VAWKUM Transactions on Computer Sciences
http://vfast.org/journals/index.php/VTCS@ 2022, ISSN(e): 2308-8168, ISSN(p): 2411-6335
Volume 10, Number 2, July-December 2022 pp:45-60

Performance Prediction for Undergraduate Degree Programs Using Machine Learning Techniques - A Preliminary Review

Waqar Un Nisa[1], **Mudasser Naseer**[2*], **Muhammad Atif**[2], **Salwa Muhammad Akhtar**[3], **Mehr Un Nisa**[4]

[1]Faculty of CS and IT, Virtual University of Pakistan, Lahore, Pakistan; [2]Department of and IT, The University of Lahore, Lahore, Pakistan; [3]INS Department, HSM, University of Management and Technology, Lahore, Pakistan; [4]Ph.D. Scholar, Faculty of Science, Information Technology University, Lahore, Pakistan

Keywords:
Performance Prediction, Grade prediction, Traditional Education, Supervised Learning

Journal Info:
Submitted:
November 19, 2022
Accepted:
December 22, 2022
Published:
December 20, 2022

Abstract Academic Performance prediction for undergraduate students is considered as one of the hot research areas since last couple of decades. An accurate and timely prediction of the student's performance can directly influence the three participants; learner, instructor and the institution. This study presents a brief, preliminary review to explore existing literature from 2010 to 2022 in the context of performance prediction for Undergraduate Degree Programs (UDP). This review is organized according to Online and Traditional Education Systems (TES), and granularity level of performance output i.e., Degree program (Final CGPA), Next-semester, and the Course level grades. Aggregate analysis of the extracted data reveals that course level prediction is highly worked area deploying classification and regression techniques using data from academic domain. Existing empirical studies are mostly evaluated using accuracy, precision, recall and F1-measure and are validated with 10-fold cross validation. Contribution of this study is the novel categorical distribution of studies with respect to education system and granularity levels. Another important finding was the Success ratio of different Machine learning (ML) techniques used for these prediction studies. It is concluded that further research is required for TES to discover interdependent group of courses and Course Clusters for a certain degree program and then to develop prediction models for those course clusters.

***Correspondence Author Email Address:**
mudasser.naseer@cs.uol.edu.pk

1 Introduction

Education serves as a pillar for the sustainability and progress of any society. Higher Education Institutions (HEIs) provide a reliable workforce that contribute to the overall development of a healthy society. The Horizon and Educause reports [1],[2] ,[3], [4] have been realizing the importance of educational data analysis, predictive learning analytics and student's success since 2011. It becomes very important to understand and utilize the educational data to improve educational policies and institutional outcomes.

Student performance prediction history can be traced back to years ago using statistical models [5, 6] but use of Machine Learning (ML) and Data Mining (DM) techniques began to use about two decades ago [7, 8] for the prediction of academic grades, pattern classification and other mining activities.

Two major educational environments exist in higher educational institutions based on the student-teacher mode of interaction. First one is, Online Education System (OES), and is mainly consists of Massive Open Online Courses (MOOC) and Virtual Degree Programs (VDP). The second one is the Traditional Education System (TES) with Traditional Degree Programs in face-to-face physical classroom environment. Now a days TES also include blended systems where lectures and exams are conducted in physical classroom but the academic and teaching activities are available on LMS or other student resources.

The motivation behind the study was to explore the use of different ML technique implemented by different researcher for the purpose of grade prediction of undergraduate students, especially in the TES, and to some extend in the online education system. Keeping in view the motivation, this study presents a macroscopic review of work for undergraduate student's performance prediction systems for the period from 2010 to 2022. We follow the concept of Systematic Mapping Study (SMS) proposed by Kitchenham [9]. The authors in [8] suggested to use SMS for scope building and discovering areas for future primary studies. Some of the past reviews [10, 11] have used Systematic Literature Review (SLR) to perform an un-biased meta-analysis of primary studies. These SLR's evaluate and present their findings according to a specific and well-defined research question(s). This study, however, address the issue of performance prediction at an abstract level. Prior to proceed, it is important to describe the word performance in this research context. Performance prediction here refers to prediction of any academic performance factor for undergraduate students, like course grade, Cumulative Grade Point Average (CGPA), pass/fail assessment, prediction of at-risk (of failure) and drop-outs etc. Performance prediction systems have been deployed using Statistical, Data Mining or Machine learning algorithms. To predict efficiently and accurately, researchers used predictor(s) from different data domains namely Course Specific (CS), Demographic (Demo), Academic (Acad), Past Academic (P-Acad), Pedagogical, and Financial etc. To evaluate the prediction performance of an algorithm usually statistical evaluation measures are used e.g., precision, accuracy, recall, F-1 measure, Mean Square Error (MSE), Mean Absolute Error (MAE), Area Under Curve (AUC) etc.

The contribution of this work is the summarization of the latest trends or applying different ML and Data Mining algorithm for the purpose of grade prediction of undergraduate students along with introduction of the term success ratio of any algorithm. The success ratio depicts that how many studies apply a specific algorithm and in how many studies that particular algorithm outperforms all the other algorithms used. The rest of paper is organized as Section 2 describes the methodology followed to organize the complete this review work, Section 3 covers Discussion and Future Work, whereas the Section 4 presents Conclusion of the study.

2 Methodology

Student's performance prediction in UDP is a prevalent research area these days. Our intent is to conduct a review of the work published during the past almost twelve years to discover and aggregate the facts presented in these publications. Moreover, objective of this review is to reveal possible gaps that need to be addressed for a future primary study. We follow the review strategy proposed by Kitchenham [12]. Figure-1 shows the work flow of our study.

2.1 PLANNING

It is imperative and very useful to define research questions, list of resources to execute search strings, selection criteria and the list of features prior to perform actual search.

2.1.1 Research Question

Objective of this review is to explore the studies to find answers of the following research questions;

1. What are the types of education systems used for UDP?
2. What is the granularity of performance prediction output in each education system categorized in question no 1?
3. What are the effective prediction algorithms, predictors and evaluation measures for each category identified in the question no 1 and 2?
4. What are possible research areas for future primary studies regarding performance prediction?

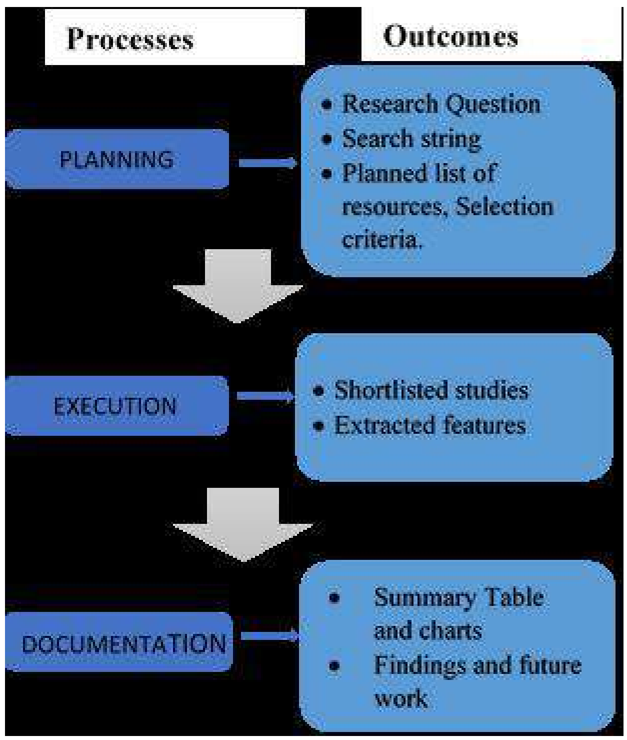

Figure 1. Review Methodology

2.1.2 List of Resources

To get the answer of first question, an abstract search is required using general key words. Therefore, search strings are defined using combination of following key words; “Undergraduate students”, “Grades Prediction”, “Performance prediction”, “Educational Data Mining”, “Online Education”, “Traditional Education”, and “Conventional Education”. In order to execute search strings, the digital libraries; IEEE Xplore, Springer Link, Science Direct, ACM Digital Library, Elsevier and Google Scholar are selected.

2.1.3 Selection Criteria and List of Features

Lastly, selection criteria for the studies and list of features that are required to be extracted from the studies are identified and listed below. Selection Criteria is identified as;

• Studies should be relevant to search strings and research question defined.

• Studies could be one of the Journal papers, conference paper or Thesis published between 2010-2022 and must not be a review or survey paper. Finally, it is beneficial to identify list of features required for analysis. These are;

• Author and year of study.
• Target education system.
• Data source use.
• Output class to be predicted.
• ML algorithm(s) used.
• Most efficient algorithm identified by the authors of the respective study.
• Predictors and Data domains used.
• Evaluation metrics used for each study.

2.2 Execution

Search is executed for title and keywords using search string planned in planning phase from digital libraries mentioned above. Initial scanning of search results reveals that OES have been widely studied for performance prediction leaving very little scope for future exploration. In order to discover the future research direction for TES we diverted the main focus of our review study towards the TES. The search strings were also updated accordingly. Iteratively we searched, select and classify the studies according to selection criteria defined in Planning phase. Finally, 82 papers were collected and their abstracts and conclusion sections were scanned to validate our selection criteria. This screening helped out in finding education system and granularity level of output class. In terms of granularity, three levels were discovered; Final CGPA of degree program, Next-term or Future performance, and Course Level (Current course the student is enrolled in). According to selection criteria all irrelevant and duplicate studies were excluded and 47 papers were shortlisted. These papers were thoroughly studied for quality assessment, keeping in mind the planned list of features. We prefer those studies that used real-time data (not benchmark data) and the outcomes are unambiguous. After this stage, 40 research articles were finalized. At the end of this study, however, some review papers were also included for meta-analysis. In the next section, the finally selected research papers are briefly explained under different categories.

2.2.1 Performance Prediction for Online Education System

Tremendous research work has been done for OES including Distance learning Degree programs and Online Open courses with millions of learners enrolled. Possible reasons for researcher's inclination towards OES might be the availability of online data and easy intervention. Although the main focus of this study is not OES, however, we refer few OES studies because of obvious research gap between OES and TES. Most of the OES based studies used lecture watching behavior and online activities performed for the course.

We generalize these course related activities as Course Specific (CS) data for this review. Like the authors in [13] analyze online learning behavior of a student for prediction of future weeks of study using Long-short term memory Recurrent Neural Network (RNN). Another Artificial Neural Network (ANN) based study utilized Multi-layer Perceptron (MLP) and Radial Basis Function (RBF) to predict pass/fail rate [14]. The authors use final exam score for a basic IT course of an Open Education System of Anadolu University, Turkey. In the same stream [15] proposed RNN based prediction model for information science course using course specific data of Kyushu University, Japan.

Few regression based studies to predict course level grades for Open Online Courses are [16], [[17], [18], [19], [20] and [21]. Yang et al. [16] used Multi Linear Regression (MLR) by incorporating Principal Component Analysis MLR-PCA to predict calculus grades using lecture watching behavior and course assessments data. Whitehill [17] evaluate reliable dropout predictors based on clickstream feature using L2-regularized logistic Regression on 40 MOOCs. The authors in [18] studied Logistic Regression (LogR) based model to identify at-risk student and improve course completion rate for Coursera MOOCs. Both [19] and [20] conducted their research under Open Academic Analytics Initiative (OAAI) using online open educational resources from Marist College, New York. Bainbridge et.al. used dataset from three domains, Demo, academic and behaviors. LogR based method identified at-risk individuals for their study. While goal of Jayaprakash was to increase the student's performance by incorporating intervention strategies and to control drop-out ratio. The authors in [21] used six regression-based techniques for final assessment prediction of the informatics course of Hellenic Open University and found Regression based M5Rule algorithm very effective. Finally classification based analytics model for University of Phoenix was developed by Barber [22] to identify at-risk students in bachelors and master's degree programs by using features from Demo, academic and CS domains. LogR and Naïve Bayes (NB) based models were used for their study proving NB model more accurate.

2.2.2 Performance Prediction for Traditional Education System

The TES is a major means of higher education in many developing and underdeveloped countries and is far less studied for performance prediction. This study, therefore, explores TES to reveal further research opportunities in perspective of performance prediction. This category includes all education systems in which lecture is delivered in physical classroom while other class activities (assessment, labs, exams) can be taken in traditional way or via learning management system. We identify three classes of research articles in terms of granularity of "prediction output"; Final GPA (CGPA) Prediction, Next-term or Future-Term performance prediction and Course Level performance prediction.

Final CGPA Prediction Underlying prediction techniques and predictor's correlation with the output class differ for different education system and target output variable. Like predicting Degree's final GPA mostly consider entry test results, some P-academic records and on-going semester grades as set of predictors(s).

Decision Tree (DT) based algorithms are mostly used for this class of performance prediction. Like Xu et al. [23] used Random Forest (RF) to predict gradual progress during each semester as well as final CGPA. Al-Barrak et al. [24] performed study for King Saud University, Riyadh using J48 to predict final GPA of students. Similarly, Ogunde Ajibade [25] proposed a DT based Iterative Dichotomiser 3 (ID3) algorithm. It uses entry test results and secondary school grades to estimate the Final CGPA.

In two other studies, Alharbi [26] identify students who are at-risk of degree completion using combination of nine classification techniques deploying demographic and general performance data, and Adam Anthony [27] predicted time to complete the degree using Bayesian network.

Next-term's Grades Prediction The next-term grade prediction models have been developed to assist students for future course enrollments and successful and timely degree completion. Moreover, these models are also helpful in analyzing the study patterns and controlling retention rate of the institution. Mostly Matrix factorization (MF), Collaborative Filtering (CF) and Regression techniques were employed by the researchers using Acad data domain for predicting future grades. Like Rovira et al. [28] developed an automated tool using data of University of Barcelona. They used five classification techniques to build drop-out prediction model at course level in which LR outperformed other methods and predict future course GPA and course recommendation using CF Recommendation System. Bydžovská [29] developed a future enrollment recommender system using item-to-item CF and computed course similarities using information from the course catalogue by making course clusters. Regression models have also been used for predicting future grades, like Polyzou Karypis [30] used Linear regression using CS data from the University of Minnesota. Another study by Sweeney et al. [31] using Personalized regression model based on Random Forest and Matrix factorization used the data of George Mason University.

It is also observed that in recent years ANN based techniques are also successfully being employed for all level of performance prediction systems. For example Iqbal [32] used Restricted Boltzmann Machines (RBM) to predict future semester grades successfully, using data from Information Technology University (ITU), Pakistan.

Course-level Grade Prediction This category of performance prediction (course level) is frequently researched. Studies under consideration mostly used classification and regression algorithms to predict. Yang et al. [33] conducted a study for a blended education system using eight classification techniques for University of Tartu in Estonia. Student's procrastination data (students' homework submission behavior) was extracted from Moodle system and Linear-Support Vector Machine (L-SVM) proved to be successful to predict multi-class grades for the course. Pelaez et al [34] merge latent class analysis (LCA) and RF to develop a Latent Forest Model to identify at-risk students in Psychology course at San Diego State University. Demo., admission and academic attributes were utilized. Two Deep Learning (DL) based studies were conducted by Kumar Garg [35] and Patil et al. [36]. First one compares deep learning methods with five other classification techniques using schooling marks and CS assessments (P-Acad, CS) from DIT University, Dehradun, India. While the second one, conducted by Patil et al., compares Recurrent neural network (RNN) and feed forward neural networks FF-ANN with other classification techniques. Their results showed RNN gives more promising results using CS data.

Some more regression-based studies are performed by Meier et al. [37] and Marbouti et al. [38]. The contribution of Meier et al. was to predict student's final grades at an optimal time. They demonstrated regression setting to predict grades and classification setting to check whether a student needs additional help, if he is at-risk. Internal course assessments data of University of California Los Angeles for introductory digital signal processing course was used. In their first paper Marbouti et al. [38] performed logistic regression-based study to identify at-risk students in an engineering course based on course-specific predictor. Later in [39] Marbouti et al. performed a comparative study of seven classification models to identify at-risk students and designed an ensemble model by merging three models Naïve Bayes Classifier (NBC), SVM and K-Nearest Neighbors (KNN).A few more studies build classification models for course level performance prediction. Shakeel and Butt [40] analyzed the data of BS students enrolled in the year 2011 at the University of Gujrat, Pakistan. They used the data for prediction of students' performance and identification of weaker students. They apply different variations of Decision Tree and Bayes algorithms and concluded that Naïve Bayes achieved the highest accuracy of 91.94 %. Sulaiman, Akinbowale and Ronke [41] apply different variations of Decision Tree Algorithms for the prediction of computer programming course. For the analysis they took the data of 131 students from computer science programme at Kwara State University Nigeria. Hamoud, Hashim and Awadh [42] also compare the performance of different variation of decision tree algorithm using a demographic and educational data set of 161 students from College of Computer Science and Information Technology, University of Basrah, Iraq. They recommended the use of J48 for the prediction of a course GPA. Mingyu et al. [43] proposed an interpretable prediction method for university student academic crises warning consisting of K-prototype-based student portrait construction and Catboost-SHAP-based academic achievement prediction. They collected data from 13,613 students from a University in Dalian China. The data includes the basic information, study behavior, internet behavior and living behavior. The authors apply several clustering and prediction methods and suggested K-prototype for student portrait construction and Catboost-SHAP for academic achievement prediction based on MSE, MAE and R2 measures.

The authors in [44] explore the applications of different ML algorithms for the prediction of the academic performance of the students in higher education as high or low. They suggested to use ANN after performing the experiments on a dataset of 162,030 students. The data was collected from private and public universities in Colombia. Mustafa Yağcı [45] predict the final exam grades of the students of a Turkish public sector university based on the information of mid-term exams, department data and the faculty data. He used several ML algorithms, among those RF and Neural Network (NN) provides highest classification accuracy. M. Bucos and Drăgulescu [46] developed a model for Romanian education system to predict pass/fail for a course. Classification techniques with CS attributes from Polyethnic University Timisoara, Romania were used suggesting Logistic Regression Classifier the best one. Rovira et al. [28]developed an automated tool using data of University of Barcelona. For course level prediction, they predicted course grade using five classification techniques and then predict drop-outs. Drop-out Classification achieved highest F1 score using LR. In his PhD research Bydžovská [29] predict course grades, drop-outs and future grades using social behavioral data (student disclosure and friends' grades) along with past academic and demographic data. Most efficient techniques for the drop-out predictions were meta-classifier (J48 and PART), while regression algorithms were used to predict course grade. SVM outperformed for both classification and regression. Majeed Junejo [47] predict student's course grades for PAF-Karachi Institute of Economics and Technology, Pakistan. CS and few pedagogical attributes were used for the study achieving highest accuracy using Rule induction (DT) method. Jishan et al. [48] deployed three classification techniques to predict course grades at North South University, Bangladesh using CS features using three classification techniques. Their study suggested to use NB as it is proved to be computationally faster than ANN in their study. Huang et al. [49] compared four classification models to predict the students' grades in an engineering dynamics course using P-academic and CS data. Multiple-LR gives good results to predict pass/fail percentage while SVM gives highest accuracy for predicting performance.

2.3 Documentation

To provide a better insight of existing literature and to elaborate findings, all of the referenced work is summarized here w.r.t to their predictor's usage, evaluation metrics and implemented ML-techniques for different categories. Complete summary table of the literature review is provided in Appendix A.1 for reference. In this section we have also compared aggregate analysis outcome of this review with couple of previous survey studies to check agreement of our findings with the previous studies.

2.3.1 Categories Analysis

We have shortlisted 40 primary studies for aggregate analysis. First of all, to answer Research Question-1 (RQ-1), we classify studies into two types of education systems. We finalized 30 TES based research papers and 10 OES based. Keeping in mind the RQ-2, these two education systems were further categorized for granularity of performance prediction. Three levels were identified to answer RQ-2. All of the selected OES studies targeted Course level prediction while in TES there were 17 Course level studies, 7 Final CGPA level and 6 for the next semester GPA or next term's performance, see Table 1. To analyze the frequency of each granularity level for both education systems accumulatively, a Pie chart is drawn, Figure 2. This chart depicts 77 % of the studies target Course level performance and hence it is highly worked area. This classification of target education system and granularity level is a new concept for future exploratory surveys.

Table 1. Categorical distribution of Studies

Granularity	OES	TES	Total	Percentage (Total)
Final GPA	-	7	7	17.5%
Next Semester GPA	-	6	6	15%
Course level	10	17	27	67.5%
Total	10	30	40	100%

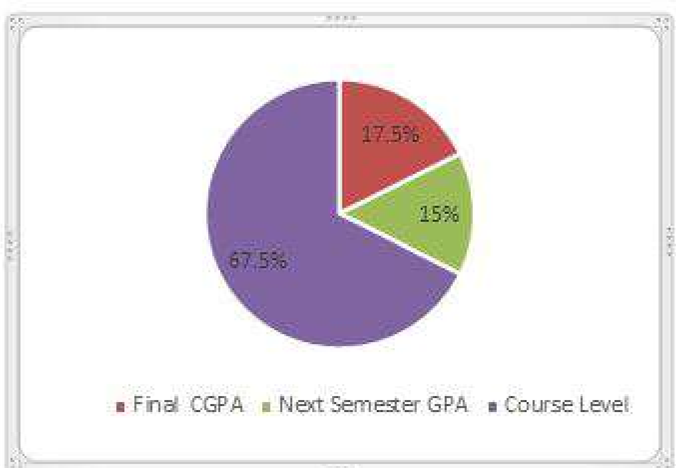

Figure 2. Percentage Distribution for Granularity Levels

Year wise distribution of the selected primary studies is provided in Figure 3 revealing most of the selected studies were published between 2014 and 2017.

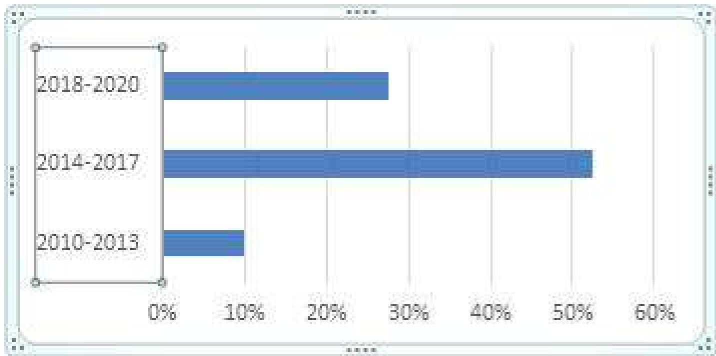

Figure 3. Distribution of selected studies according to year of publishing

2.3.2 Algorithm Usage Analysis

All of the shortlisted studies referenced for this review used ML algorithms to build their prediction models. Some past review papers revealed that classification and regression based algorithms are most commonly used in the prediction models, followed by the clustering algorithms and Association rule mining [50], [10], [11], [51], and [52]. In classification scenario ANN and Deep learning techniques are rather new and the survey presented in [53] emphasized that the usage frequency of ANN and Deep learning methods were low in the last decade while higher tendency is observed in current decade, especially in the past five years.

In this study we distribute ML techniques in seven categories for both OES and TES. For aggregate analysis ANN represents all neural network algorithms including DL algorithms. Decision Tree is generalized for all DT based algorithms including RF. Regression is used for any LR and LogR algorithms. Analyzing the referenced literature here, it is observed that Regression and DT are most widely used by the researchers for performance prediction systems followed by SVM, ANN, NB and others. While KNN is least frequently used for performance prediction. Usage frequency of classification techniques is summarized in Table 2 and Figure 4 below.

Table 2: Usage Distribution of ML Techniques

Techniques		ANN	Reg	NB	DT	SVM	KNN	others
34emOES	Usage	4	8	2	2	4	0	0
	success	3	6	1	0	0	0	0
34emTES	Usage	8	15	10	24	9	5	13
	success	5	7	4	11	3	1	4
34emTotal	Usage	12	23	12	26	13	5	13
	success	8	13	5	11	3	1	4
Success Ratio		66.7 %	56.5 %	41.7 %	42.3 %	23.1 %	20 %	30.8 %

By analyzing the usage frequency of each ML algorithm and their respective success frequency, an important factor, success ratio, is discovered. It is the ratio of success of the technique w.r.t. its usage (see Figure 5). Here we can see that Regression has the highest success ratio followed by ANN. While DT has only 40% success ratio, although its usage frequency, in this study as well as in some past review studies, is very high.

Some previous surveys showed similar results for usage frequency. Like in [10], the authors identified DT and ANN as the highly used technique and in [11] Decision Trees, Naive Bayes and ANN were identified as mostly used techniques. Usage frequency of ML techniques is discussed in almost every past survey but none of the previous studies mentioned the term success ratio of used techniques. Concluding this, a very useful finding comes across that Regression has highest success ratio and thus can be claimed as efficient ML technique for performance prediction.

2.3.3 Predictor's Usage Analysis

This study aggregates data domains into three broad classes;

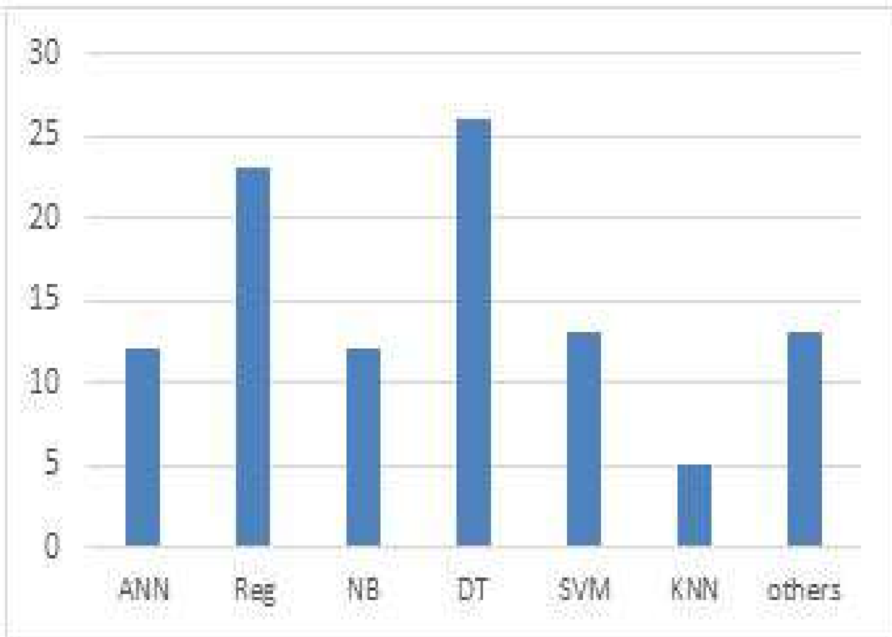

Figure 4. Usage Frequency of ML Techniques

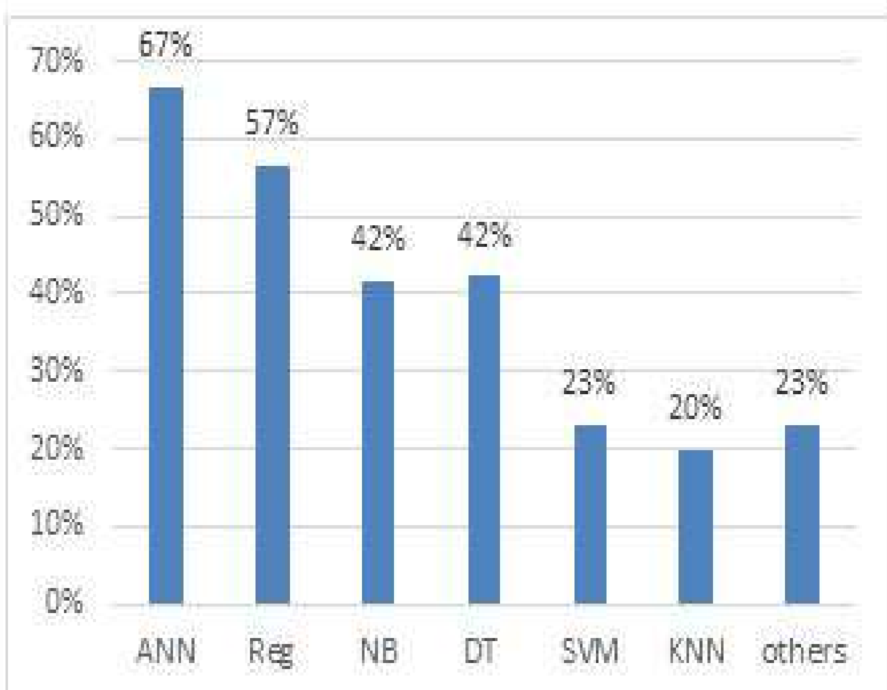

Figure 5. Success Ratio of ML Techniques

- Course Specific (include internal course assessments, online lecture watching behavior and other course related activities for OES)
- Academic (current CGPA, other academic features related to degree and P-Academic including high school, college grades, and Entry test marks etc.)
- Demographic (Gender, age, Traveling time, Employment status, financial status etc.)
- Others (include pedagogical and social behavioral data considered by only a few studies)

Table 2. : Usage Frequency of Data Domains

Data Domain	CS	Acad	Demo	Others
OES	10	5	5	0
TES	17	14	10	4
Total	27	19	15	4
% age	41%	29%	23%	6%

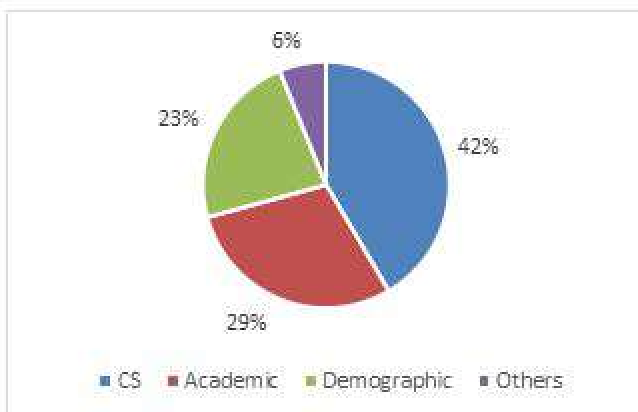

Figure 6. Usage Frequency of Data Domains

It is observed that CS data is most widely used data domain followed by academic and demographic data respectively. Usage frequency of these data domains for both education systems is provided in Table 3. The usage percentage of each of these data domain is shown in Figure 6.

It is also argued in the literature that CS data alone is not reliable for prediction [37]. As students are motivated for study in the beginning of semester and course contents are relatively easy in the beginning than post-midterm. Another important issue with using only CS data is that it considers initial assessments and midterm score thus it becomes too late to control failure after many assessments have been done. Therefore, combining previous academic records (both from previous institutions and past semesters (if applicable), especially grade report of pre-requisites is more important and reliable for effective prediction [49], [52]. Some studies claimed that good instructor-student communication level is a much stronger factor for earning good grades [54, 55]. Many researches, however, emphasized the importance of demographic data [37] and some used sensitive data like psychometric and behavioral data [29, 33] but using these data domains is not a practical approach in most of the cases because of their availability and validity.

Comparing our findings with few past reviews shows that they also possess similar distribution in terms of predictor's usage [52], [10] and [11]. Past surveys reported previous grades and CS data as the two most important predictors. Demographic and other academic parameters have second preference while psychometric parameters (student interest, study behavior, engage time etc.) are applied very rarely in predicting students' performance.

2.3.4 Evaluation Metrics Analysis

Going through the literature we identified different metrics for evaluation of prediction techniques. It was observed that most of the studies used Precision, Recall, Accuracy and F-Measure for evaluation of prediction model. Therefore, for convenience we collectively call them Standard. We can observe from the Table 4 and Figure 7 that Standard (Precision, Recall, Accuracy, F-Measure) have highest usage frequency for both education system i.e., 77%. While error rates (Root Mean Square Error (RMSE), Mean Absolute Error (MAE) etc.) used by 13% studies and 23% researchers used other metrics such as ROC (receiver operating characteristic) curve, and AUC (Area under the ROC Curve) etc.

Table 3. Usage Freq. of Evaluation metrics

Metrics	Standard	Error rate	Others
OES	5	3	2
TES	23	2	5
Total	28	5	7
Percentage	70%	12%	18%

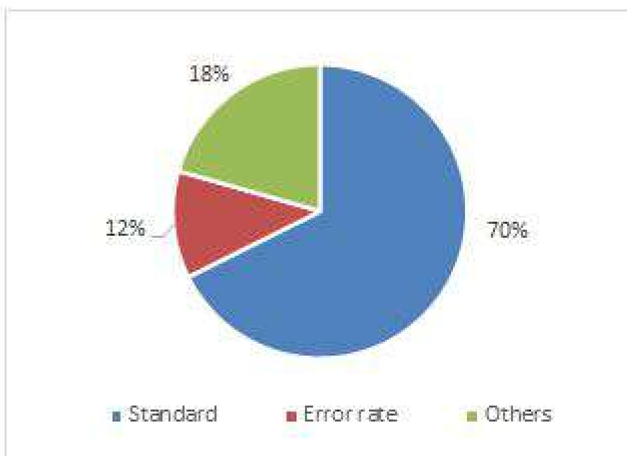

Figure 7. Usage Frequency of Evaluation metrics

These results help us in answering RQ-3 where Regression and DT proved to have highest usage frequency but regression possess highest success ratio to predict efficiently. In terms of predictors, the CS predictors for both education systems show highest usage, whereas the standard evaluation metrics are widely used to evaluate the ML techniques.

3 Discussion and Future work

This review presents the existing literature (from 2010 to 2022) in a new perspective of target education systems and granularity level of performance for undergraduate degree programs. It is observed that if student's performance is accurately predicted at lowest granularity level i.e., course level then it would be beneficial to improve student performance in current semester, upcoming semesters and finally for the complete degree program. By analyzing the hypothetical relevance of predictors, it can be concluded that predictors for every degree program differ. Like predictor for medical science, mathematics, computer science, law or humanities programs will not be the same. Another important observation is to discover critical courses in each degree program. Interestingly, for each degree program there exist group of courses that are interrelated in terms of concepts and learning and hence earning grades. Sometimes they are specified as Pre-requisites by the institution and sometimes they are not. Therefore, it is suggested to form course clusters for critical courses for a specific degree program and develop prediction models for each course cluster. Thus, further research work should be conducted by focusing on prediction models for critical course clusters rather than predicting individual course grades.

One of such course cluster for undergraduate Computer Science (BSCS) program is discovered as Programming Course Cluster. There are eight to ten compulsory programming courses that are required to pass to earn undergraduate computer science degree. Programming courses are considered challenging and authors in studies [56], [57], [58], [59] have emphasized the importance of failure rates of programming courses towards successful graduation. By looking at these facts, programming course cluster can be formed for CS degree programs. In the same pattern there exist potential course clusters for every degree program. No remarkable work is done (as per our knowledge) towards this perspective of performance prediction.

4 Conclusion

Objective of this review was to distinctively analyze performance prediction researches from a new perspective. Consequently, existing works is presented in a novel way by classifying target education systems and granularity level of performance prediction output for undergraduate degree programs. Two education systems were classified as Online Education System and Traditional Education system. The study also identifies three granularity levels for performance prediction output; Final CGPA, Next Semester GPA and Course Level performance prediction. It is revealed that Course level prediction have been widely performed in the past. Aim of the study was to provide a concise overview of this area using aggregate charts and findings within each category. It was observed that Decision Tree and Regression both possess high usage frequency but Regression algorithms proved to be efficient by calculating success ratio of the used ML techniques. Course specific predictors and standard evaluation metrics (Precision, Recall, Accuracy and F-measures) have highest usage frequency for work referenced here. The current review also discover potential for future research by emphasizing upon the importance of course clusters within each degree program at undergraduate level. Thus, presented work is significant in exploring new aspects of data and finding the gap that can be covered in future primary studies.

5 Credit Author Statement

Waqar Un Nisa: Conceptualization, Methodology. **Mudasser Naseer**: Writing-Original draft preparation, Supervision. **Muhammad Atif**: Data curation, Investigation, Validation. **Salwa Muhammad Akhtar**: Visualization. **Meher Un Nisa**: Writing- Reviewing and Editing

6 Compliance with Ethical Standards:

It is declared that all authors don't have any conflict of interest. It is also declared that this article does not contain any studies with human participants or animals performed by any of the authors. Furthermore, informed consent was obtained from all individual participants included in the study.

References

[1] S. Grajek and J. L. Grama, "Higher Education's 2018 Trend Watch and Top 10 Strategic Technologies," 2018.

[2] K. Pelletier. (2019). Key Issues in Teaching and Learning. Available: https://www.educause.edu/eli/initiatives/key-issues-in-teaching-and-learning

[3] E.-A. W. Group, "The Predictive Learning Analytics Revolution: Leveraging Learning Data for Student Success," 2015.

[4] K. Pelletier, M. McCormack, J. Reeves, J. Robert, and N. Arbino, "2022 EDUCAUSE Horizon Report - Teaching and Learning Edition," 2022.

[5] G. Fowler and L. W. Glorfeld, "Predicting Aptitude in Introductory Computing: A Classification Model," AEDS Journal, vol. 14, 01/01 1981.

[6] T. R. Hostetler, "Predicting student success in an introductory programming course," SIGCSE Bull., vol. 15, pp. 40-43, 1983.

[7] Y. Ma, B. Liu, C. K. Wong, P. S.Yu, and S. M. Lee, "Targeting the right students using data mining.," presented at the Proceedings of the 6th International Conference on Knowledge Discovery and Data Mining., Boston, Massachusetts, USA, 2000.

[8] B. a. W. P. Bidgoli, "Using genetic algorithms for data mining optimization in an educational web-based system," in Proceedings of Genetic and Evolutionary Computational Conference, , Chicago, Illinois, USA, 2003, pp. 2252–2263.

[9] B. A. Kitchenham and S. M.Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Keele University and University of Durham joint report2007.

[10] A. M. Shahiria, W. Husaina, and N. a. A. Rashida, "A Review on Predicting Student's Performance Using Data Mining Techniques," Procedia Computer Science, vol. 72), pp. 414 – 422, 2015.

[11] A. A. Saa, M. Al-Emran, and K. Shaalan, "Factors Affecting Students' Performance in Higher Education: A Systematic Review of Predictive Data Mining Techniques," Technology, Knowledge and Learning, April 25 2019.

[12] B. Kitchenham, D. Budgen, and P. Brereton. (2016). Evidence-Based Software Engineering and Systematic Reviews.

[13] Q. Qi, Y. Liu, F. Wu, X. Yan, and N. Wu, "Temporal models for personalized grade prediction in massive open online courses," presented at the Proceedings of ACM Turing Celebration Conference - China, Shanghai, China, 2018.

[14] H. S. Y. Aybek and M. R. Okur, "Predicting Achievement with Artificial Neural Networks: The Case of Anadolu University Open Education System," International Journal of Assessment Tools in Education, vol. 5, pp. 474-490, 2018.

[15] F. Okubo, T. Yamashita, A. Shimada, and H. Ogata, "A neural network approach for students' performance prediction," presented at the Proceedings of the Seventh International Learning Analytics 38; Knowledge Conference, Vancouver, British Columbia, Canada, 2017.

[16] S. J. H. Yang, O. H. T. Lu, A. Y. Q. Huang, J. C. H. Huang, H. Ogata, and A. J. Q. Lin, "Predicting Student's Academic Performance Using Multiple Linear Regression and Principal Component Analysis," Journal of Information Processin, vol. 26, pp. 170-176, 2018.

[17] J. Whitehill, K. Mohan, D. Seaton, Y. Rosen, and D. Tingley, "MOOC Dropout Prediction: How to Measure Accuracy?," presented at the Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale, Cambridge, Massachusetts, USA, 2017.

[18] J. He, J. Bailey, B. I. P. Rubinstein, and R. Zhang, "Identifying at-risk students in massive open online courses," presented at the Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, Texas, 2015.

[19] J. Bainbridge, J. Melitski, A. Zahradnik, E. J. M. Lauría, S. Jayaprakash, and J. Baron, "Using Learning Analytics to Predict At-Risk Students in Online Graduate Public Affairs and Administration Education," Journal of Public Affairs Education, vol. 21, pp. 247-262, 2015.

[20] S. M. Jayaprakash, E. W. Moody, E. J. M. Lauría, J. R. Regan, and J. D. Baron, "Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative," Journal of Learning Analytics, vol. 1, pp. 6-47, 2014.

[21] S. B. Kotsiantis, "Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades," Artificial Intelligence Review, vol. 37, pp. 331-344, April 01 2012.

[22] R. Barber and M. Sharkey, "Course correction: using analytics to predict course success," presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada, 2012.

[23] J. Xu, K. H. Moon, and M. v. d. Schaar, "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs," IEEE Journal of Selected Topics in Signal Processing vol. 11, pp. 742-753, 2017.

[24] M. A. Al-Barrak and M. Al-Razgan, "Predicting students final gpa using decision trees: a case study," Int. J. Inf. Educ. Technol., vol. 6, 2016.

[25] A. Ogunde and D. Ajibade, "A data mining system for predicting university students? graduation grades using id3 decision tree algorithm," J. Comput. Sci. Inf. Technol, vol. 2, pp. 21-46, 2014.

[26] Z. Alharbi, J. Cornford, L. Dolder, and B. D. L. Iglesia, "Using data mining techniques to predict students at risk of poor performance," presented at the SAI Computing Conference (SAI), London, 2016.

[27] A. Anthony and M. Raney, "Bayesian network analysis of computer science grade distributions," presented at the Proceedings of the 43rd ACM technical symposium on Computer Science Education, Raleigh, North Carolina, USA, 2012.

[28] S. Rovira, E. Puertas, and L. Igual, "Data-driven system to predict academic grades and dropout," PLoS ONE vol. 12, 2017.

[29] H. Bydžovská, "Towards Prediction and Recommendation in Higher Education," PhD, Faculty of Informatics, Masaryk University, 2016.

[30] A. Polyzou and G. Karypis, "Grade prediction with models specific to students and courses," International Journal of Data Science and Analytics, vol. 2, pp. 159-171, December 01 2016.

[31] M. Sweeney, H. Rangwala, J. Lester, and A. Johri, "Next-Term Student Performance Prediction: A Recommender Systems Approach," JEDM | Journal of Educational Data Mining,, vol. 8, pp. 22-51, 2016.

[32] Z. A. Iqbal, A. Qadir, J. Mian, A. Noor, and A. Kamiran, " Machine Learning Based Student Grade Prediction: A Case Study," J CoRR, vol. 1708.08744, 2017.

[33] Y. Yang, D. Hooshyar, M. Pedaste, M. Wang, Y.-M. Huang, and H. Lim, "Predicting course achievement of university students based on their procrastination behaviour on Moodle," Soft Computing, 2020/07/09 2020.

[34] K. Pelaez, R. Levine, J. Fan, M. Guarcello, and M. Laumakis, "Using a Latent Class Forest to Identify At-Risk Students in Higher Education.," JEDM | Journal of Educational Data Mining, vol. 11(1), pp. 18-46, 2019.

[35] V. Kumar and M.L.Garg, "Comparison of Machine Learning Models in Student Result Prediction," presented at the International Conference on Advanced Computing Networking and Informatics. Advances in Intelligent Systems and Computing, Singapore, 2019.

[36] A. P. Patil, K. Ganesan, and A. Kanavalli, "Effective Deep Learning Model to Predict Student Grade Point Averages," presented at the EEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, 2017.

[37] Y. Meier, J. Xu, O. Atan, and M. v. d. Schaar, "Predicting grades," presented at the IEEE Transactions on Signal Processing, 2016.

[38] F. Marbouti, H. A. Diefes-Dux, and J. Strobel, "Building course-specific regression-based models to identify at-risk students.," presented at the In The american society for engineering educators annual conference., Seattle, WA, 2015.

[39] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," Computer Education, vol. 103, pp. 1-15, 2016.

[40] K. Shakeel and N. A. Butt, "Educational Data Mining to Reduce Student Dropout Rate by Using Classification," in 253rd OMICS International Conference on Big Data Analysis Data Mining, Kentucky, USA, 2015.

[41] S. O. Abdulsalam, A. N. Babatunde, and R. S. Babatunde, "Comparative Analysis of Decision Tree Algorithms for Predicting Undergraduate Students' Performance in Computer Programming," Journal of Advances in Scientific Research Its Application (JASRA), vol. 2, pp. 79 - 92, 2015.

[42] A. K. Hamoud, A. S. Hashim, and W. A. Awad, "Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, pp. 26-31, 2017.

[43] Z. Mingyu, W. Sutong, W. Yanzhang, and W. Dujuan, "An interpretable prediction method for university student academic crisis warning," Complex Intelligent Systems, vol. 8, pp. 323-336, 2022.

[44] C. F. Rodríguez-Hernandez, M. Musso, E. Kyndt, and E. Cascallar, "Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation," Computers and Education: Artificial Intelligence, vol. 2, 2021.

[45] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," Smart Learning Environments, vol. 9, p. 11, 2022/03/03 2022.

[46] M. Bucos and B. Drăgulescu, "Predicting Student Success Using Data Generated in Traditional Educational Environments," TEM Journal, vol. 7, pp. 617-625, 2018.

[47] E. A. Majeed and K. N. Junejo, "Grade Prediction Using Supervised Machine Learning Techniques," in e-Proceeding of the 4th Global Summit on Education 2016 (GSE 2016), 2016, pp. 222-234, .

[48] S. T. Jishan, R. I. Rashu, N. Haque, and R. M. Rahman, "Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique," Decision Analytics, vol. 2, p. 1, March 12 2015.

[49] S. Huang and N. Fang, "Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. ," Computers Education, vol. 61, pp. 133-145, 2012.

[50] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," Expert Systems with Applications, vol. 41, pp. 1432-1462, 2014/03/01/ 2014.

[51] C. A. D. Río and J. A. P. Insuasti, "Predicting academic performance in traditional environments at higher-education institutions using data mining: A review," Ecos de la Academia, vol. 4, 2016.

[52] C. Del Río and J. Pineda Insuasti, "Predicting academic performance in traditional environments at higher-education institutions using data mining: A review," Ecos de la Academia, vol. 4, p. 2016, 12/31 2016.

[53] O. B. Coelho and I. F. Silveira, Deep Learning applied to Learning Analytics and Educational Data Mining: A Systematic Literature Review, 2017.

[54] L. Kennelly and M. Monrad, "Approaches to Dropout Prevention: Heeding Early Warning Signs with Appropriate Interventions," 2007.

[55] L. L. Pallock and S. D. Lamborn, "Beyond parenting practices: Extended kinship support and the academic adjustment of African American and European American teens.," Journal of Adolescence, vol. 26, pp. 813-828, 2006.

[56] . Bennedsen and M. E. Caspersen, "Failure rates in introductory programming," SIGCSE Bull., vol. 39, pp. 32-36, 2007.

[57] R. Bornat and S. Dehnadi, "Mental models, consistency and programming aptitude," presented at the Proceedings of the tenth conference on Australasian computing education Wollongong, NSW, Australia, 2008.

[58] C. Watson and F. W. B. Li, "Failure rates in introductory programming revisited," presented at the Proceedings of the 2014 conference on Innovation 38; technology in computer science education, Uppsala, Sweden, 2014.

[59] S. Bergin and R. Reilly, The influence of motivation and comfort-level on learning to program, 2019.