

Student's Academic Performance Forecasting: Survey

Mr Prakash Varma¹, Mr L. P. Bhaiya²

^{1,2}Computer Science Engineering, Bharti College of Engineering and Technology, Durg

Abstract: Data mining (DM) gaining reputation due to its benefits within the academic environment. Most of the instructional group, now a day applied these techniques to make enhancement of their training device. By using these techniques, the tutorial overall performance of the student is analyzed and if locate something wrong with the scholar overall performance then well-timed help will be supplied to that scholar. In our schooling device, we lack in discovering those factor which commonly impacts the pupil overall performance in academics. Therefore, a scientific overview of all the author's works performed in this subject is required to apprehend the facts mining software in schooling and the way it helps to improve and predict the pupil tutorial performance. In this article, the most important attention moves around two essential factors: Firstly, to discover the maximum indispensable factors that have an impact at the pupil overall performance used thru the most researcher and secondly to discover the set of rules that are usually used.

Index Terms: Academic Performance, Educational Data Mining, Prediction, Classification.

1. DATA MINING INTRODUCTION IN EDUCATION

Education is a necessary detail for the development of any nation. So, desirable interest for its implementation must be considered. Every institution collected loads of information from the scholar within the shape of a session take a look at, assignment marks, inner and outside laboratory assessment, private attributes. But all those information used to make the ultimate end result of the scholar whether or not skip or fail the specific examination. If scholar facts are analyzed timely then genuinely it will assist the scholar to enhance their overall performance in academics. Now a day's instructional gadget is no longer restricted to any formal educating inner a have a look at room however it is going beyond that like on-line MOOC course, Web-based schooling device (WBS), a sensible tutorial device (ITS), Project-based getting to know (PBL), webinar, Seminar, Workshops and lots of more. All these coaching techniques make the training machine more attractive and grant masses of knowledge and gaining knowledge of through a distinct source. If all those strategies do not remarkable test and evaluated then it adversely influences the education gadget. So for making any training device to have a success, then the perfect comparison test ought to be maintained. All these instructing methodologies stated above generate a number of facts related to the scholar development in academics, students marks and pupil attendance and if that fact is no longer proper analyzed then all the amassed records is going to be wasted and no similar use of that records happens. So, nowadays every and every instructional group applies an assessment or assessment requirements to study the general overall performance of the college students. In this contemporary technology of schooling, there are masses of opinions or evaluation equipment are to be had, which might be beneficial to have a look at the academic overall performance of the students. So at current DM is one among the first-rate assessment or evaluation gear, that is used to investigate the scholar facts. The span of DM techniques is not only limited to the training system, but it's far going to covers almost entire companies the place information are generated. The critical application areas of DM are retail, banking, telecommunication, marketing, tourism, production, hospitality, clinic tool and many others. These complete listed groups are taking the blessings of enforcing DM to increase their performance in future by using analyzing historical information. The DM techniques are used to research the historical records of any corporation with the aid of the use of an in-built algorithm, which similarly finds the hidden facts from that reality which is not possible to discover manually. By periodically evaluating the pupil progress, timely assist and assist are supplied to these college students who're at the danger of failure or dropout, with a view to enhance their overall performance in destiny. It moreover facilitates the management and administrative personnel of the organization to take suitable action for making enhancement in the machine.

By taking all those factors into mind, an element literature survey is performed. In this survey report, the primary centre of interest is to discover the factors that have an effect at the pupil overall performance, greater frequently than now not used DM set of rules, Accuracy of the used DM set of rules.

2. THE METHODOLOGY USED FOR REVIEW

Now a day, many young researchers working in the field of prediction educational overall performance of college students, predicting dropout student, predicting pupil campus placement and so forth. So for making an effective literature evaluation, a scientific paintings plan is a must. So for that, we divide our research into unique parts. The first part of this segment is related to building research questions system for the evaluate system. The second part of this section is related to Search Strategy used

for seeking some correct research article/papers. The third a part of this section is related to some of the studies papers studied on this systematic overview process and the last part of this segment is related to the detailed class of EDM Research.

A. Research Questions method:

To start any take a look at, it must be higher to set all the targets for your examine. All the examine paintings is performed effectively and competed on time if accompanied by a work plan. So for higher implementation of this literature assessment, we proposed some goals, which further restriction our evaluation manner to a few particular points. Below is a list of proposed objectives:

- Identify students attributes influencing the performance in academics.
- Identify distinctive DM techniques for predicting the pupil's overall performance in academics.
- Identify existing gaps within the current studies paintings

As the targets for the assessment procedure is formulated, so for the higher understanding of this examine, the Main purpose at the back of this literature look at is to remember the suitability of the considered goals in hand. scholar are worked on these subjects and located some desirable result. Below is the list of all the research topic of the EDM with their numbers of studies, which are considered on these review papers:

B. Search Strategy used for looking for some accurate research article/papers:

As the studies in EDM extended for the remaining ten years, so to discover the best research papers isn't an easy task. A lot of world-magnificence database of research papers are to be had today. So to locate a few true research papers, choice of properly research magazine or database are the maximum important task. For that purpose, we finalize some studies journal or research papers database to locate some right studies papers, technical notes, technical article, assessment papers, conference intending and so forth. Below is a table-1 wherein searched database, seek items, booklet period and search techniques are mentioned:

Table-1: Literature examine contains the following terms for search on EDM paintings

Terms used for search	Possible Values used
Databases	IEEE Xplore, Springer Link, Science Direct, Reserachgate, Google Scholar, ACM Digital Library
Search Articles	Journal, Workshops, Abstract, Technical, Book Chapters, Conference
Search applied on	Full text, Chapters, Article
Publication period	2012, 2013, 2014, 2015, 2016, 2017, 2018
Search Strategies	EDM, Student academic prediction, Student performance analysis, Student performance prediction in academic, DM application in Education

Table-1 suggests that, we follow some well known while finding some accurate research papers. For locating those papers we're searching on IEEE Xplore, Springer database, Science Direct, ACM, Reserachgate and Google scholar. In papers categories, magazine articles, workshops papers, conference intending, Technical article are searched. Here, we don't forget research papers from 2012-17 for our evaluate method.

In this, we considered most papers (46 studies papers, assessment papers, technical articles and many others.) of the year 2012-2017, which is further called the ultra-modern years of the ebook of the studies article. All those papers are taken from world-elegance research database with true citation, conference intending papers, research articles, technical papers, IEEE papers, evaluate papers and so on for consideration. In Table 2, we attempted to introduce best those research objectives/questions in which the young researcher are operating since last few years for the betterment of the training gadget. In our overview, we discovered that "Student Performance prediction in Education and Predicting vital attributes associated with student overall performance is the hottest topic of studies. Around 60% of the research

Table 2: Detailed classification of EDM Research

Classification of EDM Research work	Number of studies
Student Performance prediction in Education	28
Predicting Institutional Placement	1
Predicting student Final result on the basis of the previous result	4
Predicting critical attributes related to student performance	18
Finding an association between different attributes of the student	3
Important to check the Effectiveness of the curriculum	2
Clustering similar attributes student in a class	2

In this specific section, our predominant centred become to develop a good studies question for our overview procedure. So simplest 3 primary goals are formed that is very vital for similarly research factor of view. Secondly, we follow a good seek strategy used for searching properly research article/papers. Thirdly, we're choosing 38 odd studies papers in this systematic review. In the end, we divided all of the research areas into distinctive subtopics. Below is the graphical illustration of statistics represented in desk 2 for better understanding.

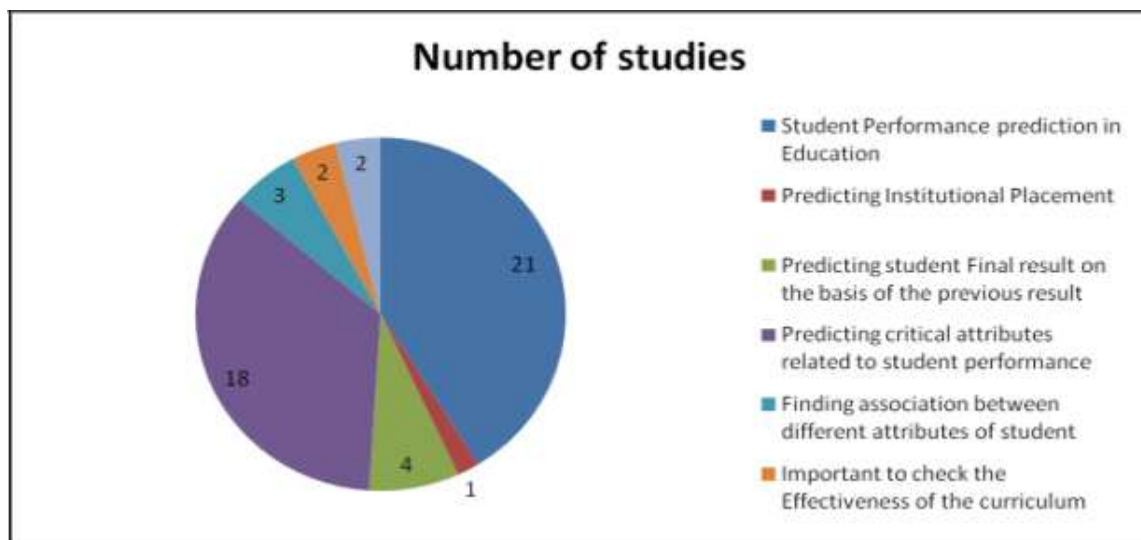


Figure 1: Graphical representation of numbers of studies inside the region of EDM

3. IMPORTANT FACTORS FOR PREDICTING STUDENT'S PERFORMANCE

Use either SI (From conclusion of phase 2, we discovered the predicting pupil performance in educational is maximum researched articles. This section will intensively speak the tremendous component in predicting scholar's instructional performance in training. Mainly two elements affect the scholar's overall performance in academics, which are pupil attributes and prediction strategies. Table 4 suggests the whole not unusual student attributes and Table 5 suggests the prediction techniques used by distinctive research student of their research papers. Section 3.1 will be targeted on the sizable student attributes utilized by the one-of-a-kind researcher for predicting pupil academic overall performance and section 3.2 will be centred on prediction strategies used in predicting students educational overall performance.

A. Important elements in predicting the overall performance of pupil's in academics

This section of literature is particularly to discover those elements which affect the pupil's educational overall performance prediction end result. Student attributes are specifically categorized into a exclusive group like students educational attributes, own family attributes, social attributes, non-public attributes, institutional attributes etc. In this manner, we nearly assessment

forty seven odd studies papers. The student attributes which maximum of the researcher are often used are cumulative grade point average (CGPA) and evaluation marks within the academics. In his observe, Mukesh Kumar and Prof. A. J Singh located that educational attributes and own family attributes are the deciding elements of pupil educational performance. They are not considered private attributes of the scholar into consideration because they sense that at the time of facts collection pupil won't fill their correct records which in addition impacts the prediction result[37]. In every other have a look at authors located that its miles possible to are expecting the commencement performance inside the fourth 12 months at university the usage of simplest pre-college marks and marks in the firth and second 12 months course. They were now not considered socio-financial or demographic functions of the students. In pre-university marks, they have taken into consideration higher or senior secondary marks of the students [17]. In any other have a look at, authors located that even as predicting the scholar performance in academics, attendance marks and cumulative grade factor average (CGPA) are the maximum influencing elements. They considered those factors are critical due to the fact those marks are deciding thing in there higher education. They similarly given extra stress on CGPA because it will end result whether or not a pupil is an academic dropout or the pupil will complete their education [20]. M. Kumar, S. Shambhu & P. Aggarwal, in there, have a look at discovered that, while predicting student overall performance attendance (ATND), the inner grade of the student (INT_GRD) is the deciding elements. They similarly gave stress on the medium of instruction, faculty location, faculty sorts are also the deciding aspect for the overall performance prediction [28]. Edin Osmanbegovi and Mirza Suljic in their take a look at discovered that university entrance exam (URK), GPA (PO), family (BCD), Material (MAT) are the attributes which affect the pupil instructional performance the most [6]. In [23] authors determined that the midterm-1 grade and midterm-2 is the robust predictor for the final failing grade. They discovered 91% accuracy model to predict the failing students in advance before their final exam.

Table 3 list all of the attributes which might be taken into consideration via the specific researcher for his or her prediction. These entire attributes are in addition divided into five distinctive classes like instructional, private, own family, social and faculty-related attributes [23]. They are the usage of simplest the girl scholar's dataset of one-yr duration, however the extra accuracy of the end result they further suggested to include male pupil's information also.

Table 3: Critical factors for predicting the scholar's overall performance in academicsStudent

Student Attributes	Possible Values Used in all research papers while implementing DM algorithm
Academic Attributes	Internal and External Assessment, Lab Marks, Sessional Marks, Attendance, CGPA, semester marks, grade, seminar performance, assignment, attendance, schools marks etc.
Personal Attributes	Age, Gender, height, weight, Student interest, Level of motivation, sports person, hobbies etc.
Family Attributes	Qualification, Occupation, Income, Status, Support, Siblings, Responsibilities etc.
Social Attributes	A number of friends, Social networking, Girls/boys friends, Movies, Travel Outing, friends parties etc.
School Attributes	Teaching Medium, Accommodation, Infrastructure, Water & Toilet facilities, Transportation system, Class Size, School Reputation, School Status, Class Size, School Type, Teaching Methodology etc.

B. Algorithm predicting academic performance of student's

In EDM, predicting pupil instructional overall performance usually, we use predictive modelling. To build a predictive modelling, lots of responsibilities are taken into consideration together with type, clustering, regression, association rule mining etc. But in our look at, we determined that classification algorithms are the maximum popular algorithm used for prediction.

C. Decision Tree algorithm used for prediction:

A selection tree is regularly implemented classification set of rules used for prediction. It is used decision trees, in which nodes represent a feature/characteristic, branches represent a decision and leaf of the tree represents an outcome. This classification algorithm belongs to the types of supervised studying algorithms. Here, whilst building a tree, the maximum influencing thing of the student's dataset is at the basis stage. There are extraordinary methods to find the foundation mode like Gini Index, Information Gain. As shown in Table 5, about 20 studies papers are used decision tree set of rules for prediction.

In [27] authors used academic and demographic students factors for making academic prediction and observed 98.86% accuracy. In [29] authors used CGPA, Attendance, twelfth marks, Engineering Cut-off, Education Medium, Board Type are influencing factors for prediction and found most accuracy of as much as 97.27%. In [41] authors (2013) used inner evaluation

records as the primary predictor and found the accuracy up to 56.25% which isn't taken into consideration as a good prediction model [41].

Table 5: Result accuracy the usage of the Decision Tree algorithm

Student Attributes used for Prediction	Author's Name	Year	Accuracy
midterm-I examination result	Mashaal A. et. al [23]	2015	70.88%
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al [17]	2015	73.08%
finance level, motivation level, gender and grades	Kolo David Kolo et.al [24]	2015	66.80%
gender, race, hometown, GPA, family income,	Fadhilah Ahmad et. al. [22]	2015	68.8%
Grade Point Average(GPA), sem-1, sem-2 grades	C. McGuinness et. al [19]	2014	65.00%
academic and demographic information	Mrinal Pandey et. al.[27]	2016	98.86%
Internal Grade Internet access and Attendance	Mukesh Kumar et. al.[28]	2016	69.34%
10th, 12th, graduation marks, parent's qualification	Mukesh Kumar et. al.[37]	2017	61.40%
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al.[33]	2017	74.78%
Int. assessment CGPA, Extra-curricular	M. Mayilvaganan et. al [47]	2014	66.00%
Activities Internal, Demographic, Extra-curricular	N. J. Aarthi et. al. [13]	2014	90.00%
External, CGPA Demographic, Extra-curricular	Natek, Zwilling et. al. [16]	2014	90.00%
Psychometric, Extra-curricular activities, soft skills	T. Mishra et. al. [15]	2014	88.00%
Demographic High attribute, school background	P. Parkavi et. al. [12]	2013	65.00%
External assessment of student	U. K. Singh et. al. [2]	2012	85.00%
Entrance GPA, examination, Materials, Family	E. Osmanbegović et. al. [6]	2012	73.93%
CGPA, Attendance, 12th marks, Eng	E.S.V. Kumar et. al.[39]	2016	97.27%

D. Neural Network algorithm used for prediction:

Neural Network algorithms are regularly used gaining knowledge of a set of rules used in Machine Learning. It is some other set of rules which is carried out in EDM prediction. It is used to detect all the viable interactions between predictors attribute. D. M. S. Anupama Kumar (2012) used scholars characteristic like Internal and External evaluation as the most influencing elements for making a prediction in self-regulated gaining knowledge of in higher schooling with most accuracy 98.00% [1]. P. M. Arsad, N. Buniyamin, J.-I. A. Manan (2013) used External evaluation of the student for scholars performance prediction with 97.00% accuracy [13]. In [33] authors (2015) used pre-university, 1st and 2nd-12 months guides marks as predominant attributes for predicting academic performance at diploma level with accuracy up to 62.50% [17]. As shown in Table 6, approximately 8 studies papers are used Neural Network algorithm for prediction.

Table 6: Result accuracy using Neural Network algorithm

Student Attributes used for Prediction	Author's Name	Year	Accuracy
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al [17]	2015	62.50%
Grade Point Average (GPA), sem-1, sem-2 grades	C. McGuinness et. al [19]	2014	69.00%
CGPA, Quiz, Midterm, Lab. and Atted, Final Grade	S T Jishan et. al. [20]	2015	75.00%
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al.[33]	2017	70.43%
Demographic attribute, High school background	P. Parkavi et. al. [12]	2013	72.00%
External assessment	N. Buniyamin et. al. [13]	2013	97.00%
Internal assessments, External assessment	Anupama Kumar et. al. [1]	2012	98.00%
Entrance examination, GPA, Materials, Family	E. Osmanbegović et. al. [6]	2012	71.20%

E. Naive Bayes algorithm used for prediction:

Naive Bayes algorithm is likewise a simple however powerful algorithm for making a prediction. It is easy to construct and useful for very huge pupil datasets. This algorithm would then be capable of classifying an unlabelled example based totally at the statistics discovered from the labeled examples. This set of rules is the second often used device learning algorithm for constructing predictive modeling. As shown in Table 7, approximately sixteen studies papers are used Naive Bayes set of rules for prediction. Mrinal Pandey and S. Taruna (2016) used academic and demographic attributes for a pupil's performance prediction and got 91.57% accuracy[27]. V. Ramesh, P. Parkavi, K. Ramar (2013) used Demographic and High faculty

background attributes for predicting scholars performance with maximum accuracy 50.00%, which is not considered as a great predictive model and need a few developments.

Table 7: Result accuracy using Naive Bayes algorithm

Student Attributes used for Prediction	Author's Name	Year	Accuracy
midterm-I examination result	Mashaal A. et. al. [23]	2015	81.01%
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al [17]	2015	83.65%
gender, race, hometown, GPA, family income, university mode entry, and SPM grades, subjects	Fadhilah Ahmad al.[22] et.	2015	67.0%
Grade Point Average (GPA), sem-1, sem-2 grades	C. McGuinness et. al [19]	2014	76.51%
CGPA, Quiz, Midterm, Lab. and Atted, Final grade	S T Jishan et. al. [20]	2015	75.00%
Qualification, Courses, class, Date_Of_Work	A.M.El-Halees et.al. [21]	2015	77.46%
Academic and demographic information	MrinalPandey et. al.[27]	2016	91.57%
Internal Grade, Internet access and Attendance	Mukesh Kumar et. al.[28]	2016	69.34%
10th, 12th, graduation marks, parent's qualification	Mukesh Kumar et. al.[37]	2017	61.40%
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al.[33]	2017	75.65%
Int. assessment, CGPA, Extra-curricular activities	M. Mayilvaganan et. al [47]	2014	73.00%
Demographic attribute, High school background	P. Parkavi et. al. [12]	2013	50.00%
Entrance examination, GPA, Materials, Family	E. Osmanbegović et. al. [6]	2012	76.65%
CGPA, Attendance, 12th marks, Eng. Cut-off, Education Medium, Board	E.S. Vinoth Kumar et. al.[39]	2016	85.92%
Sex, race, home, income, university entry mode, and GPA	AA Aziz, N H Ismail et. al. [40]	2014	63.30%
dCentrality, nMessages, nReplies, nWords, dPrestige, aEvaluation	M.I. López, J.M Luna et. al.[42]	2012	89.40%

F. K-Nearest Neighbor algorithm used for prediction

KNN algorithm used for class and regression prediction. It is normally used for its ease of interpretation and occasional calculation time. M. Mayilvaganan, D. Kalpanadevi (2014) used Internal, Extra-curricular overall performance and CGPA for prediction the performance with most accuracy 83.00% [47]. In [19] authors (2014) used Grade Point Average (GPA), sem-1, sem-2 grades for predicting learner development in tertiary training with accuracy as much as 69.00%. As proven in Table 8, approximately four studies papers are used K-Nearest Neighbor algorithm for prediction.

Table 8: Result accuracy using K-Nearest Neighbor algorithm

Student Attributes used for Prediction	Author's Name	Year	Accuracy
pre-university, 1st and 2nd-year courses marks	Raheela Asif, et. al [17]	2015	74.04%
Grade Point Average (GPA), sem-1, sem-2 grades	C. McGuinness et. [19]	2014	69.00%
Qualification, Courses, class, Date_Of_Work	A.M. El-Halees et.al. [21]	2015	77.46%
Int. assessment, CGPA, Extra-curricular activities	M. Mayilvaganan et. [47]	2014	83.00%

G. Support Vector Machine algorithm used for prediction:

In device studying, SVM is supervised gaining knowledge of that analyzes the dataset used for classification and regression analysis. There are few works that have used SVM as strategies to expect pupils educational overall performance. S.A. Oloruntoba, J. L. Akinode (2016) used the Academic end result, semester-1 and semester-2 grades for student instructional overall performance prediction the usage of guide vector gadget with the accuracy of 98.00% [43]. Mrs. M. Jamuna, Mrs. S. A. Shoba (2017) used students University Admission Score, Number of Failures on the first-year college exam attributes for Students Performance Prediction [44]. K. B. Eashwar, R. Venkatesan (2017) used a psychological, social-behavior, circle of relatives-circumstances and parental-care parameters for Students Performance Prediction [46]. As proven in Table 9, about 6 studies papers are used Support Vector Machine set of rules for prediction.

Table 9: Result accuracy using a Support Vector Machine algorithm

Student Attributes used for Prediction	Author's Name	Year	Accuracy
Grade Point Average (GPA), sem-1, sem-2 grades	C. McGuinne ss et. al [19]	2014	81.22%
Int. assessment, CGPA, Extra-curricular activities	M. Mayilvaga nan et. al [47]	2014	80.00%
Academic result, sem-1, sem-2 grades	S.A. Olorunto ba et. al. [43]	2016	98.00%
students' University Admission Score, Number of Failures at	Mrs M. Jamuna et. al. [44]	2017	97.62%
the first-year university examination academic details and conducting online test scores	A Kadamban de et. al. [45]	2017	90.00%
psychological, social-behaviour, family-circumstances and parental-care parameters	K. B. Eashwar et. al. [46]	2017	96.70%

4. RESULT ANALYSIS

Here, we're trying to talk about the distinctive classification algorithms with their maximum and minimum accuracy. From phase 3.2, we discuss DT, NN, NB, K-NN and SVM in detail. In Table 5, we examine twenty papers and found that the decision tree algorithm has the most accuracy is around 98.86% and minimum accuracy is 56.25%. In Table 6, we analyze eight papers and discovered that the Neural Network (NN) algorithm have maximum accuracy is around 98.00% and minimal accuracy is 62.50%. In Table 7, we examine 16 papers and observed that the Naive Bayes (NB) set of rules have maximum accuracy is around 91.57% and minimum accuracy is 50.00%. In Table 8, we examine 4 papers and found that the K-Nearest Neighbor (KNN) algorithm have maximum accuracy is around 83.00% and minimum accuracy is 69.00%. In Table 9, we analyze six papers and discovered that the Support Vector Machine (SVM) set of rules have the most accuracy is around 98.00% and minimal accuracy is 80.00%. In Table 10, display the detail of various set of rules with their accuracy.

Table 10: DM Classification algorithms with their maximum and minimum accuracy

DM Classification Algorithm	Maximum Accuracy	Minimum Accuracy
Decision Tree Algorithm	98.86%	56.25%
Neural Network Algorithm	98.00%	62.50%
Naive Bayes Algorithm	91.57%	50.00%
K- Nearest Neighbor algorithm	83.00%	69.00%
Support Vector Machine Algorithm	98.00%	80.00%

Fig. 2 shows the student's performance prediction accuracy that uses classification algorithm from 2012 to 2017. Decision Tree algorithm has a maximum accuracy (98.86%) followed by Neural Network (98.00%), Support Vector Machine (98.00%), Naive Bayes (91.57%) and K-Nearest Neighbor algorithm (83.00%). In the end, the algorithm that has lower prediction accuracy is the Naive Bayes algorithm by (50.00%).

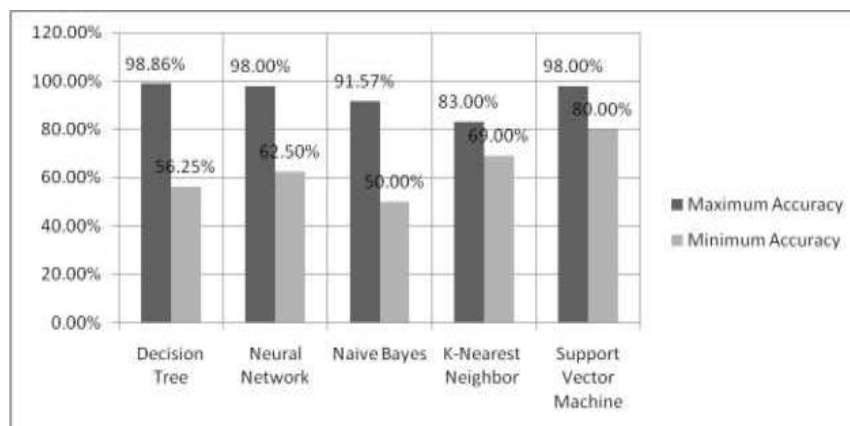


Fig 2: Graphical representation of Classification algorithms with their max and min accuracy

There are lots of software tool available in the market like WEKA, RapidMiner, MATLAB, KNIME, Rattle GUI, Orange, Apache Mahout, R, ML-Flex, Natural Language Toolkit etc. But in our analysis, WEKA is the frequently used software tool by the researcher for prediction. RapidMiner is the second most frequently used software tool. These tools are freeware and easily available for use. These software tools are supporting a lot of classification, clustering and association rule mining algorithm for use on different types of datasets.

5. CONCLUSION

EDM is considered one of the most important areas of studies today. Analyzing almost 46 papers, we determined that, predicting student's overall performance in academic is a beneficial topic of research which facilitates educator, academician, coverage makers and management for improving the teaching and mastering manner during the world. The scholars attribute taken by one-of-a-kind researchers are classed as instructional, psychological, social behavior, own family attributes, and private attributes. In most of the reviewed papers, researchers used CGPA and inner and external assessment marks for their prediction. While studying prediction algorithms we observed that Classification, Clustering, Linear regression, Association Rule mining are used for prediction. But in a maximum of the cases, Classification algorithms are often used for making students educational prediction. Under the Classification algorithms, DT, NN, NB, K-NN and SVM are frequently used by many researchers. In the end, we enthused to perform our studies on pupil's educational prediction in our instructional gadget which facilitates students, educators, control and policymaker to improve the academic gadget through continues tracking of college students.

REFERENCES

1. D. M. S. Anupama Kumar, "Appraising the significance of self-regulated learning in higher education using neural networks", International Journal of Engineering Research and Development Volume 1 (Issue 1) (2012) 09–15.
2. K. Bunkar, U. K. Singh, B. Pandya, R. Bunkar, Data mining: Prediction for performance improvement of graduate students using classification, in Wireless and Optical Communications Networks (WOCN), 2012 Ninth International Conference on, IEEE, 2012, pp. 1–5.
3. S. Parack, Z. Zahid, F. Merchant, "Application of data mining in educational databases for predicting academic trends and patterns", in Technology Enhanced Education (ICTEE), 2012 IEEE International Conference on, IEEE, 2012, pp. 1–4
4. Anuwatvisit S, Tungkasthan A, Premchaiswadi W., "Bottleneck mining and Petri net simulation in education situations. In: ICT and Knowledge Engineering (ICT & Knowledge Engineering), 10th International Conference. IEEE; Bangkok, Thailand; 2012, 244–251.
5. Mohammed M. Abu Tair, Alaa M. El-Halees, Mining Educational Data to Improve Students' Performance: A Case Study, International Journal of Information and Communication Technology Research, ISSN 2223-4985, Volume 2 No. 2, February 2012.
6. Edin Osmanbegović and Mirza Suljic, Data Mining Approach For Predicting Student Performance, Economic Review – Journal of Economics and Business, Vol. X, Issue 1, May 2012.
7. M. of Education Malaysia, National higher education strategic plan (2015). URL <http://www.moe.gov.my/v/pelan-pembangunan-pendidikan-malaysia-2013-2025>
8. U. bin Mat, N. Buniyamin, P. M. Arsad, R. Kassim, An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention, in: Engineering Education (ICEED), 2013 IEEE 5th Conference on, IEEE, 2013, pp. 126–130.
9. D. M. D. Angeline, Association rule generation for student performance analysis using apriori algorithm, The SIJ Transactions on Computer Science Engineering & its Applications (CSEA) 1 (1) (2013) p12–16.
10. K. F. Li, D. Rusk, F. Song, Predicting student academic performance, in Complex, Intelligent, and Software Intensive Systems (CISIS), 2013 Seventh International Conference on, IEEE, 2013, pp. 27–33.
11. C. Romero, M.-I. López, J.-M. Luna, S. Ventura, Predicting students' final performance from participation in on-line discussion forums, Computers & Education 68 (2013) 458–472.

12. V. Ramesh, P. Parkavi, K. Ramar, Predicting student performance: a statistical and data mining approach, *International Journal of Computer Applications* 63 (8) (2013) 35–39.
13. P. M. Arsad, N. Buniyamin, J.-l. A. Manan, A neural network students' performance prediction model (nnsppm), in *Smart Instrumentation, Measurement and Applications (ICSIMA)*, 2013 IEEE International Conference on, IEEE, 2013, pp. 1–5.
14. G. Elakia, N. J. Aarthi, Application of data mining in the educational database for predicting behavioural patterns of the students, Elakia et al,/(IJCSIT) *International Journal of Computer Science and Information Technologies* 5 (3) (2014) 4649–4652.
15. T. Mishra, D. Kumar, S. Gupta, Mining students' data for prediction performance, in *Proceedings of the 2014 Fourth International Conference on Advanced Computing & Communication Technologies, ACCT '14*, IEEE Computer Society, Washington, DC, USA, 2014, pp. 255–262. doi:10.1109/ACCT.2014.105. URL <http://dx.doi.org/10.1109/ACCT.2014.105>
16. S. Natek, M. Zwilling, Student data mining solution–knowledge management system related to higher education institutions, *Expert systems with applications* 41 (14) (2014) 6400–6407.
17. Raheela Asif, Agathe Merceron, Mahmood K. Pathan, Predicting Student Academic Performance at Degree Level: A Case Study, *I.J. Intelligent Systems and Applications*, 2015, 01, 49-61 Published Online December 2014 in MECS (<http://www.mecs-press.org/>) DOI: 10.5815/ijisa.2015.01.05.
18. A. Bogar'in, C. Romero, R. Cerezo, M. S'anchez-Santillan Clustering for improving educational process mining, in *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, ACM*, 2014, pp. 11–15.
19. G. Gray, C. McGuinness, P. Owende, An application of classification models to predict learner progression in tertiary education, in *Advance Computing Conference (IACC)*, 2014 IEEE International, IEEE, 2014, pp. 549–554.
20. S. T. Jishan, R. I. Rashu, N. Haque, R. M. Rahman, Improving accuracy of student's final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique, *Decision Analytics* 2 (1) (2015) 1–25.
21. Randa Kh. Hemaïd and Alaa M. El-Halees, Improving Teacher Performance using Data Mining, *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 2, February 2015.
22. Fadhilah Ahmad, Nur Hafieza Ismail and Azwa Abdul Aziz, The Prediction of Students' Academic Performance Using Classification Data Mining Techniques, *Applied Mathematical Sciences*, Vol. 9, 2015, no. 129, 6415 - 6426 HIKARI Ltd, www.m-hikari.com <http://dx.doi.org/10.12988/ams.2015.53289>.
23. Mashael A. Al-Barrak and Mona S. Al-Razgan, predicting students' performance through classification: a case study, *journal of Theoretical and Applied Information Technology* 20th May 2015. Vol.75. No.2.
24. Kolo David Kolo, Solomon A. Adepoju, John Kolo Alhassan, A Decision Tree Approach for Predicting Students Academic Performance, *I.J. Education and Management Engineering*, 2015, 5, 12-19 Published Online October 2015 in MECS (<http://www.mecs-press.net>) DOI: 10.5815/ijeme.2015.05.02.
25. Mohammed I. Al-Twijri and Amin Y. Noaman, A New Data Mining Model Adopted for Higher Institutions, *Procedia Computer Science* 65 (2015) 836 – 844, doi: 10.1016/j.procs.2015.09.037.
26. Mihai Dascalu and Elvira Popescu et. al., Predicting Academic Performance Based on Students' Blog and Microblog Posts, Springer International Publishing Switzerland 2016 K. Verbert et al. (Eds.): EC-TEL 2016, LNCS 9891, pp. 370–376, 2016. DOI: 10.1007/978-3-319-45153-4_29.
27. Mrinal Pandey and S. Taruna, Towards the integration of multiple classifiers pertaining to the Student's performance prediction, <http://dx.doi.org/10.1016/j.pisc.2016.04.076> 2213-0209/© 2016 Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).
28. Mukesh Kumar, Shankar Shambhu & Punam Aggarwal, "Recognition of Slow Learners Using Classification Data Mining Techniques" *Imperial Journal of Interdisciplinary Research (IJIR)* Vol-2, Issue-12, 2016 ISSN: 2454-1362, <http://www.onlinejournal.in>

29. Doleck T, Jarrell A, Poitras EG, Chaouachi M, Lajoie SP. Examining diagnosis paths: a process mining approach. In: Computational Intelligence & Communication Technology (CICT), Second International Conference. IEEE, Ghaziabad, India, 2016, 663–667.
30. Vidal JC, Vázquez -Barreiros B, Lama M, Mucientes M. Recompiling learning processes from event logs. Knowledge-Based Syst 2016, 100:160–174.
31. Julia Rudnitckaia. Process Mining: Data Science in Action. Berlin, Germany: Springer; 2016.
32. Romero C, Cerezo R, Bogarín A, Sánchez -Santillán M. Educational process mining: a tutorial and case study using moodle data sets. In: Data Mining and Learning Analytics: Applications in Educational Research. Hoboken, NJ: John Wiley & Sons; 2016, 1–28.
33. Raheela Asif, Saman Hina and Saba, "Predicting Student Academic Performance using Data Mining Methods", International Journal of Computer Science and Network Security (IJCSNS), VOL.17, 2017.
34. Ricardo Mendes And Joao P. Vilela, "Privacy-Preserving Data Mining: Methods, Metrics, and Applications", IEEE, 2017.
35. Schulte J, Fernandez de Mendonca P, Martinez- Maldonado R, Buckingham Shum S. Large-scale predictive process mining and analytics of university degree course data. In: Proceedings of the Seventh International Learning Analytics & Knowledge Conference. ACM; Vancouver, Canada; 2017, 538–539.
36. Mukesh Kumar, A.J. Singh, Disha Handa, "Literature Survey on Student's Performance Prediction in Education using Data Mining Techniques", International Journal of Education and Management Engineering(IJEME), Vol.7, No.6, pp.40-49, 2017.DOI: 10.5815/ijeme.2017.06.05
37. Mukesh Kumar, A.J. Singh, "Evaluation of Data Mining Techniques for Predicting Student's Performance", International Journal of Modern Education and Computer Science (IJMECS), Vol.9, No.8, pp.25-31, 2017.DOI: 10.5815/ijmeecs.2017.08.04
38. Febrianti Widyahastuti, Viany Utami Tjhin, "Predicting Students Performance in Final Examination using Linear Regression and Multilayer Perceptron", IEEE, 2017.
39. R. Sumitha and E.S. Vinoth Kumar, Prediction of Students Outcome Using Data Mining Techniques, International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-2, Issue-6, June 2016 ISSN: 2395-3470.
40. Azwa Abdul Aziz, Nor Hafieza Ismailand Fadhilah Ahmad, First Semester Computer Science Student's Academic Performances Analysis by Using Data Mining Classification Algorithms, Proceeding of the International Conference on Artificial Intelligence and Computer Science(AICS 2014), 15 - 16 September 2014, Bandung, INDONESIA. (e-ISBN978-967-11768-8-7).
41. Priyam, A., Gupta, R., Rathee, A., & Srivastava, S. (2013). Comparative Analysis of Decision Tree Classification Algorithms, 334–337.
42. López, M. I., Luna, J. M., Romero, C., & Ventura, S. (2012). Classification via clustering for predicting final marks based on student participation in forums. International Educational Data Mining Society. 148-151.
43. S.A. Oloruntoba, J. L. Akinode, "Student Academic Performance Prediction Using Support Vector Machine" International Journal of Engineering Sciences & Research Technology, 2016, DOI: 10.5281/zenodo.1130905, ISSN: 2277-9655.
44. Mrs M. Jamuna, Mrs S. A. Shoba, "Educational Data Mining & Students Performance Prediction Using SVM Techniques", International Research Journal of Engineering and Technology (IRJET), 2017, e-ISSN: 2395-0056.
45. A Kadambande, S Thakur et. al., 'Predict Student Performance by Utilizing Data Mining Technique and Support Vector Machine', International Research Journal of Engineering and Technology, e-ISSN: 2395 -0056, 2017.
46. K. B. Eashwar, R. Venkatesan, " Student Performance Prediction Using SVM", International Journal of Mechanical Engineering and Technology (IJMET), Volume 8, Issue 11, November 2017, pp. 649–662, Article ID: IJMET_08_11_066
47. M. Mayilvaganan, D. Kalpana Devi, Comparison of classification techniques for predicting the performance of student's academic environment, in Communication and Network Technologies (ICCNT), 2014 International Conference on, IEEE, 2014, pp. 113–118.