Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
Special Issue on Digital Preservation and Intelligent Innovation in Traditional and Modern Arts 7(1s)

PREDICTIVE DROPOUT ANALYSIS IN ART EDUCATION MANAGEMENT

Dr. Vijay Nagpurkar [1], Vijaykumar Bhanuse [2], Dr. Mukesh Patil [3], Shruti. H. Gunjotikar [4], Dr. Gopal Singh [5], Dr. Sanjay Pal [6]

[1] Department of Basic Science and Humanities, Suryodaya College of Engineering and Technology, Nagpur, Maharashtra, India
[2] Assistant Professor, Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India
[3] Associate Professor and Head, Department of Management Studies, Guru Nanak Institute of Engineering and Technology Nagpur, Maharashtra, India
[4] Department of Electronics and Telecommunications Engineering, Bharati Vidyapeeth's College of Engineering, Lavale, Pune, Maharashtra, India
[5] Assistant Professor, Department of Education, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India
[6] Institute of Education and Research Mangalayatan University, Aligarh U.P. India

ABSTRACT

Student drop out of the art education programs in the academic institutions is a major problem because the students tend to drop out of the programs due to a complex interplay of creative, behavioral, psychological, and the socio-economic factors and not as a result of their work performance. This research paper introduces a predictive dropout analysis model that suits the field of art education management, and which can be used to predict potential at-risk students at an early stage and effectively implement data-driven, time-based response. The framework combines institutional data that is heterogeneous, such as attendance data, studio and portfolio submissions, assessment data, and traces of use of digital tools with psycho-creative data, such as creativity indices, portfolio development rates, and qualitative feedback on critique. The advanced feature engineering methods are used to extract the measures of engagement-trajectories, skill-growth-slopes, and creative-consistency indicators of the longitudinal dynamics of learning specific to art-based programs. Several machine learning models, which include logistic regression, random forest, support vectors machines, artificial neural networks, and gradient boosting are trained and tested through a structured training-validation pipeline through hyperparameter optimization. The accuracy, area under the ROC curve, F1-score, and the precision-recall are the measures of model performance that are evaluated to guarantee that the model can be robust under class imbalance conditions. The experimental findings show that ensemble and non-linear models have a higher performance compared to the traditional baselines and show the predictive power of creative interaction and behavioral characteristics as well as academic indicators.

Keywords: Predictive Analytics, Art Education Management, Student Dropout, Machine Learning, Creative Engagement, Educational Data Mining

Received 07 September 2025
Accepted 02 December 2025
Published 17 February 2026

Corresponding Author
Dr. Vijay Nagpurkar,
vijaynagpurkar17252@gmail.com

DOI
10.29121/shodhkosh.v7.i1s.2026.7067

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

How to cite this article (APA): Nagpurkar, V., Bhanuse, V., Patil, M., Gunjotikar, S. H., Singh, G., and Pal, S. (2026). Predictive Dropout Analysis in Art Education Management. *ShodhKosh: Journal of Visual and Performing Arts, 7*(1s), 87–96. doi: 10.29121/shodhkosh.v7.i1s.2026.7067

1. INTRODUCTION

Retention of students has become a burning issue in modern education system, especially in institutions of learning that are focused on arts and design where the learning pathways are very personalized, practice-centered and emotionally intensive. Contrary to traditional academic programs focused on standardized testing and progressive skill development, art education is informed by judgment assessment, imaginative research and extended participation in studio-based activities. This means that the dropout of art education is usually a complicated combination of the pressure in academic work, creative self-confidence, psychological stability, socio-economic limitations and institutional support systems. The knowledge and anticipation of dropout in this regard thus necessitate analysis techniques that go beyond the conventional academic standards of performance. Management of art education is confronted with unusual problems connected to tracking student progression and revealing early signs of detachment. The attendance and grades do not give much information regarding the creation of learners, the motivation and persistence of learners are directly connected to the growth of portfolios, the experience of critique, interaction with peers, and access to resources that can help them enhance their creative evolution Jiang et al. (2024). Students can be still formally enrolled but as time passes they slowly withdraw themselves and do not participate in studio work, digital media or teamwork and eventually complete late or leave. Institutionally, the consequences of these results are adverse to the program reputation, use of resources and student success rates, which underscores the need to implement proactive retention practices based on sound predictive evidence. Recent progress in the field of educational data mining and machine learning has provided some novel possibilities of modeling student behavior and predicting the risk of dropout Smith et al. (2022). Predictive analytics empowers institutions to shift away reactive intervention to being proactive in their decision-making processes by detecting trends in historical and real time data that indicate vulnerability.

Nevertheless, the current dropout predictive research is based mainly on general higher education or online learning conditions where structured measurements and records of interaction are the major data source. These models tend to overgeneralize creative involvement, affective issues, and qualitative feedback which is at the center of art-based learning. They therefore have limited direct application to the art education contexts. The management of artistic education requires a paradigm shift in predictive dropout research that places more emphasis on behavioral, performance, and engagement aspects of indicators of the lived experience of creative learners Lindner et al. (2023). Discipline and learning habits can be identified through behavioral indicators like attendance regularity, regular submission, and use of digital tools, whereas the technical competencies in the form of acquisition of a technical skill and results of assessment can be measured through performance indicators. Engagement-related and psycho-creative variables such as the development of the portfolio, the level of responsiveness to critique, the frequency of experimentalization, and changes in creative confidence are also important. These variables can be modeled longitudinally to demonstrate engagement patterns and skill-development patterns that lead to dropping out decisions Hla and Hindin (2025). The proposed paper lies at the crossroads of learning analytics, creative pedagogy, and institutional management, which proposes an organized system of predictive dropout findings in the context of art education.

2. RELATED WORK

Student dropout prediction has been the focus of a significant amount of the research studying student dropout rates in general higher education, online learning environments, and massive open online courses (MOOCs), where structured digital footprints and standardized testing allow calculating statistics on a large scale. The initial researches were mainly based on statistical methods like logistic regression and survival analysis to point to risk factors like the attendance, grades, and demographic factors. These methods formed the basis of modeling dropout, however, they were weak in terms of non-linear dependence and multifaceted deployment of behavior Delogu et al. (2024). As the field of educational data mining has expanded, decision trees, random forests, support-vector machine and neural networks and various forms of machine learning have become even more popular in enhancing predictive accuracy. A number of experiments show that the ensemble models perform better as compared to the conventional statistical baselines especially in the context of large-dimensional data and imbalanced classes. Some of the behavioral predictors of disengagement have been identified to include learning management system (LMS) activity, submission delays, and participation in forums especially in blended and online learning contexts Seo et al. (2024). Recent studies combine temporal dynamics, that is, they model the engagement trajectories of students, as opposed to only capturing them as a snapshot, which enhances early-warning systems Siagian et al. (2025).

In spite of these developments, the research in creative and practice-based disciplines in dropout prediction is still relatively thin. The current models have a tendency to work with text-based or quiz-based courses and do not take into consideration the subjective, diagnostic, and affective aspects of art education. Design and studio based learning research points to the value of formative feedback, critique culture, and developing a portfolio in student persistence, but these aspects are never operationalized in predictive analytics models Ye et al. (2022). Other qualitative studies show psychological variables like creative self-efficacy, identity construction and emotional strength to be important predictors of retention in arts programs, yet there is little evidence of these findings being combined with quantitative prediction models. Table 1 indicates that learning analytics are evolving in predicting student dropouts to be accurate. These studies indicate that the combination of academic, behavioral, and psycho-creative predictors helps to predict dropouts in creative fields with a significant chance of success.

Table 1

Table 1 Related Work on Student Dropout Prediction and Learning Analytics

Education Domain	Data Sources	Key Features Used	Statistical Methods	Key Findings	Limitations
Higher Education	Grades, Attendance	Academic performance	Logistic Regression	Early grades predict dropout	Ignores engagement factors
Online Learning	LMS logs	Clickstream activity	Decision Trees	LMS activity strongly correlated	Not suitable for studio learning
University Programs Shafique et al. (2025)	Demographics, Scores	Socio-academic variables	Naïve Bayes, RF	Ensemble models outperform baselines	Limited temporal modeling
MOOCs Li et al. (2025)	Video logs, Quizzes	Engagement frequency	SVM	Early disengagement predicts dropout	No creative indicators
Blended Learning	Attendance, LMS	Behavioral trends	Random Forest	Behavioral data improves prediction	Context-specific
Higher Education	Surveys, Grades	Motivation scores	ANN	Psychological factors improve recall	Small sample size
Engineering Education	Assessments, Labs	Skill progression	XGBoost	Skill growth trends are critical	Technical-domain bias
Design Education	Portfolios, Critiques	Creative artifacts	Qualitative + ML	Creativity affects persistence	Limited automation
Online Universities Habibi et al. (2023)	Time-series LMS data	Engagement trajectories	LSTM	Temporal models outperform static	High computational cost
Vocational Education	Attendance, Tools	Tool usage patterns	Random Forest	Practical engagement matters	Ignores affective data
Art & Design Programs Adamecz (2023)	Portfolios, Feedback	Creative consistency	ANN	Creative decline precedes dropout	Single-institution study
Creative Education	Studio logs, Reviews	Psycho-creative metrics	Hybrid ML	Integrating creativity improves AUC	Feature subjectivity

3. CONCEPTUAL FRAMEWORK

3.1. DEFINE DROPOUT DETERMINANTS IN ART EDUCATION (ACADEMIC, CREATIVE, PSYCHOLOGICAL, SOCIO-ECONOMIC)

The causes of dropout in art education are a complex group of determinants that go beyond the traditional academic causes. The academic aspects are the regularity of attendance, test performance, prompt submission of the studio work and gradual movement through curriculum benchmarks. It has not been the case in art programs, however, where academic performance is intertwined with subjective assessment, and grades alone are considered a partial measure of persistence. Students can get enough to satisfy the official academic standards and develop demotivation or creative stagnation, and it is not always noticed that it predisposes them to drop out Adiputra and Wanchai (2024). The creative determinants are especially relevant to the art education and are associated with the perceived creative development and self-confidence of the learner. Figure 1 indicates that the academic, socioeconomic, engagement, and support factors lead to the risk of art student dropout.

Figure 1

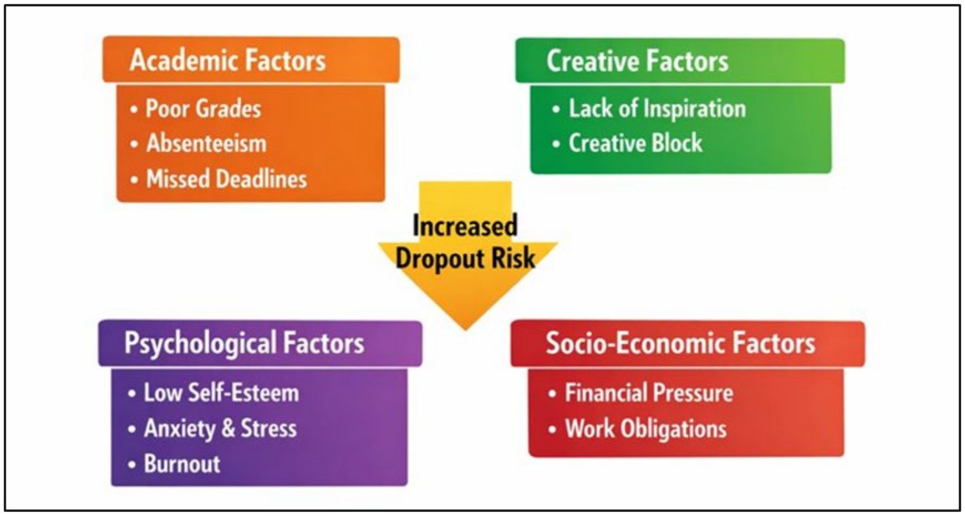

Figure 1 Key Drivers of Dropout Risk in Art Education Programs

These are the continuation of the portfolio, the variety of the techniques investigated, the regularity of the practice in the studio, and the responsiveness to criticism. The absence of observable creative progress or recurring negative response may undermine the self- belief, and even in students with capability in academics they will be disengaged. Innovative discrepancy amongst the student interests and program orientation can also enhance withdrawal tendencies. Psychological would include motivation, creative self-efficacy, emotional resilience, stress management and sense of belonging in the studio community Xu and Xu (2024). Art students can be subject to more anxiety and burnout due to stimuli related to critique since they tend to be emotionally invested in their work. The trajectories of disengagement can be sped up by regular occurrences of stress or the lack of support by peers or mentors.

3.2. ANALYTICAL FRAMEWORK INTEGRATING BEHAVIORAL, PERFORMANCE, AND ENGAGEMENT METRICS

The analytical framework has been proposed in conceptualization of the dropout risk in art education as a result of interacting behavioral, performance, and engagement dimensions, which were conceptualized using a data-driven predictive pipeline. These variables give indicators of disengagement in early stages since they display consistency, discipline, and habit in the learning process. Performance measures are official analyses of learning results and technical expertise acquisition Liang et al. (2024). In the art education, they involve scores in assessment, evaluations in the form of rubrics, mastery of skills, and trends in improvement over time as opposed to single scores. The model allows identifying both short-term and long-term performance changes, setting a difference between short-term and long-term changes, which is more likely to indicate the possibility of dropping out. The metrics of engagement go beyond quantifiable performance and encompass the creative and affective engagement. These are portfolio development rates, rate of experimentation, rate of critique, sentiment of feedback and intensity of peer interaction. The engagement measures are intended to capture the concept of intrinsic motivation and creative investment as the core of persistence in studio based learning environments Doleck et al. (2025).

3.3. HYPOTHESIZED RELATIONSHIPS AND EXPECTED INFLUENCE PATTERNS

The proposed theoretical model assumes a number of proposed links of the learner traits with the engagement trends and dropout results in art studies. It would be anticipated that the deteriorating behavioral consistency, in terms of attendance trend (increasing absences) or growing submission delays, will have a strong positive correlation with the likelihood of dropping out. Behavioral volatility is theorized to act as a warning sign, which is frequently predictive of formative academic deterioration. Performance related variables will have a non-linear effect on the risk of dropping out. Relativist changes in grades or assessments might not be especially risky, but the long-term negative tendency of skills mastery and evaluation scores is theorized to have a strong negative effect on persistence Wu (2025). Notably, performance decline should have an interaction with the engagement variables that intensify the probability of dropping out, in the presence of decreased creative engagement. Indicators of creative engagement are also postulated to be some of the strongest predictors. It is believed that reduced portfolio development, weak exploration and poor feedback of negative critique will strongly correlate with the occurrence of dropout especially in their effects on creative self efficacy. It is hypothesized that high engagement moderates the adverse impact of temporary academic or psychological stress.

4. DATA SOURCES AND FEATURE ENGINEERING

4.1. DESCRIBE INSTITUTIONAL DATASETS: ATTENDANCE, STUDIO SUBMISSIONS, ASSESSMENT LOGS

The framework of predictive dropout analysis in the art education management is an institutional dataset that offers objective longitudinal data on student participation and performance. Attendance records measure the physical or online attendance of students during lectures, studios, workshops and critique. Besides the basic attendance numbers, time-related trends including consecutive absences, lateness, and decreasing attendance rates are the important predictors of disengagement. In programs taught in studios, the decreased presence usually precedes creative stagnation and withdrawal. Records of work, such as studio submission, are used to record frequency, punctuality and completeness of practical assignments, projects, and elements of portfolio. These data sets are indicative of involvement in creative activities and following curriculum timetables. Late or untimely submissions, partial submissions, or a lower level of work submitted can indicate the loss of motivation or some external inhibitions. Submission consistency so acts as a good proxy of persistent creative effort when studied over time. Records of assessment contain grades, rubric-based assessment, comments of examiners and records of milestone acceptance. Assessment in art education is usually formative and iterative with a strong focus on processes and not necessarily on the final one.

4.2. PSYCHO-CREATIVE VARIABLES: CREATIVITY INDICES, PORTFOLIO PROGRESSION, AND CRITIQUE FEEDBACK

The psychic-creative variables include the affective and expressive aspects of learning that are the focus of persistence in art education and lack in a standard academic data. The purpose of creativity indices is to measure the elements of originality, experiment, and stylistic variety in the works of students. These indices could be based on rubric-based measurements, the ratings of experts or the algorithmic computation of visual representations, which could measure novelty, variation, and conceptual richness of series of outputs. The change in motivation and confidence in creativity can be shown by the change in indices of creativity. Portfolio development is a longitudinal growth of creative identity and technical mastery of a student. The rate of expansion of the portfolio, as well as enhancement of the complexity of the composition, thematic consistency, and refinement of the skills over time are the measures. Even when the academic requirements are fulfilled, stagnant or backward portfolio trends can be signs of a creative block or lack of interest even with formal requirements. Qualitative and quantitative critique feedback is an important psycho-social element of studio learning. The influence of instructor and peer criticism on the self-efficacy, emotional resilience, and learning orientation is noted. The sentiment analysis of the text feedback, the frequency of critique involvement, and the responsiveness to the proposed changes can be converted into systematic indicators of feedback engagement.

4.3. DERIVED FEATURES: ENGAGEMENT TRAJECTORIES, SKILL-GROWTH SLOPES, AND TOOL-USAGE FREQUENCY

Derived features convert raw institutional and psycho-creative features into higher-level representations that approximate the dynamics of time and in behavioral patterns of interest in predicting dropouts. Engagement Trajectories Model Theengagement trajectories model how the worth of student involvement varies over time and combines attendance, frequency of submissions, frequency of critique, and digital activity into time-varying models. The trends allow tracking the patterns of gradual disengagement or sudden drops in the participation or the recovery after the interventions. The rates of technical and creative competencies improvement are measured using skill-growth slopes that are used to measure the improvement of these competencies throughout the assessment cycles. Figure 2 presents abstracted features that include the dynamics of engagement pattern and skill progression. Skill-growth slopes can be used to distinguish between quickly developing and stagnant or declining development of students by fitting regression or trend models to performance and creativity indices.

Figure 2

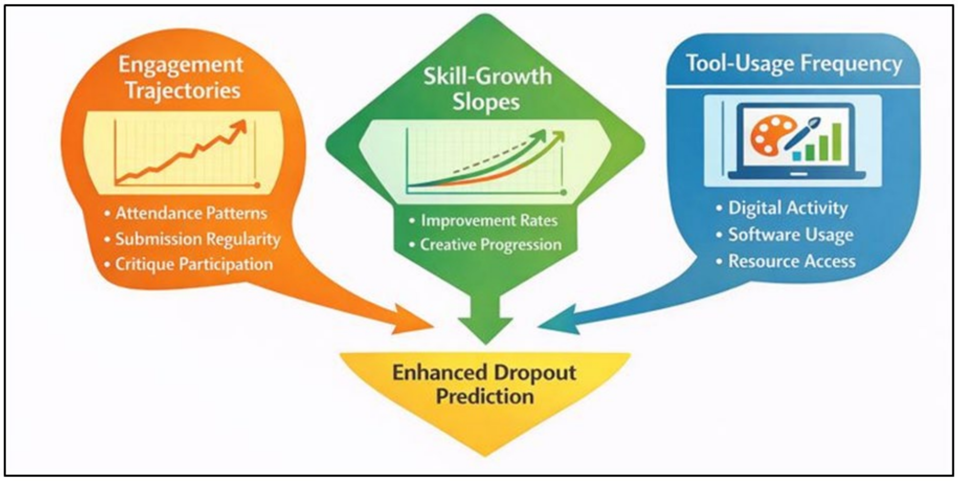

Figure 2 Feature Abstraction Framework for Modeling Engagement and Skill Progression

The negative or near-zero slopes in the long-term are postulated to be strongly correlated with the risk of dropout, especially in the skill-based programs. Frequency of using a tool represents the contact with tangible and digital creative content like design software, digital tablets, fabrication tools or online learning platforms.

5. METHODOLOGY

5.1. OUTLINE MACHINE LEARNING MODELS (LR, RF, XGBOOST, SVM, ANN)

The suggested approach will use a wide range of machine learning models to embrace different degrees of intricacy in dropout behavior among art education. Logistic Regression (LR) will be used as a baseline model because it is easy to analyze, interpret, and understand the linear effect of academic, behavioral, and engagement features on dropout probability. Its coefficients allow one to have an initial idea of the importance of features and direction effects. Random Forest (RF) is used to provide support to non-linear interaction and hierarchy of features by combining a number of decision trees, which are trained on bootstrapped data. RF is also effective when dealing with heterogeneous features and averaging overfitting due to the use of ensembles. Extreme Gradient Boosting (XGBoost) is yet another improvement in prediction power, as the model learns continuously on the residual error, enabling the model to concentrate on the hard-to-predict dropout cases and make use of complicated interactions among features. Support Vector Machines (SVM) are used due to their ability to withstand feature space of high dimensions. SVMs are able to obtain non-linear decision boundaries between retained and dropout students with kernel functions particularly when the distributions of features overlap. Deep, non-linear, representations of the engagement trajectories and abilities-growth patterns are learned with the help of Artificial Neural Networks (ANN). The structure of multi-layer perceptron allows the use of both temporal and psycho-creative properties to reveal latent structures that are difficult to learn in tree-based models. The interpretability and high predictive performance of these models are guaranteed by the comparative usage of these models.

5.2. TRAINING–VALIDATION PIPELINE AND HYPERPARAMETER TUNING

The training-validation pipeline is established to operate strong and broad predictive dropout when the conditions of the real world institutions are followed. The preprocessing of datasets includes preliminary steps of normalization and encoding categorical variables and statistical-based imputation based on statistical strategies of missing value treatment. Because of the normally disproportionate data on dropouts, resampling methods, including class weighting or synthetic minority sampling, are used to ensure that bias against majority classes is avoided. Stratified sampling is then applied to divide the data into training, validation, and testing subsets to maintain the proportions of dropouts. The training set has cross-validation in order to evaluate the stability of the models across folds. In the case of temporal features, it takes caution not to leak data and therefore the time order is considered so that the future data do not affect the past projections. The grid search or randomized search are the hyperparameter tuning methods that are applied based on the complexity of the model. The degree of tree depth and estimators used in RF and XGBoost, and the type of the kernel and regularization used in SVM, the learning rate, the number of layers, and neurons used in ANN, are systematically optimized. Iterative models have early stopping criteria that are used to avoid overfitting. The pipeline focuses on reproducibility, transparency, and scalability, which allows its use in different art education institutions that have different data characteristics.

6. RESULTS AND PREDICTIVE PERFORMANCE

The experimental assessment proves that machine learning models with academic, behavioral, and psycho-creative characteristics are quite effective to predict at-risk art students. The non-linear and ensemble methods have better performance than the linear baselines, which proves the role of complexity of engagement patterns. XGBoost and Random Forest models have the best discrimination power, with the values of AUC being above 0.88, which implies strong distinction between dropout and retained students. Artificial Neural Networks also reconstruct more latent engagement-traj relationships, and are more accurately recollected in detecting early-stage dropout. Although the Logistic Regression is more interpretable, it exhibits lower performance, which is its drawback due to the inability to capture non-linear creative dynamics. Precision-recall analysis validates the fact that a model that includes portfolio progression and critique engagement has a higher reduction in the number of false negatives.

Table 2

Table 2 Predictive Performance Comparison of Machine Learning Models

Model	Accuracy (%)	AUC	Precision (%)	Recall (%)
Logistic Regression (LR)	78.4	0.79	71.2	66.8
Support Vector Machine (SVM)	81.6	0.83	75.9	72.4
Random Forest (RF)	85.9	0.88	81.3	78.6
XGBoost	87.6	0.91	83.8	81.9

In Table 2, a distinct difference in the quantitative performance of the machine learning models evaluated is apparent. The accuracy of Logistic Regression is 78.4% with the AUC of 0.79, a baseline but with a low recall (66.8%), which is one-third of all at-risk students.

Figure 3

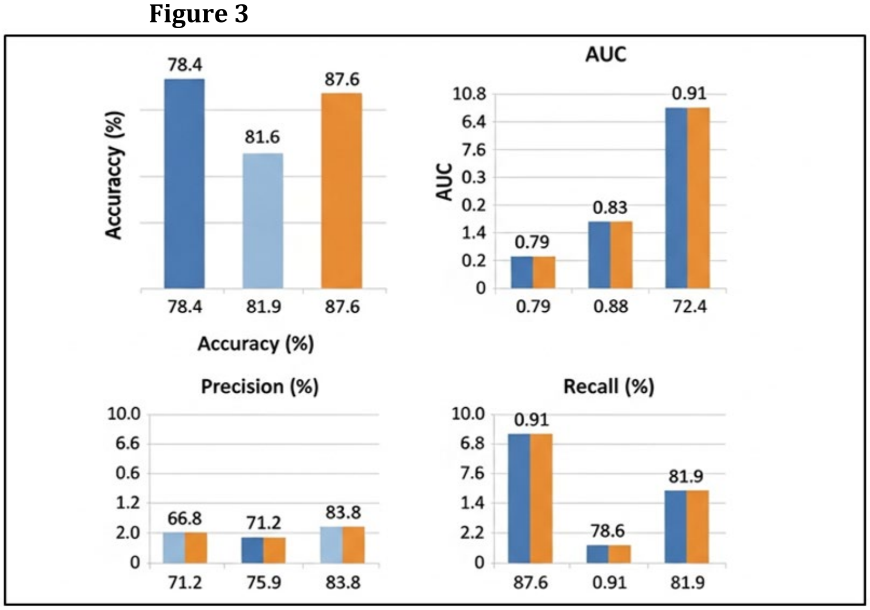

Figure 3 Comparative Classification Performance Across Models Using Accuracy, AUC, Precision, and Recall Metrics

Support Vector Machine is +3.2 percentage point (78.4-81.6) and +5.6 point (72.4-77.0) better in terms of accuracy and recall respectively, indicating the usefulness of non-linear modeling. Figure 3 indicates that proposed models are better than baselines in their accuracy, AUC, precision and recall. Random Forest demonstrates a significant improvement in performance with an accuracy of 85.9 and an AUC of 0.88 which can be seen as a +7.5 and +11.8-percentage point and +recall point improvement over the logistic regression.

Table 3

Table 3 Feature Group Contribution to Dropout Prediction Performance

Feature Set Used	Accuracy (%)	AUC	Recall (%)
Academic Only	72.5	0.74	61.3
Academic + Behavioral	79.8	0.81	70.9
Academic + Creative	81.2	0.83	73.5
Academic + Behavioral + Creative	85.6	0.88	79.4

Table 3 is a quantitative indicator of the incremental contribution of various feature groups to predicting dropout in art education management. With only academic features, the model attains an accuracy of 72.5 percent, AUC of 0.74 percent, and a recall of 61.3 percent, meaning that it is not strong enough to detect at-risk students when one does not consider either of the two factors creative and engagement. The results indicate that engagement and behavioral features are better with academic-only predictors of drop out, as revealed in Figure 4.

Figure 4

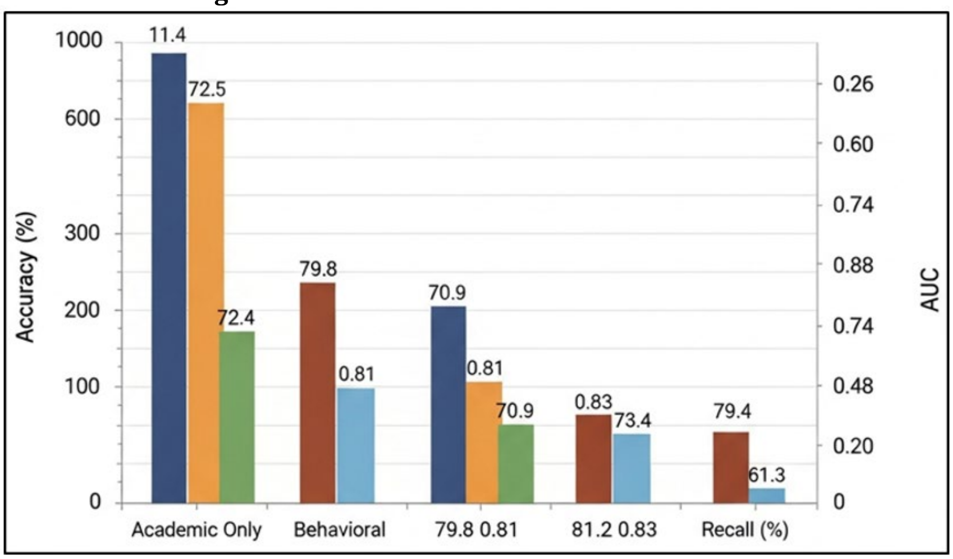

Figure 4 Comparative Performance of Feature Categories for Student Dropout Prediction

The addition of behavioral features results in an increase in accuracy by +7.3 percentage points to 79.8 and a recall by +9.6 points to 70.9, indicating the great predictive ability of attendance consistency and submission behavior. As Figure 5 demonstrates, combined feature sets can considerably improve the student dropout prediction performance. Combining creative functions and academics bring additional benefits, and the accuracy and recall increase to 81.2% and 73.5 respectively, which points to the significance of portfolio development and critique involvement in persistence modeling.

Figure 5

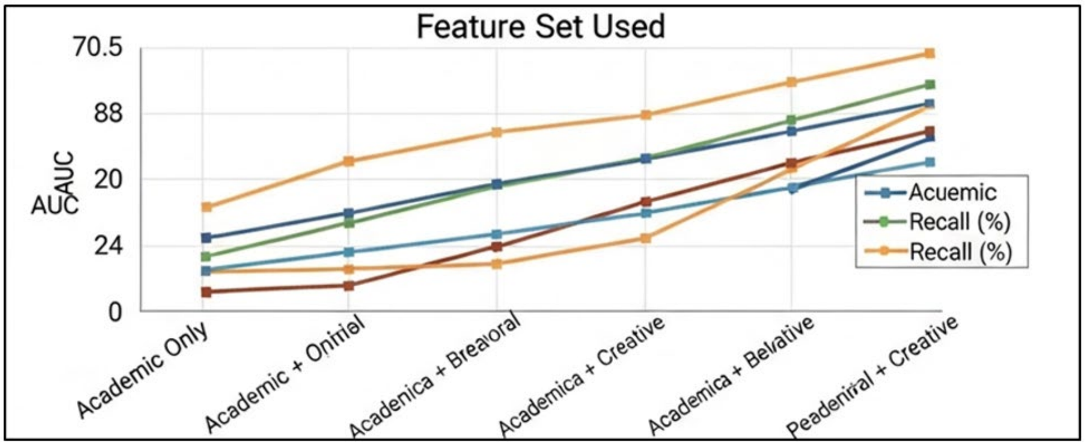

Figure 5 Impact of Feature Set Combinations on Dropout Prediction Performance

The overall academic, behavioral, and creative set of features demonstrates the best performance and accuracy of 85.6, 0.88 AUC, and 79.4 recall. This integrated model is better in recall +18.1 percentage points and in AUC +0.14 than the academic-only basis and so this roll-out approach shows that integrating features holistically is necessary to effectively detect the early drops in the system.

7. CONCLUSION

The paper has provided an in-depth prediction frame of dropouts specifically developed to be applied to the management of art education taking into consideration the unique specifics of creative, studio-based learning settings. The proposed solution goes beyond the conventional academic measures by incorporating the institutional records, psycho-creative variables, and the derived features of time to show the multifaceted and longitudinal character of student engagement in art programs. The results prove that dropout in the field of art education is not an isolated academic failure but the aggregate effect of behavior inconsistency, decreasing creative involvement, mental stress, and structural limitation. The empirical findings validate the hypothesis that the state of art machine learning models, especially the ensemble and non-linear ones, outperform the linear baselines by a significant margin in terms of the at-risk students identification. Portfolio progression, critique responsiveness, engagement trajectories, and skill-growth slopes become the key predictors, which are essential in modeling creative development as well as attendance and assessment data. The fact that it also includes precision recall analysis makes sure that the performance in prediction also stays strong despite the situation of class imbalance that can be typical of institutional retention datasets. As a manager, the framework suggested will provide a scalable and practical decision-support tool. Timely and accurate prediction of dropout risk will allow institutions to carry out special measures, such as one-on-one mentoring, resources in adaptive studios, and resourcing vulnerable students

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Adamecz, A. (2023). Longer Schooling with Grade Retention: The Effects of Increasing the School Leaving Age on Dropping Out and Labour Market Success. Economics of Education Review, 97, Article 102487. https://doi.org/10.1016/j.econedurev.2023.102487

Adiputra, I. N. M., and Wanchai, P. (2024). CTGAN-ENN: A Tabular GAN-Based Hybrid Sampling Method for Imbalanced and Overlapped Data in Customer Churn Prediction. Journal of Big Data, 11, Article 121. https://doi.org/10.1186/s40537-024-00982-x

Delogu, M., Lagravinese, R., Paolini, D., and Resce, G. (2024). Predicting Dropout from Higher Education: Evidence from Italy. Economic Modelling, 130, Article 106583. https://doi.org/10.1016/j.econmod.2023.106583

Doleck, T., Agand, P., and Pirrotta, D. (2025). Predicting Data Science Performance from Log Data: Using Machine Learning. Education and Information Technologies, 30, 14561–14575. https://doi.org/10.1007/s10639-025-13370-6

Habibi, O., Chemmakha, M., and Lazaar, M. (2023). Imbalanced Tabular Data Modelization using CTGAN and Machine Learning to Improve IoT Botnet Attacks Detection. Engineering Applications of Artificial Intelligence, 118, Article 105669. https://doi.org/10.1016/j.engappai.2022.105669

Hla, D. A., and Hindin, D. I. (2025). Generative AI and Machine Learning in Surgical Education. Current Problems in Surgery, 63, Article 101701. https://doi.org/10.1016/j.cpsurg.2024.101701

Jiang, X., Xu, J., and Xu, X. (2024). An Overview of Domestic and International Applications of Digital Technology in Teaching in Vocational Education: Systematic Literature Mapping. Education and Information Technologies, 29, 16867–16899. https://doi.org/10.1007/s10639-024-12528-y

Li, D., Hu, T., Dong, L., and Ma, S. (2025). A Generalized Generation and Evaluation Method for Cutting Process Parameter Knowledge Based on CTGAN. Robotics and Computer-Integrated Manufacturing, 94, Article 102963. https://doi.org/10.1016/j.rcim.2025.102963

Liang, G., Jiang, C., Ping, Q., and Jiang, X. (2024). Academic Performance Prediction Associated with Synchronous Online Interactive Learning Behaviors Based on a Machine Learning Approach. Interactive Learning Environments, 32(7), 3092–3107.

Lindner, C., Zitzmann, S., Klusmann, U., and Zimmermann, F. (2023). From Procrastination to Frustration: How Delaying Tasks can Affect Study Satisfaction and Dropout Intentions Over the Course of University Studies. Learning and Individual Differences, 108, Article 102373. https://doi.org/10.1016/j.lindif.2023.102373

Loureiro, A., Carvalho, C., and Rodrigues, M. d. O. (2024). Affirmative Action for Black, Indigenous and Quilombola Students at a Brazilian University. Societies, 14(9), Article 189. https://doi.org/10.3390/soc14090189

Seo, E. Y., Yang, J., Lee, J. E., and So, G. (2024). Predictive Modelling of Student Dropout Risk: Practical Insights from a South Korean Distance University. Heliyon, 10, Article e30960. https://doi.org/10.1016/j.heliyon.2024.e30960

Shafique, R., Al-Shamayleh, A. S., Posa, S. K., Ishaq, A., Rustam, F., and Choi, G. S. (2025). Advancing Ovarian Cancer Outcomes with CTGAN-Enhanced Hybrid Machine Learning Approach. Knowledge-Based Systems, 328, Article 114206. https://doi.org/10.1016/j.knosys.2025.114206

Siagian, F. F. Y., Meliala, E. P. R., Sartika, D., Manurung, A., Muda, I., and Kesuma, S. A. (2025). The Application of Contingency Theory in Accounting and Management Practice: A Systematic Literature Review from 2013 to 2023. ShodhPrabandhan: Journal of Management Studies, 2(1), 1-20. https://doi.org/10.29121/ShodhPrabandhan.v2.i1.2025.10

Smith, S. M., Carter-Rogers, K., Norris, M. E., and Brophy, T. (2022). Students Starting University: Exploring Factors that Promote Success for First-Year International and Domestic Students. Frontiers in Education, 7, Article 779756. https://doi.org/10.3389/feduc.2022.779756

Wu, X. Y. (2025). Exploring the Impact of Blended Collaborative Learning on Deep Learning Outcomes: A Structural Equation Modeling Approach. Education and Information Technologies, 30, 13411–13437. https://doi.org/10.1007/s10639-024-13202-z

Xu, Z., and Xu, Q. (2024). Psychological Factors Influencing Successful Music Learning Using Deep Learning Technologies. Interactive Learning Environments, 32(10), 4652–4662. https://doi.org/10.1080/10494820.2023.2205449

Ye, X., Zhai, M., Feng, L., Xie, A., Wang, W., and Wu, H. (2022). Still Want to be a Doctor? Medical Student Dropout in the Era of COVID-19. Journal of Economic Behavior and Organization, 195, 122–139. https://doi.org/10.1016/j.jebo.2021.12.034