

Machine Learning and Student Performance Prediction

Subjects: Computer Science, Artificial Intelligence | Anthropology

Contributor: Boran ŞEKEROĞLU

Improving the quality, developing and implementing systems that can provide advantages to students, and predicting students' success during the term, at the end of the term, or in the future are some of the primary aims of education. Due to its unique ability to create relationships and obtain accurate results, artificial intelligence and machine learning are tools used in this field to achieve the expected goals.

Keywords: student performance prediction ; AI ; machine learning ; deep learning ; education

1. Introduction

Rather than aiming to mimic human intelligence, artificial intelligence (AI) seeks to create a sequence of processes that can implement skills that are inaccessible to human intelligence. However, as the human brain is described as the most complex machine in the universe, all created AI models are consistently compared with human intelligence and efforts have increased to create machines that can think, speak, and make decisions as humans do. Nevertheless, even though the first quarter of the 21st century has almost ended, AI is still mainly used to support humans and appears to be far from being implemented comprehensively.

The application areas of AI and its sub-field, machine learning (ML), are not limited to the subjects in which people have existing knowledge, but have spread to every field that provides new knowledge and developments in different business ^[1] or scientific areas ^{[2][3]}. In almost every field of engineering, social, and medical science, AI research contributes to humanity, science, or artificial intelligence studies, resembling both large and small bricks used to construct a building. In recent years, there has been increased interest in studies on AI in Educational Sciences (ES) as this building remains incomplete.

Education is the most crucial element that facilitates the development and progress of individuals and, consequently, countries. Novel studies are continually being conducted in Educational Sciences to accelerate this development and progress, make it more efficient, and provide new services. At the same time, the importance of AI in research and applications in education is increasing, as AI and ML applications and research are being performed in different ES fields to drive education ^{[4][5]}. Updating course content for teachers ^[6], enabling students to receive personalized education ^{[7][8]}, promoting smart course selection in higher education ^[9], and exam-paper preparation ^[10] are just a few examples of these applications.

In recent decades, attempts have been made to predict student performance, both during and at the end of term, employing the classification and regression skills of ML models by using the information obtained from questionnaires, demographic information ^{[11][12]}, or stored Massive Open Online Courses (MOOCs), mobile autonomous school (MAS) and blended course information ^{[13][14][15]}. These studies aimed to increase student achievement levels, provide more effective and individualized teaching strategies for students, and improve their learning skills by predicting the performance of students at the end of term ^[16], during the active semester ^[17], or by determining the risk levels ^{[18][19][20]}, drop-out probabilities ^{[21][22]}, or the factors that are most influential on student performance ^[23].

However, different machine learning models, varied datasets, and different evaluation and validation criteria were reported in these studies, which has led to difficulties in determining the direction of future research, level of success and efficiency, and knowledge provided by the ML models for this field.

2. Machine Learning and Student Performance Prediction: Critical Gaps and Possible Remedies

Artificial intelligence and machine learning have been studied in educational sciences as well as in all areas of life. However, research differs in terms of its purpose, the models used, datasets, evaluation criteria, and validation strategy.

From the point of view of machine learning, although every study contributes to and accelerates student performance evaluations, the differences make it challenging to implement these studies in real life.

Discussions on various subjects have been performed to provide a broad perspective on student performance prediction studies based on the results presented above.

Datasets have a significant impact on the aims of studies and can directly affect the domain of the study, the choice of ML model, and the evaluation metrics used.

Questionnaires cannot generally be used for in-term estimations because of their reliability and discrete training events when they are obtained. However, they continue to be considered in research on the impact of family and personal preferences on student achievements.

Online data have been widely used in end-of-term, in-term, and drop-out studies. The recording of students' interest in the courses at a certain level and as time-series in these datasets has enabled artificial intelligence and machine learning models to learn more meaningful data and produce more successful results. In addition, they have reduced the number of people that the surveys need to reach, the effort, and the financial costs by obtaining data in digital environments in the infrastructures of educational institutions.

Wang et al. ^[24] have taken the studies one step further by using a dataset that includes observations of students' activities and educational habits on campus. Therefore, in the most basic way, where and how much time students spend, the effects of the books they buy from the library, and other aspects could be observed, and student performance predictions could be performed accordingly.

The implementation of classification and regression tasks in student performance prediction studies is also related to the characteristics of the data collected and the aims of the studies. The frequently used Student Performance Dataset ^[25] allows researchers to perform investigations in regression and classification domains with the raw exam results provided. On the other hand, in other datasets ^{[11][26]}, the results were classified based on categorized performances or letter grades of students, which complicates the implementation of a regression study on these datasets. The data content might allow researchers to perform prediction studies in the classification domain (i.e., students' withdrawal from the course, end-of-term success: pass/fail) or in the regression domain (i.e., end-of-term success: exam point).

Regardless of whether the studies are in regression or classification, when they are categorized under the five main headings identified above (estimation of course drop-outs, end-of-term success level prediction, in-term success prediction, students' risk identification, and future success estimation), different questions might be asked from an educational perspective:

- Do drop-out predictions contribute significantly to the student and their success levels?
- When does the prediction of end-of-term success contribute to students' self-development and their education?
- Do the in-term performance predictions provide sufficient time to contribute to the students?
- How early can risk predictions related to courses taken by students contribute to them?

In addition to the effect of the datasets on the studies, reduced errors and more interpretable results in classification studies make them more applicable in this field. However, the analysis of individual results in regression tasks complicates the evaluation of the results since each sample has a unique error.

The results obtained from the systematic literature review showed that all the reasons described above caused the implementation of classification research (62%) to be significantly higher than regression studies (38%).

The considered datasets and problem domain affect the ML model selection in these studies. When the ML models are considered, it can be concluded that the ability of neural-based models to process and learn a considerable amount of data and produce successful results is an essential factor for consideration in most student performance prediction studies ^{[15][16][27]}. Furthermore, it has been observed that the use of recurrent neural networks, which can learn by remembering past experiences while learning data, especially in time-series data such as online datasets, has become prevalent ^{[18][24]}.

Even though the SVM and SVR models were frequently considered in the studies, the optimization of the classification/regression process is one of their crucial characteristics, while the projection of data into another feature space becomes more effective when the data are more informative. For this reason, SVM was generally considered in classification studies with a limited number of data, ensemble/hybrid models, comparative studies, or where attributes or instances were selected using data mining techniques [23][28][29]. Therefore, the classification and regression ability of SVM and SVR were optimized with minimized and selected data.

However, the uncertainties of neural-based models and SVM–SVR regarding interpretability led researchers to implement models with successful and interpretable results. At this point, DT and RF became the focus models of the researchers. Nevertheless, DT's sensitive approach to data and the risk of obtaining low results highlighted the RF, which optimizes the DT results by producing a certain number of DTs. Thus, the researchers attempted to achieve significant results and identify the factors that directly affect student performance [30][31].

The contribution of all machine learning models to student performance prediction studies is undeniable. The direct use of models, their use in data analysis and selection, and their use in creating hybrid or ensemble models have directed each study to lead to further developments.

The literature review results showed that analysis using deep learning has gained importance in recent years [8][24]. Additionally, the success of artificial neural network models in transferring the obtained knowledge to other models has started to be investigated in student performance predictions [32].

Since predicting which model would achieve superior results in artificial intelligence and machine learning applications is difficult, it is impossible to determine a specific model that would lead to future studies. However, the data created by developing computer and data storage technologies will make artificial neural networks and deep learning methods, which can process and learn big data, more widespread for regression and classification tasks. Furthermore, other conventional and tree-based models might be considered more frequently in the data analysis and data selection stages of studies on student performance evaluation. However, the models' abilities should not be underestimated for particular datasets with limited inputs and samples [33].

As mentioned above, the dataset, aims, and study domain directly affect the determination of evaluation metrics, which are different in both domains.

The most significant problem encountered in classification experiments is the accuracy obtained in experiments using imbalanced data. In a classification study, where a class with a large number of samples can be better learned, achieving a high accuracy result does not provide information about the efficiency of the other class results, and the accuracy causes misleading results. Therefore, the vast majority of students can be identified in classes, such as pass or fail, risky or not, etc. In imbalanced data, the F1 score and ROC AUC score are evaluated as metrics that show the model's overall performance more efficiently. In contrast, recall and precision metrics show the success level for specific classes and offer insight for learning the relevant classes. For this reason, the use of the accuracy metric alone might lead to difficulties in the real-life implementation of studies.

The standardization of the F1 and ROC AUC scores in all studies would contribute significantly to the studies' analysis, the success of models, and the trend of future studies for classification studies.

Similarly, the standardization of the R_2 score, which provides a scaled and consistent evaluation of the models, would provide a more effective evaluation of the proposed models. However, using additional metrics in regression problems is essential to measure error minimization since the results might differ from obtaining high

scores. In this case, it is challenging to determine which metric is more efficient, while the aim in measured data would be the determinant. Therefore, the use of a minimum of two error metrics, such as MAE and MSE (or root MSE), could be standardized for all student performance prediction studies.

In both problem domains, the applicability of the obtained results can be defined directly proportionally to the validation techniques.

In order to optimize the learning process, a method that is widely used is to select the instances to be used in training with different techniques. This could increase training and testing accuracy while reducing computational time in big data. However, in studies where attributes are selected rather than instances, it should be considered that each different

student entry contains new and independent data for both the training and test phase. Studies have shown that other training instances could change the accuracy by more than 10% [2][33].

The hold-out method does not use all samples during the training and testing phases. It also creates problems in determining the ratio of dividing the dataset into training, testing, or validation sets. Nevertheless, even the hold-out method is preferred to reduce the computational time for big data; computers today can perform a vast number of operations in student performance prediction studies.

For this reason, in studies where instances are not analyzed and selected, k-fold cross-validation, which considers each sample during the learning and testing phase, should be a standard approach for training the model. The average number of folds results should be evaluated as the overall success rate of the model. This will provide a more objective approach for evaluating the obtained results, and the actual abilities of the models will be determined as a result of the evaluations made by considering all the data.

Many different factors, such as changes in the conditions of student admission to educational institutions in countries, education abroad, foreign language level, education in the mother tongue, cultural differences, demographic structure, personal preferences, place of residence, weather conditions of the country of residence, health conditions, etc., and factors that directly affect the university lives of students, were reported in different studies [34][35][36]. Additionally, it is known that students' skills and their inclination and interest in courses also affect the level of success. If future studies are conducted without considering all these issues, they might experience difficulties in ensuring that the developed systems are applicable. Success only in specific predictions might mislead students and provide incorrect information to the instructors and experts who are responsible for determining education policies.

The spread of flipped classrooms, online education, and distance education during the COVID-19 period enabled the more effective recording of student data, logs, homework, quizzes, exams, feedback, and grades in many educational institutions, especially in HEI.

Educational institutions could enable more comprehensive studies in this mobile era by measuring students' time on campus, their time on social media, and achievement records with volunteer students by developing special mobile software. Combining all obtained behavioral and educational data with demographic, personal, and cultural data will provide a more accurate prediction of student performance.

Therefore, it will be possible to predict students' success in in-class and online courses more accurately by combining all the data, not only with questionnaires or online information. This will enable the generalization of AI and ML for both online and in-class courses, and artificial intelligence for student success in educational institutions will become more widespread.

Several studies have demonstrated that regional, national, and cultural differences, education in a foreign language, socioeconomic effects, demographic situation, and the role of instructors could have significant effects on the same scales in predicting student performance [37][38][39]. In addition, the developments in e-learning systems provide significant outcomes [40][41]. In this context, in the globalized world, creating a global education information consortium and the acquisition of data with the same criteria in different educational institutions worldwide will make reaching the goal non-specific and spread it across the globe. Moreover, it will undoubtedly create a tremendous source of information in determining students' performance in terms of success in the semester, at the end of the semester, and in the next semester, whether it is distance education, in-class education, open education, etc.

Today's computer technology provides the infrastructure to analyze big data that are obtained. This will make deep learning and data mining more focused on student performance prediction studies and might provide more accurate and applicable results than the results obtained so far.

Therefore, the drop-out and end-of-term predictions mentioned above as research topics will be provided as support points to the students. However, the high prediction success of in-term and risk estimations and the implementation of an early warning system will provide the most significant contribution. This will allow students to observe the risk levels and expected success level at the end of the term after enrolling in the course, and the students' interest in the course will be increased.

Furthermore, instead of reaching general judgments in predicting students' performance levels, it would be possible to implement personalized and individualized performance evaluation systems and to further the meta-analysis results of Schneider and Preckel's [42] studies for HEI.

References

1. Khashman, A.; Carstea, C. Oil price prediction using a supervised neural network. *Int. J. Oil Gas Coal Technol.* 2019, 20, 360.
2. Sekeroglu, B.; Tuncal, K. Prediction of cancer incidence rates for the European continent using machine learning models. *Health Inform. J.* 2021, 27, 1460458220983878.
3. Ozcil, I.; Esenyel, I.; Ilhan, A. A Fuzzy Approach Analysis of Halloumi Cheese in N. Cyprus. *Food Anal. Methods* 2021.
4. Chen, L.; Chen, P.; Lin, Z. Artificial Intelligence in Education: A Review. *IEEE Access* 2020, 8, 75264–75278.
5. Perrotta, C.; Selwyn, N. Deep learning goes to school: Toward a relational understanding of AI in education. *Learn. Media Technol.* 2019, 45, 1–19.
6. Guan, C.; Mou, J.; Jiang, Z. Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *Int. J. Innov. Stud.* 2020, 4, 134–147.
7. Somasundaram, M.; Junaid, K.; Mangadu, S. Artificial Intelligence (AI) Enabled Intelligent Quality Management System (IQMS) For Personalized Learning Path. *Procedia Comput. Sci.* 2020, 172, 438–442.
8. Liu, J.; Loh, L.; Ng, E.; Chen, Y.; Wood, K.; Lim, K. Self-Evolving Adaptive Learning for Personalized Education; Association for Computing Machinery: New York, NY, USA, 2020; pp. 317–321.
9. Tilahun, L.; Sekeroglu, B. An intelligent and personalized course advising model for higher educational institutes. *SN Appl. Sci.* 2020, 2, 1635.
10. Wu, Z.; He, T.; Mao, C.; Huang, C. Exam Paper Generation Based on Performance Prediction of Student Group. *Inf. Sci.* 2020, 532, 72–90.
11. Yilmaz, N.; Sekeroglu, B. Student Performance Classification Using Artificial Intelligence Techniques. In *Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2020; Volume 1095.
12. Zaffar, M.; Hashmani, M.; Savita, K.; Sajjad, S.; Rehman, M. Role of FCBF Feature Selection in Educational Data Mining. *Mehran Univ. Res. J. Eng. Technol.* 2020, 39, 772–778.
13. Jiang, P.; Wang, X. Preference Cognitive Diagnosis for Student Performance Prediction. *IEEE Access* 2020, 8, 219775–219787.
14. Gitinabard, N.; Xu, Y.; Heckman, S.; Barnes, T.; Lynch, C. How Widely Can Prediction Models Be Generalized? Performance Prediction in Blended Courses. *IEEE Trans. Learn. Technol.* 2019, 12, 184–197.
15. Gamulin, J.; Gamulin, O.; Kermek, D. Using Fourier coefficients in time series analysis for student performance prediction in blended learning environments. *Expert Syst.* 2015, 33.
16. Aydogdu, S. Predicting student final performance using artificial neural networks in online learning environments. *Educ. Inf. Technol.* 2020, 25, 1913–1927.
17. Zhao, L.; Chen, K.; Song, J.; Zhu, X.; Sun, J.; Caulfield, B.; Namee, B. Academic Performance Prediction Based on Multisource, Multifeature Behavioral Data. *IEEE Access* 2021, 9, 5453–5465.
18. He, Y.; Chen, R.; Li, X.; Hao, C.; Liu, S.; Zhang, G.; Jiang, B. Online At-Risk Student Identification using RNN-GRU Joint Neural Networks. *Information* 2020, 11, 474.
19. Mengash, H. Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems. *IEEE Access* 2020, 8, 55462–55470.
20. Yang, J.; Devore, S.; Hewagallage, D.; Miller, P.; Ryan, Q.; Stewart, J. Using machine learning to identify the most at-risk students in physics classes. *Phys. Rev. Phys. Educ. Res.* 2020, 16, 020130.
21. Figueroa-Cañas, J.; Sancho-Vinuesa, T. Early Prediction of Dropout and Final Exam Performance in an Online Statistics Course. *IEEE Rev. Iberoam. Tecnol. Aprendiz.* 2020, 15, 86–94.
22. Xing, W.; Du, D. Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention. *J. Educ. Comput. Res.* 2018, 57, 073563311875701.
23. Injadat, M.; Moubayed, A.; Nassif, A.; Shami, A. Multi-split optimized bagging ensemble model selection for multiclass educational data mining. *Appl. Intell.* 2020, 50, 4506–4528.
24. Wang, X.; Yu, X.; Guo, L.; Liu, F.; Xu, L. Student Performance Prediction with Short-Term Sequential Campus Behaviors. *Information* 2020, 11, 201.

25. Cortez, P.; Silva, A. Using Data Mining to Predict Secondary School Student Performance. In Proceedings of the 5th Future Business Technology Conference (FUBUTEC 2008), Porto, Portugal, 9–11 April 2008; pp. 5–12, ISBN 978-9077381-39-7.
26. Kuzilek, J.; Hlosta, M.; Zdrahal, Z. Open University Learning Analytics dataset. *Sci. Data* 2017, 4, 170171.
27. Adejo, O.; Connolly, T. Predicting student academic performance using multi-model heterogeneous ensemble approach. *J. Appl. Res. High. Educ.* 2017, 10, 61–75.
28. Lu, H.; Yuan, J. Student Performance Prediction Model Based on Discriminative Feature Selection. *Int. J. Emerg. Technol. Learn. (IJET)* 2018, 13, 55.
29. Tran, O.; Dang, H.; Thuong, D.; Truong, T.; Vuong, T.; Phan, X. Performance Prediction for Students: A Multi-Strategy Approach. *Cybern. Inf. Technol.* 2017, 17, 164–182.
30. Shanthini, A.; Vinodhini, G.; Chandrasekaran, R. Predicting Students' Academic Performance in the University Using Meta Decision Tree Classifiers. *J. Comput. Sci.* 2018, 14, 654–662.
31. Wakelam, E.; Jefferies, A.; Davey, N.; Sun, Y. The potential for student performance prediction in small cohorts with minimal available attributes. *Br. J. Educ. Technol.* 2019, 51, 347–370.
32. Tsiakmaki, M.; Kostopoulos, G.; Kotsiantis, S.; Ragos, O. Transfer Learning from Deep Neural Networks for Predicting Student Performance. *Appl. Sci.* 2020, 10, 2145.
33. Ever, Y.; Dimililer, K.; Sekeroglu, B. Comparison of Machine Learning Techniques for Prediction Problems. In *Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2019; Volume 927.
34. Balci, S.; Ayhan, B. Internet usage patterns among university students. *J. Selcuk Commun.* 2007, 5, 174–197.
35. Bodovski, K.; Jeon, H.; Byun, S. Cultural capital and academic achievement in post-socialist Eastern Europe. *Br. J. Sociol. Educ.* 2017, 38, 887–907.
36. Richardson, M.; Abraham, C.; Bond, R. Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychol. Bull.* 2012, 138, 353–387.
37. Boz, Y.; Boz, N. Prospective chemistry and mathematics teachers' reasons for choosing teaching as a profession. *Kastamonu Educ. J.* 2008, 16, 137–144.
38. Kayalar, F.; Kayalar, F. The effects of Auditory Learning Strategy on Learning Skills of Language Learners (Students' Views). *IOSR J. Humanit. Soc. Sci. (IOSR-JHSS)* 2017, 22, 4–10.
39. Memduhoğlu, H.; Tanhan, F. Study of organizational factors scale's validity and reliability affecting university students' academic achievements. *YYU J. Educ. Fac.* 2013, X, 106–124.
40. Franzoni, V.; Pallottelli, S.; Milani, A. Reshaping Higher Education with e-Studium, a 10-Years Capstone in Academic Computing. *Lect. Notes Comput. Sci.* 2020, 12250, 293–303.
41. Franzoni, V.; Tasso, S.; Pallottelli, S.; Perri, S. Sharing Linkable Learning Objects with the Use of Metadata and a Taxonomy Assistant for Categorization. *Lect. Notes Comput. Sci.* 2019, 11620, 336–348.
42. Schneider, M.; Preckel, F. Variables Associated With Achievement in Higher Education: A Systematic Review of Meta-Analyses. *Psychol. Bull.* 2017, 143, 565–600.

Retrieved from <https://encyclopedia.pub/entry/history/show/40132>