

Application of Machine Learning Methods to Predict Student Performance: A Systematic Literature Review

Akpofure Avwersuoghene Enughwure¹, Mercy Ebitiminipre Ogbise²

¹Lecturer, Nigeria Maritime University, Okerenkoko, Delta State, Nigeria

²Nigeria Maritime University, Okerenkoko, Delta State, Nigeria

ABSTRACT: In recent times, the need for the application of machine learning in the educational frontier has become crucial. Most educational administrators and researchers are using various machine-learning methods to improve students' retention, predict students' performance, mitigate students' dropout rate as well as position them to make better decisions in curriculum design and admission policy in their learning citadel.

This study focused on the application of machine learning techniques to predict students' performance in recent times. With a systematic approach, the research identified the existing prediction methods and tools used to predict students' performance, observed the type of variables considered by the researchers in this research area. With the use of Boolean keyword searches in various journal databases and filters, eighty (80) research papers related to students' performance prediction using machine learning methods were selected and considered.

Weka was observed to be the most used tool in the reviewed articles. Most of the datasets used in the reviewed articles were on university students with the authors focusing on the performance of the students' full program. A great number of the authors applied at least one supervised learning methods in their work and reviewed articles from countries with low literacy rate. The results from the various researches will help to improve the academics of students in those countries and monitor the students' performance which would also improve their literacy rate.

Keywords: Machine Learning, Student Performance, Education

1.0 Introduction

In recent times, machine learning (ML) practices have been a big deal in various industries in the world including the educational frontier. The need to automate different tasks such as grading students, improving student retention, testing students, predicting student performance, as well as administrative tasks like material optimizations in the academic facet, has called for the application of machine learning techniques and methods. Arguably, the most important task in every learning institution is to monitor and improve their students' performance. Early prediction of student performance in the right fashion will improve student retention as well as the testing methods used for the students. This practice will also aid the educators and education policymakers by giving them better information about their students' learning ability as well as how best they can help students who are lagging in a given set.

This study focuses on reviewing previous research works on building models to predict student's performance in a learning environment in the last ten (10) years. The authors of this article developed a systematic approach to the review work. This approach is to support the objectives of this study, which are:

1. To identify the existing prediction methods and the tools employed for predicting students performance
2. To study and identify the variable type used for the predictive process.
3. To identify and study the researchers who employed these learning models to analyze students performance

In this paper, the study is structured as follows. In Section 2, we reviewed previous reviewed literature related to students' performance. Section 3, We performed the review process by using the systematic approached design for this course. Section 4, we presented our results as well as discussed the insights we gather from the results. Section 5, concludes the review as well as state our limitations and future study in this research area.

2.0 Literature Review

The need to predicting student academic performance has become a critical factor in improving the quality of academic curriculum; assist the students while they study as well as providing the tutors more options when training their students. In recent times, there are many works been published related to this subject matter. We found several literature reviews that examined student academic performance modelling in different views. The review of Kavipriya focuses on the students' performance prediction, analysis, early alert, and evaluation by using different methods of data mining. He reviewed different classifications methods like decision tree, naïve Bayes algorithms etc. He suggested since predicting student's performance is difficult due to a lot of challenges like data imbalances, there is a need to employ support vector machine that offer the best accuracy in his study [1]. Shahiria A.M. et. al. took a lot at the various data mining techniques implemented in different student's performance prediction models. They evaluated how these predicting algorithms can be used to identify the most important attributes in students' database [2]. In this work, they pointed out that cumulative grade point average (CGPA) and internal assessment as the attributes most researchers used to build their model. A review on predictive modelling technique for student academic performance monitoring by A. Mat et. al. focuses on various learning tasks employed when building the predictive models as well as the application of these models [3]. After accessing several researches, they noted that these models can be applied in various fronts like course recommendations, and career path planning. Mishra T. et.al surveyed Students' performance and employability prediction through Data mining. In this work, they focused mainly on traditional educational set-up [4]. Their findings were most predictors used in traditional educational set-up were demographic and socioeconomic factors. They also mentioned that a few research works have been done on employability prediction. Zaffar M. et. al studied different prediction models for students who enrolled for programming subjects only [5]. Their results were only a few researchers used Filter Feature selection algorithms in the pre-processing stage. The attendance of students, their performance in mathematics, and physics play an important role in influencing their performance.

3.0 Method

3.1 Research questions

The main goal of this systematic literature review is to explore the applications of machine learning methods to educational data to predict/monitor/intervene in student academic performance. In a bid to reach this above goal, we have stated the following main research questions:

RQ1. What machine learning methods and tools were applied to predict student performance?

RQ2. Who are the authors, journals, and the variable type used in this research area?

RQ3. What are the courses/subjects used to access the students' performance?

To answer the research questions, the required information was retrieved from the chosen literature by both reviewing the contents of the articles and by utilizing the meta-data properties of the article's citation. The authors of this article have classified the chosen articles concerning the machine learning algorithms used, the input data type used, journal article type, stakeholders involved in the studies.

3.2 Literature Collection Process

We follow a standard systematic literature review methodology that involves the use of specific queries search built in a Boolean format on various selected bibliographical database and a set of rules for filtering the result of the queries. The authors of the paper queried seven (7) databases that are related to computer, education, engineering, and general science. To make certain only relevant literature are captured, distinctive search keywords were introduced in this process. Logical operators like "+" and "&" were used to combine these distinctive keywords during the search. The combination of these keywords and logical operators resulted in Boolean search keywords like "Machine Learning" + "Student", "Predicting Student Performance" which were used to perform the searches on Google Scholar and Science Direct. Date filters were used to filter search results to exclude articles that were published earlier than "2010".

3.2.1 Location of affiliations

In an attempt to see how scientific contributions have been made towards this front, we classified the research articles according to the location of the corresponding authors. This provided insights on the global publication developments in this research area.

3.2.2 Type of journal source and dataset

The authors of this study attempted to see which journal database published articles within the research area as well as the dataset used in the published work. Given this is an educational research area, did the scholars worked with undergraduate or high school students?

3.2.3 Type of tools/techniques used

The authors of this study examined the tools and techniques employed for this prediction modelling. Currently in the analytics facet, several tools are available to perform various analyses. Tools like WEKA, SPSS, MATLAB, Orange, R, Python are used for data mining, machine learning as well as model deployment.

The machine learning techniques used in the reviewed studies can be categorized into three groups.

- Supervised Algorithms – Regression (Linear and Logistic), Decision Trees (C4.5, ID3, CHAID), support vector machines, naïve Bayes
- Unsupervised Algorithms – Clustering, Correlation
- Semi-Supervised Algorithms- Association Rules and the combination of any supervised and unsupervised algorithms.

4.0 Result and Discussion

This section displays the discoveries from the literature review process that was explained in the earlier sections. The dataset consisted of 98 electronic materials (research articles, review articles, thesis as well as e-books on machine learning and predictive models). The authors of this paper only considered research and review articles on applying machine learning methods to predict academic performance. The selected literature collection consists of 80 articles. The line chart below displayed the reviewed articles in this study by their year of publication.

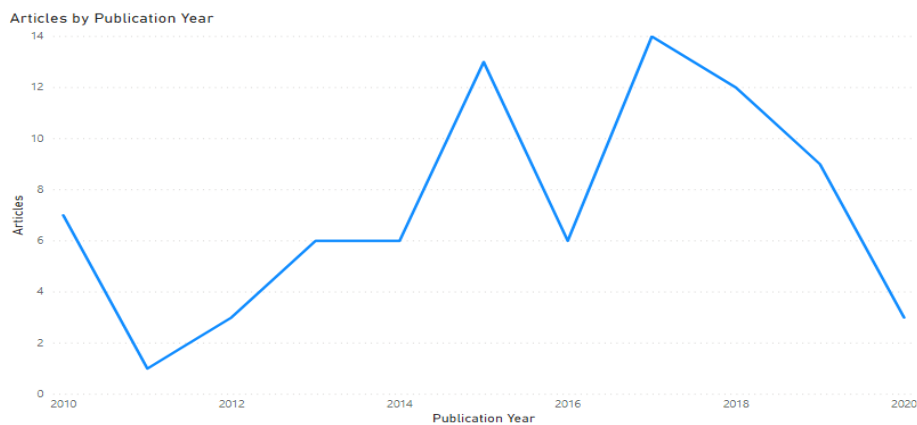


Figure 1: A Pie Chart displaying the distribution of reviewed articles by the journal database.

Figure 2 presents the distributions of the selected articles from different journal databases for this study. It shows that the research gate leads the collection with 24 articles followed by Science Direct, IEEE Xplore Digital Library, and Google Scholar with 20, 16, and 11 respectively. Semantic Scholar, ACM Digital Library, Hindawi, and Springer were sources for eight (8) articles. One article journal source is unknown [23].

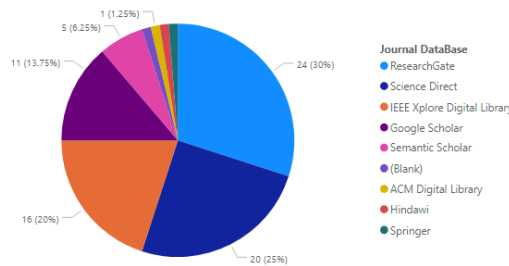


Figure 2: A Pie Chart displaying the distribution of reviewed articles by the journal database.

Most of the data used for predicting students' performance in the collection of articles reviewed in this study were University Data. Seventy-one (71) articles used university data with sixty-seven (67) focused on undergraduate students, three (3) articles were on both undergraduate and postgraduate while only one on postgraduate students [24]. Nine (9) articles focused on high school students. This illustrated in figure 3.

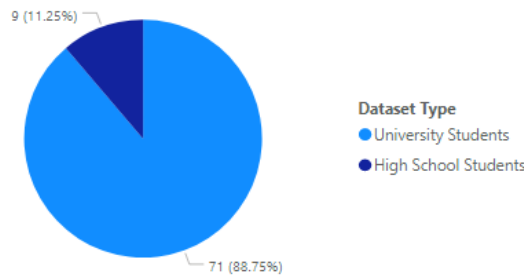


Figure 3: A pie chart of the distribution of dataset type in the reviewed articles.

Out of the selected articles, 49 articles used at least one tool with WEKA as the most-used analytical tool in the collection of reviewed articles. The other 31 articles did not clearly state the tools they employed for their analysis. SPSS, RapidMiner, Matlab were also used in more than three research papers. Figure 4 shows the list of all the tools used with its frequency of use.

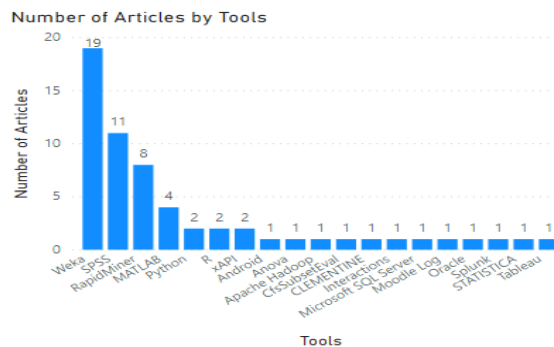


Figure 4: A bar chart displaying the number of articles by tool used

The researchers of the selected papers employed at least one machine learning method in their respective works. There are 61 selected works applied supervised machine-learning algorithms. Seventeen (17) studies applied semi-supervised (a combination of supervised and unsupervised algorithms) and two unsupervised as presented in table 1.

Table 1: A Presentation of the Reviewed Articles by machine learning method used

S/N	Machine Learning Method Type	Number of Articles
1	Supervised	61
2	Semi-Supervised	17
3	Unsupervised	2

In an attempt to improve the predictive model, some researchers implemented k-fold cross-validation (CV) in their model. K-fold CV occurs when a dataset is randomly divided into a K number of subsets where each subset is used as a testing data and the remaining subsets are used as the training set. This process is done for K times [12] [14] [18]. The final result is estimated as the average error rate on test examples [21]. Twenty-six (26) articles within the selected studies applied cross-validation to their model at some point [22]. Most of them used 10-fold cross-validation method [13][15][16][17]. C. Lye et.al introduced the cross-validation concept into their study to evaluate the best predictive models in terms of accuracy [19][20]. The Figure below as well as table 2 shows the distribution of the selected papers by their use of cross-validation in their research work.

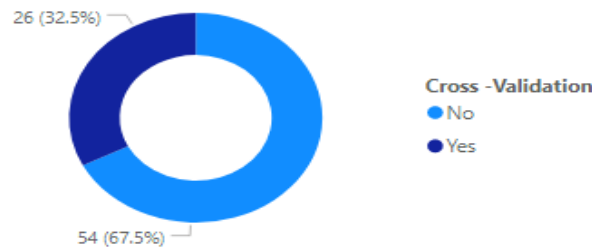


Figure 5: A Pie Chart showing the distribution of the reviewed articles by the application of cross-validation methods.

Table 2: A Presentation of the Reviewed Articles by the use of cross validation techniques

S/N	The use of Cross Validation	Reference Numbers
1	Yes	[8][12][13][14][15][16][17][20][21][22][28][41][46][47][52][59][60][65] [67][72][77][79][82][83]
2	No	[6][7][9][10][11][18][19][23][24][25][26][27][29][30][31][32][33][34][35][36][37] [38][39][40][42][43][44][45][48][49][50][51][53][54][57][58][61][62] [63][64][66][68][69][70][71][73][74][75][76][77][78][80][81][84][85]

Looking at the authors of the collected papers in this study, there are some insights the authors of this paper unravel. Using the nationality of the corresponding author and the source of the dataset used to build the models, thirteen (13) of the studies

status [58], number of vehicles [35], student’s ability to make money , . Other researchers measured students’ performance with the teacher’s contribution, qualification as well as competence to the course work [70] [16] [77]. The figure below shows the distributions of the reviewed articles in this study by the variable type used in their analysis (Student learning factor, social-economic factor, student's social life factor, and educators inputs).

Legend ● Social Economic Factors ● Social Life ● Educators Input ● Student's Learning Factor



Figure 8: A Presentation for the reviewed article showing their distribution among the variable type considered in this study.

Conclusion

The art of predicting student performance is highly beneficial to everyone especially to the educational administrators and students. The administrators can evaluate their teaching procedures to ensure their students get the best-fit learning plan from them. Students can also appraise their learning patterns to know the tailor-made learning activities that will meet their needs. The authors of this paper have accessed previous studies on predicting students’ performance using machine learning methods. Most of the authors of the reviewed articles are from countries with low literacy rate, hence, this will help academics in those countries to monitor their student performance thereby improving their literacy rate. Weka is the most-used analytical tool used by the researchers when predicting students’ performance. When considering the machine learning methods used, the supervised learning method was mostly employed in this research space. Most researchers attempted to predict the performance of students using the students’ academic data, social-economic data, and educator competence data as well. This approach proves that a student academic performance does not only dependent on his academic input. A full student’s academic program is used to predict the performance of students in most of the reviewed work.

References

1. 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 ISSN: 2277 128X.
2. A.M. Shahiri, W. Husain, and N.A. Rashid," A Review on Predicting Student's Performance using Data Mining Techniques", Proceeding Computer Science, vol. 72, (2015), pp. 414-422.
3. Asiah, M., et al.: A review on predictive modeling technique for student academic performance monitoring. In: EDP Sciences (eds.) MATEC Web of Conferences 2019, EAAIC, vol. 255, p. 03004. EDP Sciences (2019)
4. Mishra, T., Kumar, D. and Gupta, S., “Students’ Performance and Employability Prediction through Data Mining: A Survey”, Indian Journal of Science and Technology, Vol.10(24), (2017)

5. Maryam Zaffar, Manzoor Ahmed Hashmani, and KS Savita. 2018. A Study of Prediction Models for Students Enrolled in Programming Subjects. In 2018 4th International Conference on Computer and Information Sciences (ICCOINS). IEEE, 1—5
6. D. Isha Shetty, Dipshi Shetty and Sneha Roundhal, "Student Performance Prediction" International Journal of Computer Applications Technology and Research Vol 8, Issue 5, pp157-160, (2019)
7. K. Kaur and K. Kaur, "Analyzing the Effect of Difficulty Level of a Course on Students Performance Prediction using Data Mining", International Conference on Next Generation Computing Technologies (2015) doi.org/10.1109/NGCT.2015.7375222
8. R. Mohd Hanis, E. Abdullah, "Predicting Student Performance in Object-Oriented Programming Using Decision Tree: A Case at Kolej Poly-Tech Mara Kuantan", *International Conference on Software Engineering & Computer Systems*, pp. 11-8, 2013
9. Zafar Iqbal, Junaid Qadir, Adnan Noor Mian and Faisal Kamiran, "Machine Learning-Based Student Grade Prediction: A Case Study", 2017
10. GOWRI, G.Sh., THULASIRAM, R., and BABURAO, M. (2017) Educational Data Mining Application for Estimating Students Performance in Weka Environment. 14th ICSET-2017. IOP Conference Series: Materials Science and Engineering, 263(3), pp. 1-9. DOI:10.1088/1757-899X/263/3/032002
11. S. E. Sorour, T. Mine, K. Goda, S. Hirokawa, "A predictive model to evaluate student performance", *Journal of Information Processing*, vol. 23, no. 2, pp. 192-201, 2015.
12. S. M. M. Rubiano, J. A. D. Garcia, (2015), "Formulation of a predictive model for academic performance based on students' academic and demographic data", IEEE Frontiers in Education Conference, 2015
13. N. Z. Zacharis, "A multivariate approach to predicting student outcomes in web-enabled blended learning courses", *Internet High. Educ.*, vol. 27, pp. 44-53, Oct. 2015.
14. Z.J. Kovacic (2010)," Early Prediction of Student Success: Mining Students Enrolment Data", Paper presented at Proceedings of Informing Science & IT Education Conference (InSITE), Casinio Italia, June 19-24, 2010
15. C. Márquez-Vera, C.R. Morales, S.V. Soto, "Predicting School Failure and Dropout by Using Data Mining Techniques", *IEEE Journal of Latin-American learning technologies*, vol. 8, no. 1, pp. 7-14, February 2013.
16. Hayes, D., Bernacki, M., Hong, W., Markle, J., and Voorhees, N., 2017. Using LMS Data to Provide Early Alerts to Struggling Students. In: First Year Engineering Experience (FYEE) Conference. Daytona Beach, FL: American Society for Engineering Education.
17. Xiaofeng Ma and Zhurong Zhou. 2018. Student pass rates prediction using optimized support vector machine and decision tree. In Computing and Communication Workshop and Conference (CCWC), 2018 IEEE 8th Annual. IEEE, 209–215.
18. 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.
19. C.-T. Lye, L.-N. Ng, M. D. Hassan, W.-W. Goh, C.-Y. Law, & N. Ismail, Predicting Pre-university Student's Mathematics Achievement, *Procedia - Social and Behavioral Sciences*, 8 (2010), 299–306. <http://dx.doi.org/10.1016/j.sbspro.2010.12.041>
20. Hajra Waheed, Saeed-Ul Hassan, Naif Radi Aljohani, Julie Hardman, Raheel Nawaz, "Predicting Academic Performance of Students from VLE Big Data using Deep Learning Models", *Computers in Human Behavior*, pp. 106189, 2019 <https://doi.org/10.1016/j.chb.2019.106189>.
21. E. A. Amrieh, T. Hamtini and I. Aljarah, "Mining Educational Data to Predict Student's academic Performance using Ensemble Methods" International Journal of Database Theory and Application, vol. 9 No. 8 (2016), pp 119-136
22. E.N. Azizah, U. Pujianto, E. Nugraha and Darusalem, "Comparative performance between C4.5 and Naïve Bayes classifiers in predicting student academic performance in a Virtual Learning Environment" International Conference on Education and Technology (ICET) (2018) <https://doi.org/10.1109/ICEAT.2018.8693928>
23. F. T. Matthew, O. Olumide, A. I. Adepoju, O. Olatayo, E. Adebimpe, O. Bolaji, E. Funmilola, O. Adeyele, "Development of Mobile-Interfaced Machine Learning-Based Predictive Models for Improving Students' Performance in Programming Courses" International Journal of Advanced Computer Science and Applications, Vol. 9 No. 5, 2018
24. V. K. Pal, V. K. K. Bhatt, "Performance Prediction for Post Graduate Students using Artificial Networks" International Journal of Innovative Technology and Exploring Engineering (IJITEE) Vol. 8, Issue 7S2, 2019
25. M. Ramaswami and R. Bhaskaran, "A CHAID Based Performance Prediction Model in Educational Data Mining" International Journal of Computer Science Issues 2010
26. M. N. Quadri, N.V. Kalyankar, "Drop Out Feature of Student Data for Academic Performance Using Decision Tree Techniques" Global Journal of Computer Science and Technology 2010

27. Anal Acharya, Devadatta Sinha, "Early Prediction of Students Performance using Machine Learning Techniques" *International Journal of Computer Applications* 2014
28. S.K Yadav, S. Pal, "Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification" *World of Computer Science and Information Technology Journal (WCSIT)* 2012
29. Mukesh Kumar, A. J. Singh, "Performance Analysis of Students Using Machine Learning & Data Mining Approach" *International Journal of Engineering and Advanced Technology* 2019 Vol 8, Issue 3, (2019)
30. Raju Ranjan, Jayanthi Ranjan, "Predicting Parameters Importance in Higher Education Using Data Analysis: A Regression Model" *International Conference on Computational Intelligence & Communication Technology* 2016
31. A.Dinesh Kumar, R.Pandi Selvam, V.Palanisamy, "Prediction of Student Performance using Hybrid Classification" *International Journal of Recent Technology and Engineering* 2019
32. H. Hamsa, S. Indiradevi, J. J. Kizhakkethottam, "Student academic performance prediction model using decision tree and fuzzy genetic algorithm" *1st Global Colloquium on Recent Advancements and Effectual Researches in Engineering, Science and Technology* pp 326-332, (2016). DOI: 10.1016/j.protcy.2016.08.114
33. H. Agrawal, H. Mavani, "Student Performance Prediction using Machine Learning" *International Journal of Engineering Research & Technology* Vol 4, Issue 3 (2015)
34. Sana, Siddiqui, I. F. & Arain, Q. A. (2019). Analyzing Students' Academic Performance through Educational Data Mining. *3C Tecnología. Glosas de innovación aplicadas a la pyme. Special Issue, May 2019*, pp. 402-421.
35. A. Daud, N. R. Alijohani, R.A. Abbasi, M. D. Lytras, F. Abbas, J. S. Alowibdi, "Predicting Student Performance using Advanced Learning Analytics" *International World Wide Web Conference Committee* 2017
36. A. Ali, U. Ali, "Predictability of engineering students' performance at the University of Engineering and Technology, Peshawar from admission test conducted by educational testing and evaluation agency (ETEA), NWFP, Pakistan" *World Conference on Educational Sciences* 2010 pp 976-982 DOI:10.1016/j.sbspro.2010.03.137
37. M. Wasif, H. Waheed, N. R. Aljohani, S. Hassan, "Understanding Student Learning Behavior and Predicting Their Performance" In *Cognitive Computing in Technology-Enhanced Learning* 2019
38. S. Hassan, H. Waheed, N. R. Aljohani, M. Ali, "Virtual Learning Environment to Predict Withdrawal by Leveraging Deep Learning" *International Journal of Intelligent Systems* 2019 doi.org/10.1002/int.22129
39. R. Asif, A. Merceron, S. A. Ali, N. G. Haider, "Analyzing undergraduate students' performance using educational data mining" *Computers & Education*, 2017 <https://doi.org/10.1016/j.compedu.2017.05.007>
40. R. Asif, A. Merceron, M. K. Pathan, "Predicting Student Academic Performance at Degree Level: A Case Study" *International Journal of Intelligent Systems and Applications* DOI: 10.5815/ijisa.2015.01.05 pp 49-61 (2015)
41. Affendey, L., Paris, I., Mustapha, N., Sulaiman, M., & Muda, Z. (2010). Ranking of influencing factors in predicting students' academic performance. *Information Technology Journal*, 9(4), 832-837.
42. M. Smith, L. Therry, J. Whale "Developing a Model for Identifying Students at Risk of Failure in a First-Year Accounting Unit" *Higher Education Studies*, vol,2 No 4, (2012), pp 10.5539/hes.v2n4p91
43. M. Mgala, A. Mbogho "Data-Driven Intervention-Level Prediction Modeling for Academic Performance" *International Conference on Information and Communication Technologies and Development* (2015)
44. K. A. Capao, A. D. Cantara, A. M. Ceniza, P. M. J. Eduardo, S. B. Polinar, and J. M. Tero, "Predicting Academic Performance with Intelligence, Study Habits and Motivation Factors using Naive Bayes Algorithm," *International Journal of Engineering Research & Technology*, Vol. 5, No. 3, pp. 182-185, 2016.
45. P. Shayan, M. van Zaanen "Predicting Student Performance from Their Behavior in Learning Management Systems" *International Journal of Information and Education Technology*, vol 9, Issue 5 (2019)
46. R. Conijn, C. Snijders, Ad Kleingeld, U. Matzat "Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS" *IEEE Transactions On Learning Technologies* (2017). DOI: 10.1109/TLT.2016.2616312
47. Ya-Han Hu, Chia-Lun Lo, Sheng-Pao Shih "Developing early warning systems to predict students' online learning performance" *Computers in Human Behaviour*, vol 36, pp 469 - 478 (2014)
48. J. Schell, B. Lukoff, C. Alvarado "Using Early Warning Signs to Predict Academic Risk in Interactive, Blended Teaching Environments" *Internet Learning*, vol 1, Issue 2 (2014), pp 249-263. DOI: <https://doi.org/10.1109/TLT.2019.2912167>
49. Ji Won You "Identifying significant indicators using LMS data to predict course achievement in online learning" *Internet and Higher Education*, vol 29, pp 23-30, (2016)
50. Y. Altujjar, W. Altamimi, I. Al-Turaiki, M. Al-Razgan "Predicting Critical Courses Affecting Students Performance: A Case Study" *Symposium on Data Mining Applications*, vol 82, pp 65-71 (2016) DOI: 10.1016/j.procs.2016.04.010

51. Alana M. de Morais, Joseana M. F. R. Araújo, Evandro B. Costa "Monitoring Student Performance Using Data Clustering and Predictive Modelling" IEEE Frontiers in Education Conference"2014)
52. Camilo E. López G., E. L. Guzmán, Fabio A. González "A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining" IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, vol 10, Issue 3 (2015)
53. O. Adejo, T. Connolly "An Integrated System Framework for Predicting Students' Academic Performance In Higher Educational Institutions" International Journal of Computer Science & Information Technology,(2017)
54. F. T. Matthew, A. I. Adepoju, O. Olatayo, O. Ayodele "Development of Multi-Factored Predictive Models for Students' Performance in Programming Courses in a Nigerian Tertiary Institution" Transylvanian Review, vol XXVI No 25(2018)
55. Oyerinde O. D, Chia P. A "Predicting Students' Academic Performances – A Learning Analytics Approach using Multiple Linear Regression" International Journal of Computer Applications, vol 157, Issue 4 (2017). DOI: 10.5120/ijca2017912671
56. K. D. Koloa, S. A. Adepojob, J. K. Alhassan "A Decision Tree Approach for Predicting Students Academic Performance" International Journal of Education and Management Engineering, vol 5, (2015). DOI: 10.5815/ijeme.2015.05.02
57. M. Nasiri, B. Minaei, F. Vafaei "Predicting GPA and Academic Dismissal in LMS Using Educational Data Mining: A Case Mining" 6th National and 3rd International Conference of E-Learning and E-Teaching,(2012), <https://doi.org/10.1109/ICELET.2012.6333365>
58. E. P. I. García, P. M. Mora "Model Prediction of Academic Performance for First-Year Students" Mexican International Conference on Artificial Intelligence, (2011). DOI 10.1109/MICAL.2011.28
59. G. Gray, C. McGuinness, P. Owende "An application of classification models to predict learner progression in tertiary education" IEEE International Advance Computing Conference (2014).
60. S. Ikbal, A. Tamhane, B. Sengupta, M. Chetlur, S. Ghosh, J. Appleton "On early prediction of risks in academic performance for students" IBM Journal of Research and Development, vol 59, Issue 6 (2015), <https://doi.org/10.1147/JRD.2015.2458631>
61. Á. F. Agudo-Peregrina, S. Iglesias-Pradas, M. Ángel Conde-González, Á. Hernández-García "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning" Computers in Human Behaviour, (2014) <http://dx.doi.org/10.1016/j.chb.2013.05.031>
62. C. K. Hoe, M. S. Ahmad, T. C. Hooi, M. Shanmugam, S. S. Gunasekaran, Z. C. Cob, A. Ramasamy "Analyzing students records to identify patterns of students' performance" International Conference on Research and Innovation in Information Systems, (2013) [virhttps://doi.org/10.1109/ICRIIS.2013.6716767](https://doi.org/10.1109/ICRIIS.2013.6716767)
63. S. Rizvia, B. Rienties, S. A. Khoja "The role of demographics in online learning; A decision tree-based approach" Computers & Education (2019). DOI: <https://doi.org/10.1016/j.compedu.2019.04.001>
64. M. Hussain, W. Zhu, Wu Zhang, "Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores" Computational Intelligence and Neuroscience, vol 2018, <https://doi.org/10.1155/2018/6347186>
65. H. Liu, Z. Wang, P. Benachour, P. Tubman "A Time Series Classification Method for Behaviour-Based Dropout Prediction" International Conference on Advanced Learning Technologies" (2018), DOI 10.1109/ICALT.2018.00052
66. M. Hlosta, Z. Zdrahal, J. Zendulka "Ouroboros: Early identification of at-risk students without models based on legacy data" International Learning Analytics & Knowledge Conference, (2017), pp 6-15 <http://dx.doi.org/10.1145/3027385.3027449>
67. H. Heuer, A. Breiter "Student Success Prediction and the Trade-Off between Big Data and Data Minimization" E-Learning Fachtagung Informatik, (2018)
68. S. Palmer "Modelling engineering student academic performance using academic analytics" International Journal of Engineering Education, International Journal of Engineering Education, 29(1):132-138, 2013
69. F. Okubo, T. Yamashita, A. Shimada, H. Ogata "A Neural Network Approach for Students' Performance Prediction" Proceedings of the Seventh International Learning Analytics & Knowledge Conference, (2017). DOI: <https://doi.org/10.1145/3027385.3029479>
70. U. Ashfaq, B. P. M., R. Mafas "Managing Student Performance: A Predictive Analytics using Imbalanced Data" International Journal of Recent Technology and Engineering, vol 17, Issue 3 (2020). DOI: 10.1186/s41239-020-0177-7
71. M. Tan, P. Shao, "Prediction of Student Dropout in E-Learning Program Through the Use of Machine Learning Method" International Journal of Emerging Technologies in Learning, (2015),<http://dx.doi.org/10.3991/ijet.v10i1.4189>

72. S. Helal, Jiuyong Li, L. Liu, E. Ebrahimie, S. Dawson, D. J. Murray, Qi Long, "Predicting academic performance by considering student heterogeneity" Knowledge-Based Systems, (2018)
73. S. Huang, N. Fang "Regression Models of Predicting Student Academic Performance in an Engineering Dynamics Course" American Society for Engineering Education(2010)
74. T. M. S. Ahamed, N. T. Mahmood, R. M Rahman "An intelligent system to predict academic performance-based on different factors during adolescence" Journal of Information and Telecommunication, (2017), <https://doi.org/10.1080/24751839.2017.1323488>
75. W. M. Aly, O. F. Hegazy, H. M. N. Rashad "Automated Student Advisory using Machine Learning" International Journal of Computer Applications, vol 81, No. 19 (2013)
76. M. Adewale, A. O. Bamidele, U. O. Lateef "Predictive modelling and analysis of the academic performance of secondary school students: Artificial Neural Network approach" International Journal of Science and Technology, vol "9", Issue 1 (2018), pp 1-8
77. Y. K. Saheed, T. O. Oladele, A. O. Akanni, W. M. Ibrahim, " Student Performance Prediction Based On Data Mining Classification Techniques" Nigerian Journal of Technology, vol 37, No. 4 (2018), pp 1087-1091, <http://dx.doi.org/10.4314/njt.v37i4.31>
78. R. Asif, S. Hina, S. I. Haque "Predicting Student Academic Performance using Data Mining Methods" International Journal of Computer Science and Network Security, vol 17, No. 5 (2017).
79. S. Bergin, A. Mooney, J. Ghent, K. Quille "Using Machine Learning Techniques to Predict Introductory Programming Performance" International Journal of Computer Science and Software Engineering, vol 4, Issue 12 (2015), pp 323-328
80. N. A. Yassein, R. G. M. Helali, S. B. Mohomad "Predicting Student Academic Performance in KSA using Data Mining Techniques" Journal of Information Technology & Software Engineering, vol 7, Issue 5, 10.4172/2165-7866.1000213, (2017)
81. K. Hamoud, A. S. Hashim, W. A. Awadh "Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis" International Journal of Interactive Multimedia and Artificial Intelligence, vol 5, No. 2 (2018)
82. E. B. Belachew, F. A. Gobena "Student Performance Prediction Model using Machine Learning Approach: The Case of Wolkite University" International Journal of Advanced Research in Computer Science and Software Engineering (2017). DOI: 10.23956/ijarcsse/V7I2/01219
83. M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, S. Ali "Using machine learning to predict student difficulties from learning session data" Artificial Intelligence Review,(2018)
84. Literacy Rate by Country Population. (2020-02-17). Retrieved 2020-04-20, from <http://worldpopulationreview.com/countries/literacy-rate-by-country/>
85. K. Alaslani and M. Alandejani, "Identifying factors that influence students performance through social networking sites: An exploring case study" Heliyon vol 6 (2020) <https://doi.org/10.1016/j.heliyon.2020.e03686>
86. Chaffai, Abdelmajid & el Abdouli, Abdeljalil & Anoun, Houda & Hassouni, Larbi & Khalid, Rifi. (2015). Student Performance Analysis using Large-Scale Machine Learning. International Journal of Computer Science and Information Security. Vol 13, No.12