# Student Performance Prediction Using Data Mining Techniques

Dr.K.Venkata Subramanian[1], Dr. S.Jayalakshmi[2], Dr.M.Suresh [3], Dr. D Jebakumar immanuel[4], Sk.Nasreen[5]

[1, 3, 4, 5] Department of Computer Science and Engineering,
[2]Department of Electrical and Electronics Engineering
[1, 2, 3, 5]QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India
[4]SNS College of Engineering, Coimbatore 641107, Tamilnadu
subramanian.k@qiscet.edu.in [1], jaya.lakshmi@qiscet.edu.in [2], csehod@qiscet.edu.in [3], jebakumarimmanuel@snsce.ac.in[4], nasreen.sk@qiscet.edu.in [5]
Corresponding Author Mail: qispublications@qiscet.edu.in

**Article Info**
**Page Number: 297-308**
**Publication Issue:**
**Vol. 69 No. 1 (2020)**

**Article History**
**Article Received:**
**12** September 2020
**Revised: 16** October 2020
**Accepted: 20** November 2020
**Publication: 25** December 2020

Abstract

Universities must have a way to pick students who will do well in school based on objective criteria used in the admissions process. This study looks at how data mining tools can help colleges decide who to let in by making it easier to predict how well students will do once they are there. From 2016 to 2019, 2,039 students at a Saudi state university's Computer Science and Information College were used to test the new method. outcomes show that prospective students' early success in college can be predicted by looking at data collected before they are admitted like average grade points percentage etc. outcomes also show that a student's score on the probabilistic test is the best predictor of how well they will do in the future. Because of this, this score needs to be given more weight during the selection process. Sometimes the ANN method was more accurate (by more than 79%) than the other ways we looked at to group things.

**Keywords:** Decision Trees, Support Vector Machines, and Naive Bayes.

## I. INTRODUCTION

Today, getting into any university or college is hard, but getting into schools of computer science and engineering is even harder. Universities should set up admissions systems that are based on valid and reliable criteria so that they can choose students who have the best chance of doing well in their programmes. Also, before letting someone in, each university should use the most advanced ways to predict how well they will do in school. This would help policymakers at schools come up with good standards for who can get in. All most all the educational institution face the difficulty to handle the huge students database. This is because they only use traditional statistical methods and don't use new, highly predictive methods like (EDM) Educational Data Mining, which is the industry standard for evaluating and predicting how well students will do.

Educators use a method called "educational data mining" to predict how well their students will do in school (EDM). Better data makes it easier to plan ways to help students do better in school. The goal of this research is to help colleges make better decisions about who to let in by various techniques to handle the data to predict an applicant's academic performance better before letting them in.

Applications include improving the learning experience, making it easier for students to finish courses successfully, helping students choose courses, and making profiles of students.

1.2 Predicting how students will do: Categorization is the most common assignment used to predict how well a student will do. There are two main ways to predict a student's success: characteristics and prediction algorithms. According to research in, professors use students' cumulative grade point averages more than any other measure to figure out how well they are doing in school. This tool has been used in a lot of studies.

Other things that researchers often use to predict how well a student will do in college are tests, MCQ, CLA, and assessment grades.. Extracurricular activities, student demographics, and the web of friendships between students are just some of the things that are looked at.

1.3 Motivation

This work adds to the body of literature in more than one way. To do this, we first use techniques for data mining categorization to make four models that use applicants pre-admission poles to predict how well they will do in their early classes. Quiz scores, internal exam results, extracurricular activities, student demographics, cumulative GPA, and social network interactions are all examples of data that can be found in these different roles and are often used to predict a student's success in higher education. However, pre-admission test scores and other indicators of how well a student will do are rarely used in the process of admissions. We also figure out which of our strategies for grouping students is best at predicting their success, as measured by F1-Measuremetrics and other tools.

Most universities find it hard to figure out how students will do by looking at huge educational databases. This is because they only use old-fashioned statistical methods to make their predictions instead of more modern methods.

In order to choose students who will do well in school, it is important to have the criteria to validate the admission process. This is the main point of this article.

1.6 Scope of Project:

The goal of this research is to help universities make better decisions about who to let in by figuring out how smart potential students are. As ways to make predictions, people came up with and built an ANN, a Decision Tree, a Support Vector Machine (SVM), and a Naive Bayes model. For this study, PNU, one of the largest universities in KSA, gave a dataset of 2,039 student records. The ways things are done are not unique to any one type of university.

II. What It Adds to What We Know

ML algorithms have mostly been used in the past to predict how well students will do in school, but they have rarely been used to solve classification problems. The results of this study help make things work better. Our problem statement is based on nature, and the four models we've looked at help us figure out how to measure evaluation metrics. This study has a big impact on how things turn out for students.

## PROPOSED SYSTEM

To answer the first research question, we used a model built with the Linear Regression method. This method is often used to figure out how a set of independent variables (the predictors) are related to a set of dependent variables (i.e., response). We used the model to see how the HSGA, SAAT, and GAT related to the CGPA after the first two semesters of college. The linear relationship between each admission criterion and the CGPA was described using the correlation coefficient, which is typically employed in statistics to quantify the strength and direction of linear relationships between two variables. In order to calculate the impact of each entrance requirement on the students' first-year GPA, we also used the coefficient of determination.

3.1.1 Algorithms

3.1.1 SVM

Both classification and regression issues can be resolved using the supervised machine learning technique known as the Support Vector Machine (SVM). However, its principal application is for categorization issues. The value of each feature is represented by the value of a certain coordinate, and each piece of data is represented by a point in an n-dimensional space (where n is the number of features). Finding the hyper-plane that splits the data into two categories most effectively is the next stage in classifying the data.

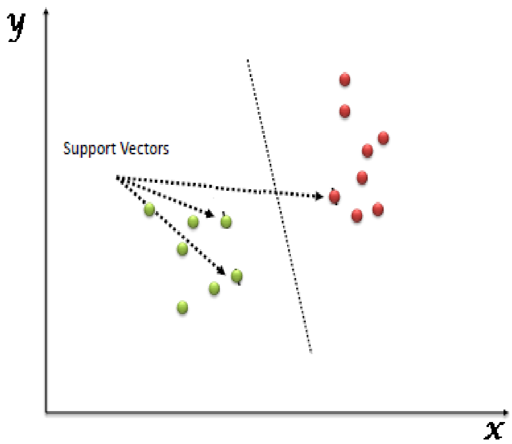



Fig.1. SVM classifier segregates the two classes (hyper-plane/ line).

The points where each observation was made are called "support vectors." The support vector machine (SVM) classifier is the best way to separate the two groups (hyperplane and line).

Naive Bayes classifiers (3.1.2) are a type of classification algorithm that can be used for problems with two classes or more. It is easiest to explain a method by using binary or categorical input values.

This method is often called "naive Bayes" or "stupid Bayes" because the probabilities for each hypothesis are made simple so that they can be calculated. P(d1, d2, d3|h) is calculated as P(d1|h) * P(d2|H), since it is expected that the attribute values are conditionally independent given the goal value.

With real data, this is a very unlikely assumption to make, since the attributes are likely to affect each other. But even if this assumption is wrong, the method still works in surprising ways.

The Random Forests Algorithm:

It is a key part of modern machine learning algorithms and is popular in the field of supervised learning. It can be used in ML to help with Classification and Regression. It uses ensemble learning, which is when different classifiers are mixed together to solve a hard problem and make the model more accurate.

It's a classifier that, as its name implies, averages the outcomes of numerous decision trees used to analyse various subsets of the supplied dataset in order to increase the precision of the predictions made using the dataset. The output of a random forest is predicted by a combination of smaller trees, not by a single large decision tree.

3.1.3 Methodology Decision Trees as a reference

By going from the decision tree's root node to one of its leaf nodes, which contains the group, objects are organised into groups. To categorise an instance, you must first examine the attribute displayed by the tree's root node and then proceed along the branch based on the attribute's value. The offspring tree, whose root is now at the new node, goes through this process once more.

3.1.4 Network of Neurons (ANN) Artificial:

Neural networks are a type of parallel computer that can simulate how the human brain works. The main goal is to make a system that can do a variety of computer tasks faster than what is currently possible. In this course, the main ideas and terms of Artificial Neural Networks are explained. In this course, you will also learn a lot about the training algorithm and network design of the different types of networks used in ANN.

Every neuron has something called a "activation signal" that tells it how to react to something outside of itself. Once the input signals and the activation rule have been put together, the output signals can be sent to other devices.

3.1.5 SDLC (Software Development Life Cycle) (Software Development Life Cycle)

An SDLC is sometimes used to describe the process of making software. The phases of the Software Development Life Cycle are a way to keep track of the whole process of making and maintaining software.

The Software Development Life Cycle (SDLC) is a set of steps for making or improving software. Each step is carefully planned and carried out.

Software Development Life Cycle Payoffs

The goal of any SDLC process is to make it easier to make a product that meets or exceeds all of the above criteria, stays within budget, and meets or beats all of the deadlines. The SDLC is a plan for how to use and get rid of legacy software once an application has been made. Most of the time, the SDLC process is made up of the following steps: Analysis (requirements and design), development, testing, deployment, and maintenance are all parts of the process (response). Using the cloud-based technology that Veracode offers, automated security testing can be added to the SDLC process.

In the first step, called "requirements gathering," we make a complete list of the client's needs. This list includes questions like, "What inputs and outputs does the client need?"

During the second stage, "Analysis," we make a document called the "High Level Design Document" based on what the client wants. The document has sections called "Abstract," "Functional Requirements," "Non-Functional Requirements," "Existing System," "Proposed System," and "Design."

Since it's hard for everyone to understand the High Level Design Document, we made a "Low Level Design Document" to make it easier to understand. This work was made with the help of UML (Unified Modeling Language). There is a use case, a sequence, and cooperation.

In the fourth stage, which is called "coding," code is made in small pieces called "modules." After all the modules have been built, we put them all together.

During testing, we make sure that all of the client's needs have been met. In any case, we're growing and changing again.

In the sixth and final stage, implementation, we make sure that all of the client's needs have been met before moving on with the project. That is, the programme has to be put on a server.

After the programme is released, we will be available to fix any client-side problems that come up so that you can keep using it without interruption.

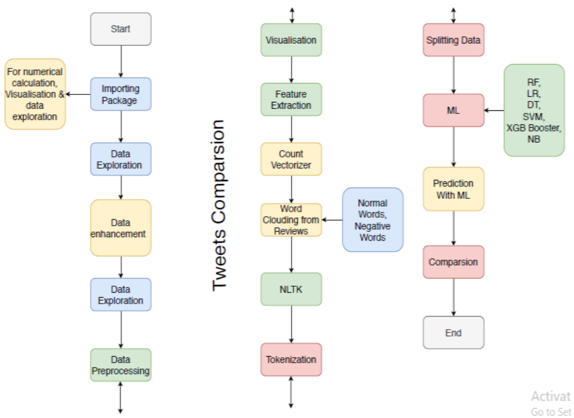



Fig.2. **Architecture**

## IV. Implementation And Outcome

4.1 Regarding the Data Students from PNU's CCIS were used in this study because enrollment data was accessible. However, there is no one type of university that does things in a particular way. The data was obtained from the Admitted and Registration Deanship's electronic academic database. The Institutional Review Board at Penn State granted the ethical approval (Number 19-0152). PNU, the largest college for women in the world, is the sole public university for women in the Kingdom of Saudi Arabia. More than 50,000 students, as well as more than 5,000 academic and staff members, are housed on the 32 million-square-foot campus. The CCIS was established in 2007 to address the rising demand for higher education in computer science and technology among women in Saudi Arabia. Computer science, information systems, and IT make up CCIS. The three departments of CCIS each have four years to complete their prerequisites. There are two semesters in a year.

In the first part of this research, we gathered information from 902 students in 2016–2017 and 667 students in 2017–2018 from all three departments. All of the Saudi women who took part in this study were (the central region of the Kingdom of Saudi Arabia). Also, the government gives all college students money every month so they don't have to work while they're in school. Students from both groups were mostly accepted based on three entrance requirements: the HSGA, the SAAT, and the GAT, with 60%, 20%, and 20% of the weight given to each, respectively.

In line with what was found and suggested in the first part of this study, PNU has given the SAAT score a bigger role as an admissions factor. This is to help predict student success better. A new weighting system has been put in place, giving 30% to HSGA, 40% to SAAT, and 30% to GAT.

Second, we looked at 470 student records from the 2018-2019 school year, when the new weightings were used to decide who was accepted. There were things going on in other student groups that were similar to what was found in the first part of the investigation. We used this set of data to compare the cumulative grade point averages of new students to those of new students who were admitted under the old weighting scheme.

A Look at Some Ways to Measure Success 4.2

These ideas can be used to figure out how well data mining models work:

How accurate are predictions of good things happening? The Real Success Rate (TP).

The rate of wrongly positive results is called the False Positive Rate (FP).

This is shown by the True Negative Rate (TN), which is the percentage of times a negative outcome was correctly predicted.

• False Negative Rate (FN): The number of times a wrong negative prediction is made.

The accuracy is found by dividing the number of correctly predicted values (TP + TN) by the total number of values (TP + TN + FP + FN).

We call it "recall" when we can accurately predict that something good will happen.

Accuracy is measured by the ratio of true positives to true positives plus false positives.

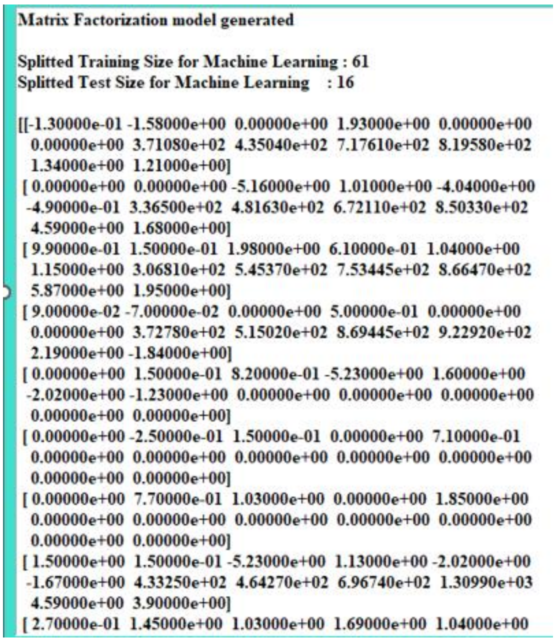
SpreadsheetPiece(TP + FP)

The F1-Measure puts the most weight on how well a classifier does with common categories. It also shows how well recall and precision are balanced (2 Recall Precision/(Recall + Precision)).

Outcomes Expected in Section 4.3: The methods described in this work were put into action using base classifiers like Random Forest, SVM, Logistic Regression, and KNN. The results of the base classifier's predictions will be fed into the ensemble classifier, which will then be used to make more accurate predictions about current and future classes.

In above diagram click on 'Upload UCLA Students Dataset' button to upload dataset

In above diagram I am uploading 'dataset.txt'' as student dataset. After uploading will get below diagram


Matrix Factorization model generated

Splitted Training Size for Machine Learning : 61
Splitted Test Size for Machine Learning : 16

```
[[-1.30000e-01 -1.58000e+00  0.00000e+00  1.93000e+00  0.00000e+00
   0.00000e+00  3.71080e+02  4.35040e+02  7.17610e+02  8.19580e+02
   1.34000e+00  1.21000e+00]
 [ 0.00000e+00  0.00000e+00 -5.16000e+00  1.01000e+00 -4.04000e+00
  -4.90000e-01  3.36500e+02  4.81630e+02  6.72110e+02  8.50330e+02
   4.59000e+00  1.68000e+00]
 [ 9.90000e-01  1.50000e-01  1.98000e+00  6.10000e-01  1.04000e+00
   1.15000e+00  3.06810e+02  5.45370e+02  7.53445e+02  8.66470e+02
   5.87000e+00  1.95000e+00]
 [ 9.00000e-02 -7.00000e-02  0.00000e+00  5.00000e-01  0.00000e+00
   0.00000e+00  3.72780e+02  5.15020e+02  8.69445e+02  9.22920e+02
   2.19000e+00 -1.84000e+00]
 [ 0.00000e+00  1.50000e-01  8.20000e-01 -5.23000e+00  1.60000e+00
  -2.02000e+00 -1.23000e+00  0.00000e+00  0.00000e+00  0.00000e+00
   0.00000e+00  0.00000e+00]
 [ 0.00000e+00 -2.50000e-01  1.50000e-01  0.00000e+00  7.10000e-01
   0.00000e+00  0.00000e+00  0.00000e+00  0.00000e+00  0.00000e+00
   0.00000e+00  0.00000e+00]
 [ 0.00000e+00  7.70000e-01  1.03000e+00  0.00000e+00  1.85000e+00
   0.00000e+00  0.00000e+00  0.00000e+00  0.00000e+00  0.00000e+00
   0.00000e+00  0.00000e+00]
 [ 1.50000e+00  1.50000e-01 -5.23000e+00  1.13000e+00 -2.02000e+00
  -1.67000e+00  4.33250e+02  4.64270e+02  6.96740e+02  1.30990e+03
   4.59000e+00  3.90000e+00]
 [ 2.70000e-01  1.45000e+00  1.03000e+00  1.69000e+00  1.04000e+00
```

**Fig.3. records converted to feature**

In above diagram we can see all records converted to feature vector and in above diagram in first 3 lines we can see from above matrix application using 61 records to train machine learning model and 16 records to test accuracy or to calculate Mean Square Error of classifier. If algorithm prediction result is high then accuracy will be more and Mean Square Error (MSE) will be less. Now we got matrix and data to train and test classifier. Now click on 'Run SVM Algorithm' to train SVM classifier and to get it accuracy and MSE value

In above diagram uploading new student records as test file and below are the prediction results

```
[2.70000e-01 1.45000e+00 1.03000e+00 1.69000e+00 1.04000e+00 8.00000e-02
 3.99670e+02 4.41830e+02 7.04345e+02 0.00000e+00 0.00000e+00 0.00000e+00]
Predicted New Course GPA Score will be : High

[ 1.3100e+00  3.8000e-01 -1.7800e+00  1.1300e+00 -6.4000e-01 -1.2300e+00
  0.0000e+00  4.7721e+02  0.0000e+00  6.3017e+02  3.7800e+00  0.0000e+00]
Predicted New Course GPA Score will be : Low

[ 5.10000e-01  1.45000e+00 -2.40000e-01  2.00000e+00  4.40000e-01
 -4.90000e-01  3.30890e+02  4.69750e+02  6.88650e+02  1.10304e+03
  3.26000e+00  3.13000e+00]
Predicted New Course GPA Score will be : High

[ 1.06   1.45   1.03   1.2    2.77   0.     0.    490.38  592.565
  0.     2.85   0.  ]
Predicted New Course GPA Score will be : High
```

**Fig.4. student performance in current subjects**

The student grades for ongoing courses are shown in square brackets in the above diagram. These grades are transformed to matrix factorization and then used with the EPP train model to forecast whether a student's GPA would be low or high. After each test record is shown in the diagram above, the anticipated result value is shown. Click now on “Mean Square Error Graph’ button to get below graph

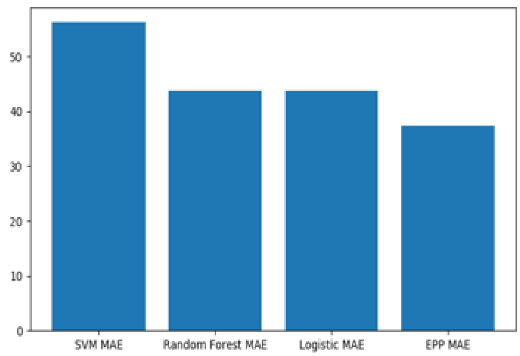
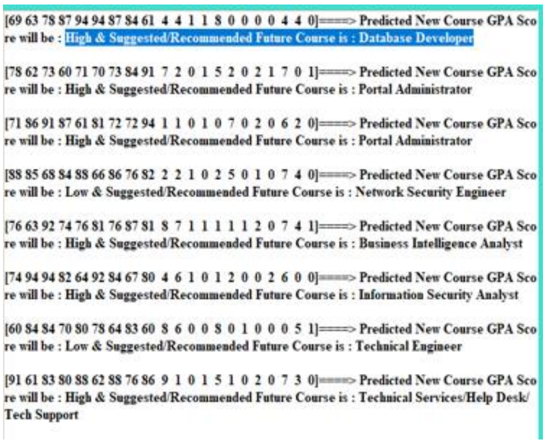



**Fig.5. algorithm name, MSE**

X- axis  in the graph above reflects algorithm name, while Y- axis  is MSE (mean square error). The proposed algorithm has a lower MSE error and higher accuracy when compared to existing algorithms, as seen in the following graph. We can deduce from the following graph that the proposed EPP performs better in terms of prediction compared to other algorithms. **Extension Outcomes:**

In this project in propose work we were predicting student performance as HIGH or LOW but not predicting based on performance which future course is suitable for him. So in extension work we changed dataset and algorithms prediction to predict not only student performance grade but also predict future suitable courses for him based on past performance.

ISSN: 2094-0343
2326-9865



```
[69 63 78 87 94 94 87 84 61  4  4  1  1  8  0  0  0  0  4  4  0]====> Predicted New Course GPA Sco
re will be : High & Suggested/Recommended Future Course is : Database Developer

[78 62 73 60 71 70 73 84 91  7  2  0  1  5  2  0  2  1  7  0  1]====> Predicted New Course GPA Sco
re will be : High & Suggested/Recommended Future Course is : Portal Administrator

[71 86 91 87 61 81 72 72 94  1  1  0  1  0  7  0  2  0  6  2  0]====> Predicted New Course GPA Sco
re will be : High & Suggested/Recommended Future Course is : Portal Administrator

[88 85 68 84 88 66 86 76 82  2  2  1  0  2  5  0  1  0  7  4  0]====> Predicted New Course GPA Sco
re will be : Low & Suggested/Recommended Future Course is : Network Security Engineer

[76 63 92 74 76 81 76 87 81  8  7  1  1  1  1  1  2  0  7  4  1]====> Predicted New Course GPA Sco
re will be : High & Suggested/Recommended Future Course is : Business Intelligence Analyst

[74 94 94 82 64 92 84 67 80  4  6  1  0  1  2  0  0  2  6  0  0]====> Predicted New Course GPA Sco
re will be : High & Suggested/Recommended Future Course is : Information Security Analyst

[60 84 84 70 80 78 64 83 60  8  6  0  0  8  0  1  0  0  0  5  1]====> Predicted New Course GPA Sco
re will be : Low & Suggested/Recommended Future Course is : Technical Engineer

[91 61 83 80 88 62 88 76 86  9  1  0  1  5  1  0  2  0  7  3  0]====> Predicted New Course GPA Sco
re will be : High & Suggested/Recommended Future Course is : Technical Services/Help Desk/
Tech Support
```

**Fig.19. student performance data**

Below the arrow symbol in the preceding diagram, we can see the expected performance of the student as HIGH or LOW, along with a suggested or recommended next course.

Therefore, students can not only anticipate their grades but also plan their academic careers with the help of extension activities.

## CONCLUSION

The purpose of this research is to help universities improve their admissions processes by forecasting students' future academic success. Artificial Neural Networks (ANNs), Decision Trees, Support Vector Machines (SVMs), and Naive Bayes were used to propose and create four different types of prediction models. One of the largest universities in KSA, PNU, provided a dataset consisting of 2,039 student records for this study. However, the approaches taken are not specific to any particular type of university. The results of the study provide credence to the usefulness of predictive modelling in universities, where it may be used by decision-makers to better use the limited resources available to them. In addition, the findings demonstrate that, with adequate pre-admission data, a high-performance model to predict students' early performance may be built. It was found that the ANN model achieved an accuracy of about 79.22 percent in this study. Furthermore, this study demonstrates that the ANN technique surpasses the others in accuracy and precision metrics, while the Decision Tree technique outperforms the others in recall and F1-Measure. Results were poorest for Naive Bayes. According to the results of this research, SAAT score is the entrance criterion that best predicts future academic achievement; as a result, it should be given greater weight. The admissions committee at PNU, where the research was done, considered the advice and chose to change the weights of the three entrance factors specified in the study, giving more weight to the student's performance on the Scholastic Aptitude Test (SAT).

This study compares the cumulative grade point averages of first-year PNU students who were admitted under the former admissions weighted scheme with those who were admitted under the new system. After implementing the revised entrance weighted system, 31% more students earned an excellent or very good grade point average in their first year, while 18% fewer students earned a grade of "acceptable" or "poor."

BIBILOGRAPHY


**1.** D. Kabakchieva, "Predicting Student Performance by Using Data Mining Methods for Classification," Cybernetics and Information Technologies, vol. 13, Mar. 2013**.**

**2.** S. K. Mohamad and Z. Tasir, "Educational data mining: A review", *Procedia Social Behav. Sci.*, vol. 97, pp. 320-324, Nov. 2013.

**3.** A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works", *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1432-1462, Mar. 2014

**4.** H. Aldowah, H. Al-Samarraie and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: A review and synthesis", *Telematics Informat.*, vol. 37, pp. 13-49, Apr. 2019.

**5.** C. Anuradha and T. Velmurugan, "A comparative analysis on the evaluation of classification algorithms in the prediction of students performance", *Indian J. Sci. Technol.*, vol. 8, no. 15, pp. 974-6846, Jan. 2015.

**6.** V. L. Miguéis, A. Freitas, P. J. V. Garcia and A. Silva, "Early segmentation of students according to their academic performance: A predictive modelling approach", *Decis. Support Syst.*, vol. 115, pp. 36-51, Nov. 2018.

**7.** M. Mayilvaganan and D. Kalpanadevi, "Comparison of classification techniques for predicting the performance of students academic environment", *Proc. Int. Conf. Commun. Netw. Technol.*, pp. 113-118, Dec. 2014.

**8.** S. Natek and M. Zwilling, "Student data mining solution–knowledge management system related to higher education institutions", *Expert Syst. Appl.*, vol. 41, no. 14, pp. 6400-6407, Oct. 2014.

9. T. M. Christian and M. Ayub, "Exploration of classification using NB tree for predicting students' performance", *Proc. Int. Conf. Data Softw. Eng. (ICODSE)*, pp. 1-6, Nov. 2014.

10. J. Albishri, S. Aly and Y. Alnemary, "Admission criteria to Saudi medical schools. Which is the best predictor for successful achievement?", *Saudi Med. J.*, vol. 33, pp. 1222-1228, 2012.

11. K. V. J.K. Jothi Kalpana, "Intellectual performance analysis of students by using data mining techniques", International Journal of Innovative Research in Science, Engineering and Technology, vol. 3, 2014.

12. V. Ramesh, "Predicting student performance: A statistical and data mining approach", International Journal of Computer Applications, vol. 63, no. 8, 2013.

13. D. A. M. Dr. Abdullah AL-Malaise and M. Alkhozae, "Students performance prediction system using multi agent data mining technique", International Journal of Data Mining and Knowledge Management Process, vol. 4, no. 5, 2014.

14. P. Kavipriya, "A Review on Predicting Students‟ Academic Performance Earlier, Using Data Mining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 12, December 2016.

15. Shruthi P, Chaitra B P, "Student Performance Prediction in Education Sector Using Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 3, March 2016.

16. Humera Shaziya, et.al. "Prediction of Students Performance in Semester Exams using a Naïve bayes Classifier", International Journal of Innovative Research in Science,

Engineering and Technology, Vol. 4, Issue 10, October 2015.

17. C. Romero, S. Ventura, P. G. Espejo, C. Hervas, ´ Data mining algorithms to classify students, in: Educational Data Mining 2008, 2008.
18. B. M. Bidgoli, D. Kashy, G. Kortemeyer, W. Punch, Predicting student performance: An application of data mining methods with the educational web-based system lon-capa, in: Proceedings of ASEE/IEEE frontiers in education conference, 2003.
19. V. Oladokun, A. Adebanjo, O. Charles-Owaba, Predicting students academic performance using artificial neural network: A case study of an engineering course, The Pacific Journal of Science and Technology 9 (1) (2008) 72–79.
20. A. Bogar´ın, C. Romero, R. Cerezo, M. Sanchez-Santill ´ an, ´ Clustering for improving educational process mining, in: Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, ACM, 2014, pp. 11–15.
21. C. Coffrin, L. Corrin, P. de Barba, G. Kennedy, Visualizing patterns of student engagement and performance in moocs, in: Proceedings of the fourth international conference on learning analytics and knowledge, ACM, 2014, pp. 83–92.
22. 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.
23. 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.
24. S. S. Meit, N. J. Borges, B. A. Cubic, H. R. Seibel, Personality differences in incoming male and female medical students., Online Submission.
25. A. Simsek, J. Balaban, Learning strategies of successful and unsuccessful university students., Online Submission 1 (1) (2010) 36–45.
26. 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
27. C. Romero, M.-I. Lopez, ´ J.-M. Luna, S. Ventura, Predicting students' final performance from participation in on-line discussion forums, Computers & Education 68 (2013) 458–472.
28. N. Thai-Nghe, T. Horvath, ´ L. Schmidt-Thieme, Factorization models for forecasting student performance, in: Educational Data Mining 2011, 2010.
29. P Ramprakash, M Sakthivadivel, N Krishnaraj, J Ramprasath. "Host-based Intrusion Detection System using Sequence of System Calls" International Journal of Engineering and Management Research, Vandana Publications, Volume 4, Issue 2, 241-247, 2014
30. N Krishnaraj, S Smys."A multihoming ACO-MDV routing for maximum power efficiency in an IoT environment" Wireless Personal Communications 109 (1), 243-256, 2019.
31. N Krishnaraj, R Bhuvanesh Kumar, D Rajeshwar, T Sanjay Kumar, Implementation of energy aware modified distance vector routing protocol for energy efficiency in wireless

sensor networks, 2020 International Conference on Inventive Computation Technologies (ICICT),201-204
32. Ibrahim, S. Jafar Ali, and M. Thangamani. "Enhanced singular value decomposition for prediction of drugs and diseases with hepatocellular carcinoma based on multi-source bat algorithm based random walk." Measurement 141 (2019): 176-183. https://doi.org/10.1016/j.measurement.2019.02.056
33. Ibrahim, Jafar Ali S., S. Rajasekar, Varsha, M. Karunakaran, K. Kasirajan, Kalyan NS Chakravarthy, V. Kumar, and K. J. Kaur. "Recent advances in performance and effect of Zr doping with ZnO thin film sensor in ammonia vapour sensing." GLOBAL NEST JOURNAL 23, no. 4 (2021): 526-531. https://doi.org/10.30955/gnj.004020 , https://journal.gnest.org/publication/gnest_04020
34. N.S. Kalyan Chakravarthy, B. Karthikeyan, K. Alhaf Malik, D.Bujji Babbu,. K. Nithya S.Jafar Ali Ibrahim , Survey of Cooperative Routing Algorithms in Wireless Sensor Networks, Journal of Annals of the Romanian Society for Cell Biology ,5316-5320, 2021
35. Rajmohan, G, Chinnappan, CV, John William, AD, Chandrakrishan Balakrishnan, S, Anand Muthu, B, Manogaran, G. Revamping land coverage analysis using aerial satellite image mapping. Trans Emerging Tel Tech. 2021; 32:e3927. https://doi.org/10.1002/ett.3927
36. Vignesh, C.C., Sivaparthipan, C.B., Daniel, J.A. et al. Adjacent Node based Energetic Association Factor Routing Protocol in Wireless Sensor Networks. Wireless Pers Commun 119, 3255–3270 (2021). https://doi.org/10.1007/s11277-021-08397-0.
37. C Chandru Vignesh, S Karthik, Predicting the position of adjacent nodes with QoS in mobile ad hoc networks, Journal of Multimedia Tools and Applications, Springer US,Vol 79, 8445-8457,2020