Like no other attributes affects the prediction process any more. It is considered to be an efficient algorithm [14, 15].

Table 3: Accuracy based results with Naive Bayes algorithm

Methods	Attributes	Accurate	Papers
		Results (%)	
Naive Bayes Algorithm	Background and demographic info	50	Ramesh and his colleagues (2013) [19]
	CGPA	75	Jishan and his colleagues (2015) [17]
	Exam results, CGPA, and Extra activities	73	Mayilvaganan and his colleagues (2014)[18]

F. *Multilayer Perceptron Algorithm (MLP)*: The Neural network uses this algorithm mostly. An input layer is formed by some elements, which contains the sensory information, for purpose of approximation in classification function. An advance multilayer perceptron algorithm is used. It consists of neurons, which are known as nodes. It is used to approximate the nonlinear functions.

Table 4: Prediction parameter results with neural network

Methods	Attributes	Accurate	Papers
		Results (%)	
Neural Network	Exam data (external)	97	Arsad and his colleagues (2013) [22]
	Background and demographic info	72	Ramesh and his colleagues (2013) [19]
	Psychometric	69	Gray and his colleagues (2014) [23]
	Exam results	75	Jishan and his colleagues (2015) [17]

G. *Logistic Regression*: By use of the logic function, we can also build a model and estimate the results. It is basically a statistical method used for prediction and the estimations [14].

H. *K-Nearest Neighbor (KNN)*: It is considered one of the best simple and accuracy based algorithm for performance-based prediction. Support Vector Machine (SVM) is also used for classification purpose [14].

Table 5: Prediction parameter results with K-Nearest and SVM

Methods	Attributes	Accurate Results (%)	Papers
Support Vector MachineK-Nearest Algorithm (SVM)	Exam(internal), CGPA, Extra activities	83	Mayilvaganan, and his colleagues (2014) [18]
	Psychometric attributes	69	Gray and his colleagues (2014) [13]
	CGPA, exam (internal) , Extra activities	80	Mayilvaganan, and his colleagues (2014) [18]

According to prediction attributes and classification techniques, we have concluded some useful results as in Figure 2 we have seen that major accurate results are analyzed from the neural network algorithms based on final exam and CGPA, which is 97% accurate.

The best results are obtained from decision tree algorithm by use of the CGPA, demographic, background and scholarship related information, which is 91% and 90% based on exam (internal), demographic, and the extra activities. On the other side psychological factors, extra activities, and knowledge skills are also considered as the important factors. The prediction of the results with decision tree algorithm is 88% accurate. The naive Bayes provides 75% accurate results based on CGPA. K-nearest neighbour provides 83% results, because of these factors; Exam (internal), CGPA, Extra activities. The support vector machine provide 80% results accurate, on basis of CGPA, exam (internal) and Extra activities. We have concluded from this survey that education system should use Exam (Internal, External), CGPA, background information and extra activities for prediction of the performance to get best results. The neural network and decision tree algorithms are to be preferred. To predict the performance in advance will help the education system to improve the quality and results as well as reduce the failure rates. The Figure 2 shows the graphical representation of accuracy, which is analyzed on multiple attribute for prediction of performance. This section is about details of the prediction results of the work done in recent studies for determining the performance of students by using data mining

algorithms and techniques as described in Table 6 as well as in Figure 2 and 4.2. All the major attributes, which are used for prediction, are mentioned clearly. We have achieved our goal of prediction of the best algorithms based on accuracy results of previous studies. The highest prediction accuracy is seen in Table 6 work [3] by use of ADTree algorithm, the classification results provides 97.3% accurate which will work efficiently in future for prediction purpose. This prediction is based on all attributes. It includes internal and external Exam information, the background information, extra-curricular activities, social interaction, and psychological behavior of student as well as CGPA etc. If we considered the most important attributes then the classification results are improved to 99% for ADTree algorithm. The 2nd highest accuracy is obtained on a datasets of 77 attributes and 670 students is used, in which the best attributes results are shown in Table 6 is obtained by ADTree is 97.2% in [19].

Ridor provides 95.4% accuracy and JRip, NNge provide 96% to 97% accuracy. The 3rd highest accuracy is achieved in [9] by comparison of classifiers by use of twenty two features. In which the best accuracy is achieved by Logistic algorithm is 83.88% and Bayes Net, SMO and Lazy algorithms also provides nearly 80% of accurate results of classifiers. If we reduce the features to eight, the results are vary and 83% accuracy has been seen in many algorithm results, if we use the most influential features for comparison then the accuracy is reduce in mostly algorithms to 70%. It does not provide efficient results.

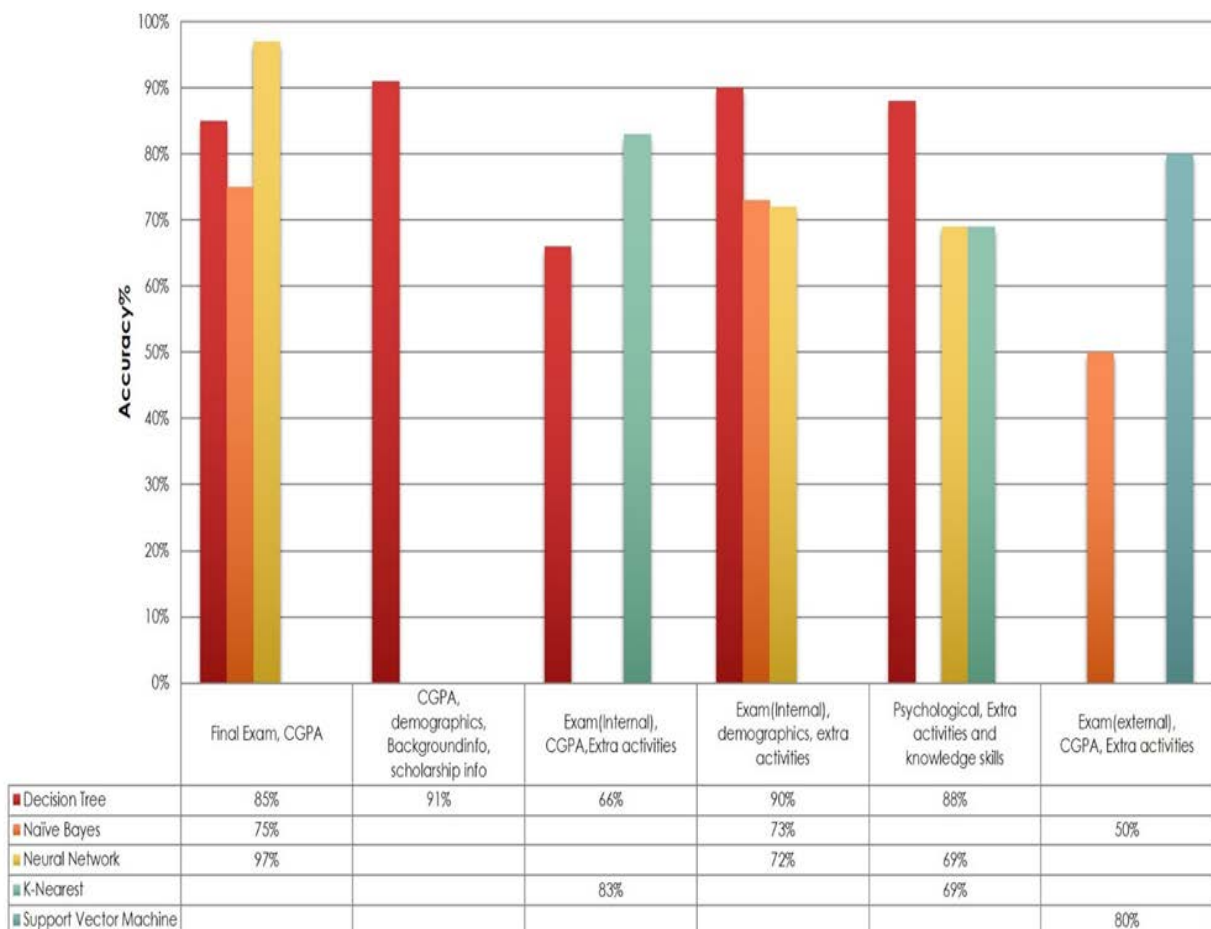


Figure 2: Prediction Accuracy with Algorithms

Table 6: Comparisons of Evaluation Parameters and their impact on Techniques

Ref	Techniques	Accuracy %	TP Rate	TN Rate	Execution Time	Precision/Recall	Number/Size of Rules
[2] 2012	Classifiers						
	NB	77%	0,500 0,851	0,149 0,500		0,500 0,851	197
	MLP	71%	0,371 0,821	0,397 0,804	Null	0,371 0,821	183
	J48	74%	0,290 0,882	0,118 0,710		0,290 0,882	190
[24] 2012	Classifiers						
	ID3 algorithm	72%			0.12	68.2% 06.4%	432 students
	C4.5	74%	Null	Null	0.08	70.4% 09.6%	
ADT	72%			0.06	82.8% 11.4%		
[9] 2013	Algorithms						
	ID3	75.145%	Null	Null	47.6	Null	182 students
	C4.5	75.145%			39.1		
[4] 2013	Classification Algorithms						
	JRip	94.8	97.7 98.7	65.0 78.3	Null	Null	Null
	NNge	96.9	88.8	88.3			
	OneR	88.8	99.8	37.1			
	Prism	94.7	97.9	70			
	Ridor	95.4	98.2	86.7			
	ADTree	97.2	96.7	75			
	J48	94.8	96.1	68.3			
	RandomTree	93.6	96.5	75			
	REPTree	94.6	96.4	76.7			
	SimpleCart	94.6					
[13] 2013	(Based on Best attributes)						
	JRip		97.7	78.3	Null	Null	8.0
	NNge	95.7	98.5	73.3			31.0
	OneR	96.1	98.9	41.7			2.0
	Prism	93.7	99.5	25.0			76.0
	Ridor	93.7	96.6	65.0			4.0
	ADTree	93.1	99.7	76.7			21.0
	J48	97.3	97.4	53.3			31.0
	RandomTree	93.9	95.7	48.3			212.0
	REPTree	94.9	98.0	56.7			44.0
	SimpleCart	94.5	97.7	65			5.0
	ICRM v1	95.1	94.3	90.0			2.0
	ICRM v2	92.1	97.5	75.0			7.6
	ICRM v3	94.9	84.4	93.3			4.0
	76.7						
[7] 2014	Classifiers						
	Naïve Bayes	Null	90.5	23.8	Null	Null	Null
	Decision Tree		90.0	24.7			
	Decision Table		90.7	28.8			
Logistic	90		20				

	Regression						
[14] 2015	Algorithms(22 features)	83.47					
	Lazy	79.76					
	Multi. perception	83.88					
	Logistic	82.60	Null	Null	Null	Null	Null
	Tree(J48)	81.32					
	Random Test	83.26					
	Bayes Net	73.99					
	Naïve Bayes	83.38					
	SMO						
[25] 2016	Rule Induction Algorithms						
	JRip	80					
	OneR	83.3					
	Ridor	79					
	Decision Tree Algorithms		Null	Null	Null	Null	Null
	ADTree	83.4					
	J48	82.3					
	Random Tree	75.7					
	REPTree	77.9					
	SimpleCart	80					

An overall 79.2% accuracy is seen in paper [25], which is achieved by use of the student performance in university, social behavior, personal data, background information, educational background and the current status of student so the results are gained by rule induction and decision tree algorithms with original data is 79.2% by use of JRip algorithm and standard deviation is 4.354. The accuracy is achieved 80% by the SMOTE on original dataset. In work of [7] predict the students risks by use of Naïve Bayes, decision tree and logistic regression the ROC area. The maximum gained is 1.0 by logistic regression and minimum by Naïves bayes, which is 0.5. For prediction of student performance used of [2] NB, MLP and J48 algorithms, the best results had been obtained from MLP algorithm. They provided 74% accurate prediction results. This method takes 4.13 sec to build a model. The work of [24] used the ID3, C4.5 and ADT algorithms. It also provides a survey based research work results and best accuracy is obtained from C4.5, which is 74.416 %, and precision for this algorithm is 70.4.

Another [9] approach in which comparison of two algorithms is presented by use of ID3 and C4.5 algorithms. The both algorithms give the accuracy of 75.145% but the execution time varies from 39.1 to 47.6 mile-seconds. They get from 173 to 130 are the corrected prediction about the performance of students. After the deep study of literature the best result are presented in graphical representation by using Figure 3.

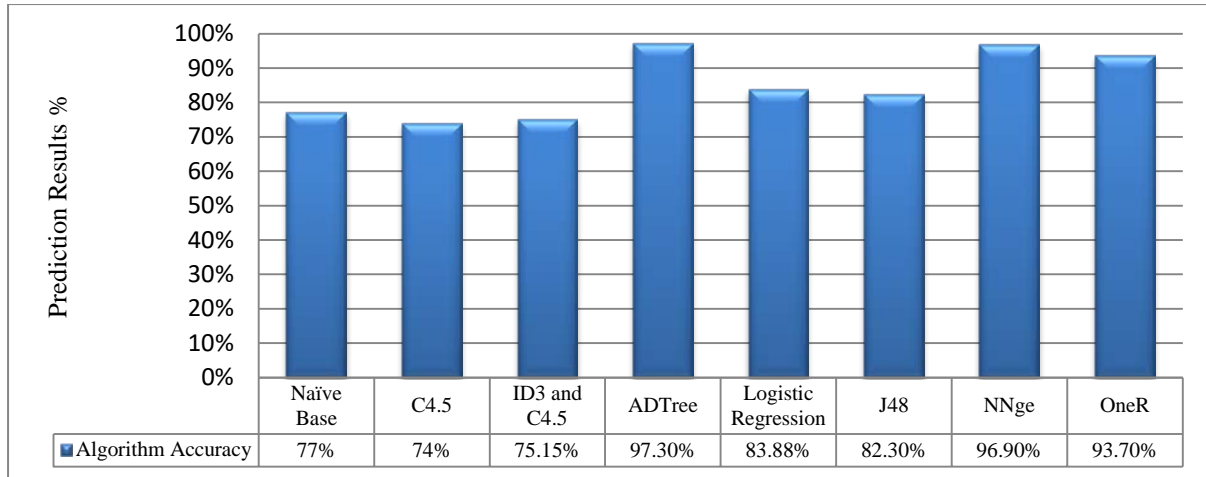


Figure 3: Algorithm accuracy based graphical representation 2012-2017

5. Conclusion and Future Work

The prediction of student’s performance in advance is very important issue. We concluded after deep studies that various datasets of student provides different results with different attributes. This is the reason that the results are vary with different evaluation measures like accuracy, precision and geometric mean. We concluded after these studies that every data mining approach and algorithm results are varied according to the dataset and variable attribute used for prediction. However, if we use the decision tree algorithms, ADTree, JRip, Ridor, logistic regression and neural network approach, according to our requirements these algorithms provide extra ordinary accurate results for future prediction and help in the betterment of education system. In this way, we can improve the prediction methods and performance of education system. This research will implement in future by use of real datasets of Fast University and take the student’s attributes. In order to determine the effects of best predictive algorithm (Decision Tree/Neural Network) and other techniques will evaluate by statistical and empirical studies. In this way, we can compare the results of students with previous semester results. The best techniques are selected based on accuracy. We will find more efficient techniques based on other execution measure like recall and other in future.

References

- [1]. Kabakchieva, D., Predicting student performance by using data mining methods for classification. Cybernetics and information technologies, 2013. 13(1): p. 61-72.
- [2]. Osmanbegović, E. and M. Suljić, Data mining approach for predicting student performance. Economic Review, 2012. 10(1).
- [3]. Käser, T., N.R. Hallinen, and D.L. Schwartz. Modeling exploration strategies to predict student performance within a learning environment and beyond. in Proceedings of the Seventh International Learning Analytics & Knowledge Conference. 2017. ACM.
- [4]. Marquez-Vera, C., C.R. Morales, and S.V. Soto, Predicting school failure and dropout by using data mining techniques. Tecnologías del Aprendizaje, IEEE Revista Iberoamericana de, 2013. 8(1): p. 7-14.
- [5]. Rovira, S., E. Puertas, and L. Igual, Data-driven system to predict academic grades and dropout. PloS

- one, 2017. 12(2): p. e0171207.
- [6]. Ahmed, A.B.E.D. and I.S. Elaraby, Data Mining: A prediction for Student's Performance Using Classification Method. *World Journal of Computer Application and Technology*, 2014. 2(2): p. 43-47.
- [7]. Tamhane, A., et al. Predicting student risks through longitudinal analysis. in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2014. ACM.
- [8]. Daud, A., et al. Predicting Student Performance using Advanced Learning Analytics. in *Proceedings of the 26th International Conference on World Wide Web Companion*. 2017. International World Wide Web Conferences Steering Committee.
- [9]. Adhatrao, K., et al., Predicting Students' Performance using ID3 and C4. 5 Classification Algorithms. arXiv preprint arXiv:1310.2071, 2013.
- [10]. Aher, S.B. and L. Lobo. Applicability of data mining algorithms for recommendation system in e-learning. in *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*. 2012. ACM.
- [11]. Ikbali, S., et al., On early prediction of risks in academic performance for students. *IBM Journal of Research and Development*, 2015. 59(6): p. 5: 1-5: 14.
- [12]. Ktana, A., D. Khaja, and I. Ninka. Extracting Relationships between Students' Academic Performance and Their Area of Interest Using Data Mining Techniques. in *Computational Intelligence, Communication Systems and Networks (CICSyN), 2014 Sixth International Conference on*. 2014. IEEE.
- [13]. Márquez-Vera, C., et al., Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data. *Applied intelligence*, 2013. 38(3): p. 315-330.
- [14]. Mgala, M. and A. Mbogho. Data-driven intervention-level prediction modeling for academic performance. in *ICTD*. 2015.
- [15]. Shahiri, A.M. and W. Husain, A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 2015. 72: p. 414-422.
- [16]. Bunkar, K., et al. Data mining: Prediction for performance improvement of graduate students using classification. in *Wireless and Optical Communications Networks (WOCN), 2012 Ninth International Conference on*. 2012. IEEE.
- [17]. Jishan, S.T., et al., Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. *Decision Analytics*, 2015. 2(1): p. 1.
- [18]. Mayilvaganan, M. and D. Kalpanadevi. Comparison of classification techniques for predicting the performance of students academic environment. in *Communication and Network Technologies (ICCNT), 2014 International Conference on*. 2014. IEEE.
- [19]. Ramesh, V., P. Parkavi, and K. Ramar, Predicting student performance: a statistical and data mining approach. *International journal of computer applications*, 2013. 63(8).
- [20]. Elakia, G. and N.J. Aarthi, Application of data mining in educational database for predicting behavioural patterns of the students. Elakia et al./(*IJCSIT*) *International Journal of Computer Science and Information Technologies*, 2014. 5(3): p. 4649-4652.
- [21]. Mishra, T., D. Kumar, and S. Gupta. Mining Students' Data for Prediction Performance. in *Advanced*

- Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on. 2014. IEEE.
- [22]. Arsad, P.M., N. Buniyamin, and J.-I.A. Manan. A neural network students' performance prediction model (NNSPPM). in *Smart Instrumentation, Measurement and Applications (ICSIMA)*, 2013 IEEE International Conference on. 2013. IEEE.
- [23]. Gray, G., C. McGuinness, and P. Owende. An application of classification models to predict learner progression in tertiary education. in *Advance Computing Conference (IACC)*, 2014 IEEE International. 2014. IEEE.
- [24]. Yadav, S.K., B. Bharadwaj, and S. Pal, Mining Education data to predict student's retention: a comparative study. *arXiv preprint arXiv:1203.2987*, 2012.
- [25]. Meedeche, P., N. Iam-On, and T. Boongoen, Prediction of Student Dropout Using Personal Profile and Data Mining Approach, in *Intelligent and Evolutionary Systems*. 2016, Springer. p. 143-155.