Comparative Analysis of Various Techniques used for Predicting Student's Performance

Amita Dhankhar[a], Kamna Solanki[b]

[a] University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India
[b] University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India

Abstract

Digitization is transforming all aspects of education. Learner's interactions with their online and offline learning environment lead to a trail of data that can be used for the purpose of analysis. Learning analytics (LA) and Educational data mining and (EDM) are emerging fields that attempt to develop methods to confront an abundance of data from the educational domain in order to optimize learning and leveraging decisions related to learning, teaching, and educational management. EDM/LA techniques interpret such enormous data and turn it into useful action. It provides insight to teachers to improve teaching, to understand learners, to identify difficulties faced by learners, and to provide meaningful feedback to learners thereby improving the learner's performance. This paper aims to compare different EDM/LA techniques and to identify their potential strength and weaknesses that are applied in the educational domain to predict the student's performance.

Keywords 1

Educational data mining, learning analytics, machine learning, supervised learning, unsupervised learning.

1. Introduction

Technology is evolving rapidly [1]. This technological advancement leads to the generation of tremendous amounts of data and it becomes an integral part of all sectors [2]. The educational sector is no exception. Big data in the field of the education sector provides unprecedented opportunities for teachers and educational institutes. The exploration and analysis of an enormous amount of data so that significant patterns can be discovered is called Data mining (DM). It can also be defined as "a non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns from data" [3]. The DM techniques when applied to the data gathered from the educational domain to extract knowledge is called Educational data mining [4]. One of the significant areas of interest for researchers in EDM is the prediction of student's performance. Timely predicting student's performance helps in identifying poorly performing students thereby helping teachers to provide early intervene. EDM/LA techniques like classification, clustering, association analysis, prediction are used to transform raw data into significant information. Computational advancements in data mining and learning analytics have helped this effort significantly [5]. Considering the importance of various techniques for predicting student's performance detailed comparative analysis of these techniques would be valuable. The sections that follow are listed as methodology is described in Section-2; Results are summarized in section 3; the conclusion is summarized in section 4.

2. Methodology

This paper performed a comparative analysis of various techniques used for predicting student performance.

WTEK-2021: Workshop on Technological Innovations in Education and Knowledge Dissemination, May 01, 2021, Chennai, India.
EMAIL: amita.infotech@gmail.com (Amita Dhankhar)
ORCID: 0000-0002-9305-4088 (Amita Dhankhar)

For this purpose, relevant articles were identified, selected, evaluated critically using several criteria, and then finding were integrated. Few Research questions were formulated to streamline our contribution, which are:

RQ-1 What EDM/LA techniques are used for predicting student performance?
RQ-2 Comparative analysis of various techniques on the different facet that includes their strength, weaknesses, and accuracy.

To assess and address the above-mentioned Research Questions, we have adopted the PICO model [6] that consists of 4 key components namely population, intervention, comparison, and outcomes. Details of the PICO components of this paper are given in the Table 1. We have searched three databases namely Scopus, IEEE, and Science Direct for the articles published from 2016 to 2020.

Population	Articles predicting student's performance
Intervention	EDM/LA techniques
Comparison	Comparative analysis of EDM/LA techniques
outcomes	Effectiveness, the accuracy of the techniques

The search string used for the search is
(Prediction **OR** forecast **OR** predict) **AND** (techniques **OR** methods **OR** framework) **AND** (student's performance **OR** retention **OR** at-risk) **AND** (Engineering **OR** Higher education) **AND** (data mining **OR** machine learning **OR** Learning analytics)

To obtain relevant results, the syntax of the string was modified slightly for each database. The articles identified through database searching were evaluated using inclusion and exclusion criteria. Inclusion criteria included articles that explicitly predict student's performance/predictive models/techniques/methods, considered only journal articles, full text is available for analysis, focus on empirical studies, articles in the domain of higher education. Articles not written in English, conference articles, full text not available were excluded.

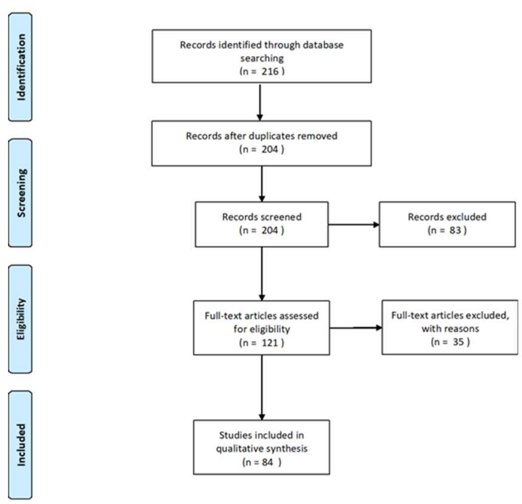

Figure 1: PRISMA flow chart of methodology [7].

3. Results

In this section, we describe the details of the reviewed articles, EDM/LA techniques used for predicting student's performance, and comparative analysis of various techniques on the different facets that include their strength, weaknesses, and accuracy. Regression and Classification techniques are the most commonly used techniques in educational data mining and learning analytics. It is the supervised learning method that analyzes a set of data and classifies data into a different predefined set of classes. In the context of higher education, this approach has been used to determine or predict student's success or failure by identifying the patterns from the student's learning activities with online learning resources. Classification techniques can be used to predict student's performance, to predict students at-risk or retention [8-10], students dropout prediction [11,12], predict student's achievement [13], predict which students would likely submit their assignments [14], assessing student's engagement during the course [15]. In this section, we have discussed various techniques used for predicting student's performance.

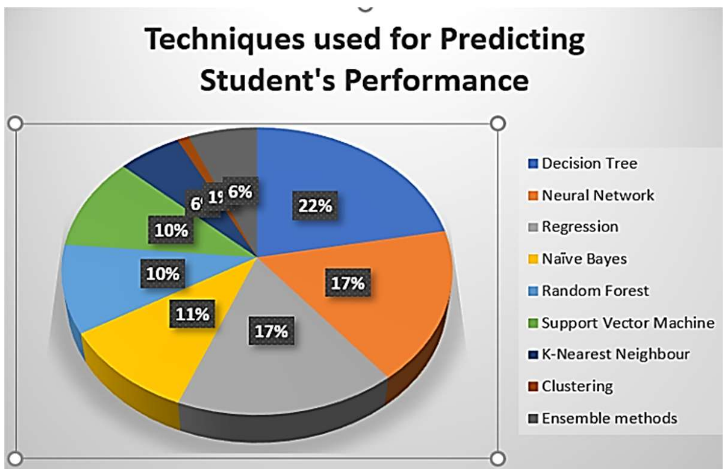

Figure 2: Distribution of techniques used for Predicting students' performance

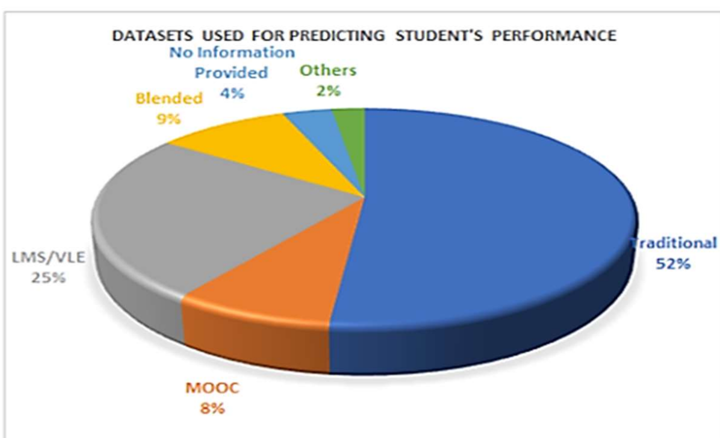

Figure 3: Distribution of datasets used for predicting student's performance

3.1. k-NN

K- nearest neighbor is supervised machine learning algorithm. It is the simplest yet powerful technique that can be used for both classification and regression predictive problems. The basic concept of KNN is to classify the test data in a given dataset by using feature similarity. It calculates the distance (closeness or proximity) between the test data and each training data in the dataset. Then it performs the majority voting and classifies the test data by the majority votes of neighbor classes. The distance can be calculated by using various distance functions like Euclidean, Cosine, Chi-square, Minkowsky, etc [38-42].

3.2. Naive Bayes

Naive Bayes is a classification algorithm that assumes that the predictor variables are independent of each other. The base of the naive Bayes is the Baye's theorem which is derived from the conditional probability. Bayesian theorem gives an equation for computing posterior probability P1(c1|x1) from P1(c1), P1(x1), and P1(x1|c1).

$$p1(c1|x1) = \frac{p1(x1|c1)p1(c1)}{P1(c1)}$$

P1(c1|x1): the posterior probability of type (c, target) provided predictor (x, attributes), P1(c1): the previous probability of a class, P1(x1|c1): the perspective, which is the probability of predictor given class, P1(1x): the previous probability of predictor. It classifies the test data by computing conditional probability with feature vectors x1, x2…., xn which belong to particular class Ci. Naive Bayes algorithms can be applied in recommendation system spam filtering, sentiment analysis [43-48].

3.3. Logistic Regression

LR is a statistical method that can be used for binary classification problems. It assumes that classes are almost linearly separable. It uses a logistic function also called the sigmoid function which is used to map predicted values to probabilities. It utilizes a logit function for predicting the probability of occurrences of a binary event [49-53].

3.4. Linear Regression

It is a supervised learning process. It finds the function which predicts for given X predicts Y where Y is continuous.

F(X)→ Y

Many types of functions can be used. The simplest type of function is a linear function. X can comprise a single feature or multiple features. The basic concept of linear regression is to find a line that best fits data. The best fit line means the total prediction error for all data points is as small as possible. The error is the distance between the point to the regression line [54-58].

3.5. Support Vector Machine

It is a very popular machine learning technique. It can be used to perform both classification and regression. The core idea of SVM is that it tries to find out a hyperplane that separates two classes as widely as possible. In other words, it finds the hyperplane that maximizes the margin. As margin increases the generalization accuracy increases. The points through which the hyperplane passes are called support vectors. The variations to SVM are linear SVM, Polynomial kernel SVM, Radial Basis Function SVM [24][25][38][58][59].

3.6. Decision Trees

A decision tree is not a distance-based method. It can be used for both regression and classification both. Though, it is mostly used for classification. DT naturally extended to do multi-class classification. The structure of DT is in the form of a tree. Decision nodes and leaf nodes are the two types of nodes in DT. Starting with the root node, it checks the conditions and accordingly goes to the matching branch and continues till it reaches the leaf node. The predicted value will be at the leaf node [60-69].

3.7. Random Forest

Random Forest is basically a bagging technique. In this, some of the row samples and feature samples are taken and given to one of the many base learners. In a random forest base, learners are decision trees. This step is basically bootstrap. After this aggregation is done by using majority voting [70-73].

Table 1: Papers on prediction of student's performance

Paper No.	Objective	Predictive Model/Technique /Method	Evaluation	Data Set used	Mode
[9]	Identifies the students who are at-risk of a course failure, early prediction of the students who are at-risk and withdrawal from the course and identifies patterns of students who pass with distinction	Logistic Regression SVM Deep ANN classification model	Deep ANN classification model achieved 93% accuracy.	Open University Learning Analytics (OULA)	Online (VLE)
[11]	The objective is to predict whether a student will drop out of a course	LOGIT_Act knowledge discovery system. It uses logistic regression modeling and classification.	LOGIT_Act Knowledge System achieves an accuracy of 97.13%	Activity data from Moodle DB of Madrid Open University	MOODLE
[12]	Predict dropout by using an integrated framework with feature selection, feature generation.	FSPred Framework which uses FEATURE SELECTION + logistic regression model	F1 score of FSPred is 84.69	XuetangX for KDD CUP 2015	MOOC
[13]	The objective is to design a student achievement predicting framework using A layer-supervised multi-layer perceptron (MLP) Neural Network-based method.	SVM, NB, LR, MLP, MLP- Neural Network-based method.	F1 score of MLP Neural Network based method is 81.3%	University	Traditional

[24]	An innovative two-stage approach is proposed and evaluated the effectiveness of it by applying the approach using two different but complementary datasets.	Gaussian RBF kernel and the polynomial kernel were applied to the RF, Deep Neural Network, SVM.	95.53% accuracy achieved by Deep Neural Network	Higher education data set	Moodle learning manage ment system
[25]	Simple model Gradual At-risk (GAR) is presented, to identify at-risk students.	Support Vector (SV), K-Nearest Neighbors (KNN), Decision Tree (DT)-CART, Naïve Bayes (NB)	SVM achieved an accuracy of 92.41%	Universita t Oberta de Catalunya	UOC LMS
[26]	Two models have proposed naming the learning achievement model and the at-risk student model	Generalized Linear Model (GLM) and Gradient Boosting Machine (GBM)AdaBosst algo, Multi-Layer Perceptron (NNET2), Feedforward Neural Network with a single hidden layer (NNET1), Random Forest (RF).	Gradient Boosting Machine (GBM)AdaB osst algo achieved the highest accuracy that is 89.4%	Harvard University and Massachu setts Institute of Technolog y online courses, Open University online courses.	VLE
[27]	Predict the possibility of drop out students by implementing machine and statistical learning method using deep neural network	logistic regression, a multilayer perceptron algorithm	Accuracy=7 7%,	University in Taiwan	Universit y's Institutio nal Research Database ;
[28]	The aim is to discover the impact of online activity data and assessment grades in the LMS on student's performance	Sequential minimal optimization (SMO), logistic regression, multilayer perceptron (MLP), decision tree (J48), random forest	Random Forest achieved the highest accuracy i.e 99.17%	Deanship of E-Learning and Distance Education at King Abdulaziz University	LMS
[29]	Use of DM techniques to predict students' academic performance and to help to advise students	Decision tree, Naive Bayes	J48 achieve the highest accuracy that is 84.38%	Umm Al-Qura University in Makkah	Tradition al
[30]	Developed "University Students Result Analysis and Prediction System"	decision tree algorithms: J48,	Accuracy of J48 is	university student database,	Tradition al

		REPTree, and Hoeffding Tree	highest i.e 85.64%	from students through Google doc survey	
[31]	Proposed a Multi-task learning framework finding out the performance of students and "mastery of knowledge points" in MOOCs using online behavior based on assignments.	"Multi-task multi-layer LSTM with cross-entropy as the loss function", M-S-LSTM, M-F-LSTM standard multi-layer perceptron (MLP), LSTM, standard logistic regression (LR), naïve Bayes (NB).	The proposed model achieved F1-score=93.59	University	MOOC
[32]	Proposed deep LSTM to find out students at-risk by converting the problem into a sequential weekly format.	deep LSTM model, SVM, Logistic Regression, ANN	The proposed model achieved 90% accuracy	OULA	VLE
[33]	Aim to analyze various EDM techniques for improving the accuracy of prediction in a university course for student academic performance.	Random Forest (RF), k-Nearest Neighbour (k-NN), Logistic Regression Naïve Bayes.	Random forest achieved the highest accuracy i.e 88%	University	Tradition al
[34]	Applied ML methods to find out the final grades of students using their previous grades.	Decision tree algorithm	Accuracy is 96.5%	engineeri ng degree at an Ecuadoria n university	Tradition al
[35]	Behavioral data analyzed based on a learning management system used for distance learning courses in a public University. Predictive models have been developed, analyzed, and compared.	Naïve Bayes (NB), Support Vector machine (SVM), Logistic regression (LR), CART-Decision Tree	Logistic Regression achieved the highest accuracy that is 89.3%	University of Pernambu co Distance Learning Departme nt (NEAD/UP E)	Moodle LMS platform
[36]	Predicting student academic performance using "multi-model heterogeneous ensemble" approach	Decision tree (DT), (ANN) artificial neural network, and (SVM) Support Vector Machine,	Ensemble method the hybrid model achieved the highest	The University of the West of Scotland	LMS and (SRS)Stud ent record system

		an Ensemble method hybrid model	accuracy that is 77.69%		question naire
[37]	Predict the performance of students before the completion of the course. Analyzed the progress of the students throughout the course and combine them with prediction results.	Decision Tree, 1-Nearest Neighbour, Naive Bayes, Neural Networks, Random Forest Trees	Naive Bayes achieved the highest accuracy that is 83.6%,	Informatio n Technolog y Engineeri ng University , Pakistan.	Tradition al

Table 2: Advantages and Disadvantages of various techniques used in predicting student's performance

Predicting Techniques	Advantages	Disadvantages
k-NN [16] [38-42]	Simple algorithm and easy to understand, interpret & implement. As no assumption of data therefore helpful for nonlinear data. A versatile algorithm as it can be used for both regression & classification both.	As it stores all training data it becomes a computationally expensive algorithm and requires high memory storage. When the size of N increases the prediction becomes slow. k-NN fails if data points in the dataset are randomly spread. If the data point is far away from the points in the dataset then it is not sure for its class label. Not good for low latency systems.
Naïve Bayes [17]	Simple to understand and implement. If conditional independence of features is true then Naïve Bayes performs very well. Useful algorithm for high dimensions for example text classification, email spam. Extensively used when we have categorical features Run time complexity, training time complexity, run time-space complexity are low. Interpretability is good.	If conditional independence of features is False then Naïve Bayes performance degrades. Seldom is used for real-valued features. Easily overfit (means if data slightly changes model changes drastically) if you don't use Laplace smoothing.

Logistic Regression [18]	Perform well if classes are almost linearly separable. Model interpretability is easy as we can determine feature importance. For small dimensionality, it performs very well, Memory efficient and it has less impact on outliers because of a sigmoid function.	If classes are not almost linearly separable then logistic regression fails. If dimensionality is large then it is prone to overfit and has to apply L1 regularize.
Linear Regression [19]	Simple to implement. Model Interpretability is easy. Perform very well for a linearly separable dataset. The impact of Overfitting can be reduced by using regularization.	The high impact of outliers. Multicollinearity must be removed before applying LR. Prone to underfitting.
Support Vector Machine [20]	The real strength of SVM is the kernel trick, with the right kernel/ appropriate kernel function SVM solves complex problems. Very effective when the dimensionality is high. Can do linearly inseparable classification with global optimal.	Not easy to find the right kernel/ appropriate kernel function. Training time complexity is high for a large dataset. Difficult to interpret and understand the model as we cannot find feature importance directly from the kernel. For RBF with small sigma, outliers have a huge impact on the model.
Decision Tree [21]	High Interpretability Need not to perform feature standardization or normalization. Feature logical interaction is inbuild in DT. DT naturally extended to do multiclass classification. Feature importance is straightforward in DT. Space efficient.	In case of imbalanced data, we have to balance the data and then apply DT. For large dimensionality time complexity to train DT increases dramatically. If a similarity matrix is given, then DT does not work as DT needs the features explicitly. As depth increases the possibility of overfitting increases, interpretability

		decreases, and the impact of outliers can be significant.
Random Forest [22]	Robust to outliers. Need not to perform feature standardization or normalization Feature logical interaction is inbuild in RF. RF naturally extended to do multiclass classification. Feature importance is straightforward in RF.	Does not handle large dimensionality very well. Does not handle categorical features with many categories effectively. Train time complexity is high.
Ensembled Methods [84-91]	Captures linear and nonlinear relationships in data. Robust and stable model. It minimizes noise, bias, and variance.	Interpretability of the model reduces due to increased complexity. Train time is more. Difficult to select a model to ensemble.
Neural Network [23] [74-83]	Non-linear program. Operates with insufficient data. Capable of updating and reasoning.	The required large information for training. Do not assist mixed variables. Black box nature.

4. Critical Analysis

- The Comparative analysis shows that the techniques used to find out the student's performance are quite indecisive as different authors present different results.
- It is also evident from the comparative analysis of the data that mostly the authors have used supervised learning techniques whereas a few authors have chosen the unsupervised learning techniques for predicting the performance of the students. So, there should be more emphasis on the use of unsupervised learning techniques by the researchers.
- It shows that the Decision tree is a mostly used technique by authors followed by neural network and regression.
- It is also evident from the comparative analysis that most authors predicted student's performance at the university level.

5. Conclusion

In this paper, we have reviewed EDM/LA techniques and their strengths and weaknesses for predicting student performance. From the analysis of these papers, we can draw some conclusions.

The comparative analysis indicates ambivalent results on techniques that can best predict student's performance. Asif *et al.,* [37] showed that for predicting student's performance Naïve Bayes achieved the highest classification accuracy at 83.6%. However, Rodrigues *et al.,* [35] noted that logistic regression outperformed the decision tree (CART), support vector machine, Naïve Bayes with 89.3% prediction accuracy. Moreover, Adejo *et al.,* [36] indicated that the ensembled hybrid model achieved the highest prediction accuracy at 77.69% as compared to DT, ANN, SVM. According to Ramaswami *et al.,* [33] Random Forest outperformed NB, LR, K-NN with 88% prediction accuracy. Baneres *et al.,* [25] noted that SVM achieved the highest prediction accuracy with 92.41% as compared to however it is SV, KNN, CART, NB. Hung *et al.,* [24] noted that deep NN achieved 95.53% prediction accuracy and outperformed RF, SVM. However, it is indecisive which technique predicts the student's performance more accurately as different authors present different results. It is evident from the reviewed papers that DT (22%) is a mostly used technique by the authors for predicting student's performance followed by neural network and regression. In addition to Random Forest, SVM, NB, Ensemble methods have also been used. Moreover, it is evident from the data collected for this paper that most authors used supervised learning techniques whereas only a few authors (2%) used unsupervised learning techniques for the prediction of student's performance. It is an opportunity for the researchers to conduct further research in unsupervised learning techniques. Also, 52% of the papers reviewed have predicted student's performance at the university level. It would be encouraging for the researcher to apply the same working line of predictive techniques on Blended, VLE, LMS, MOODLE, MOOC environments.

6. References

[1] Chae, B. K. (2019). A general framework for studying the evolution of the digital innovation ecosystem: The case of big data. *International Journal of Information Management*, *45*, 83–94.

[2] Dhankhar A., Solanki K. (2019). A Comprehensive Review of Tools & Techniques for Big Data Analytics. International Journal of Emerging Trends in Engineering Research, vol 7, No.11, pp: 556-562.

[3] Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (Eds.). (1996, February). Advances in knowledge discovery and data mining. American Association for Artificial Intelligence.

[4] Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(6), 601-618.

[5] Dhankhar A., Solanki K., Dalal S., Omdev (2021) Predicting Students Performance Using Educational Data Mining and Learning Analytics: A Systematic Literature Review. In: Raj J.S., Iliyasu A.M., Bestak R., Baig Z.A. (eds) Innovative Data Communication Technologies and Application. Lecture Notes on Data Engineering and Communications Technologies, vol 59. Springer, Singapore. https://doi.org/10.1007/978-981-15-9651-3_11

[6] Petersen, K.; Vakkalanka, S.; Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update. Inf. Softw. Technol. **2015**, 64, 1–18.

[7] Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Prisma Group. Preferred reporting items for systematic reviews and metaanalyses: The PRISMA statement. BMJ **2009**, 6, 1–8.

[8] Chui, K. T., Fung, D. C. L., Lytras, M. D., & Lam, T. M.: Predicting at-risk university students in a virtual learning environment via a machine learning algorithm. *Computers in Human Behavior*, *107*, 105584 (2020).

[9] Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R.: Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, *104*, 106189 (2020).

[10] Xing, W., Chen, X., Stein, J., & Marcinkowski, M.: Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization. *Computers in human behavior*, *58*, 119-129 (2016).

[11] Burgos, C., Campanario, M. L., de la Peña, D., Lara, J. A., Lizcano, D., & Martínez, M. A.: Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers & Electrical Engineering*, *66*, 541-556 (2018).

[12] Qiu, L., Liu, Y., & Liu, Y.: An integrated framework with feature selection for dropout prediction in massive open online courses. *IEEE Access*, *6*, 71474-71484 (2018).

[13] Qu, S., Li, K., Zhang, S., & Wang, Y.: Predicting achievement of students in smart campus. *IEEE Access*, *6*, 60264-60273 (2018).
[14] Olive, D. M., Huynh, D. Q., Reynolds, M., Dougiamas, M., & Wiese, D.: A quest for a one-size-fits-all neural network: Early prediction of students at risk in online courses. *IEEE Transactions on Learning Technologies*, *12*(2), 171-183 (2019).
[15] Ramesh, A., Goldwasser, D., Huang, B., Daume, H., & Getoor, L.: Interpretable Engagement Models for MOOCs using Hinge-loss Markov Random Fields. *IEEE Transactions on Learning Technologies* (2018).
[16] https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761.
[17] https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c
[18] https://machinelearningmastery.com/logistic-regression-for-machine-learning/
[19] https://towardsdatascience.com/linear-regression-detailed-view-ea73175f6e86
[20] https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47
[21] https://towardsdatascience.com/decision-tree-in-machine-learning-e380942a4c96
[22] https://towardsdatascience.com/understanding-random-forest-58381e0602d2
[23] https://towardsdatascience.com/understanding-neural-networks-19020b758230
[24] Hung, J. L., Shelton, B. E., Yang, J., & Du, X.: Improving predictive modeling for at-risk student identification: A multistage approach. *IEEE Transactions on Learning Technologies*, *12*(2), 148-157 (2019).
[25] Baneres, D., Rodríguez-Gonzalez, M. E., & Serra, M.: An early feedback prediction system for learners at-risk within a first-year higher education course. *IEEE Transactions on Learning Technologies*, *12*(2), 249-263 (2019).
[26] Al-Shabandar, R., Hussain, A. J., Liatsis, P., & Keight, R.: Detecting At-Risk Students With Early Interventions Using Machine Learning Techniques. *IEEE Access*, *7*, 149464-149478 (2019).
[27] Tsai, S. C., Chen, C. H., Shiao, Y. T., Ciou, J. S., & Wu, T. N.: Precision education with statistical learning and deep learning: a case study in Taiwan. *International Journal of Educational Technology in Higher Education*, *17*, 1-13 (2020).
[28] Alhassan, A., Zafar, B., & Mueen, A.: Predict Students Academic Performance based on their Assessment Grades and Online Activity Data. International Journal of Advances Computer Science and Applications, 11(4) (2020).
[29] Alhakami, H., Alsubait, T., & Aliarallah, A.: Data Mining for Student Advising. International Journal of Advanced Computer Science and Applications, 11(3) (2020).
[30] Hoque, M. I., kalam Azad, A., Tuhin, M. A. H., & Salehin, Z. U.: University Students Result Analysis and Prediction System by Decision Tree Algorithm. Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 115-122 (2020).
[31] Qu, S., Li, K., Wu, B., Zhang, X., & Zhu, K.: Predicting Student Performance and Deficiency in Mastering Knowledge Points in MOOCs Using Multi-Task Learning. *Entropy*, *21*(12), 1216 (2019).
[32] Aljohani, N. R., Fayoumi, A., & Hassan, S. U. (2019). Predicting at-risk students using clickstream data in the virtual learning environment. *Sustainability*, *11*(24), 7238, (2019).
[33] Ramaswami, G., Susnjak, T., Mathrani, A., Lim, J., & Garcia, P. (2019). Using educational data mining techniques to increase the prediction accuracy of student academic performance. *Information and Learning Sciences*.
[34] Buenaño-Fernández, D., Gil, D., & Luján-Mora, S.: Application of machine learning in predicting performance for computer engineering students: A case study. *Sustainability*, *11*(10), 2833 (2019).
[35] Rodrigues, R. L., Ramos, J. L. C., Silva, J. C. S., Dourado, R. A., & Gomes, A. S.: Forecasting Students' Performance Through Self-Regulated Learning Behavioral Analysis. *International Journal of Distance Education Technologies (IJDET)*, *17*(3), 52-74 (2019).
[36] Adejo, O. W., & Connolly, T.: Predicting student academic performance using multi-model heterogeneous ensemble approach. *Journal of Applied Research in Higher Education* (2018.)
[37] Asif, R., Merceron, A., Ali, S. A., & Haider, N. G.: Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, *113*, 177-194 (2017).

[38] Rubiano, S. M. M., & Garcia, J. A. D.: Analysis of data mining techniques for constructing a predictive model for academic performance. *IEEE Latin America Transactions*, *14*(6), 2783-2788 (2016).
[39] Wakelam, E., Jefferies, A., Davey, N., & Sun, Y.: The potential for student performance prediction in small cohorts with minimal available attributes. *British Journal of Educational Technology*, *51*(2), 347-370 (2020).
[40] Guerrero-Higueras, Á. M., Fernández Llamas, C., Sánchez González, L., Gutierrez Fernández, A., Esteban Costales, G., & González, M. Á. C.: Academic Success Assessment through Version Control Systems. *Applied Sciences*, *10*(4), 1492 (2020).
[41] Al-Sudani, S., & Palaniappan, R.: Predicting students' final degree classification using an extended profile. *Education and Information Technologies*, *24*(4), 2357-2369, 2019.
[42] Zhou, Q., Quan, W., Zhong, Y., Xiao, W., Mou, C., & Wang, Y.: Predicting high-risk students using Internet access logs. *Knowledge and Information Systems*, *55*(2), 393-413.
[43] Injadat, M., Moubayed, A., Nassif, A. B., & Shami, A.: Systematic ensemble model selection approach for educational data mining. *Knowledge-Based Systems*, 105992 (2020).
[44] Ashraf, M., Zaman, M., & Ahmed, M.: An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches. *Procedia Computer Science*, *167*, 1471-1483 (2020).
[45] Huang, A. Y., Lu, O. H., Huang, J. C., Yin, C. J., & Yang, S. J.: Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs. *Interactive Learning Environments*, *28*(2), 206-230 (2020).
[46] Francis, B. K., & Babu, S. S.: Predicting academic performance of students using a hybrid data mining approach. *Journal of medical systems*, *43*(6), 162 (2019).
[47] Adekitan, A. I., & Noma-Osaghae, E.: Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, *24*(2), 1527-1543, 2019.
[48] Livieris, I. E., Tampakas, V., Karacapilidis, N., & Pintelas, P.: A semi-supervised self-trained two-level algorithm for forecasting students' graduation time. *Intelligent Decision Technologies*, *13*(3), 367-378 (2019).
[49] Gershenfeld, S., Ward Hood, D., & Zhan, M.: The role of first-semester GPA in predicting graduation rates of underrepresented students. *Journal of College Student Retention: Research, Theory & Practice*, *17*(4), 469-488, 2016.
[50] Strang, K. D.: Predicting student satisfaction and outcomes in online courses using learning activity indicators. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, *12*(1), 32-50 (2017).
[51] Ellis, R. A., Han, F., & Pardo, A.: Improving learning analytics–combining observational and self-report data on student learning. *Journal of Educational Technology & Society*, *20*(3), 158-169 (2017).
[52] Christensen, B. C., Bemman, B., Knoche, H., & Gade, R.: Pass or Fail? Prediction of Students? Exam Outcomes from Self-reported Measures and Study Activities. *IxD&A*, *39*, 44-60 (2018).
[53] Yang, S. J., Lu, O. H., Huang, A. Y., Huang, J. C., Ogata, H., & Lin, A. J.: Predicting students' academic performance using multiple linear regression and principal component analysis. *Journal of Information Processing*, *26*, 170-176 (2018).
[54] B. Raveendran Pillai, Gautham. J.: Deep regressor: Cross subject academic performance prediction system for university level students "International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-11S (2019).
[55] Sothan, S. (2019). The determinants of academic performance: evidence from a Cambodian University. *Studies in Higher Education*, *44*(11), 2096-2111.
[56] Moreno-Marcos, P. M., Pong, T. C., Muñoz-Merino, P. J., & Kloos, C. D.: Analysis of the factors influencing learners' performance prediction with learning analytics. *IEEE Access*, *8*, 5264-5282 (2020).
[57] Zhang, X., Sun, G., Pan, Y., Sun, H., He, Y., & Tan, J.: Students performance modeling based on behavior pattern. *Journal of Ambient Intelligence and Humanized Computing*, *9*(5), 1659-1670 (2018).

[58] Gitinabard, N., Xu, Y., Heckman, S., Barnes, T., & Lynch, C. F.: How widely can prediction models be generalized? performance prediction in blended courses. *IEEE Transactions on Learning Technologies*, *12*(2), 184-197 (2019).
[59] Moreno-Marcos, P. M., Pong, T. C., Muñoz-Merino, P. J., & Kloos, C. D.: Analysis of the factors influencing learners' performance prediction with learning analytics. *IEEE Access*, *8*, 5264-5282 (2020).
[60] Dhankhar, A., Solanki, K., Rathee, A., & Ashish.: Predicting Student's Performance by using Classification Methods. International Journal of Advanced Trends in Computer Science and engineering. 8(4), 1532-1536, 2019.
[61] Evale, D.: Learning management system with prediction model and course-content recommendation module. *Journal of Information Technology Education: Research*, *16*(1), 437-457 (2016).
[62] Tran, T. O., Dang, H. T., Dinh, V. T., & Phan, X. H.: Performance prediction for students: a multi-strategy approach. *Cybernetics and Information Technologies*, *17*(2), 164-182 (2017).
[63] Seidel, E., & Kutieleh, S.: Using predictive analytics to target and improve first year student attrition. *Australian Journal of Education*, *61*(2), 200-218 (2017).
[64] Kostopoulos, G., Kotsiantis, S., Pierrakeas, C., Koutsonikos, G., & Gravvanis, G. A.: Forecasting students' success in an open university. *International Journal of Learning Technology*, *13*(1), 26-43, (2018).
[65] Jastini Mohd. Jamil, Nurul Farahin Mohd Pauzi, Izwan Nizal Mohd. Shahara Nee.: An Analysis on Student Academic Performance by Using Decision Tree Models, The Journal of Social Sciences Research ISSN(e): 2411-9458, ISSN(p): 2413-6670 Special Issue. 6, pp: 615-620 (2018)
[66] Bucos, M., & Drăgulescu, B.: Predicting student success using data generated in traditional educational environments. *TEM Journal*, *7*(3), 617 (2018).
[67] Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q.: Predicting academic performance by considering student heterogeneity. *Knowledge-Based Systems*, *161*, 134-146 (2018).
[68] Yaacob, W. F. W., Nasir, S. A. M., Yaacob, W. F. W., & Sobri, N. M.: Supervised data mining approach for predicting student performance. *Indones. J. Electr. Eng. Comput. Sci*, *16*, 1584-1592, (2019).
[69] Mimis, M., El Hajji, M., Es-Saady, Y., Guejdi, A. O., Douzi, H., & Mammass, D.: A framework for smart academic guidance using educational data mining. *Education and Information Technologies*, *24*(2), 1379-1393, 2019.
[70] Huang, A. Y., Lu, O. H., Huang, J. C., Yin, C. J., & Yang, S. J.: Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs. *Interactive Learning Environments*, *28*(2), 206-230 (2020).
[71] Gutiérrez, L., Flores, V., Keith, B., & Quelopana, A.: Using the Belbin method and models for predicting the academic performance of engineering students. *Computer Applications in Engineering Education*, *27*(2), 500-509 (2019).
[72] Crivei, L. M., Ionescu, V. S., & Czibula, G.: An analysis of supervised learning methods for predicting students' performance in academic environments. *ICIC Express Lett*, *13*, 181-190 (2019).
[73] Sadiq, H.M., & Ahmed, S.N.: Classifying and Predicting Students' Performance using Improved Decision Tree C4.5 in Higher Education Institutes, Journal of Computer Science, 15(9), 1291-1306.
[74] Jorda, E. R., & Raqueno, A. R.: Predictive Model for the Academic Performance of the Engineering Students Using CHAID and C 5.0 Algorithm. International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 6, pp. 917-928 (2019).
[75] Vora, D. R., & Rajamani, K.: A hybrid classification model for prediction of academic performance of students: a big data application. *Evolutionary Intelligence*, 1-14, (2019).
[76] Kokoç, M., & Altun, A.: Effects of learner interaction with learning dashboards on academic performance in an e-learning environment. *Behaviour & Information Technology*, 1-15 (2019).
[77] Ramanathan, L., Parthasarathy, G., Vijayakumar, K., Lakshmanan, L., & Ramani, S.: Cluster-based distributed architecture for prediction of student's performance in higher education. *Cluster Computing*, *22*(1), 1329-1344 (2019).

[78] Adekitan, A. I., & Noma-Osaghae, E.: Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, *24*(2), 1527-1543, 2019.
[79] Pal, V. K., & Bhatt, V. K. K.: Performance Prediction for Post Graduate Students using Artificial Neural Network. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2278-3075 (2019).
[80] Guerrero-Higueras, Á. M., Fernández Llamas, C., Sánchez González, L., Gutierrez Fernández, A., Esteban Costales, G., & González, M. Á. C.: Academic Success Assessment through Version Control Systems. *Applied Sciences*, *10*(4), 1492 (2020).
[81] Yang, T. Y., Brinton, C. G., Joe-Wong, C., & Chiang, M.: Behavior-based grade prediction for MOOCs via time series neural networks. *IEEE Journal of Selected Topics in Signal Processing*, *11*(5), 716-728 (2017).
[82] Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R.: Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, *104*, 106189 (2020).
[83] Coussement, K., Phan, M., De Caigny, A., Benoit, D. F., & Raes, A.: Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model. *Decision Support Systems*, 113325 (2020).
[84] Wan, H., Liu, K., Yu, Q., & Gao, X.: Pedagogical Intervention Practices: Improving Learning Engagement Based on Early Prediction. *IEEE Transactions on Learning Technologies*, *12*(2), 278-289 (2019).
[85] Xu, J., Moon, K. H., & Van Der Schaar, M.: A machine learning approach for tracking and predicting student performance in degree programs. *IEEE Journal of Selected Topics in Signal Processing*, *11*(5), 742-753 (2017).
[86] Bhagavan, K. S., Thangakumar, J., & Subramanian, D. V.: Predictive analysis of student academic performance and employability chances using HLVQ algorithm. *Journal of Ambient Intelligence and Humanized Computing*, 1-9 (2020).
[87] Kamal, P., & Ahuja, S.: An ensemble-based model for prediction of academic performance of students in undergrad professional course. *Journal of Engineering, Design and Technology* (2019).
[88] Adekitan, A. I., & Noma-Osaghae, E.: Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, *24*(2), 1527-1543, 2019.
[89] Shanthini, A., Vinodhini, G., & Chandrasekaran, R. M.: Predicting Students' Academic Performance in the University Using Meta Decision Tree Classifiers. *J. Comput. Sci.*, *14*(5), 654-662 (2018).
[90] https://towardsdatascience.com/ensemble-methods-in-machine-learning-what-are-they-and-why-use-them-68ec3f9fef5f
[91] Dhankhar, A., & Solanki, K.: State of the art of learning analytics in higher education. *International journal of emerging trends in engineering research, 8(3), 868-877 (2020).*