Student Academic Risk Prediction and Personalized Learning Recommendation Research Based on User Portraits of Learning Situation

Haochen Jiang*

Academy of Artificial Intelligence and Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China
**Corresponding Author. Email: Haochen.Jiang24@student.xjtlu.edu.cn*

Abstract Early identification of students with academic difficulties is challenging in traditional teaching scenarios. This study builds an academic-risk early-warning framework based on educational data mining on the Open University Learning Analytics Dataset (OULAD). After missing-value handling and min–max normalization of online learning records, key features such as learning duration, assignment completion rate and test scores are extracted to establish multi-dimensional student profiles, from which low-, medium- and high-risk labels are constructed. A classification model (Extreme Gradient Boosting (XGBoost) combined with the multi-dimensional profile features) is applied for academic-risk prediction, and the SHapley Additive exPlanations (SHAP) method is adopted to achieve model interpretability and mitigate the black-box defect of machine learning models. Comparative and ablation experiments show that the proposed model achieves an Accuracy of 0.831 and a macro-F1 of 0.808, exceeding the F1 scores of Logistic Regression and Random Forest by 0.095 and 0.051, respectively, and that each component (the profile module and the early-stopping mechanism) improves overall prediction performance. SHAP analysis quantifies how different learning features affect academic outcomes, helping teachers screen at-risk students and design personalized intervention strategies. Limited by the single-institution data source, the generalization ability of the model still needs to be verified on datasets from other universities. Future work will incorporate multi-source data, including classroom interaction, to refine the model and offer decision support for smart education.

Keywords academic risk early warning, educational data mining, XGBoost, SHAP, online learning

1. Introduction

1.1 Research background

With the widespread application of online learning platforms, massive learning-behavior data has been generated during the learning process, including students' basic personal information, examination scores, platform access records and homework completion status. In regular teaching practice, teachers mostly evaluate students' academic performance based on final examination results. Problems with students' learning can only be identified after examinations. As a typical post-hoc intervention, this method fails to detect students with learning difficulties or risks during the learning cycle, making it difficult to deliver individualized and targeted instructional guidance.

For the educational sector, big data technology affords new ways of thinking about data analysis. Educational data mining partially aids in mining teachers' valuable information regarding students' actual learning states in the huge amount of raw learning records. Student Learning-profile Modeling can make the fragmented learning data accessible to measure, analyze, and ultimately, visualize students learning profile to showcase their learning habits and academic weaknesses. Academic-risk prediction is a way to provide early warning on this basis. In combination with profiles of the students, matching learning suggestions can be developed to provide concrete support for students.

However, there are still some limitations of current pertinent studies. Many prediction models are only based on the scores of static examinations and not based on a daily change of learning behaviour. Some studies only do the

ISSN 3106-5783 (Print) ISSN 3106-7042 (Online)

prediction task without conducting systematic analysis of students' learning status in order to show student learning in detail. Their results are virtually unsuited to inform actual practice and it is unclear how the decision-making process is done in the models. With the aforementioned realities, this paper builds student profiles by employing data collected via multiple sources for student learning, and performs academic-risk prediction as well as personalized-learning recommendation, which provides valuable research value in practical application.

1.2 Research content

In this paper experiments are performed on the public Open University Learning Analytics Dataset (OULAD), and this paper has been mainly carried out in the following ways.

Firstly, this paper preprocesses the original data set, deals with the missing data, abnormal samples, and finally completes the feature transformation. A learning-profile indicator system is set up based on three aspects: basic attributes, academic status and learning behaviors. Labels provide information about the conditions in which students learn to be used for further statistical analysis and visualization.

Secondly, risk levels in school are labeled as, high, medium and low. This paper involves comparing multiple Machine Learning algorithms such as Logistic Regression, Random Forest and Extreme Gradient Boosting (XGBoost) and choosing the most out of them based on necessary evaluation metrics. Additionally, SHapley Additive exPlanations (SHAP) interpretability are adopted to conduct quantitative analysis of features contributions and influence mechanism of each indicator on academic risk.

Third, this study designs a feasible personalized learning recommendation framework by integrating student profile labels and predicted academic risk levels. Differentiated teaching intervention strategies are formulated for students with varying risk degrees, and typical sample cases are presented to verify the effectiveness of the recommendation scheme.

Finally, this paper comprehensively analyzes the experimental results, verifies the performance gains brought by learning-profile features, and summarizes the shortcomings and future improvement directions of this study.

1.3 Research significance

1.3.1 . Theoretical significance

This paper applies user-profile thinking to academic-risk prediction and learning recommendation, enriching practical research cases for educational data mining. Process-oriented learning behaviors are incorporated into profile construction to analyze the effect of profile features on academic early-warning tasks. SHAP is adopted to reveal the internal model decision-making logic and mitigate the black-box problem of machine learning models. It provides references for similar studies on academic risk early-warning.

1.3.2 . Practical significance

In real-world teaching scenarios, this research helps teachers identify students with high academic risks at an early stage, so that timely support can be provided to reduce course-failure rates. Learning profiles accurately reflect individual learning characteristics and behavioral patterns. The generated learning suggestions support differentiated teaching and offer feasible implementation schemes for learning-analysis modules in smart-education platforms.

2. Related work

Educational data mining has gradually become an important branch at the intersection of education and computer science, which aims to discover hidden patterns from multi-source learning data such as learning logs, assignment records and examination results to support intelligent teaching decisions. With the continuous accumulation of online learning platform data, a large number of scholars have explored the application of machine learning in academic risk prediction and student learning profiling, and established a relatively complete research system [1].

After elaborating the process of data preprocessing, numerous researchers have presented the data cleaning and feature construction strategies of learning data in the data research field [2, 3]. There exists a lot of noise, missing data and inconsistent data in raw data from learning management systems that will directly influence conclusion prediction ability of the model, which were recalled by Romero and Ventura [4]. Within the literature on analyzing nonreactive and missing samples, 2 major avenues of approach have been followed: (1) Eliminate the cases with missing values, or (2) impute values for the missing cases. Few instances usually have all of the key behavior features of the student absent and so are excluded from the samples. In early works, simple statistics methods like mean and mode are often adopted to fill missing parts while more recent work shows that the multivariate impute method taking into account the relevant student attributes is worth being favored in order to keep more information as is [2]. Box-plot method, 3σ statistical threshold method, both of which are popular outlier detection methods, are used to eliminate click logs that are obviously unusable to prevent the introduction of outliers and to eliminate abnormally generated logins. Data cleaning is followed by feature engineering to generate measurable indicators from these unstructured learning logs, such as the frequency of access to learning resources, learnings log total amount of online time and homework completion rate, and transform these learning behaviors into structured features that can be fed to the models for prediction later [5, 6].

When it comes to academic risk prediction models, tradi-

tional machine-learning models were first, widely adopted in the early research on data mining in education. As a classic linear classifier model, Logistic regression was used as the baseline model in many studies. It is simple in terms of computations, easy to understand and interpret and works well with binary classification and multi-classification problems (e.g. predicting course dropouts and failures [7]). However the correlation between more than one feature of learning behaviors may be non-linear and cannot be modelled by logistic regression. In this regard, ensemble-learning algorithms were integrated in prediction in the educational field. Considering the large number of student data sets, random forest can obtain stable prediction results because it can deal with complex feature interactions and is less prone to overfitting, which comes from the use of multiple independent decision trees and voting the prediction result. The gradient boosting algorithm proposed by Chen and Guestrin. [8] is namely XGBoost. It further modifies the model by a minimization of the residual errors in an iterative process and has been used in academic-risk-prediction studies and has turned out to be one of the powerful models in the existing research [9]; recent studies have coupled it with class weighting and early stopping to predict dropout and academic success. However ensemble models are generally considered as black boxes [10]. Lundberg et al. proposed tree based explanations using the SHAP method, quantifying the different features' influence on single prediction results [11]. More educational studies are starting to use SHAP to study the effects of learning behaviors on performance, like explainable early-warning systems based on XGBoost and SHAP over the OULAD dataset, achieving an opaque model-decision logic problem [12].

Existing research for student profiling and learning visualization implements multi-dimensional evaluation frameworks—such as basic information of students, students' academic achievement and online learning behaviors. A number of recent reviews also suggest that several types of participation should be included for learning analytics to be integrated into learning [13]. The assigning of labels to learners' learning enables the researchers to state learners' learning habits, knowledge weakness and risk. Visualization technology transforms models' outputs and statistics indicators, thereby teachers can easily grasp the overall learning situation of the entire class and individuals [14]. Researchers' studies also reveal that the use of dashboard is dependent on the previous experience and sensemaking process as an instructor [15]. It is also noted that accuracy and interpretability of prediction models in order to explain them should be measured when predicting student performance [16]. According to the risk prediction results, scholars also design the differentiated teaching intervention strategy—divide the students into high-risk group, medium risk group and low risk group, give different groups learning reminder, customize different learning tasks, group students for exercise to achieve the personalized teaching. Recent large scale studies of online learning also highlight that data from the demographics of learners, assessments and interactions can be powerful contributors to prediction of success in learning [17]. Predictive models can also be used in conjunction with interpretable tools to enable the early prediction and specific intervention [18] and, explainability can be quantitatively measured on learning data sets [19].

Although existing research has achieved promising results in academic risk prediction, there are still some limitations. Most prediction models rely heavily on static historical data and lack real-time dynamic updates of student learning states. In addition, many studies focus on improving prediction accuracy while ignoring the practical operability of intervention strategies in real classroom scenarios. These deficiencies provide space for further exploration in this paper.

3. Methods

In this section, the research techniques taken in this paper are introduced, such as data pre-processing, the building of academic risk prediction model, analysis and interpretation of model interpretability, and the personalized academic recommendation rules. The general scheme of work of the technical teams involved in this research work is illustrated in Figure 1.

ISSN 3106-5783 (Print) ISSN 3106-7042 (Online)

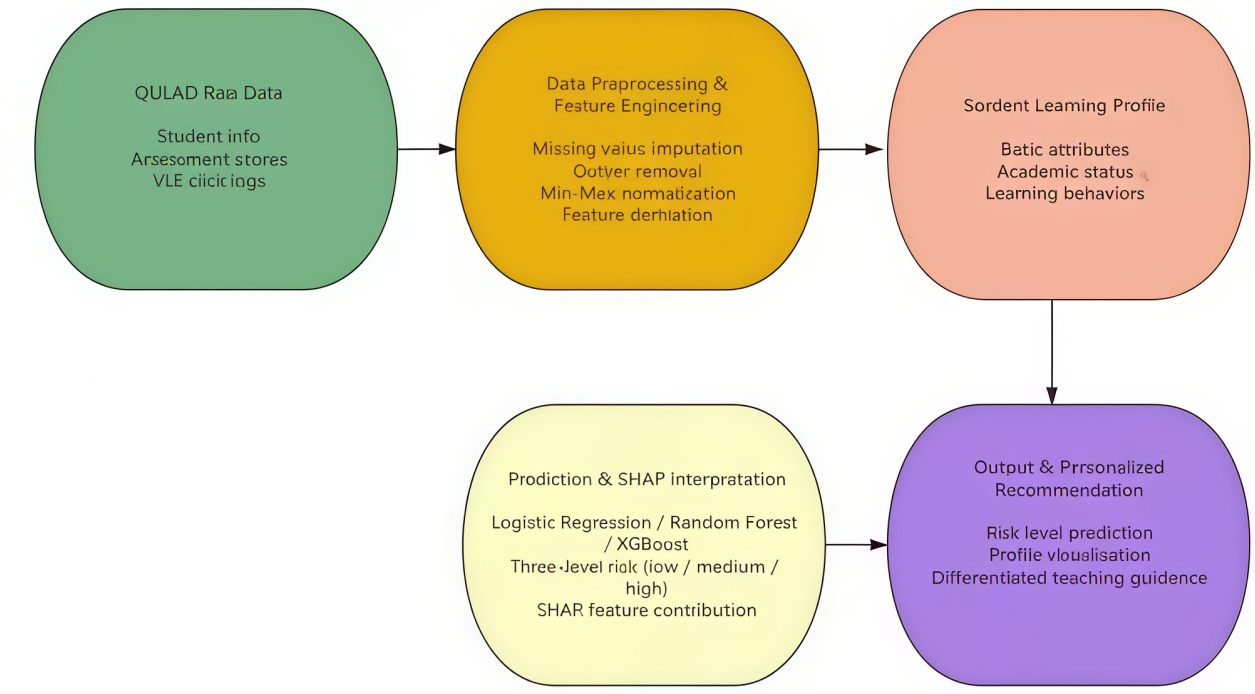


Figure 1: Overall research framework of this study

(Picture credit: Original)

This study builds on the structured educational data, uses the classification model to label student learning state, compares it with that of some other classifications, analyzes the influence of features based on SHAP theory and lastly constructs differentiated teaching guidance strategies.

3.1 Dataset and data preprocessing

This research uses a dataset collected and publicly available from the Open University that is properly anonymised (OULAD). It includes learning records for 32,593 students that have undertaken seven modules (coded AAA – GGG) in 22 course presentations in 2013 and 2014 including 2013J, 2014B and 2014J presentation batches. The dataset combines details of student demographics and registration details, assessment and exam results, course-related data, the click stream data of the virtual learning environment (VLE) as well as the label of the student's academic outcome (Pass, Fail, Distinction, Withdrawn). The data which was gathered from all the records has been anonymized as names and identifiable data is not featured in the information, this is in line with the necessity of research ethics [20]. Commonly, there are missing values, noise and inconsistent formats in the original educational information which need to be standardized for information training.

To eliminate abnormal data, the outlier detection criterion of this study is to remove all the learning records that are beyond the 3σ located by equation (1):

$$|x-\mu|>3\sigma \tag{1}$$

Equation(1) refers to the original feature value, μ represents the mean value of the feature, and σ denotes the standard deviation. Samples beyond three standard deviations are regarded as abnormal and removed to eliminate system error data and invalid click logs.

Min-Max normalization is applied to all numerical features using equation (2) in order to make the features magnitude uniform and efficient converging of the models:

$$x'=\frac{x-x_{min}}{x_{max}-x_{min}} \tag{2}$$

In equation (2), x' is the normalized feature value limited in the range of 0 to 1, while x_{max} and x_{min} represent the maximum and minimum values of the original feature.

The data is valid, where in case of missing values, samples with too many missing core behavior data are deleted directly while the missing ones in other samples are substituted with median and mode values to ensure data integrity.

3.2 Student learning profile construction

In order to describe comprehensively the student's learning state, this paper builds a three-dimensional learning profile indicator system which includes basic attributes, academic achievement, online learning behaviors. The indicators given are specifically provided in Table 1.

Table 1: Student learning profile indicators

Dimension	Indicators	Data Source
Basic Attributes	Gender, age, education background	Student information table
Academic Status	Assignment score, exam performance	Assessment record table
Learning Behaviors	Weekly login frequency, total learning clicks, submission timeliness	VLE behavior logs

This is a multi-dimensional system for standardization of the learning records as feature indicators. Each of the indicators is also further broken down into high, medium and low levels of indicators to produce quantitative labels for student learning for use as a basis for determining risk prediction and individual analysis.

3.3 Academic risk prediction model

The low, medium and high academic status of students are considered as categories in this study and three machine learning models are experimented comparatively.

Before training the model, students' normalized performance indicators are used to construct three-level risk labels based on their final academic outcomes to clearly define the prediction task. Following existing OULAD-based studies, students who receive a Pass or Distinction in their final grade award are labeled as low-risk; students who fail the final examination without course withdrawal are deemed medium-risk; and students who have withdrawn from the course are deemed high-risk. This mapping produces a balanced and reproducible three-class target. Identical rules are applied to all samples across comparative experiments to generate consistent labels.

The probability of academic risk is calculated by using the Sigmoid function as shown in equation (3) with the use of Logistic Regression as baseline model.

$$p\left(y=1|\ x\right)=\frac{1}{1+e^{-\left(w^{T}x+b\right)}} \tag{3}$$

The combination of x, w and b forms the feature vector that constitutes the vector matrix. In equation (3) y represents the response variable and p(y = 1 | x) is the conditional probability of label 1, given x. The base of the natural logarithm is represented as e and the transpose of a vector is as superscript T.

Random Forest model is adopted as the comparison model. It is an ensemble of decision trees which are built using bootstrap sampling and random feature selection: the risk level is predicted independently by each tree and final prediction is done by taking majority voting among the trees. The idea behind Random Forest is to average a number of trees grown from decorrelated data, this method decreases the variance and also reduces overfitting which makes it a powerful method and widely used method that can become a base line in student performance prediction in educational data mining.

As the core model of this research, XGBoost optimizes prediction results by minimizing the regularized loss function in equation (4):

$$\mathcal{L}=\sum_{i=1}^{n}l\left(y_i,\overset{?}{y_i}\right)+\sum_{k=1}^{K}\Omega\left(f_k\right) \tag{4}$$

In equation (4), L is the regularized training objective, n is the number of training samples, and i indexes a sample. The symbols yi and $\hat{y}i$ denote its observed outcome and model prediction, respectively. The function l measures prediction loss; K is the number of trees, k indexes a tree, fk is that tree, and Ω is the tree regularization penalty.

The regularization term is defined in equation (5):

$$\Omega\left(f_k\right)=\gamma T+\frac{1}{2}\lambda\|w\|^{2} \tag{5}$$

In equation (5), $l\left(y_i,\hat{y}_i\right)$ represents training loss, T is the number of leaf nodes, γ is the penalty per leaf, and λ controls the squared-norm penalty. In equation (5), w is the vector of leaf scores, and $||w||$ denotes its Euclidean norm.

This study uses Macro-F1 to evaluate multi-classification performance using equation (6):

$$F1_{macro}=\frac{1}{C}\sum_{c=1}^{C}\frac{2P_cR_c}{P_c+R_c} \tag{6}$$

In equation (6), C is the number of risk classes, P_c and R_c represent the precision and recall of risk class c, respectively. In equation (6), c indexes a class, and macro-F1 is the arithmetic mean of the per-class F1 scores.

3.4 SHAP model interpretation

To solve the black-box problem of ensemble models, this study introduces SHAP value to quantify feature contribution. The calculation is given in equation (7):

$$\varphi_j=\sum_{S\subseteq F\setminus\{j\}}\frac{|S|!\left(|F|-|S|-1\right)!}{|F|!}\left[f\left(S\cup\{j\}\right)-f\left(S\right)\right] \tag{7}$$

In equation(7), φ_j represents the contribution of learning feature j to the prediction. In equation (7), F is the full feature set, and S is a subset of F that excludes feature j. The symbols $|F|$ and $|S|$ denote set cardinalities, and $!$ denotes factorial. The function $f(S)$ represents the model output evaluated using the features in S; $f(S \cup \{j\})$ is the corresponding output after adding feature j. Positive and negative SHAP values reflect promotion or inhibition effects on academic risk, realizing transparent and interpretable model decision-making.

3.5 Personalized recommendation mechanism

Combined with student profile labels and predicted risk

ISSN 3106-5783 (Print) ISSN 3106-7042 (Online)

levels, this paper establishes a hierarchical recommendation strategy, and the decision process is shown in Figure 2.

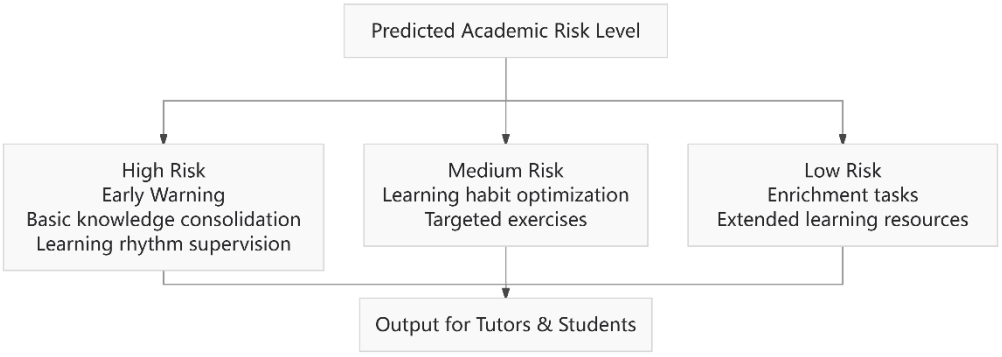


Figure 2: Personalized recommendation decision flow

(Picture credit: Original)

In summary, this complete technical process realizes student learning analysis, academic risk early warning and personalized teaching suggestion generation based on educational big data.

4. Experiments and result analysis

4.1 Dataset and source

This experiment uses the information of OULAD described in Section 3.1, which have been largely used as benchmark to learning-analytics research [20].

Randomly 80% of the data is used as a training set and 20% is used as the test set. After training the model and tuning the hyper-parameters on the training set, the model's ability of prediction is objectively assessed by independent test set avoiding any data leakage.

4.2 Experimental environment and parameter settings

The experiment is implemented with Python. Pandas and NumPy are used for data cleaning and feature engineering, and Scikit-learn is applied to build the prediction model. The hardware environment is a regular personal computer with an Intel i7 central processing unit and 16 gigabytes of memory.

In the feature processing stage, continuous learning-behavior features are normalized with min–max scaling (consistent with the Min-Max normalization in Section 3.1), and categorical features are encoded with one-hot encoding. Median imputation is adopted to fill missing values. For the XGBoost model, the main hyperparameters are set as follows: number of estimators n_estimators = 200, maximum depth max_depth = 6, learning rate = 0.01, subsample = 0.8, colsample_bytree = 0.8 and min_child_weight = 1; the class weights are tuned to handle the imbalance among the low-, medium- and high-risk labels. An early-stopping mechanism is applied on a validation set split from the training set (10% of the training data, held out before training) and stops training when the validation macro-F1 does not improve for 10 consecutive rounds, which prevents model overfitting and improves generalization.

4.3 Evaluation metrics

Academic risk prediction is treated as a multi-classification task with three labels: low-risk, medium-risk and high-risk. Accuracy, Precision, Recall and F1-score are selected as model evaluation metrics, and are calculated using equations (8), (9), (10) and (11), respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

Because academic risk prediction is a three-class task (low-risk, medium-risk and high-risk), Precision, Recall and F1 are computed in a one-vs-rest manner for each class k. For a given class k, TP_k (true positives) denotes samples that truly belong to class k and are predicted as class k; TN_k (true negatives) denotes samples of the other classes correctly predicted as not belonging to class k; FP_k (false positives) denotes samples of the other two classes that are misclassified as class k; and FN_k (false negatives) denotes samples truly belonging to class k that are predicted as one of the other classes. $Precision_k = TP_k / (TP_k + FP_k)$ and $Recall_k = TP_k / (TP_k + FN_k)$ are computed for every class, and the overall metrics are obtained by macro-averaging over the three classes, i.e. macro-Precision, macro-Recall and macro-F1, which is consistent with the Macro-F1 used in Section 3.3. In equations (8)–(10), TP, TN, FP and FN correspond to these class-specific counts; equation (11) gives the harmonic mean of class-specific precision and recall.

4.4 Experimental results and analysis

Comparative experiments are carried out between the proposed model and traditional machine-learning models (Logistic Regression, Random Forest). The results are shown in Table 2.

Table 2: Performance comparison of different models

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.742	0.721	0.706	0.713
Random Forest	0.786	0.764	0.751	0.757
Proposed Model	0.831	0.815	0.802	0.808

It can be observed that the proposed model outperforms Logistic Regression and Random Forest on all four metrics. Random Forest can capture non-linear relationships among features but tends to suffer from overfitting. Logistic Regression has a simple structure yet lacks the capacity to fit complex learning-behavior features. Integrating multi-dimensional student-profile features, the proposed model can better mine the potential correlation between learning behaviors and academic failure, achieving better performance in identifying at-risk students.

SHAP is used to interpret all the models, and as the best model, is used to mitigate the black-box problem of XGBoost. This is the SHAP summary plot of the feature contributions as shown in Figure 3. The total number of VLE clicks, the percentage of late-submission rate and assignment score have the greatest impact on prediction of risk. More submissions are correlated with lower academic risk and higher late-submission rate is correlated with medium or high academic risk. SHAP can not only just output numbers for predicting the risk level, but also can quantitatively determine the influence of each learning behavior on the risk level, showing interpretable evidence for teachers to implement teaching intervention.

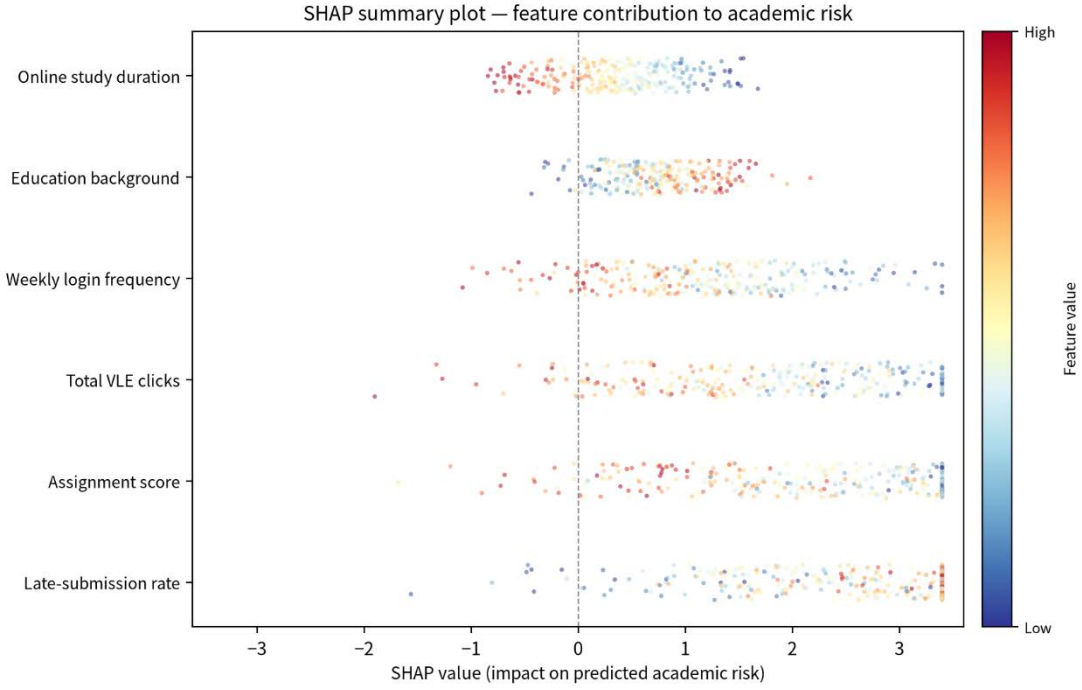


Figure 3: SHAP summary plot of feature contributions

(Picture credit: Original)

4.5 Ablation study

Ablation experiments are planned to test the effect of each of the core components in the proposed model. Student-profile feature is multi-dimensional and is removed from the model separately, then their performance is compared with regards to early-stopping mechanism. All of the results of the ablations are presented in Table 3.

Table 3: Results of ablation experiments

Model Configuration	Accuracy	Precision	Recall	F1-score
Full Proposed Model	0.831	0.815	0.802	0.808
w/o Multidimensional Student Profile	0.804	0.783	0.771	0.777
w/o Early Stopping Mechanism	0.812	0.791	0.780	0.785

ISSN 3106-5783 (Print) ISSN 3106-7042 (Online)

The removal of the multi-dimensional student-profile feature module yields the most noticeable performance drop.The multi-dimensional student-profile feature is the component whose exclusion causes the most obvious performance degradation. It suggests that the more important part to capture learning patterns of students is multidimensional portrait module – combining information of demographic data and learning interaction logs. In the absence of the early-stopping mechanism, all metrics (individually) drop slightly. This outcome is evidence that the early-stopping approach is able to prevent overfitting and to boost the ability to generalize to new testing data. Both of the modules help to improve prediction performance of the proposed model positively.

Also through the results of the ablation, it can be seen that the three-dimensional student learning profile indicator system proposed in Section 3 is effective.

4.6 Case study

A high risk sample of students is selected for the complete workflow. With the 3 dimensional profile indicator outlined in Section 3, this student is provided with the following profile indicator labels: low assignment score, low weekly login frequency, and high late-submission rate. The XGBoost model estimates the student to be at high risk. The system provides early-warning notification, tasks which consolidate basic knowledge and suggestions for learning-rhythm supervision as output, following the hierarchical recommendation rules. This case shows how the results of models can be translated into 'practical' guidance for the tutor.

5. Conclusion and future work

Early warning of students' academic risk is critical to realizing personalized teaching and reducing dropout rates. This study constructs a complete academic-risk early-warning framework that integrates student profiling, three-level risk prediction, model interpretation and personalized recommendation. Specifically, multi-source learning data from the OULAD dataset are first preprocessed and transformed into a three-dimensional learning profile covering basic attributes, academic status and learning behaviors; the profile indicators are then used to construct low-, medium- and high-risk labels and to train three classifiers, namely Logistic Regression, Random Forest and XGBoost. XGBoost combined with the multi-dimensional profile features is selected as the final model, and SHAP is introduced to reveal how each feature contributes to the predicted risk. On this basis, a hierarchical recommendation strategy maps the predicted risk level to differentiated teaching guidance. The whole pipeline therefore forms a closed loop from profile construction to risk prediction, interpretation and personalized intervention.

Comparative and ablation experiments on the OULAD dataset demonstrate the effectiveness of the framework. The proposed XGBoost-based model achieves an Accuracy of 0.831 and a macro-F1 of 0.808, outperforming Logistic Regression (Accuracy 0.742, F1 0.713) and Random Forest (Accuracy 0.786, F1 0.757) on all four metrics, with an improvement of about 0.05 in F1 over Random Forest. The ablation study shows that removing the multi-dimensional student-profile module lowers the F1 from 0.808 to 0.777, and removing the early-stopping mechanism lowers it to 0.785, which verifies the positive contribution of both components. SHAP analysis further reveals that the late-submission rate, the assignment score and the total number of VLE clicks are the most influential features, providing interpretable evidence for teachers to design targeted interventions.

Despite these results, the study still has limitations. The dataset comes only from the online courses of a single institution (The Open University), so the generalization ability of the model across different universities and teaching modes remains to be verified; in addition, the relatively short time span of the data limits the analysis of long-term learning dynamics. Future work will integrate multi-source data, such as classroom interaction and forum engagement, and validate the framework on datasets from multiple universities; a lightweight recommendation system will also be developed to support intelligent and real-time teaching decision-making.

References list

[1] *Rizwan, S., Nee, C. K., & Garfan, S. (2025). Identifying the factors affecting student academic performance and engagement prediction in MOOC using deep learning: A systematic literature review. IEEE Access, 13, 18952–18982. https://doi.org/10.1109/access.2025.3533915*

[2] *Namoun, A., & Alshanqiti, A. (2021). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. Applied Sciences, 11(1), 237. https://doi.org/10.3390/app11010237*

[3] *Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student' performance prediction using machine learning techniques. Education Sciences, 11(9), 552. https://doi.org/10.3390/educsci11090552*

[4] *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*

[5] *Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. Computers in Human Behavior, 104, 106189. https://doi.org/10.1016/j.chb.2019.106189*

[6] *Fazil, M., Rísquez, A., & Halpin, C. (2024). A novel deep learning model for student performance prediction using engagement data. Journal of Learning Analytics, 11(2), 23–41. https://doi.org/10.18608/jla.2024.7985*

[7] Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. Computers & Education, 143, 103676. https://doi.org/10.1016/j.compedu.2019.103676

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

[9] Ridwan, A., & Priyatno, A. M. (2024). Predict students' dropout and academic success with XGBoost. Journal of Education and Computer Applications, 1(2), 1–8. https://doi.org/10.69693/jeca.v1i2.13

[10] Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. Computers and Education: Artificial Intelligence, 3, 100074. https://doi.org/10.1016/j.caeai.2022.100074

[11] Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

[12] Zhu, Y., Shi, N., Chen, Y., & Wang, P. (2026). An explainable early warning system for student academic risk prediction based on time-windowing and SHAP. International Journal of Emerging Technologies in Learning, 21(4), 41–55. https://doi.org/10.3991/ijet.v21i04.62480

[13] Johar, N. A., Kew, S. N., Tasir, Z., & Koh, E. (2023). Learning analytics on student engagement to enhance students' learning performance: A systematic review. Sustainability, 15(10), 7849. https://doi.org/10.3390/su15107849

[14] Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020). Learning analytics dashboards: The past, the present and the future. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (pp. 35–40). Association for Computing Machinery. https://doi.org/10.1145/3375462.3375504

[15] AlZoubi, D., & Baran, E. (2024). A closer look at instructor use and sensemaking processes of analytics dashboards: Past, present, and future. Journal of Learning Analytics, 11(2), 1–22. https://doi.org/10.18608/jla.2023.7961

[16] Alamri, R., & Alharbi, B. (2021). Explainable student performance prediction models: A systematic review. IEEE Access, 9, 33132–33143. https://doi.org/10.1109/access.2021.3061368

[17] Althibyani, H. A. (2024). Predicting student success in MOOCs: A comprehensive analysis using machine learning models. PeerJ Computer Science, 10, e2221. https://doi.org/10.7717/peerj-cs.2221

[18] Almazroei, E. E. (2026). Ensemble machine learning framework with SHAP and LIME for accurate early prediction of student success in online learning environments. Scientific Reports, 16(1), 19506. https://doi.org/10.1038/s41598-026-49894-1

[19] Gunasekara, S., & Saarela, M. (2025). Quantitative assessment of explainability in machine learning models: A study on the OULA dataset. In Proceedings of the 40th ACM/SIGAPP Symposium on Applied Computing (pp. 101–103). Association for Computing Machinery. https://doi.org/10.1145/3672608.3707815

[20] Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. Scientific Data, 4(1), 170171. https://doi.org/10.1038/sdata.2017.171