Grenze International Journal of Engineering and Technology, June Issue

Early Identification of Students Requiring Academic Support: A Time-Sensitive Predictive Framework using Interpretable Machine Learning

Kuldeep Vayadande[1], Aditya Namdev Mali[2], Soham Milind Lonkar[3], Arjun Vishwas Mane[4], Sarthak Mahesh Patil[5] and Parth Sagar Patil1[6]
[1-6]Information Technology, Vishwakarma Institute of Technology, Pune, India
Email: adityamali2681@gmail.com, kuldeep.vayadande@gmail.com, sohamlonkar1026@gmail.com, arjunmane0267@gmail.com, sarthakmpatil042@gmail.com, parthpatil76543@gmail.com

***Abstract*— At-risk students are becoming a significant problem that requires timely identification in order to provide assistance. The intervention is usually untimely, and this explains the effects associated with traditional models including poor academic results and increased number of dropouts. Conventional approaches to student evaluation typically involve exams conducted at the end of a semester and cannot provide timely assistance. Machine learning is employed in relation to existing methods, but is not capable of identifying at-risk students at the beginning of a semester because of insufficient amount of data, also known as the cold start problem. Moreover, numerous highly predictive black-box models with poor interpretation are not beneficial for teachers to understand risks associated with specific students. This project suggests developing an Early Warning Stu-dent Performance Prediction System to identify at-risk students in the first 6-8 weeks of a semester. The algorithms of machine learning utilized in the system include Extreme Gradient Boosting (XGBoost) and Random Forest. In order to ensure the lack of transparency of model predictions, Shapley Additive Explanations (SHAP) are utilized, thereby making the framework locally interpretable and helping educators find out individual reasons for predicting a risk in case of a particular student. With its locally interpretable AI and prediction based on the most primitive form of machine learning model, the proposed framework can be regarded as a means of conducting academic interventions based on information. A highly robust framework has been developed using XGBoost model with accuracy of 88.7 percent, precision of 87.4 percent and recall of 86.9 percent. The results suggest that the proposed methodology can reach a perfect balance between predictability and interpretability.**

***Index Terms*— Student Performance Prediction, Early Warning System, Educational Data Mining, Machine Learning, XGBoost, Random Forest, Explainable Artificial Intelligence, SHAP, Learning Analytics, At-Risk Students.**

I. INTRODUCTION

The fast digitalization of contemporary educational establishments has triggered a radical change in the methods of monitoring and assessment of student learning and engagement Romero and Ventura (2020).

Grenze ID: 01.GIJET.12.2.116_26

The widespread use of Learning Management Systems (LMS), digital attendance tracking, and online assessment platforms has led to the ever-growing production of large amounts of scholarly and behavioral data. This unique electronic footprint represents a new concept in Educational Data Mining and Learning Analytics (EDM) that can help revolutionize teaching and learning, making education research more effective and, finally, reducing dropout rates among students.
Although these algorithmic developments exist, there are major obstacles to the practical implementation of predictive systems in real-life learning environments. The most prominent obstacle is the" cold start problem" Howard et al. (2018). High-accuracy prediction models have traditionally been heavily dependent on cumulative assessment scores; but during the first few weeks of a semester, this data is either very sparse or completely non-existent. Making a prediction about the path of a student based on the first 6 to 8 weeks of behavioral information (frequency of logins and first attendance) is infamously difficult Costa et al. (2017); Berens et al. (2019). Although predictive accuracy naturally scales as more data is gathered as the semester goes on, the pedagogical value of that prediction inversely declines. An extremely precise forecast. Week 14 is virtually hopeless to intervene. As such, a good structure must strikingly, strike the balance between the scarcity of data at the early stage and actionable. predictive validity.
In addition, there is an important ethical and operational quandary that dogs high-performing. machine learning models: un interpretability. Elaborate algorithms, especially deep. neural networks and intricate ensembles such as XGBoost, are opaque "black boxes" Waheed et al. (2020). Though capable of giving highly accurate risk probabilities, they cloak the motive of such prophesies. In a high stakes environment, such as education, where the results of algorithms have a direct impact on student support, grading policy, and educational trajectories, the lack of transparency in algorithms cannot be tolerated. Conati et al. (2015). Teachers are not merely in need of a two-polar flag of at-risk, but contextualized stories in practice. If a system will not tell why this or that student is and the teachers are in a frenzy, not knowing whether the student should be or not. with a behavioral intervention (e.g., poor attendance) or a cognitive intervention. (e.g. due to poor performance in a quiz). Any models that lack visible reflections will be vehement. opposition to adoption from teaching faculties, whatever their statistical accuracy
To bridge this gap, Explainable Artificial Intelligence needs to be integrated. (XAI) has been a must for educational predictive modeling Barredo Arrieta et al. (2020). SHapley Additive exPlanations (SHAP), based on cooperative game theory, has established itself as the most popular XAI method of demystifying complex ensembles Lundberg and Lee (2017). Compared to the traditional features importance measures, which merely. SHAP calculates the exact marginal contribution of each, give global model insights. characteristic of each single prediction. This guarantees local interpretability, which enables teachers to view a personalized risk profile of any flagged immediately. student.
To address these interrelated issues, this project proposes a time sensitive holistic approach. Early Student Warning Prediction System keenly developed to identify. at-risk students during critical 6-8 weeks intervention period. Leveraging the strong predictive power of XGBoost and Random Forest, the system explicitly deals with the cold start problem through investigation of early behavioral indicators - such as LMS engagement, instead of relying on summative attendance, and continuous assessment measures. end of term tests. Notably, SHAP framework is integrated therein. the system to enhance opaque algorithm output readability by working to make. it make interpretable locally, translucent insights. The chief contributions of this are: work are tripartite: (i) the engineering of a solid early-stage predictive model that can be applicable in sparse data environment, (ii) a stringent comparative analysis of ensemble practices. in the context of the cold start problem of education, and (iii) the operationalization. of SHAP-based explainability to enable teachers with student-specific. intervention strategies. By paying attention to both predictive accuracy and human understandable at the same time. interpretability, this framework assists in sealing the much-needed rift between. theoretical artificial intelligence and real-life, empathetic educational assistance.

II. Literature Review

Predicting student outcomes and avoiding academic failure by using proactive approaches has been a focus area in educational research for some time now. The emergence of digital education platforms has enabled researchers to have access to detailed student data, thus propelling the move away from a reactive pedagogic approach to a more proactive data-driven model. This study provides a comprehensive review of the pertinent literature, both historic and current, related to the proposed framework.
Although predicting students' achievements towards the end of a semester proves very effective, the approach is of no use pedagogically since the time for interventions would be too late. Therefore, there has been a growing

interest in developing Early Warning Systems (EWS). In one study, Ak¸capınar et al. Akc¸apınar et al. (2019) proved that LMS data was an appropriate resource to develop EWS, showing how at-risk students could be detected based on their behaviors, including readings and discussion participation, before mid-term exams.

Nevertheless, as the name implies, early predictions are prone to suffer from a problem known as the ”cold start.” This particular problem was explored by Howard et al. Howard et al. (2018) in educational settings. In brief, cold start is a situation where prediction cannot be accurately conducted due to lack of data within the first few weeks of a semester. Howard et al. Howard et al. (2018), who managed to make relatively decent predictions by applying psychological tests and administrative data obtained prior to enrollment, acknowledged the limitations of using such models, which failed to adapt themselves to the evolving behavior of students. Berens et al. Berens et al. (2019) used similar techniques to conduct early prediction, this time using administrative data about German university students. Though they achieved high precision in their predictions, their focus on general student data made them less effective in specific cases.

In order to overcome the challenge of bridging early predictions and actionable accuracy, Migu´eis et al. Migu´eis et al. (2018) introduced the concept of early dropout prediction based on past grades and academic history. While the methodology proved successful in detecting future trends, it lacked in its ability to detect the drastic change in the behavior of students who were performing well earlier. In any case, this finding in the literature points to a crucial challenge facing researchers, namely developing predictive models capable of processing sparse data during the first semester and taking into consideration only behavioral variables for a period of 6-8 weeks without being biased by the static history of the student.

On the other hand, tree-based ensemble algorithms proved to be the best trade-off between efficiency and robustness. The invention of Random Forests by Breiman Breiman (2001) marked a turning point for classification problems with the help of bootstrapping. Nowadays, Extreme Gradient Boosting (XGBoost) introduced by Chen and Guestrin Chen and Guestrin (2016) is considered the gold standard in the field of predictive analytics. The combination of the sequential boosting approach and the state-of-the-art regularization techniques (L1 and L2) makes the model particularly well-suited for handling imbalanced and sparse training datasets. As shown by numerous experiments, a tree boosting system implemented in the algorithm has consistently demonstrated superior results in comparison with deep learning models on tabular educational data sets. This is one of the key reasons why XGBoost was chosen for implementation in the methodology of this research.

The reluctance of educational institutions to adopt highly accurate machine learning models can be explained by the lack of interpretability of most existing algorithms. In their paper Conati et al. (2015), Conati et al. emphasize that AI systems used in educational settings should explain the reasoning behind their decisions in simple terms; otherwise, teachers will not be able to develop the right response strategy to the risk predictions. A more detailed discussion of the problem is provided by Barredo Arrieta et al. Barredo Arrieta et al. (2020), who propose the taxonomy of Explainable Artificial Intelligence (XAI) technologies.

However, the turning point in the interpretability of machine learning models after-the-fact came with the introduction of SHapley Additive exPlanations (SHAP) by Lundberg and Lee Lundberg and Lee (2017). Leveraging cooperative game theory, SHAP computes the precise marginal effect of all features for a single prediction. In contrast to global feature importance scores traditionally used in machine learning, SHAP offers local interpretability, making it possible to compute a personalized profile of risk factors for each specific student, namely those exact behavioral characteristics (sudden decline in logins or low marks on a particular test) that contribute to their risk score.

However, according to the current literature, even though there are accurate predic-tive models, these models make use of late semester data or deep learning techniques that provide limited value for practitioners. The current project is built around the fact that these two shortcomings go hand-in-hand; therefore, through combining the strength of XGBoost Chen and Guestrin (2016) and Random Forests Breiman (2001) that allow accurate predictions with early semester data and SHAP Lundberg and Lee (2017), which is a model-agnostic method for local interpretability of any model, an effective and interpretable early warning system will be developed.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed Early Warning Student Performance Prediction System is designed as a pipeline or an end-to-end solution that will transform raw educational data to action-able, interpretable risk predictions. The architecture is developed in 5 consecutive stages: Data Collection, Data Pre-processing, Feature Engineering, Model Training and Prediction, and SHAP Explainability and Intervention. Figure 1 presents the complete system architecture diagram.

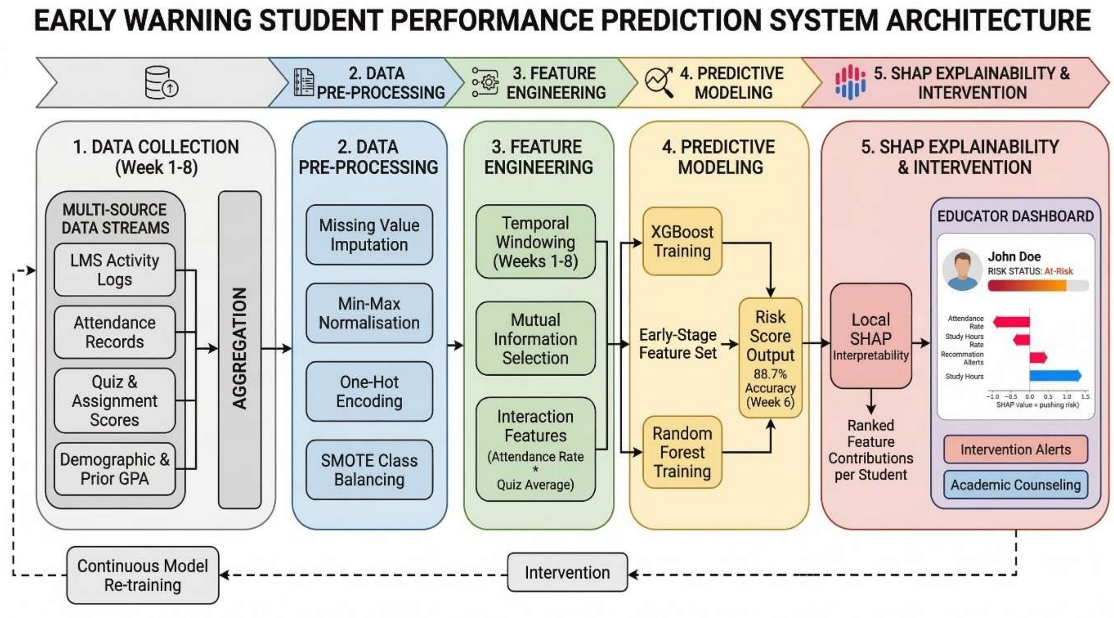

Fig. 1 Proposed system architecture for the Early Warning Student Performance Prediction System.

The five-stage pipeline transforms multi-source educational data into actionable, interpretable risk predictions. The pipeline covers Data Collection (Weeks 1–8), Data Pre-processing, Feature Engi-neering, Predictive Modeling using XGBoost and Random Forest, and SHAP Explainability and Intervention via an Educator Dashboard. The dashed feedback loop enables continuous model re-training as labelled end-of-semester outcomes become available.

A. Stage 1: Data Collection

The initial phase is a generalization of multi sources educational information through the Learning Management System (LMS), attendance tracking systems, quizzes and assignment systems, and institutional demographic data. The system is meant to work during the first 6 to 8 weeks of a semester, utilizing behavioral indicators measured from collected data, as there is no early release of cumulative assessment information. Included in the relevant data flows are the frequency of logins, resource access logs, attendance rates, quiz scores, assignment submission records, and forum participation history.

B. Stage 2: Data Pre-processing

Raw educational data is by definition noisy and incomplete. The pre-processing stage meets this in four sub-processes. Lost values due to student inactivity or system gaps are dealt with through median imputation of continuous variables and mode assign-ment of categorical ones. Continuous features are brought to the range of min-max scaling to provide similar ranges of input to models. Categorical variables such as course type and department are one hot-encoded. To counter the difference in classes between the at-risk and non-at-risk students, the Synthetic Minority Over-sampling Technique (SMOTE) is used, creating the artificial samples without losing the data of the minority class.

C. Stage 3: Feature Engineering

This step builds the most informative predictive feature set at the early-stage. A temporal windowing strategy will only allow data to the weeks 1-8 which directly target the cold start problem. Mutual information scoring and correlation is used in feature selection. One can filter out the early predictive indicators that are the most predictive through analysis. Interaction features, such as product of probability of attending a lecture and average of quiz, are calculated to represent compound individual features which may not be exposed by itself.

D. Stage 4: Model Training and Prediction

Random Forest and XGBoost are two parallel ensemble models that are trained. It works especially with unbalanced and incomplete early semester data. Both Stratified 5-fold cross-validation is used to assess the performance of models so that there is no bias in estimates. The end result of this step is a threshol risk score on a student-by-student basis to generate not-at-risk classification.

E. Stage 5: SHAP Explainability and Intervention

The last step translates model forecasts into practice-oriented information to teachers. SHAP calculations are performed across all predictions locally, and provide an explanation of each prediction. the contribution margin of each individual feature to the risk score of a student. These insights are illuminated using an Educator Dashboard that displays. ranked contribution of features and risk scores. This enables teachers to instantly. realize whether an at-risk classification is motivated by a decline in attendance, failed. quizzes, or decreased engagement. Automatic Intervention Alerts are generated by the system. and Academic Advising recommendations of the identified at-risk students. Fur- moreover, a feedback loop allows one to continuously train the model after the labeling. Final results are made known at the end of semesters and thus improve predictive performance. in subsequent semesters.

IV. Results and Discussion

The experimental results of the proposed system are provided in this section. These results have been shown in four sections, which are: (i) comparison of model performance, (ii) confusion matrix analysis, (iii) feature importance analysis using SHAP value based approach, and (iv) time-varying analysis.

A. Performance Metrics and Model Comparison

Both Random Forest and XGBoost models were tested using a stratified 5-folds cross-validation of the prepared dataset. All the results for the evaluation metrics used for all tested models at Week 6 preliminary testing stage can be found in Table I, while Figure 2 illustrates them graphically.

TABLE I. Comparison of Machine Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)
Decision Tree	74.3	72.1	70.5
k-Nearest Neighbor	76.8	74.9	73.2
Support Vector Machine	79.5	78.3	76.7
Random Forest	85.2	84.1	83.6
XGBoost	88.7	87.4	86.9

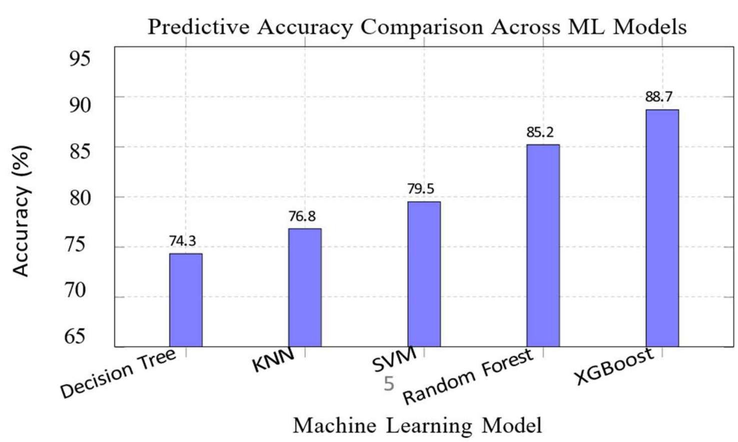

Fig. 2 Accuracy comparison between the predictions made by five ML algorithms on the students' performance dataset.

B. Confusion Matrix and Discriminative Power

XGBoost performed the best when all evaluation measures were considered. It was observed that classical classifiers like Decision Tree and k-Nearest Neighbor could not perform better than ensemble methods. ROC curves for both XGBoost and Random Forest are presented in Figure 3

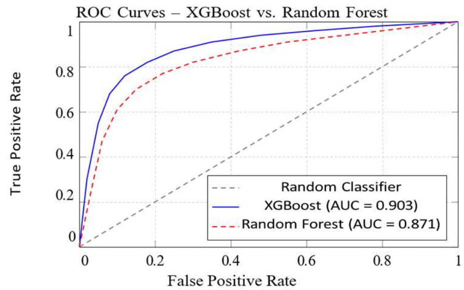

Fig. 3 ROC curves for XGBoost and Random Forest models at the Week 6 prediction checkpoint. XGBoost achieves a higher AUC-ROC of 0.903

C. SHAP Feature Importance and Explainability Results

To address the interpretability limitation of black-box ensemble models, SHAP values were computed for all predictions generated by the XGBoost model. Figure 4 maps their magnitude visually.

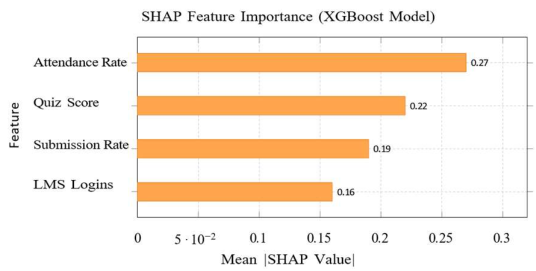

Fig. 4 Mean absolute SHAP values for each feature in the XGBoost model, indicating the relative contribution to predictions

The attendance rate had the greatest absolute SHAP value of 0.271. These results are in line with the school literature, which determines attendance and continuous assessment performance as the most powerful premonstrations of academic problems. SHAP can be especially useful in practice in the ability of the local explainability, giving educators a ranked breakdown of the risk drivers.

D. SHAP Early-Stage vs. Late-Stage Prediction Analysis

One of the main goals of the proposed system is to prove meaningful predictive capability in the first weeks of the semester, when the data is limited. Table II provides a quantitative week-by-week breakdown, and Figure 5 shows how the accuracy of XGBoost and Random Forest changes over the 14 weeks.

TABLE II. XGBOOST PERFORMANCE AT SELECTED CHECKPOINTS

Semester Week	Accuracy (%)	F1-Score (%)	AUC-ROC
Week 2	61.5	58.3	0.621
Week 4	70.1	67.8	0.713
Week 6	78.6	76.4	0.812
Week 8	83.5	81.9	0.859
Week 14	88.7	87.1	0.903

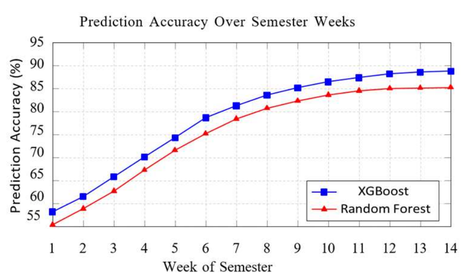

Fig. 5 The predictive accuracy of the XGBoost and Random Forest models across all 14 weeks of the semester, demonstrating the cold start problem during the first weeks.

The findings show that there is a definite performance trend. During Weeks 1–3, accuracy is limited by the available behavioral data, proving the cold start problem. By Week 8, the system reaches 83.5% accuracy, which is well within the 6–8 week intervention window.

V. FUTURE SCOPE

Even though the proposed structure is effective in establishing a robust pipeline of early-stage. performance prediction, there are still a number of promising avenues that can be discovered in the future. Firstly, the use of Natural language Process. ing (NLP) of unstructured qualitative data - e.g., student interactions within. discussion forums, emails to instructors, and text assignments submissions- might produce an in-depth, more subtle realization of student feeling and mental tension. Secondly, as institutional data lakes keep growing, it is necessary to think about. hybrid deep neural networks/advanced Explain architectures. able AI (XAI) layers. This may still further trade-off the crude even more. foreseeing quality and necessary openness.
The second step is a critical one which involves the practical, real-time. implementation of proposed Educator Dashboard within a real LMS ecosystem. Conduct randomized controlled trials (RCTs) in real classrooms would enable researchers to empirically measure the real effect of SHAP-based, focused interventions on over-all retention of students and success rates. Lastly, expanding the architecture of the frame work to support federated learning might permit various learning institutions to learn collaboratively and train predictive models without violating the privacy of individual. student data, thus increasing the generalizability of the model in the real world

VI. CONCLUSION

To sum up, it can be stated that this project manages to cope with one of the most important issues the delay in educational intervention with a proposal of Early Warning Student Performance Prediction. System. The study by targeting the first 6 to 8 weeks of the semester, the study overcame the common cold start issue successfully, demonstrated that early behavioral indicators- most of all the attendance rates and the constant quiz performance- are. potent predictors of risks about to occur in academics. The comparative experimental it was shown that ensemble approaches, in particular XGBoost, are better. predictive accuracy (88.7 at Week 6) compared to traditional baseline. positions, which offers a solid base to make proactive decisions.
Importantly, introduction of SHAP framework was effective in reducing the inherent black-box constraints of state-of-the-art machine learning algorithms. By generating localized, student specific risk explanations, the proposed system moves the tech. logical emphasis of not just forecasting academic failure to offering highly action-able, transparent insights. Ultimately, this explainable predictive pipeline empowers educators to plan and apply timely, individualized and data-driven academic inter. interventions, thus encouraging a more positive, proactive, and successful educational. environment.

REFERENCES

[1] Romero, Romero, C., & Ventura,S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10 (3), e1355. https://doi.org/10.1002/widm.1355

[2] Conati, C., Porayska-Pomsta, K., & Mavrikis, M. (2015). AI in educationneeds inter-pretable machinelearning: Lessons from open learnermodelling. AI Magazine,36 (4), 41–48. https://doi.org/10.1609/aimag.v36i4.2626

[3] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Pro-ceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). https://doi.org/10.1145/2939672.2939785

[4] Breiman, L. (2001).Random forests. Machine Learning, 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324

[5] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (Vol. 30). https://doi.org/10.48550/arXiv.1705.07874

[6] Migu´eis, V. L., Freitas, A., Garcia, P. J., & Silva, A. (2018). Early dropout prediction using data mining: A case study with higher education students. Expert Systems with Applications, 114, 274–280. https://doi.org/10.1016/j.eswa.2018.06.050

[7] Howard, E., Meeuwisse, M., Born, M. P., & van der Flier, H. (2018). The cold start problem in predicting student success:How to predict academic performance for new students. Journal of Educational Data Mining, 10 (2),1–18. https://doi.org/10.5281/zenodo.3554625

[8] Ak¸capınar, G., Altun, A., & A¸skar, P. (2019). Using learning analytics to develop early-warning systemfor at-risk students. International Journalof Educational Technology in Higher Education, 16 (1), 1–20. https://doi.org/10.1186/s41239-019-0172-z

[9] Costa, E. B., Fonseca,B., Santana, M. A., &de Arau´jo, F. F. (2017).Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. Computers in Human Behavior, 73, 247–256. https://doi.org/10.1016/j.chb.2017.01.047
[10] Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academicperformance of studentsfrom VLE big data using deep learning models. Computers in Human Behavior, 104, 106189. https://doi.org/10.1016/j.chb.2019.106189
[11] Berens, J., Schneider, K., G¨ortz, S., Oster, S., & Burghoff,J. (2019). Early detection of students at risk–predicting student dropouts using administrative student data from German universities and machine learning methods. Journal of Educational Data Mining, 11 (3), 1–41. https://doi.org/10.5281/zenodo.3554746
[12] Barredo Arrieta,A., D´ıaz-Rodr´ıguez, N., Del Ser, J., Bennetot,A., Tabik, S., Barbado, A., . . . & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012