IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A REVIEW OF STUDENT PERFORMANCE PREDICTION TECHNIQUES IN VIRTUAL LEARNING ENVIRONMENT

[1] Diksha A. Bansod, [2] Prof. Amit D. Shah

[1] P.G Student, Department of Computer Science & Engineering, Sipna C.O.E.T, Amravati, India,
[2] Assistant Professor, Department of Computer Science & Engineering, Sipna C.O.E.T, Amravati, India.

Abstract: Digitization is transforming all aspects of education. Student performance is an important research topic because a lack of student performance affects the student's final grade, retention of material, and the course dropout rate. In virtual learning environment (VLE), a student's degree of performance in educational learning is lower than that in traditional education systems. Access to online VLE activities is used as a measurement of student performance. Because the course involves VLE, often, no face-to face interaction occurs between students and the instructor. Hence, in VLE, it is difficult to measure a student's performance using traditional methodologies, because many of these predictors are not directly available in e-learning systems. Therefore, investigating students' performance in web-based learning is a challenging task [1].

This paper aims to compare different techniques and to identify their potential strength and weaknesses that are applied in the VLE domain to predict the student's performance. In this paper we have shown a framework of a general student performance prediction methodology and explained how this prediction system works. The general flow of the student performance prediction process is presented. The different predictive models like k-NN, Naive Bayes, Logistic Regression, Liner Regression, Support Vector Machines (SVM), Decision Trees and Random Forest are explained along with their advantages and disadvantages. Later, we have critically analyzed thirty research papers addressing this issue. We have summarized their methodology, problem focus, prediction model used, sample size, and evaluation in a systematic tabular form for comparison. From the detailed literature review it is found that the meta learning techniques provides the best accuracy in prediction of student performance in virtual learning environment.

Index Terms – Virtual Learning Environment (VLE), Web Based Learning, Student Performance Prediction, k-NN, Naive Bayes, Logistic Regression, Liner Regression, Support Vector Machines (SVM), Decision Trees, Random Forest,

I. Introduction

Technology is evolving rapidly. This technological advancement leads to the generation of tremendous amounts of data and it becomes an integral part of all sectors. The educational sector is no exception [2]. Rapid advancements in Technology-Enhanced Learning platforms have shown a tremendous increase in online educational data. This led to higher opportunities to optimizing users' performance with technological platforms to enhance the learning experience [3]. The progression of the accumulated educational data has stimulated the emergence of several research communities. Educational data, a by-product of the interaction between learners and instructors, has been substantiated as a multidisciplinary field of study.

In the traditional approach to education, teachers take various steps to appraise students' levels of performance. However, in web-based platforms, there are no face-to-face meetings, and it is difficult to determine student performance levels in online activities. Therefore, in web-based systems, student data represent the only source through which instructors can assess student performance and performance. Due to the absence of face-to-face meetings, web-based systems face some challenges that need to be addressed. In web-based systems, 78% of students fail to complete their courses [4]. The main reason students drop an online course is the lack of student performance, and the second most common reason is their inability to locate the requisite activities and materials for the next assessment. An important element in reducing student dropout rates in a virtual learning environment (VLE) is to understand the performance of students in meaningful activities. As student participation in course activities increases, the experiences become more engaging, and the probability of a student achieving a high assessment score and completing the e-learning course increases.

1.1. Virtual Learning Environment (VLE).

In the current study, we used data from the VLE to investigate student performance. The VLE stores course lectures, materials, and assessment information. Nearly 5crores students are enrolled in different courses across India. Students enroll in a course through the VLE, and the VLE delivers different lectures, assignments, and materials to the students. One advantage of the VLE is that it allows an instructor to see the activities in which their students participate in the VLE and helps the instructor to analyze those activities to understand student behavior. The students interact with the VLE to watch lectures, complete assignments, and read materials. Finally, student interactions with the VLE are recorded and stored in log files [5]. The logs contain student behavioral data, such as their interactions with the VLE system. An instructor can utilize these data to understand student behavior. The students are generally divided into groups, and an instructor is assigned to each group. The instructor can guide these student groups through courses, for example, by answering their questions and grading their assignments. Additionally, the instructor can use various types of intervention to support and motivate weaker students. However, the sheer number of students in VLE makes it increasingly difficult for the university to engage students in its courses via face-to-face meetings. Moreover, the number of instructors is limited, and it is not possible to contact all students in all courses. Therefore, an intelligent data system that predicts student performance by analyzing logged student data is needed.

1.2. Significance of Predicting Student Performance.

A predictive model can help instructors guide students in succeeding in a course, and be used to determine which activities and materials are more important to the course assessment. Such models also enable instructors to engage students in different activities through the VLE, thereby encouraging the students to participate in the VLE course. Instructors must invest time discerning why student performance in particular course activities and material is attenuated. A predictive system enables an instructor to automatically identify low-performance students during a course based on activities from that online course. When a student receives an advisory e-mail from an instructor (i.e., an e-mail asking about any difficulty), on a weekly basis, the student is more likely to work hard and increase their performance. Such communication is important because it assesses student workloads and addresses issues at an early stage of the course [5]. Apt advice will also improve student retention and decrease the course dropout rate. Acquiring feedback is a challenge for instructors in an e-learning system after redesigning a course and related materials. The instructor can more effectively redesign a course and student materials using a predictive model of the progress of student and the finding can be used to improve the course and materials and increased performance levels of students. Furthermore, teachers receive feedback on the courses they teach via e-learning systems and feedback focuses on the difficulty level, burden, and illustrative richness. Tracking student performance in different educational learning activities encourages high-quality learning, and comprehensive analysis of student performance can help to minimize course dropout rates.

1.3. Meta Learning Techniques Used in a Predictive Model.

We've studied several Meta Learning algorithms as analytical learning approaches intended to predict student performance during a VLE course and compared the resulting performance. Meta Learning is a field of artificial intelligence. Meta Learning algorithms can automatically find complex patterns from features extracted from existing data, enabling them to make smart decisions about current data. The main tasks of learning analytics in education are to collect data, analyze these data and provide appropriate suggestions and feedback to students to improve their learning. With the help of predictive analytics, an instructor can also discover what students are doing with the learning material and how a student's assessment scores are related to that student's performance level. The cognitive ability of computers in some fields is still below that of humans, but due to Meta Learning algorithms, computer abilities are increasing quickly in domains such as e-learning, recommendation, pattern recognition, image processing, medical diagnosis, and many others. Meta Learning algorithms are trained using sample data as inputs and then tested with new data. Instructors can use Meta Learning algorithms to obtain student related information in real time, which helps them intervene during early course stages. Meta Learning is often used to build predictive models from student data; Meta Learning techniques can address both numerical and categorical predictor variables. Decision trees (DTs) are often used to construct trees and find predictive rules based on available data.

II. A General Student Performance Prediction Process

This paper performed a comparative analysis of various techniques used for predicting student performance. For this purpose, relevant articles were identified, selected, evaluated critically using several criteria, and then finding were integrated. We utilized various techniques to study student performance in different VLE activities. The selected techniques were suitable for both domain and categorical educational attributes. The main steps in the current study are presented in Figure 1.

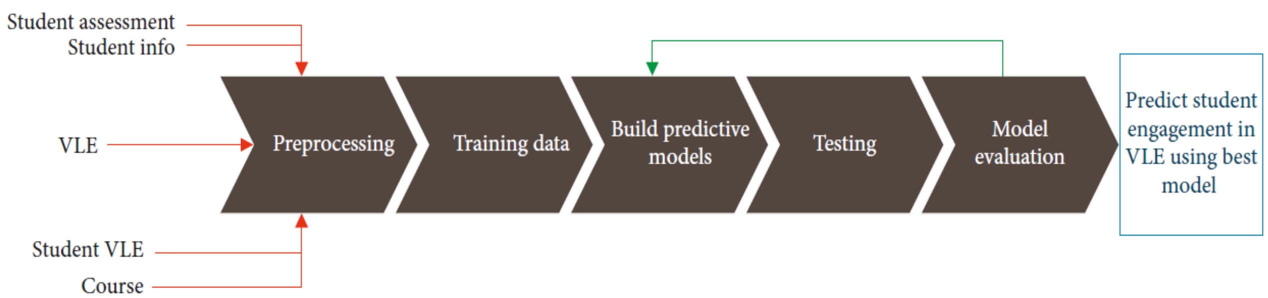

Figure 1: Flow of the student performance prediction process.

2.1 Preprocessing:

The data provided by the universities could not be directly used as inputs in the Meta Learning model. Initially, the input features of the current study are needed to be integrated into a single table. Hens exome preprocessing is needed. In many papers, the various preprocessing steps on the data were performed using MATLAB to format the raw data into a form acceptable for Meta Learning algorithms.

2.2 Training:

The training set in the study is an N-dimensional input vector that contains the input features. These features include the number of clicks on the VLE activities up to the student's completion of the first course assessment. After preparing the data, we need the trained models using the student training data. In this step, various decision tree algorithms like DT, JRIP, J48, GBT, and CART and the NBC algorithm can be used. The study shows that the classifiers were trained using a dataset constructed from the VLE data. The input features used in training were the total number of clicks on the VLE activities completed in the VLE course, and the target variable was the predicted student level of engagement (high or low) throughout the course. In the training phase, the inputs and the corresponding data classes were given to the ML classifier to allow the classifiers to discover the patterns between the input and output. Finally, the trained models used these patterns to classify unseen data.

2.3 Build Predictive Models:

Classification techniques can be used to predict student's performance, to predict students at-risk or retention, students dropout prediction, predict student's achievement, predict which students would likely submit their assignments, assessing student's engagement during the course. In this section, we have discussed various techniques used for predicting student's performance.

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.

Advantages:

- 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.

Disadvantages:

- 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.

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. It classifies the test data by computing conditional probability with feature vectors which belong to particular class. Naive Bayes algorithms can be applied in recommendation system spam filtering, sentiment analysis.

Advantages:

- 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 timespace complexity are low.
- Interpretability is good.

Disadvantages:

- 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.

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

Advantages:

- 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.

Disadvantages:

- 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.

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.

Advantages:

- 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.

Disadvantages:

- The high impact of outliers.
- Multicollinearity must be removed before applying LR.
- Prone to underfitting.

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

Advantages:

- 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.

Disadvantages:

- 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 huge impact on the model

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

Advantages:

- 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.

Disadvantages:

- 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.

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

Advantages:

- 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.

Disadvantages:

- Does not handle large dimensionality very well.
- Does not handle categorical features with many categories effectively.
- Train time complexity is high.

2.4 Testing:

A 10-fold cross-validation method to train and test the current student models is used in many research works. Cross-validation is primarily utilized to assess model performance. In k-fold cross-validation, the data are divided into k different subsets, the model is trained using k-1 subsets, and the remaining subset is used for testing. The average performance obtained from this method provides a good estimation of model performance.

2.5 Model Evaluation:

After training the classifiers in the current study, the performance of the learning models can be accessed using previously unseen data. The prediction results for the models can be obtained with the test data and counted the number of true positives, true negatives, false positives, and false negatives that were used to evaluate performance. Through this process, the numbers of true positives (low-engagement) and true negatives (high engagement) can be obtained, as well as the number of false positives and false negatives. Our main goal in this study is to minimize the false negatives rate (i.e., the number of low-engagement students incorrectly identified as high-engagement students). Therefore, we can select the model with the highest recall. We used the performance metrics like accuracy, recall, Kappa, area under the curve, etc. to measure the quality of the ML model predictions.

An engagement prediction system can be designed based on the results of the current study. Figure 2 demonstrates the interaction between the student engagement prediction system and the VLE.

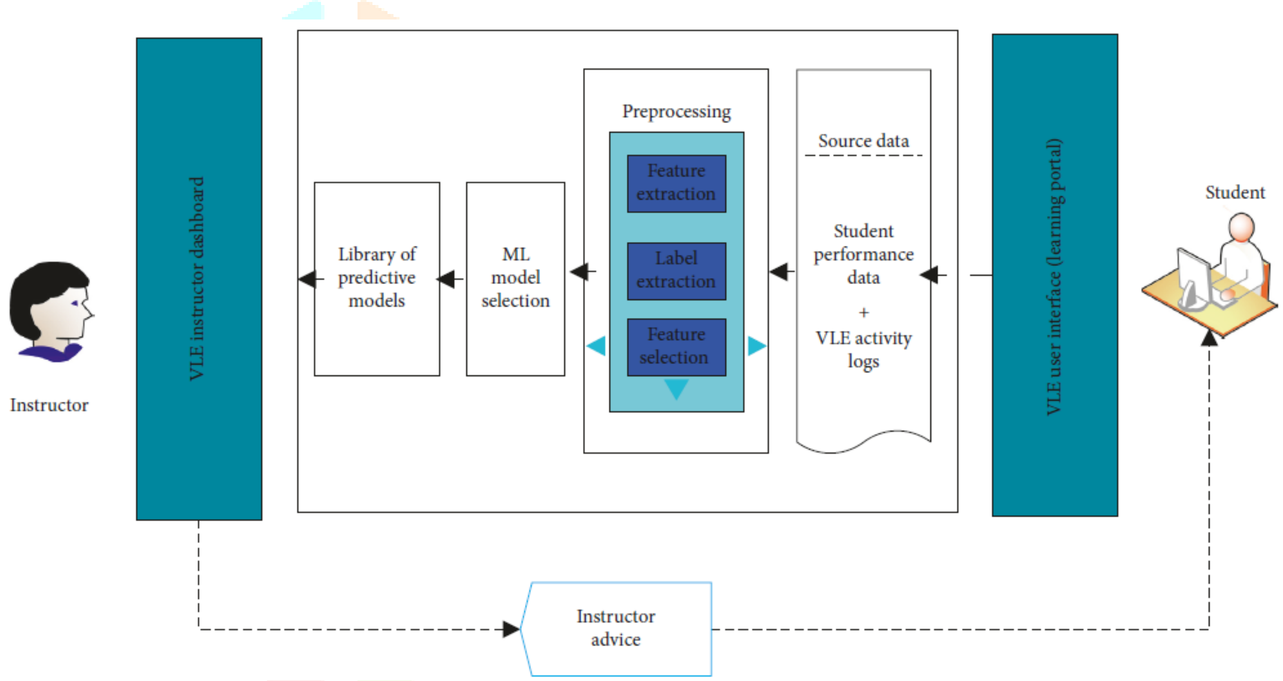

Figure 2: Structure of the student engagement prediction system.

III. Literature Review

Although there are many research works on student engagement prediction system, here we have critically analyzed and summarized around thirty research works and projects addressing this issue. It is observed that most the the recent works uses ML and ANN as their platform. Most of the works have the same working principle as we have already discussed above.

Table 3.1 shows the comparative analysis of the existing methods. The emphasis is on the title of the paper, the methodology and problem focus, the classifier model used in their experimentations, the sample size and their performance evaluation.

Table 3.1: Comparative Analysis of the existing methods

Ref. No.	Name of Authors, Paper Title, Year	Methodology and Problem Focus	Classifier Model Used	Sample Size	Performance Evaluation
[1]	Costa, E.B., et al., Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. (2017)	An investigation of the efficiencies of four educational data mining techniques used to envisage those students that may under perform in a programming module.	Neural Networks, Decision Trees (J48), Support Vector Machine (SVM) and Naive Bayes	161	92% Accuracy
[2]	Sivasakthi, M. Classification and prediction-based data mining algorithms to predict students' introductory programming performance. in (2017)	The application of data mining algorithms such as multilayer perceptron, Naive Bayes, SMO, J48, REPTree on student related data to determine those students that may require additional support.	Multilayer Perceptron, Naïve Bayes, SMO, J48 and REPTree Survey cum experimental methodology	300	93% Accuracy
[3]	Đambić, G., M. Krajcar, and D. Bele, Machine learning model for early detection of higher education students that need additional attention in introductory programming courses. (2016)	A model to identify students who might have problems passing an Introduction to programming course.	logistic regression, simple quadratic model.	181	81% Accuracy
[4]	Bergin, S., et al., Using machine learning techniques to predict introductory programming performance. (2015)	A study of six machine learning algorithms to determine student success in computer programming.	Naïve Bayes	26	78.3% Accuracy
[5]	Pathan, A.A., et al. Educational data mining: A mining model for developing students' programming skills. (2014).	The creation of a decision tree (DT) mining model for improving students programming ability in C.	Decision Tree algorithms	70	87% Accuracy
[6]	Vihavainen, A. Predicting Students' Performance in an Introductory Programming Course Using Data from Students' Own Programming Process. (2013)	An investigation into how students" behaviour during the programming process affects the course outcome.	Bayesian network classifier	200	78% Accuracy
[7]	Rizvi, S., Rienties, B., & Khoja, S. A. The role of demographics in online learning; A decision tree based approach. (2019).	Identifying impact of demographics on academic performance	Decision Tree algorithms	12	83.14% accuracy
[8]	Azizah, E. N., Pujianto, U., & Nugraha, E. Comparative performance between C4. 5 and Naive Bayes classifiers in predicting student academic performance in a Virtual Learning Environment. (2018).	Identifying academic performance based on web pages visited	Naïve Bayes, C4.5 Tree	25	63.8% accuracy
[9]	Wasif, M., Waheed, H., Aljohani, N. R., & Hassan, S.-U. Understanding Student Learning Behavior and Predicting Their Performance. (2019).	Identifying students at-risk of low performance	SVM, Logistic Regression, Random Forest, Naïve Bayes	74	89% accuracy
[10]	Haiyang, L., Wang, Z., Benachour, P., & Tubman, P. A Time Series Classification Method for Behaviour-Based Dropout Prediction. (2018).	Early prediction of dropouts by converting data into day-wise sequences	Decision Trees (10 fold crossvalidation)	26	90% accuracy
[11]	Heuer, H., & Breiter, A. Student Success Prediction and the Trade-Off between Big Data and Data Minimization. (2018).	Predicting student success through daily activities	Decision Tree, Random Forest, Logistic Regression, and Support Vector Machine	156	90.85% accuracy
[12]	Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. (2018)	Identifying students at-risk of low engagement in a VLE	DT, J48, CART, JRIP, Gradient Boosting Trees, NB	130	88.52% accuracy
[13]	Hlosta, M., Zdrahal, Z., & Zendulka, J. Ouroboros: Early identification of at-risk students without models based on legacy data. (2017)	Identifying at-risk students based on first assessment	Machine learning algorithms SVM, Naïve Bayes, Logistic Regression, Random Forest	85	71.31% accuracy
[14]	Sorour, S. E., El Rahman, S. A., & Mine, T. Teacher interventions to enhance the quality of student comments and their effect on prediction performance. (2016).	To investigate the impacts of Teacher Intervention (TIs) on students attitudes and successes involved by analyzing their freestyle data after every lesson, which were classified into AM (Attribute-based Method) and TM (Topic-based Method)	SVM	12	83.14% accuracy
[15]	Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. (2017).	To investigate self-regulation strategies in MOOC, and builds on self-regulated learning (SRL) theory, which describes ways for learners to take control of their learning process	Logistic Regression	25	63.8% accuracy
[16]	Fernández-Delgado, M., Mucientes, M., Vázquez-Barreiros, B., & Lama, M.. Learning analytics for the prediction of the educational objectives achievement. (2014)	To Predict if students will accomplish the task of the subject in light of the past outcome, to empower the educators to adjust learning plan of the instructing learning process.	SVM	74	89% accuracy
[17]	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. (2020).	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	52	93% accuracy
[18]	Burgos, C., Campanario, M. L., de la Peña, D., Lara, J. A., Lizcano, D., &	The objective is to predict whether a student will drop out of a course	Logistic Regression	84	97.13% accuracy

	Martínez, M. A.: Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. (2018).				
[19]	[19] Qiu, L., Liu, Y., & Liu, Y.: An integrated framework with feature selection for dropout prediction in massive open online courses. (2018).	Predict dropout by using an integrated framework with feature selection, feature generation.	Logistic Regression	78	84.69% accuracy
[20]	Qu, S., Li, K., Zhang, S., & Wang, Y.: Predicting achievement of students in smart campus. (2018).	The objective is to design a student achievement predicting framework using A layer-supervised multilayer perceptron (MLP) Neural Network-based method.	SVM, Naïve Bayes, Logistic Regression, MLP, MLP- Neural Network-based method.	25	81.30% accuracy
[21]	Hung, J. L., Shelton, B. E., Yang, J., & Du, X.: Improving predictive modeling for at-risk student identification: A multistage approach. (2019).	An innovative two-stage approach is proposed and evaluated the effectiveness of it by applying the approach using two different but complementary datasets.	RF, Deep Neural Network, SVM.	54	95.53% accuracy
[22]	Baneres, D., Rodríguez-Gonzalez, M. E., & Serra, M.: An early feedback prediction system for learners at-risk within a first-year higher education course. (2019).	Simple model Gradual Atrisk (GAR) is presented, to identify at-risk students.	Support Vector (SV), K-Nearest Neighbors (KNN), Decision Tree (DT)-CART, Naïve Bayes (NB)	45	92.41% accuracy
[23]	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. (2020).	Predict the possibility of drop out students by implementing machine and statistical learning method using deep neural network	logistic regression, a multilayer perceptron algorithm	12	70% accuracy
[24]	Alhassan, A., Zafar, B., & Mueen, A.: Predict Students Academic Performance based on their Assessment Grades and Online Activity Data. (2020).	The aim is to discover the impact of online activity data and assessment grades in the LMS on student's performance	Logistic regression, multilayer perceptron (MLP), decision tree (J48), random forest	23	99.17% accuracy
[25]	Alhakami, H., Alsubait, T., & Aliarallah, A.: Data Mining for Student Advising. (2020).	Use of DM techniques to predict students' academic performance and to help to advise students	Decision tree, Naive Bayes	25	84.38% accuracy
[26]	Ramaswami, G., Susnjak, T., Mathrani, A., Lim, J., & Garcia, P. Using educational data mining techniques to increase the prediction accuracy of student academic performance. Information and Learning Sciences. (2019).	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.	28	88% accuracy
[27]	Buenaño-Fernández, D., Gil, D., & Luján-Mora, S.: Application of machine learning in predicting performance for computer engineering students: A case study. (2019).	Applied ML methods to find out the final grades of students using their previous grades.	Decision tree algorithm	48	96.5% accuracy
[28]	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. (2019).	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	75	89.3% accuracy
[29]	Adejo, O. W., & Connolly, T.: Predicting student academic performance using multi-model heterogeneous ensemble approach. (2018.)	Predicting student academic performance using "multi-model heterogeneous ensemble" approach	Decision tree (DT), (ANN) artificial neural network, and (SVM) Support Vector Machine, an Ensemble method hybrid model	88	77.69% accuracy
[30]	Asif, R., Merceron, A., Ali, S. A., & Haider, N. G.: Analyzing undergraduate students' performance using educational data mining. (2017).	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.	Nearest Neighbour, Naïve Bayes, Neural Networks, Random Forest Trees	102	83.6% accuracy

IV. 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.

V. DISCUSSION

In this paper, we have reviewed various meta Learning 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. It becomes indecisive which technique predicts the student's performance more accurately as different authors present different results. This indicates that there is an opportunity for the researchers to conduct further research in meta learning techniques.

REFERENCES

[1] Costa, E.B., et al., Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. Computers in human behavior, 73: p. 247-256. (2017)

[2] Sivasakthi, M. Classification and prediction-based data mining algorithms to predict students' introductory programming performance. in 2017 International Conference on Inventive Computing and Informatics (ICICI). 2017.

[3] Đambić, G., M. Krajcar, and D. Bele, Machine learning model for early detection of higher education students that need additional attention in introductory programming courses. International Journal of Digital Technology & Economy, 2016. 1(1): p. 1-11.

[4] Bergin, S., et al., Using machine learning techniques to predict introductory programming performance. International Journal of Computer Science and Software Engineering (IJCSSE), 2015. 4(12): p. 323-328.

[5] Pathan, A.A., et al. Educational data mining: A mining model for developing students' programming skills. in The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014). 2014.

[6] Vihavainen, A. Predicting Students' Performance in an Introductory Programming Course Using Data from Students' Own Programming Process. in 2013 IEEE 13th International Conference on Advanced Learning Technologies. 2013.

[7] Rizvi, S., Rienties, B., & Khoja, S. A. The role of demographics in online learning; A decision tree based approach. Computers & Education, 137, 32–47. (2019).

[8] Azizah, E. N., Pujianto, U., & Nugraha, E. Comparative performance between C4. 5 and Naive Bayes classifiers in predicting student academic performance in a Virtual Learning Environment. 2018 4th International Conference on Education and Technology (ICET), 18–22. IEEE. (2018).

[9] Wasif, M., Waheed, H., Aljohani, N. R., & Hassan, S.-U. Understanding Student Learning Behavior and Predicting Their Performance. In Cognitive Computing in Technology-Enhanced Learning (pp. 1–28). IGI Global. (2019).

[10] Haiyang, L., Wang, Z., Benachour, P., & Tubman, P. A Time Series Classification Method for Behaviour-Based Dropout Prediction. 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT), 191–195. IEEE. (2018).

[11] Heuer, H., & Breiter, A. Student Success Prediction and the Trade-Off between Big Data and Data Minimization. DeLFI 2018-Die 16. E-Learning Fachtagung Informatik. (2018).

[12] Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. (2018). Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. Computational Intelligence and Neuroscience, 2018

[13] Hlosta, M., Zdrahal, Z., & Zendulka, J. Ouroboros: Early identification of at-risk students without models based on legacy data. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (pp. 6-15). ACM. (2017)

[14] Sorour, S. E., El Rahman, S. A., & Mine, T. Teacher interventions to enhance the quality of student comments and their effect on prediction performance. Paper presented at the Frontiers in Education Conference (FIE), 2016 IEEE. (2016).

[15] Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. Computers & education, 104, 18-33.34 (2017).

[16] Fernández-Delgado, M., Mucientes, M., Vázquez-Barreiros, B., & Lama, M.. Learning analytics for the prediction of the educational objectives achievement. Paper presented at the Frontiers in Education Conference (FIE), IEEE. (2014)

[17] 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).

[18] 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).

[19] 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).

[20] Qu, S., Li, K., Zhang, S., & Wang, Y.: Predicting achievement of students in smart campus. IEEE Access, 6, 60264-60273 (2018).

[21] 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).

[22] 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).

[23] 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).

[24] 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).

[25] Alhakami, H., Alsubait, T., & Aliarallah, A.: Data Mining for Student Advising. International Journal of Advanced Computer Science and Applications, 11(3) (2020).

[26] 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.

[27] 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).

[28] 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).

[29] Adejo, O. W., & Connolly, T.: Predicting student academic performance using multi-model heterogeneous ensemble approach. Journal of Applied Research in Higher Education (2018.)

[30] Asif, R., Merceron, A., Ali, S. A., & Haider, N. G.: Analyzing undergraduate students' performance using educational data mining. Computers & Education, 113, 177-194 (2017).