Predicting Student Burnout Risk Using Academic Behavior Logs: A Data Mining Approach

Faria Ahmed Richi  
*Computer Science & Engineering   
American International University Bangladesh* [*22-46371-1@student.aiub.edu*](mailto:22-46371-1@student.aiub.edu) *fariaahmedrichi15@gmail.com*

Md Mostafijur Rahaman Bipul

*Computer Science & Engineering American International University Bangladesh* [*22-46762-1@student.aiub.edu*](mailto:22-46762-1@student.aiub.edu)

*mostafijur1688@gmail.com*  
Horish Das Priyo   
*Computer Science & Engineering  
American International University Bangladesh* [*21-44816-1@student.aiub.edu*](mailto:21-44816-1@student.aiub.edu)

*Priyodas0177@gmail.com*

Victor Stany Rozario  *Assistant professor, Special Assistant  
American International University Bangladesh* [stany@aiub.edu](mailto:stany@aiub.edu)

***Abstract***— **Without showing any early signs, numerous university students face academic burnout in recent times. This research aimed to predict those students who are at risk of burnout using academic behavior data such as attendance percentage, assignment records, and the like. The Open University Learning Analytics Dataset (OULAD) was used and machine learning models like Logistic Regression, Decision Tree, Random Forest and XGBoost were applied to achieve the target of building a burnout prediction system where all the models performed exceptionally well within a narrow range of 95-97% accuracy. Our goal is to help the universities in early detection of struggling students so that they can offer necessary support before the student moves toward dropping out.**

**Keywords**— Academic Burnout, Learning Analytics, Behavioral Data, Machine Learning, Burnout Prediction, Early Detection, OULAD Dataset.

# Introduction

Academic burnout is a crucial but rarely addressed issue in higher education. Emotional drain, lack of drive and feelings of unproductiveness are the way of exhibiting it, which may become an enormous barrier in the way of a student’s learning and engagement [1]. Dropout is the last occurrence of academic disinvolvement whereas burnout starts earlier but not easy to detect with traditional performance metrices single-handedly [11]. Students having the signs of burnout may keep the continuation of attending classes, but struggle unspoken which leads to declining of performance, mental health, or ultimate withdrawal.

Many institutions have taken this issue earnestly and implementing prior alert systems to identify those students are at risk of failure or withdrawal by evaluating online based learning data [2]. Nevertheless, these systems do not address psychological exhaustion rather focus on academic performance only. In addition, most of the burnout detection research rely on surveys by self-reporting such as Maslach Burnout Inventory [1], which are neither perfectly suitable in terms of scalability nor consistency for large student populations.

To address this gap, our research proposes a data-driven, scalable strategy to detect burnout risk of students by evaluating their academic behavior logs specially, their engagement with the learning management system (LMS), assessment records and registration status. We target to build a prior intervention system for the institutions using machine learning models trained on behavioral data which will create the scope of easily identifying those students who are struggling before burnout escalates to academic failure or dropout.

# RESEARCH QUESTIONS

1. Can data from students' academic behavior on a Learning Management System (LMS) be used to predict the risk of academic burnout using machine learning?

1. Which machine learning models are most effective in predicting academic burnout based on LMS activity data?

1. What types of student behaviors recorded in the LMS are most useful for identifying those at risk of academic burnout?

# Related work

## Burnout and Mental Health Assessment

A study formulated burnout as a three-dimensional concept in academic contexts which are emotional exhaustion, doubt and reduced academic efficacy [1]. Bresó et al. pointed out that even though the academic involvement looks unchanged, the students are vulnerable to burnout who are facing low self-efficacy. Academic burnout and stress levels increased significantly during the COVID-19 pandemic when students had to face isolation and had no other choice but adapting the online learning environment. Basri et al. noticed that burnout which is combined with problematic internet use, negatively impacted students’ mental well-being and perceived learning [11].

## LMS Log Analytics and Engagement Detection

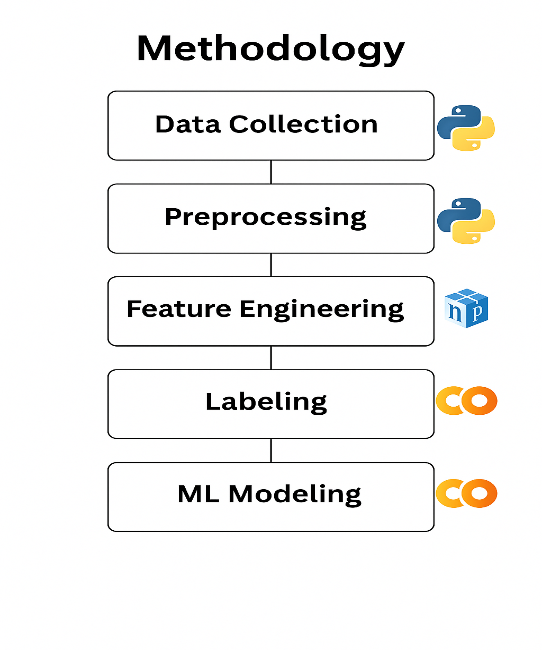
In recent years, visibility has increased in using LMS logs to detect student risk. Macfadyen and Dawson [2] described earlier that in academic failure detecting, learning analytics based on LMS interactions pattern might be useful to build a prior alerting system. Machine learning models like Random Forest was applied by Fahd et al. [3] to predict student performance using behavioral and demographic data in combined education environments. Bessadok et al. [4] additively evaluated LMS data by doing cluster of student activity and found out that for academic outcomes, engagement levels were strong indicator. Likewise, Hu et al. [5] explored that student success was highly associated with consistency in involvement in LMS environments, highlighting the predictive strength of behavioral logs.

## Dropout Prediction Using OULAD

A few researchers explored the Open University Learning Analytics Dataset (OULAD) to forecast student dropout. Jawthari and Stoffa [9] is one of them who addressed a scheme using OULAD to predict student withdrawal and described the effectiveness of the dataset in comprehensive predictive models. Their observations indicates that disinvolvement trends can be pointed out efficiently by analyzing LMS engagement and registration patterns, supporting its compatibility for burnout prediction as well.

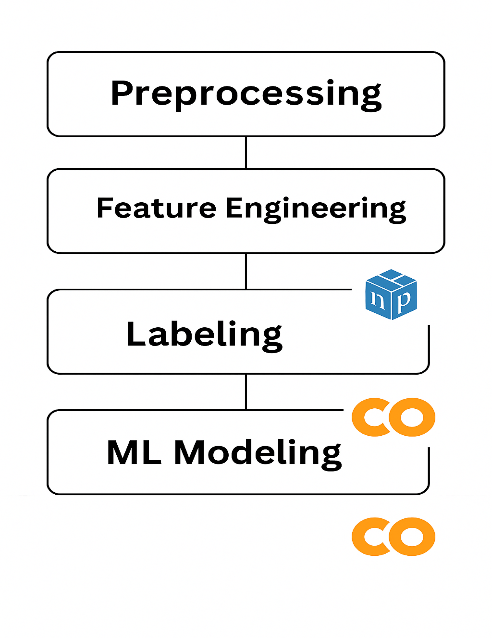
# Methodology

The methodology employed in this research is systematic and follows a set procedure to detect students who may suffer from academic burnout. The process has been broken down into five milestones: collection of relevant datasets, preprocessing which includes cleaning and organization of data, feature engineering, creation of labels for the target variable, and finally application of the model.

A logo of a university

AI-generated content may be incorrect.





A screenshot of a computer

AI-generated content may be incorrect.

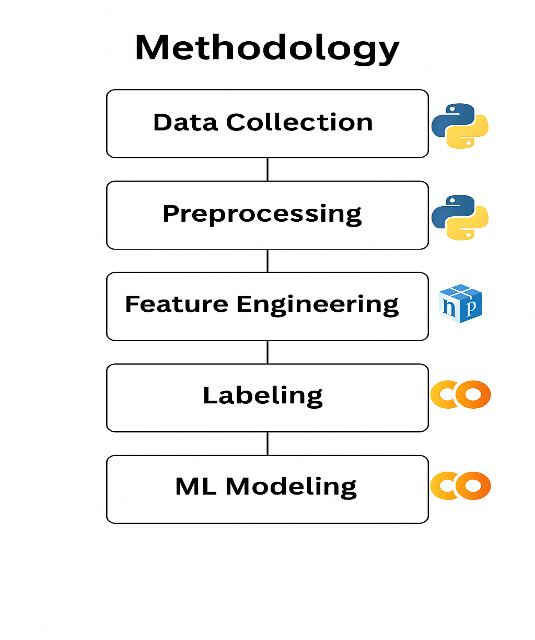


Figure 1: Methodology Flowchart

## Dataset

We utilized a publicly available dataset the (OULAD) [13]. The dataset comprises of a demographic academic performance, and interaction with the Virtual Learning Environment (VLE). The OULAD dataset is organized into seven interlinked CSV files. To predict the burnout risk using predicting burnout risk using academic behavior logs, we focused primarily on data from following files:

* studentInfo.csv: Contains demographic data for each student, such as gender, age band, highest education level, region, disability status, and final course outcomes.
* studentAssessment.csv: Records student submissions, including scores and submission dates.
* studentRegistration.csv: Captures the registration and unregistration dates for each student-module pair.
* studentVle.csv: Provides interaction logs of student activities within the VLE. Each record includes the date, resource accessed, and the number of clicks, enabling detailed temporal analysis of engagement patterns.

## Data Preprocessing

To enable effective prediction of student burnout risk, several preprocessing steps were applied to the Open University Learning Analytics Dataset (OULAD) [1]. The conditions used to identify students at risk of burnout were developed specifically for this study and are not part of the original OULAD dataset. These criteria were derived from behavioral and academic indicators available in the data. For instance:

* Students whose interaction with the Virtual Learning Environment (VLE) fell within the lowest 25% were flagged for low engagement.
* Those who missed or did not pass most of their assessments were considered academically at risk.
* A final course result marked as “Withdrawn” or “Fail” was interpreted as a potential sign of disengagement or struggle.
* A noticeable and consistent reduction in VLE activity over time was also treated as a warning signal.

These indicators were combined to create a binary label, burnout risk, where a value of 1 denotes a student potentially experiencing burnout, and 0 indicates otherwise.

## Feature Engineering and Model Setup

We extracted engagement features from the OULAD dataset containing rich log and assessment data from LMS. Specifically, we computed the following features:

**Weekly LMS click counts:** Weekly click counts on LMS include every scheduled interaction with the learning management system (VLE) logs which includes a maximum of resource clicks per week. A count click is taken as affordance access to a resource. Thus, higher counts signal resource access and more active engagement.

**Assessment submission gaps:** The time gaps between deadlines of assessments and submission timelines (or inter-submission intervals). Longer delays suggest disengagement during submission; timely delayed submissions indicate some effort.

**Grade decline trends:** Grade decline trends illustrate the changes in student scores over sequential assignments or quizzes (for example, slope-average of a linear fit through their grades). A downward trend in achievement suggests declining performance or understanding. These trends accompany features based on raw grade value (which are also available) by capturing dynamic shifts related to achievement.

These engineered features serve as proxies for academic engagement. These blended features represent proxies’ indicators for academic engagement. Active posting on forums tends to be accompanied by timely homework resubmissions reflecting steady progression towards set goals. Submission lags coupled paired with falling grades raise red flags. (For example, prior work has found that students with prompt submissions and frequent LMS logins tend to perform better.)

For evaluating feature relevance, we looked at all the numeric attributes and calculated the Pearson correlation matrix. Certain features like weekly LMS clicks, submission delays, and declining grades showed notable correlation with the burnout label and were kept for model training.

A screenshot of a computer

AI-generated content may be incorrect.

Fig 2: Pearson correlation heatmap

The heatmap captures the Pearson correlation coefficients for the numerical features in the dataset, thus visually summarizing the degree of linear relationship between them. As highlighted in the last paragraph, each cell of the heatmap quantifies to what extent two variables are related linearly, with values near +1 indicating strong positive correlation, close to -1 strong negative correlation, and around 0 little to none correlation. This type of visualization is particularly useful in spotting redundant features that are highly correlated which assists in deciding feature selection, multicollinearity evaluation, and model interpretation, enhancing the trust and accuracy of models built.

In this analysis, we adopted the Pearson correlation method because it is useful in determining the linear relationships between numerical features that are continuous, like LMS click counts, grades over time, and the gaps between submissions. All these features are measured on ratio or interval scales which are compatible with the assumptions of Pearson. Pearson correlation delivers a straightforward, optimal, and understandable measure of association with -1 to +1 denoting the two extremes of a linear relationship's strength and direction. Its performance is best when the dataset is normally distributed or has minimal outliers.

Other correlation methods, such as Spearman or Kendall Tau, apply better to ordinal or nonlinear relationships. They are less efficient and intuitive, especially in the context of large datasets. In contrast, Pearson remains one of the most preferred and accepted methods to use in exploratory data analysis and feature selection, especially in educational data mining. It aids in confirming the usefulness of the engagement features that have been created and informs modeling decisions by revealing possible multicollinearity or strong predictive signals which may be present. Considering that a good number of academic engagement behaviors tend to be linear, in this situation Pearson correlation is both reasonable and theoretically appropriate.

Logistic regression, decision trees, and random forest [14] models were run in Python through Co-Lab while gradient boosting was implemented using the XGBoost package. This tool is well-known for its streamlined implementations as well as their provisions on class weighting. Explicitly, we tried to handle the class imbalance problem due to the severe paucity of samples from the burnout risk class label by SMOTE oversampling or a subclass weight parameter setting in the classifiers so that they do not overlook the at-risk class.

The following classifiers were evaluated, chosen for their relevance in educational data mining:

**Logistic Regression:** A linear baseline model; its simplicity is both appreciated and a drawback since it makes weak assumptions. It predicts probabilities but needs to be bounded within [0, 1] by a sigmoid transformation.

**Decision Tree:** Single tree model used to cut data based on feature threshold values which partition data into homogenous subsets. The method captures non-linear dependencies without scaling requirements for features and interpretation is straightforward.

**Random Forest:** Collection of selection-based grouped ensembles of bootstrapped sampled decision trees. Its accuracy and robustness increase after the aggregation of many randomized trees, while also allowing good capture of sophisticated interactions. Its feature subsampling makes it boost and known as good under class imbalance.

**XGBoost (Extreme Gradient Boosting):** A scalable gradient-boosted tree model. XGBoost often achieves top performance in tabular classification tasks and has built-in mechanisms (e.g. regularization) to control overfitting. In prior EDM work it has outperformed both random forests and simpler classifiers when optimized properly. To train and evaluate the models, the preprocessed dataset was split into training, validation, and testing sets using a stratified split to preserve class distribution. **Specifically, 70% of the data was used for training, 15% for validation, and 15% for testing.** This approach ensured reliable tuning and unbiased performance evaluation of the models. Throughout the training process, we implemented industry-standard procedures to avoid information leakage, like applying SMOTE only on the training portion.

# Evaluation

Model performance was measured using standard classification metrics: Accuracy, Precision, Recall, F1-score, and ROC–AUC. Accuracy captures the overall correct classification rate of classifier including true positives and true negatives. Precision and Recall on the other hand only focus on positive class which is Burnout in our case. Moreover, F1 score balances both Precision and Recall providing a singular metric. Also, ROC AUC or area under the receiver-operating curve has been used as well because it assesses discrimination ability for all classification thresholds and is robust to imbalance. Most notably in our context high recall is favorable especially for proactive predictive alerts aimed at identifying as many students who are potentially at risk as early as possible.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Accuracy | Precision | Recall | F1-score | ROC–AUC |
| Logistic Regression | 96.69% | 0.97 | 0.97 | 0.97 | 0.99 |
| Decision Tree | 95.50% | 0.96 | 0.96 | 0.96 | 0.96 |
| Random Forest | 96.90% | 0.97 | 0.97 | 0.97 | 0.99 |
| XGBoost | 96.58% | 0.97 | 0.97 | 0.97 | 0.99 |

Table 1: Evaluation Metrices

A graph of different colored bars

AI-generated content may be incorrect.In Table 1, we see the evaluation metrics of four classification models side by side. As expected, ensemble methods Random Forest and XGBoost performed strongly, but Logistic Regression also emerged as a top performer. Logistic Regression achieved a high accuracy of 96.69% with excellent precision, recall, and F1-score all at 0.97, along with the highest ROC–AUC of 0.99. Random Forest slightly outperformed Logistic Regression in accuracy with 96.90%, and matched its precision, recall, F1-score of 0.97, and also achieved a robust ROC–AUC of 0.99. XGBoost followed closely with an accuracy of 96.58%, precision, recall, and F1-score of 0.97, and a similarly excellent ROC–AUC of 0.99. Meanwhile, the Decision Tree lagged the other models with an accuracy of 95.50%, precision, recall, and F1-score all at 0.96, and a lower ROC–AUC of 0.96. This supports previous research in educational and medical data mining where it is well known that ensemble methods such as Random Forest and XGBoost tend to yield accurate, stable predictions due to their ability to model complex relationships between variables while reducing overfitting.

|  |  |  |  |
| --- | --- | --- | --- |
| Model Name | Test Accuracy (%) | Cross-validation Accuracy (%) | Std Deviation (%) |
| Decision Tree | 95.50 | 95.19 | 0.35 |
| Logistic Regression | 96.69 | 96.81 | 0.33 |
| Random Forest | 96.90 | 96.77 | 0.36 |
| XGBoost | 96.58 | 96.43 | 0.44 |

Table2: Model Accuracy

Table 2 displays the results achieved by four different machine learning models—Decision Tree, Logistic Regression, Random Forest, and XGBoost—on predicting the risk of burnout in students. Random Forest tested as the highest performer with a test accuracy of 96.90%, followed closely by Logistic Regression (96.69%) and XGBoost (96.58%). The Decision Tree still performed well with an accuracy of 95.50%. All models’ cross-validation accuracy scores were consistent with their respective test results, indicating strong generalization ability. The standard deviations for cross-validation accuracy ranged from 0.33% to 0.44%, reflecting reliable performance across folds despite potential feature complexity or class imbalance.It is interesting to observe that while ensemble methods like Random Forest and XGBoost are known for adeptly handling complex feature relationships and class distribution issues, they did not dramatically outperform Logistic Regression in this context. Logistic Regression remained highly interpretable and robust, making it a strong candidate for educational burnout prediction. In summary, academic behavioral features proved to be valuable predictors of student burnout risk, as evidenced by the consistent, high performance across all applied models.

# Results

After applying the selected machine learning models to the processed and feature-engineered dataset, we evaluated their performance using a range of standard classification metrics. These metrics—accuracy, precision, recall, F1-score, and ROC–AUC—collectively assessed each model’s ability to identify students susceptible to academic burnout. All four models, namely Logistic Regression, Decision Tree, Random Forest, and XGBoost, achieved excellent classification accuracy above 95% on the test set. Random Forest achieved the highest test accuracy of 96.90%, followed closely by Logistic Regression at 96.69% and XGBoost at 96.58%. Decision Tree, while slightly lower at 95.50% accuracy, still demonstrated firm predictive capacity.Cross-validation showed the models were stable, with average accuracies consistent with their respective test results and standard deviations ranging from 0.33% to 0.44%, highlighting strong generalization. Precision, recall, and F1-scores were high across all models, with most values at or near 0.97. Furthermore, ROC–AUC scores were excellent: Logistic Regression, Random Forest, and XGBoost each achieved ROC–AUCs of 0.99, demonstrating outstanding discriminatory power between students at risk and those not at risk. The Decision Tree achieved a solid ROC–AUC of 0.96, reinforcing its predictive ability.

Fig 3: Accuracy Comparison of Classification Models

A graph of a curve

AI-generated content may be incorrect. To visually compare the performance of each data mining model, we plotted a bar chart illustrating the classification accuracy of Decision Tree, Random Forest, Logistic Regression, and XGBoost. Graphically, Random Forest slightly outperformed all other models, followed closely by Logistic Regression and XGBoost. Notably, all tested models achieved more than 95% accuracy, which is impressive. Such visuals enhance understanding and support informed decisions about the most robust model for burnout prediction.

Fig 4: ROC Curve Comparison chart

For assessing the accuracy of the applied models, Receiver Operating Characteristic (ROC) curves were generated using their AUC (Area Under the Curve) scores. All models significantly surpassed random guessing (AUC = 0.5), as shown in Figure 4, demonstrating strong classification abilities. The highest AUC score of 0.99 was achieved by Logistic Regression, Random Forest, and XGBoost, each demonstrating excellent discriminatory power in detecting burnout risk. The Decision Tree’s AUC score of 0.96, although slightly lower, still indicated strong performance. It is clear from these results that ensemble-based method like Random Forest offer high accuracy to regression models while providing greater reliability and generalization. All models demonstrated a strong ability to distinguish between students at risk and not at risk, emphasized by the steep ROC curves and high AUC scores.

# Expected Contribution

This research makes several unique contributions towards educational data mining and mental health forecasting in an academic context. To begin with, it presents a new proxy-based framework for estimating burnout risk based on behavioral proxies from learning management systems (LMS) and assessment records. This method characterizes burnout as a behavioral deficit syndrome composed of objective indicators such as low engagement activity on the LMS, missed or failed assessments, and long-term declining engagement patterns quite unlike traditional survey methods which depend on subjective self-reporting. The study also advances understanding of elementary behavioral attributes by applying interpretable machine learning models like Random Forest or XGBoost to expose what academic behaviors (like assessment gap inactivity or reduced weekly activity levels) most strongly predict burn out in a student's forward-looking data driven support approach. Finally, the system's cross case high accuracy prediction performance (more than 95% for all models used) validates its theoretical scalability alongside practical deployment value for preemptive intervention standpoints. It can be used within institutional frameworks to automatically identify at-risk students who require timely academic intervention and proactive psychological support before dis-engagement results in dropout separation.

# Limitation and Futire Work

Despite achieving high accuracy, the model has several limitations.  It can be vulnerable to slight class imbalance and lacks interpretability owing to the nature of tree-based models. Moreover, it has not been validated using multiple datasets, which limits its generalizability. Also, critical aspects such as lifestyle as well as socio-demographic variables are missing. The model also requires explaining its decisions by implementing SHAP or LIME and expanding the dataset across multiple organizations. Future research should investigate causal relationships rather than relying solely on association-based models. Burnout progression can be better tracked through incorporating time-series data and deploying the model as a real-time tool adds considerable practical value.

# Conclusion

A predictive data mining model is presented for assessing student burnout risk from behavioral learning analytics.Utilizing machine learning classifiers namely Logistic Regression, Decision Tree, Random Forest, and XGBoost we analyzed the OULAD dataset to assess predictive performance. Results indicate that ensemble models, particularly Random Forest yield superior accuracy and reliability compared to single-model approaches. Incorporating behavioral features such as assessment scores, study duration, and course interaction was crucial for identifying students at risk. These insights emphasize the value of educational data mining in not only improving academic outcomes but also fostering timely interventions for student mental well-being.

##### References

[1] E. Bresó, J. Salanova, and W. B. Schaufeli, “In search of the ‘third dimension’ of burnout…,” Appl. Psychol., vol. 56, no. 3, pp. 460–478, 2007.

[2] L. P. Macfadyen and S. Dawson, “Mining LMS data to develop an ‘early warning system’…,” Comput. Educ., vol. 54, no. 2, pp. 588–599, 2010.

[3] K. Fahd, S. J. Miah, and K. Ahmed, “Predicting student performance in a blended learning environment…,” Appl. Comput. Inform., 2021.

[4] A. Bessadok et al., “Exploring students’ digital activities and performances through LMS…,” Int. Technol. Smart Educ., vol. 20, no. 1, pp. 58–72, 2023.

[5] Y. H. Hu, C. L. Lo, and S. P. Shih, “Learning management system log analytics…,” J. Comput. Educ., 2025.

[6] C. Romero and S. Ventura, “Educational data mining and learning analytics: An updated survey,” arXiv, 2024.

[7] A. Abedi and S. S. Khan, “Detecting disengagement in virtual learning as an anomaly…,” arXiv, 2022.  
[8] O. Mohamad Nezami et al., “Automatic recognition of student engagement using deep learning…,” arXiv, 2018.

[9] M. Jawthari and V. Stoffa, “OULAD learners’ withdrawal prediction framework,” in Recent Innovations in Computing, LNEE 855, 2022.

[10] Student Burnout: A Review… factors contributing to burnout…, MDPI, 2024.

[11] S. Basri et al., “Do academic stress, burnout and problematic internet use affect perceived learning?…” Sustainability, vol. 14, no. 3, 2022.

[12] S. Ahmed, M. Hossain, and F. Rahman,“Predicting Student Dropout from MOOC Data Using Machine Learning Techniques,”*IEEE Access*, vol. 8, pp. 211547–211558, 2020.

[13] J. Kuzilek, M. Hlosta, and Z. Zdrahal, “Open University Learning Analytics dataset,” \*Scientific Data\*, vol. 4, no. 170171, pp. 1–8, 2017.

|  |
| --- |
| [14] L. Breiman, “Random forests,” \*Machine Learning\*, vol. 45, no. 1, pp. 5–32, 2001. |