Variables identification for Students Performance Prediction

Vandana Bharadi
Poornima University
Jaipur, India
vandana.bharadi@gmail.com

Dr Satya Prakash Awasthi
Poornima University
Jaipur, India
satya.prakash@poornima.edu.in

***Abstract*— Student's performance analysis has taken a leap of faith in past two years when the delivery mode was shuttling between online and offline. Various factors which are significantly affecting student's performance are now newly to be researched and identified. Its very important to not only consider and study the effect of various academic factors but also socio-economic factors are needed to analyzed. Predictive analytics has shown its capabilities in efficiently predicting results in wide areas of application including academics. This analysis and prediction is most crucial in the developing country like India, where the published rate of retention of students at university level considered very low. In this research, the academic and socio-economic details collected from student through survey. Further efficacy of various machine-learning algorithms assessed by running these algorithms on survey data. The findings demonstrate that some machine learning algorithms may create accurate predictive models using historical data on student retention.**

Keywords—Student Performance Analysis, Predictive Modeling, Data Mining

I. INTRODUCTION

In India, there are now 3,45,84,781 students enrolled, of which 2,74,20,450 (79.3 percent) are enrolled in undergrad courses, 39,20,000 are pursuing postgraduate degrees, and 132204 are hoping to earn their PhDs. Additionally, just 7.4% of students are working toward a graduation. [57] With such a large database at our disposal, we can create predictive models that link student performance to a variety of factors, like age, gender, parent's educational history, and so on.

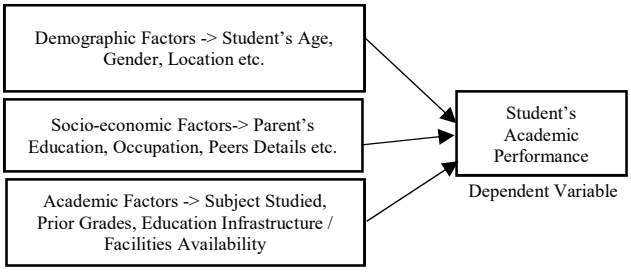
According to figures released by the Ministry of Human Resource Development[58], fewer undergraduate students are enrolling in graduate programmes each year. To enhance the educational climate in the nation, this reluctance to pursue higher education must be addressed. In order to determine the numerous elements influencing students' performance, the focus of this research is on gathering and analysing data from students. Additionally, it focuses on figuring out what corrective actions may be taken to help children perform better who are struggling with academic issues.

Grades are considered to measure of student's excellency in their academic carrier. Students labeled as ranking student, first class student, average student etc based on their respective grades. So far the academicians, researchers were focusing on building a curriculum, understanding the strength and weaknesses of the courses, developing the assessment tools etc. But after understanding the scenario of retention of students in higher academic studies, it is very much needed to understand the reasons of failure and low performance of students. With this intention, this research aims at developing a predictive model for early identification of at-risk students. Successful implementation of this model will not only targeting the low performance students but also will be benefitting other academic activities such as placements.

Following are the objectives outlined for proposed work.

1. Identification of various factors (independent variables) significantly affecting student performance (dependent variables) which can be further used for predicting student's grades most accurately. (Refer Figure 1)

Figure 1: Independent and dependent Variables of the research



2. Collection of student's data based on the identified factors (independent and dependent variables).
3. Confirming the impactful independent variables using T Test referring to survey data.

In this paper Section II listing summaries of relevant papers and common variables affecting students performance. Section III briefs on the survey design and data collection process. Section IV provides necessary figures for confirming the variables proved having significant impact on student performance.

II. RELATED WORK

Research carried out by [SEOW, 2014] supports the practice of identifying the factors that influence a student's academic performance, such as athletics, dance, and music, to mention a few. It also identifies current students who are at risk of academic failure and reduces the possibility of accepting individuals with skill sets that are unsuitable for an accounting degree programme. However, this study shows that participating in extracurricular activities has a positive impact on a student's academic performance.[10]
[Mishra, et-al, 2013] The goal of this research is to see how participation in extracurricular activities affects academic growth, social skills, and high school graduation. According to research, both the type of extracurricular initiatives and the amount of money invested might have an impact on people's progress. The purposive sample approach was employed to choose the samples. Two private schools in Lucknow were chosen to represent the study's population. For this project, 60 high school children between the ages of 13 and 15 were chosen. For data collection, a self-developed poll was used. Finally, extracurricular activities are an important component

978-1-6654-9291-1/22/$31.00 ©2022 IEEE

of any student's life. They have a huge impact on kids' life. They have an impact on students' life by increasing their conduct, school performance, school fulfilment, and positive outlooks in order to produce fruitful adults. [23]

[Ahmad.M, et-al, 2019] The results of final terminal examinations and class performance were the dependent variables. Gender, involvement, and non-participation were among the independent variables. Students who participate in co-curricular activities had superior academic outcomes, more educational attainment, and a clearer self-concept, according to the study's findings. A variety of talents, such as teamwork and leadership, have also been identified among them.[43]

[Ashwin et. al] The authors investigated and performed research to discover elements that influence student performance. In a sense, this research serves two purposes: first, it investigates variables in a developing country such as class attendance, study hours, and family income, and second, it attempts to quantify the impact of these household characteristics on performance. This helped us realize that these characteristics really have a major impact on students' academic performance and efficiency, and that they are generally disregarded qualitative factors.[1]

The first stage in developing any model for prediction is to understand and identify the factors that can be changed and their impact on students' performance monitored. Independent variables are the variables that will be changed in order to measure their influence. The variable to determine or predict known as dependent various. The dependent variable in this study will be the student's performance or their result or score.[44]

In the field of education data mining, researchers who are working on early prediction of student are implementing various methods of data mining for classifying students into pass and fail categories. While designing and implementing these algorithms the independent or the input variables that they have identified are varying based on the courses, the subjects, level of education, location of college etc. Therefor in the current study the research work in variety of such research papers have been reviewed and compiled to list the commonly affecting factors as listed in Table 1.

TABLE I. INDEPENDENT VARIABLE LISTING WITH SOURCE

	Independent Variables	References Sources
1	Age	[2,4,7,10,11,13,14,15,16,63,66, 68,70, 71,72,73,74,75,77,80]
2	Gender	[2,4,7,10,11,13,14,15,16,59,63, 66,68, 70,71,72,73,74,75,76,77,78,79, 80,81,83]
3	Family (Father/Mother) occupation	[9,15,68,71,74]
4	Family (Father/Mother) Education	[9,14,15,63,68,70,74]
5	Family Structure	[9,14,15,68]
6	Family's economic status / Financial Aid	[2,4,9,76,83]
7	Employment	[2,18]
8	Parental Encouragement	[9,15]
9	Students Habits (smoking/bad)	[9]
10	Physical Disabilities	[9]
11	Marital Status	[9,18,72,73,80]
12	Number of Children	[73]
13	City (Residing)	[9,63,66,70,71,74,75,76]
14	City (College)	[77,79,80]
15	Social Life Involvement	[15]
16	Friends	[9]
17	Prior Schooling	[2,9,10,16,59,60,62,68,72,73,74 ,75,79,80,83]
18	Medium of Instruction	[59,72,79]
19	Prior Academic Performance	[9,10,11,13,14,15,58,59,60,62,6 3,66, 67,68,72,74,76,77,78,80,81,83]
20	Area of school	[59]
21	Exam/ICA Performance	[3, 7,9,10,15,59,63,68,69,70,73,75, 76,78,79,85,87]
22	Current Course	[4,69,70,71,75,87]
23	Course Structure	[18,62,63,67,69,71,72,87]
24	Course Difficulty	[62,70]
25	Attendance	[6, 10, 16,70,79]
26	Hours Of Study	[9,15,18,76]
27	BackLog/Gap	[75,87]
28	Resource Availability (space/Material/Transport/P hone/Internet)	[9,79,81,83]
29	Faculty Training/Expertise/Involve ment	[9,66]
30	Extra-Curricular Activities	[15,68]
31	Motivation / Self Grooming	[9,10,13,15,16,18,76,77,78,79,8 1,83]
32	Student's Class Position	[17]
33	Pre-requisite	[58,62]

Among the 33 variables listed in above table following factors were assumed to be having higher impact student's academic performance and these factors were further confirmed with the faculties by conducting personal interviews.

TABLE II. INDEPENDENT VARIABLE LISTING WITH SOURCE

Sr No	Independent Variables
1	Gender
2	Prior Academic Performance
3	Age
4	Exam/ICA Performance
5	Prior Schooling
6	Motivation / Self Grooming
7	City (Residing)
8	Course Structure
9	Extra-Curricular Activities

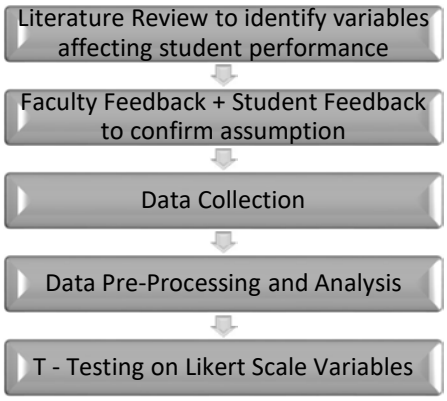
Sr No	Independent Variables
10	Family (Father/Mother) Education
11	Current Course
12	Family (Father/Mother) occupation
13	family's economic status / Financial Aid
14	Marital Status
15	Attendance
16	Family Structure
17	Hours Of Study
18	Resource Availability (space/Material/Transport/Phone/Internet)
19	City (College)
20	Medium of Instruction

The basic approach of this research study is to identify factors significantly affecting students' academic performance considering students belonging to various environmental and academic factors. This is important for identification of students having higher chances of failing or of poor performance. After identification if extra attention is given to such students so that the chances of passing the examination will be improved.

The workflow is as follows:

i. Literature Review to identify variables affecting student performance
ii. Student's data collection for identified factors will be done in order to confirm the assumptions derived through review literature regarding student's performance prediction.
iii. Faculty Feedback to further confirm these findings.

Figure 2: Workflow



III. DATA COLLECTION AND DATA ANALYSIS

A. Survey Design - student's opinion analysis

Among the 33 variables listed in above table, 20 variables (mentioned in Table 2 : Confirmed Independent Variables) were the factors which were very commonly identified independent variables while analysing student's academic performance by various researcher in the same domain. The final conclusion in identification process was not only based on the literature survey done for almost 98 papers. Further, a delphi technique was implemented in which the experience faculties from various domains were interviewed. These factors were also listed and confirmed by the domain expert in the interview process. After identifying the variables to be having possibly higher impact on student's academic performance, the survey was designed to collect information from students. This questionnaire in the survey was consisting of two sections where in section 1 was collecting demographic related details and section 2 was collecting specific details regarding educational background and academics.

While framing these questions it is confirmed that all the factors listed in Table 2 were included after identifying proper scale. This scale identification process was equally considering the further analysis process. Table 3 is listing the distinct values and the coding scheme for all the questions framed for the independent variables identified in Table 2.

TABLE III. DATA CODING FOR ALL THE QUESTIONS (FACTORS) IN SURVEY

Factor	Distinct Values	Coding
Gender	Female	0
	Male	1
Negative and Positive Factor	Strongly Disagree	1
	Disagree	2
	Neither Agree or Disagree	3
	Agree	4
	Strongly Agree	5
Staying at hostel / Rented Apartment (away from family)	Yes	0
	No	1
Family Income (Yearly)	Below 5 lakhs	0
	> = 5 lakhs and < 10 lakhs	1
	> = 10 lakhs and < 20 lakhs	2
	> = 20 lakhs and < 30 lakhs	3
	More than 30 lakhs	4
Father's/Mother's Highest Education Degree		
	Graduate (Bachelor's Degree)	1
	Post-Graduate (Master's Degree)	2
	Ph.D.	3
Current Course you studying	Engineering	0
	Pharmacy	1
	Business Administration (BBA / MBA)+F45	2
	Science	3
	Commerce	4
Current Education Level	Under-Graduate (Studying in High School)	0
	Graduate (Pursuing Bachelor's degree)	1
	Post-Graduate (Pursuing Master's degree)	2
	Ph.D	3

For understanding students and teachers opinion on various factors that can be possibly affecting student's performance in positive or negative direction Likert scale questions were framed for factors listed in Table 4 and 5.

TABLE IV. FACTORS WITH POSITIVE IMPACT ON STUDENT'S PERFORMANCE

	Factors
1	Trained Faculty guidance
2	Resource Availability (Labs, Library etc)
3	Continuous Assessment (Tests) During Course
4	Student's focus on career and self-grooming
5	Use of New Educational Tools and Techniques in Teaching (Smart boards, Portals etc)
6	Prior in-depth knowledge about current subjects (Prerequisites)
7	Score/Performance in Prior Course

TABLE V. FACTORS WITH NEGATIVE IMPACT ON STUDENT'S PERFORMANCE

	Factors
1	Extra-Curricular Activities Participation
2	Social Media Platforms utilization
3	Peer Pressure
4	Parent's involvement in student's study
5	Prior Education from Vernacular Medium
6	Prior Education from state boards (not from CBSC, ICSC, IB etc)
7	Low Attendance

Student's Data Survey circulated to students from different courses, region, and universities. 540 students responses were collected.

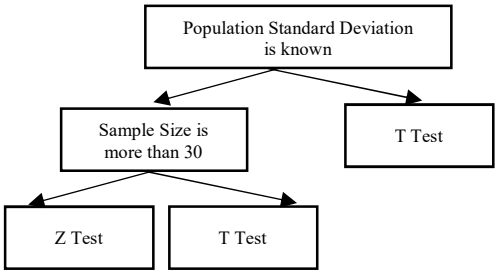
B. Survey Design - Teacher's opinion analysis

In the current study, apart from confirming the independent variables through literature review and student's survey data analysis, another survey data collected from 33 faculties of various universities, teaching different subject in different courses at different level. This survey was the feedback which was taken from 33 faculties.

C. Hypothesis Testing

In order to understand, what the entire population inferences is with respect to their agreement on positive and negative factors responsible for student's performance, t-test analysis is carried out here, as the population standard deviation is unknown as mentioned in Figure 3 .

Figure 3: Selection of T Test or Z Test



In figure 1 Refer to table 3 and 4 for the Null and alternative hypothesis for T-Test to assess factors assumed to be having positive and negative impact. As an academician and researcher, it is to be proven that these are the factors which are having impact (either favorable or non-favorable). There statement are put in alternative hypothesis. In the survey this is mapped to opinion related question using likert scale on 1 to 5. In which 1-Strongly disagree, 2-Disagree, 3-Neutral, 4 - Agree, and 5 - Strongly Agree. Example, to prove Trained Faculty Guidance has a positive impact on student performance following will be null and alternative hypothesis.

H0: u <= 3 i.e. Trained Faculty Guidance does not have positive impact on student performance

Ha: u > 3 i.e. Trained Faculty Guidance has positive impact on student performance

TABLE VI. NULL AND ALTERNATIVE HYPOTHESIS FOR T TESTING

Hypothesis	Statement	Description
Null Hypothesis	H0:u<=3	H0: Students Survey data indicates that these factors have no positive impact on their academic performance
Alternative Hypothesis	Ha:u>3	Ha: Students Survey data indicates that these factors have positive impact on their academic performance

IV. EXPERIMENTAL ANALYSIS

With the aim of understanding the positive and negative factor affecting the student performance, with significance level 0.05, the factors were shortlisted with calculated value < 0.05. These positive factors are Trained Faculty guidance, Resource Availability, Continuous Assessment (Tests) During Course, Student's focus on career and self-grooming, Use of New Educational Tools and Techniques in Teaching, Prerequisites, and Score/Performance in Prior Course. Negative factors with calculated p value < 0.05 are Social Media Platforms utilization, Peer Pressure, and Low Attendance.

Refer table 7, all the 7 positive factors having hypothesis testing p-value extremely less than 0.05. Means we failed to accept null hypothesis. Hence we accept alternative hypothesis and this indicates students are of the opinion with value > 3 that they agreeing on these factors affecting their studies in positive direction.

TABLE VII. P VALUE COMPARISON FOR POSITIVE FACTORS

Positive Factors	P value	Is Calculated P Value < 0.05 (95% CI) Factor has Positive Impact on Student's Academic Performance
Trained Faculty guidance	9.55557E-20	Yes
Resource Availability	7.63805E-23	Yes
Continuous Assessment (Tests) During Course	2.63499E-14	Yes
Student's focus on career and self-grooming	1.25342E-17	Yes
Use of New Educational Tools and Techniques in Teaching	9.97925E-12	Yes
Prerequisites	6.5692E-17	Yes
Score/Performance in Prior Course	7.29849E-13	Yes

Refer table 8, among all the 7 listed negative factors, social media platforms utilization, peer pressure and low attendance are having hypothesis testing p-value extremely less than 0.05. Means we failed to accept null hypothesis. Hence, we accept alternative hypothesis and this indicates students are of the opinion with value > 3 that they agreeing on these factors affecting their studies in negative direction.

TABLE VIII. P VALUE COMPARISON FOR NEGATIVE FACTORS

Negative Factors	**P value**	**Is Calculated P Value < 0.05 (95% CI) Factor has Negative Impact on Student's Academic Performance**
Extra-Curricular Activities Participation	0.999908058	No
Social Media Platforms utilization	0.000243027	Yes
Peer Pressure	7.16664E-10	Yes
Parent's involvement in student's study	0.203315294	No
Prior Education from Vernacular Medium	0.70099818	No
Prior Education from state boards	0.807602036	No
Low Attendance	4.13798E-08	Yes

V. CONCLUSION

The rising dropout rate among students is a major source of concern for academics and parents alike. Accurate examination of multiple factors for student failure, as well as further early prediction, can lead to an increase in student success rates. Early prediction of such pupils will also aid universities in fine-tuning the placement process. Trained Faculty guidance, Resource Availability, Continuous Assessment (Tests) During Course, Student's Focus on Career and Self-Grooming, Use of New Educational Tools and Techniques in Teaching, Prerequisites for the given course, and Score/Performance in Prior Course are the factors identified and to be focused on to improve students' performance in the current study. At the same time, kids must be taught and counselled on how to avoid being negatively impacted by the use of social media platforms and peer pressure. Attendance has also been highlighted as a significant element, and students who have a poor attendance rate are more likely to fail.

REFERENCES

[1] Ashwin Satyanarayana, Mariusz Nuckowski, "Data Mining using Ensemble Classifiers for Improved Prediction of Student Academic Performance" Spring '2016' Mid. Atlantic 'ASEE' Conference, 'April' 8.9,'2016' GWU.

[2] Predictive Analytics in Higher Education Data-Driven Decision-Making for the Student Life Cycle, January 2013 [white paper], Eduventures, Inc.

[3] Márquez-Vera, C., Romero Morales, C., & Ventura Soto, S. (2013). Predicting school failure and dropout by using data mining techniques. Revista Iberoamericana de Tecnologias Del Aprendizaje, 8(1), 7–14. https://doi.org/10.1109/RITA.2013.2244695

[4] Guleker, R. (2014). The Effect of Attendance on Academic Performance. Mediterranean Journal of Social Sciences. https://doi.org/10.5901/mjss.2014.v5n23p961

[5] Sharker, S., & Rahman, M. (2015). Determinants of Academic Performance-A Multinomial Logistic Regression Approach. International Journal of Scientific & Engineering Research, 6(11). Retrieved from http://www.ijser.org

[6] Arumugam, Shanthini & G, Vinodhini & Chandrasekaran, Dr. (2018). Predicting Students' Academic Performance in the University Using Meta Decision Tree Classifiers. Journal of Computer Science. 14. 654-662. 10.3844/jcssp.2018.654.662.

[7] Goga, M., Kuyoro, S., & Goga, N. (2015). A Recommender for Improving the Student Academic Performance. Procedia - Social and Behavioral Sciences, 180, 1481–1488. https://doi.org/10.1016/j.sbspro.2015.02.296

[8] Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A review on predicting student's performance using data mining techniques. In 3rd Information Systems International Conference, 2015 (Vol. 72, pp. 414–422). Shenzhen: Elsevier. https://doi.org/10.1016/j.procs.2015.12.157

[9] Lukkarinen, A., Koivukangas, P., & Seppälä, T. (2016). Relationship between Class Attendance and Student Performance. Procedia - Social and Behavioral Sciences, 228, 341–347. https://doi.org/10.1016/j.sbspro.2016.07.051

[10] Kamal, A. M. M. (2016). Association of student's position in a classroom and student's academic performance using ANOVA. In Proceedings - 2015 5th International Conference on e-Learning, ECONF 2015 (pp. 392–395). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ECONF.2015.56

[11] Aramburo, V., Boroel, B., & Pineda, G. (2017). Predictive Factors Associated with Academic Performance in College Students. Procedia - Social and Behavioral Sciences, 237, 945–949. https://doi.org/10.1016/j.sbspro.2017.02.133

[12] Chaurasia, V., & Pal, S. (2014). A Novel Approach for Breast Cancer Detection using Data Mining Techniques. International Journal of Innovative Research in Computer and Communication Engineering, 2(1), 2456–2465.

[13] P. Swathi Baby, T. Panduranga Vital, "Statistical Analysis and Predicting Kidney Disease Using Machine Learning Algorithms", International Journal of Engineering Research and Technology (IJERT) ISSN: 2278-018, Vol 4, Issue 07, July -2015, Pg 206-210.

[14] X. X. Ai, H. Jia and L. Xin, "SVM-based Cancer Incidence Forecasting of Patients," 2016 9th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, 2016, pp. 281-284. doi: 10.1109/ISCID.2016.2074

[15] Mohammad, M. N., Sulaiman, N., & Muhsin, O. A. (2011). A novel Intrusion Detection System by using intelligent data mining in WEKA environment. In Procedia Computer Science (Vol. 3, pp. 1237–1242). https://doi.org/10.1016/j.procs.2010.12.198

[16] Datla, M. V. (2015). Bench marking of classification algorithms: Decision Trees and Random Forests - a case study using R. 2015 International Conference on Trends in Automation, Communications and Computing Technology (I-TACT-15), 1–7. https://doi.org/10.1109/ITACT.2015.7492647

[17] All India Survey on higher education (2015-16), Ministry of Human Resource Development.

[18] Huang, S., & Fang, N. (2012). Work in progress: Early prediction of students' academic performance in an introductory engineering course through different mathematical modelling techniques. In Proceedings - Frontiers in Education Conference, FIE. https://doi.org/10.1109/FIE.2012.6462242

[19] M. Sivasakthi, "Classification and prediction based data mining algorithms to predict students' introductory programming performance," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 346-350. doi: 10.1109/ICICI.2017.8365371

[20] J. Xu, K. H. Moon and M. van der Schaar, "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs," in IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 5, pp. 742-753, Aug. 2017. doi: 10.1109/JSTSP.2017.2692560

[21] D. Baneres, M. E. Rodríguez-Gonzalez and M. Serra, "An Early Feedback Prediction System for Learners At-Risk Within a First-Year Higher Education Course," in IEEE Transactions on Learning

Technologies, vol. 12, no. 2, pp. 249-263, 1 April-June 2019. doi: 10.1109/TLT.2019.2912167

[22] A. Polyzou and G. Karypis, "Feature Extraction for Next-Term Prediction of Poor Student Performance," in IEEE Transactions on Learning Technologies, vol. 12, no. 2, pp. 237-248, 1 April-June 2019. doi: 10.1109/TLT.2019.2913358

[23] Mishra, Tripti & Kumar, Dharminder & Gupta, Sangeeta. (2014). Mining Students' Data for Prediction Performance. International Conference on Advanced Computing and Communication Technologies, ACCT. 255-262. 10.1109/ACCT.2014.105.

[24] M. Agaoglu, "Predicting Instructor Performance Using Data Mining Techniques in Higher Education," in IEEE Access, vol. 4, pp. 2379-2387, 2016.doi: 10.1109/ACCESS.2016.2568756

A. Elbadrawy, A. Polyzou, Z. Ren, M. Sweeney, G. Karypis and H. Rangwala, "Predicting Student Performance Using Personalized Analytics," in Computer, vol. 49, no. 4, pp. 61-69, Apr. 2016.doi: 10.1109/MC.2016.119

P. M. Arsad, N. Buniyamin and J. A. Manan, "Prediction of engineering students' academic performance using Artificial Neural Network and Linear Regression: A comparison," 2013 IEEE 5th Conference on Engineering Education (ICEED), Kuala Lumpur, 2013, pp. 43-48.doi: 10.1109/ICEED.2013.6908300

[25] P. A. Patil and R. V. Mane, "Prediction of Students Performance Using Frequent Pattern Tree," 2014 International Conference on Computational Intelligence and Communication Networks, Bhopal, 2014, pp. 1078-1082.

doi: 10.1109/CICN.2014.226

[26] C. L. Sa, D. H. b. Abang Ibrahim, E. Dahliana Hossain and M. bin Hossin, "Student performance analysis system (SPAS)," The 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M), Kuching, 2014, pp. 1-6.

doi: 10.1109/ICT4M.2014.7020662

A. Jamil, M. Ahsan, T. Farooq, A. Hussain and R. Ashraf, "Student Performance Prediction Using Algorithms of Data Mining," 2018 International Conference on Computing, Engineering, and Design (ICCED), Bangkok, Thailand, 2018, pp. 244-249.doi: 10.1109/ICCED.2018.00055

[27] I. A. A. Amra and A. Y. A. Maghari, "Students performance prediction using KNN and Naïve Bayesian," 2017 8th International Conference on Information Technology (ICIT), Amman, 2017, pp. 909-913. doi: 10.1109/ICITECH.2017.8079967

[28] Costa, Francisco & Bispo, Marcelo & Pereira, Rita. (2018). Dropout and retention of undergraduate students in management: a study at a Brazilian Federal University. RAUSP Management Journal. 53. 10.1016/j.rauspm.2017.12.007.

[29] Kabra R. R., Bichkar R.S. (2011). Performance Prediction of Engineering Students using Decision Trees, International Journal of Computer Applications (0975- 8887), Vol-36-No.11.

[30] Pandey, Mrinal & Sharma, Vivek. (2013). A Decision Tree Algorithm Pertaining to the Student Performance Analysis and Prediction. International Journal of Computer Applications. 61. 1-5. 10.5120/9985-4822.

[31] Liga Paura, Irina Arhipova, Cause Analysis of Students' Dropout Rate in Higher Education Study Program, Procedia - Social and Behavioral Sciences, Volume 109, 2014, Pages 1282-1286, ISSN 1877-0428, https://doi.org/10.1016/j.sbspro.2013.12.625.

[32] Pathak, P.S., Bansal, N., & Singh, S. (2015). Mulyankan: A prediction for student's performance using Neural Network. 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 46-49.

[33] Parneet Kaur, Manpreet Singh, Gurpreet Singh Josan, Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector, Procedia Computer Science, Volume 57, 2015, Pages 500-508, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2015.07.372.

[34] Gbolagade, M & Hambali, Moshood & Akinyemi, A. (2015). Predicting Postgraduate Performance Using Resample Preprocess Algorithm and Artificial Neural Network. 8. 145-158.

[35] Elgamal, Amany. (2013). An Educational Data Mining Model for Predicting Student Performance in Programming Course. International Journal of Computer Applications. 70. 22-28. 10.5120/12160-8163.

[36] Thai-Nghe, Nguyen & Drumond, Lucas & Horvath, Tomas & Krohn-Grimberghe, Artus & Nanopoulos, Alexandros & Schmidt-Thieme, Lars. (2011). Factorization Techniques for Predicting Student Performance. Educational Recommender Systems and Technologies: Practices and Challenges. 10.4018/978-1-61350-489-5.ch006.

[37] Osmanbegović, Edin & Suljic, Mirza. (2012). DATA MINING APPROACH FOR PREDICTING STUDENT PERFORMANCE. Journal of Economics & Business/Economic Review. 10. 3-12.

[38] Thakar, Pooja. (2015). Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue.

[39] Sales, A., Balby, L., & Cajueiro, A. (2016). Exploiting Academic Records for Predicting Student Drop Out: a case study in Brazilian higher education. J. Inf. Data Manag., 7, 166-180.

[40] Eddin, Mustafa & Khodeir, Nabila & Elnemr, Heba. (2018). A Comparative Study of Educational Data Mining Techniques for skill-based Predicting Student Performance. International Journal of Computer Science and Information Security, 16. 56.

[41] SEOW, Poh Sun and PAN, Gary. A Literature Review of the Impact of Extracurricular Activities Participation on Students' Academic Performance. (2014). Journal of Education for Business. 89, (7), 361-366. Research Collection School Of Accountancy.

[42] Singh Annu, Mishra Sunita (2013). Impact of Extracurricular Activities on Students in Private School of Lucknow District. (Student1 & Professor2 Department of Human Development and Family Studies, School for Home Sciences, Babasaheb Bhimrao Ambedkar University, Lucknow-226025, Uttar Pradesh, India).

[43] Ahmad, M., Rahman, M. F., Ali, M., Rahman, F., & Al Azad, M. A. (2019). Effect of Extra Curricular Activity on Student's Academic Performance. Journal of Armed Forces Medical College, Bangladesh, 11(2), 41-46. https://doi.org/10.3329/jafmc.v11i2.39822

[44] Vandana Bharadi, Madan Lal Saini, (2020) Review on Predictive Analytics for Student Performance Prediction. International Journal of Advanced Science and Technology, 29(5s), 2057 - 2067