City University eJournal of Academic Research (CUeJAR)
e-ISSN : 2682-910X
CUeJAR Homepage: https://www.city.edu.my/CUJAR

OPEN ACCESS

Research on User Experience Optimization of Mobile Learning Video Conference System Based on Machine Learning Prediction and Prototype Evaluation

Lu Lei [1], Nazean Jomhari[2]

[1]*City University of Malaysia, Malaysia*
[2]*Department of Software Engineering, Faculty of Computer Science and Information Technology, Universiti Malaya, Malaysia*

Abstract
This study aims to optimize the user experience of mobile learning video conferencing systems by combining machine learning methods with user centered design. Based on 517 valid questionnaire data, the study analyzed user experience from three dimensions: usability, user emotion, and user value, and used five classifiers including logistic regression, support vector machine, random forest, XGBoost, and LSTM for prediction. Three video conferencing system prototypes were further designed and evaluated through usability testing by 20 teachers, combined with machine learning prediction results and user feedback, to determine the optimal prototype. The results showed that logistic regression achieved perfect predictive performance (accuracy of 1.0000) for all three user experience dimensions under the 90:10 data partition, while prototype A performed the best in usability and user value, and was recommended as the basic framework for iterative development.

Keywords: *user experience; Machine learning; Prototype design; Mobile learning; Video conferencing system*

1. Introduction

In 2020, the World Health Organization declared COVID-19 a pandemic, leading to widespread social distancing measures and school closures. This severely disrupted global education, forcing institutions to adopt e-learning and mobile learning (m-learning) video conferencing platforms such as Zoom and Google Meet as primary instructional tools (Liguori & Winkler, 2020; Demuyakor, 2020; Ratten, 2020).

While m-learning enabled educational continuity, it also exposed challenges in usability, emotional engagement (affect), and perceived value for students and educators (Dhawan, 2020; Almaiah et al., 2022). Issues like technical barriers, stress, and limited engagement underscored the need to optimize m-learning platforms.

This study integrates usability, affect, and value through a machine learning (ML) predictive model, using real-time user data to enhance the user experience in m-learning environments.

CUeJAR
Received: 17th October 2025
Revised: 14th November 2025
Accepted: 5th December 2025

1.1 Content of Study

The pandemic lockdowns revealed significant deficiencies in digital education infrastructure, specifically within mobile learning (m-learning) video conferencing. According to Alodwan (2021), user satisfaction among students and teachers was highly influenced by three key factors: platform usability, emotional engagement, and perceived learning value.

Usability: Issues like complex interfaces and technical glitches disrupted instructional flow.

Affect: A lack of interaction led to disengagement and reduced user motivation.

User Value: Doubts about the effectiveness of online learning compared to in-person classes impacted long-term adoption.

Addressing the distinct needs of educational users is crucial. Students need intuitive, interactive tools to maintain engagement, while teachers require seamless content delivery and manageable classroom controls. Predicting user experience across these dimensions is key to developing adaptive m-learning platforms tailored for education.

1.2 Research Focus

This study employs Machine Learning classification algorithms to predict user experience (UX) in mobile learning video conferencing platforms. The research focuses on three core dimensions—Usability, User Affect, and User Value—to forecast and enhance the overall UX for learners and educators. These dimensions uniquely and dynamically shape user perception and interaction within mobile learning environments.

Usability: This dimension assesses the ease of use, interface clarity, and functional efficiency of the tools. It directly influences user satisfaction and learning continuity by determining how intuitively users navigate, access materials, and engage.

User Affect: This component explores users' emotional and psychological responses, including engagement, frustration, and motivation. Understanding affect is essential for creating emotionally supportive environments that foster sustained participation.

User Value: This dimension measures the perceived benefit and relevance of the platform in achieving educational goals. It reflects the platform's contribution to learning outcomes and is tied to the long-term adoption and strategic success of the technology.

1.3 Research Gap

Despite the widespread use of mobile learning video conferencing apps, research overlooks the UX of educators and administrators, who are critical to platform success. Additionally, while ML techniques have been applied in UX analysis, their potential in predicting and classifying usability, emotional impact, and perceived value in mobile learning video conferencing remains underexplored.

This study bridges the gap by employing advanced machine learning classification algorithms to analyze survey responses from students, focusing on ordinal data patterns in usability, affect, and user value.

2. Literature review

Table 2.1: Critical Analysis Table

No	Author, Year	Title	Research Focus	Significance	Limitation
1	(Kwok & Ng, 2023)	Looking through the fog of remote Zoom teaching: a case study of at-risk student prediction	Predict at-risk students in remote Zoom-based learning using behavioural and engagement data.	RFchosen for handling non-linear relationships and feature importance analysis.	No hyperparameter tuning
2	(Khoudi et al., 2025)	Leveraging machine learning and clickstream data to improve student performance prediction in virtual learning environments	Evaluate the utility of clickstream data and machine learning algorithms in predicting student performance and enhancing online learning experiences.	Advanced time series GRU model outperformed six baseline models, achieving an accuracy of 90.13%	Limited comparison with deep learning
3	(Al Noman Bhuyan et al., 2025)	A Machine Learning Approach for Predicting Student Engagement with Online Learning Platform	Predict student engagement (behavioral data and demographic factors) in online learning platforms using ML techniques.	XGBoost outperformed all the ML algorithm	Simpler models (LR) were favored for interpretability, while NN/SVM were limited by computational costs

4	(Althibyani, 2024)	Predicting student success in MOOCs: a comprehensive analysis using machine learning models	Predict student success in (MOOCs) by analyzing behavioral, demographic, and interaction data.	XGBoost achieved the highest accuracy	Predictions are static; real-time adaptation (LSTMs) was not explored
5	(Nguyen et al., 2025)	Early prediction of student performance based on behavioral data in blended learning	Predict student performance in a blended learning course by analyzing behavioral data, including online activity, offline engagement	Gradient Boosting achieved the highest accuracy (~88-92%) in predicting final grades	Feature selection bias where Did not account for external factors such as socioeconomic status)
6	(Tong and Zhan, 2023)	An evaluation model based on procedural behaviors for predicting MOOC learning performance : students' online learning behavior analytics and algorithms construction	Develops a procedural behavior-based evaluation model to predict student performance in MOOCs by analyzing temporal sequences of learning behaviors (clickstream data, interaction patterns, time-based metrics)	LSTM Networks outperformed traditional ML models (~89-93% accuracy) by leveraging sequential behavior patterns.	LSTM training requires significant resources vs. traditional ML.
7	(Miranda et	Machine learning's	Evaluates interpretable ML techniques for	LightGBM achieved 91.2% accuracy (AUC:	LightGBM struggled with sparse

	al., 2024)	model-agnostic interpretability on the prediction of students' academic performance in video-conference-assisted online learning	predicting academic performance in video-conference-based online learning (e.g., Zoom classes) during COVID-19.	0.94) in predicting pass/fail outcomes.	participants (students who rarely turned on cameras/mics).
8	(J.-E. Lee et al., 2023)	A comparison of machine learning algorithms for predicting student performance in an online mathematic s game	evaluates ML algorithms to predict student performance (e.g ., mastery level, problem-solving speed) in an online math game environment with behavioral patterns, game interaction data, demographic factors	XGB achieved the highest accuracy (91%) in predicting mastery levels, outperforming other algorithms	Less interpretable than simpler models (though SHAP values help).
9	(Alsayat and Ahmadi, 2022)	A Hybrid Method Using Ensembles of Neural Network and Text Mining for Learner Satisfaction Analysis from Big Datasets in Online Learning	Hybrid AI framework combining neural networks and text mining to analyze learner satisfaction in online education platform by processing structured behavirol data and unstructured text data	Hybrid Ensemble (LSTM + BERT + XGBoost) achieved 94.2% accuracy in satisfaction classification, outperforming single-model approaches.	Text analysis focused on post-hoc feedback rather than real-time interaction. BERT and LSTM require heavy GPU resources, limiting real-time

		Platform			deployment.
10	(Alhothali et al., 2022)	Predicting Student Outcomes in Online Courses Using Machine Learning Techniques: A Review	Systematic review analyzes machine learning (ML) techniques used to predict student outcomes (e.g., dropout, grades, engagement) in online course	Ensemble methods (XGBoost, Random Forest) were most common (~45% of studies), achieving average accuracy of 85–92% for dropout prediction.	Data Heterogeneity: Varied metrics (e.g., accuracy, F1-score) made cross-study comparisons difficult.
11	(Tan et al., 2022)	An intelligent tool for early drop-out prediction of distance learning students	Early warning system to predict student drop-out risk in distance learning programs with Engagement metrics, Academic performance, Demographic data (e.g., age, prior education)	XGBoost achieved the highest performance (AUC-ROC = 0.94), outperforming other ML models.	Hyperparameter tuning needed
12	(Nguyen, 2024)	Applying Learning Analytics to Predict the Student's Learning Outcome Based on	Develops a machine learning (ML) framework to predict student learning outcomes in online courses with behavioural	XGBoost achieved the highest prediction accuracy (91.2% AUC-ROC), outperforming other ML models.	Qualitative factors is excluded in feature selection, Less interpretable than linear models

		Online Learning Activities	engagement data , temporal patterns, and performance metrics		
13	(Singh and Renuga Devi, 2022)	Analysis of Student Sentiment Level using Perceptual Neural Boltzmann Machine Learning Approach for E-learning Applications	Analyze student sentiment levels in e-learning platforms by processing textual feedback, behavioural data, and multimodal inputs	PNBM achieved 92.3% accuracy in sentiment classification, outperforming traditional models (LSTM, SVM).	PNBM requires high-end GPUs and extensive training time (~5× longer than LSTM).
14	(Yu et al., 2022)	Analysis of student e- learning engagement using learning affect:Hybrid of facial emotios and domain model	Hybrid AI framework to measure student engagement in e-learning by combining Facial emotion analysis, Domain-specific behavioral metrics, Temporal patterns to identify at-risk students early and provide actionable insights for instructors.	Hybrid CNN-LSTM achieved 93.5% accuracy in engagement prediction, Outperforming unimodal approaches	Hardware dependency where Requires high-quality webcams and GPUs; fails with poor lighting/angles.

15	(Ayouni et al., 2021)	A new ML-based approach to enhance student engagement in online environment	Develops a machine learning framework to pre dict and enhance student engagement in online learning by instructor responsiveness. To identify disengaged students early and recommend personalized interventions	Hybrid CNN-GRU achieved 94.1% accuracy in engagement classification (High/Medium/ Low).	High computation al cost (require d GPU clusters). Complex deployment for non-technical institutions.
16	(Luo et al., 2020)	Students' Online Behavior Patterns Impact on Final Grades Prediction in Blended Courses	investigates on online behavior patterns in blended courses impact final grade prediction.	Random Forest achieve d 90.3% accuracy (AUC: 0.93) in grade prediction	Overfitting risk with noisy LMS data.
17	(Rawat et al., 2021)	A systematic analysis using classificatio n machine learning algorithms to understand why learners drop out of MOOCs	Evaluation of ML algorithms to pre dict and analyze dropout causes in MOOCs by exam to Identify the most accurate dropout predictors and o ptimal ML models for early intervention	XGBoost achiev ed the highest accuracy (93.4%) and AU C-ROC (0.96) in dropout prediction, outperforming other classifier	Temporal Gaps where Did not model real-time dropout triggers

18	(Alencar and Netto, 2020)	Measuring Student Emotions in an Online Learning Environment	ML-based framework to detect and classify student emotions (engagement, frustration, confusion) in online learning environments	Hybrid CNN-LSTM achieved 89.5% accuracy in emotion classification, outperforming single-modality approach	Computationally expensive (required GPUs). Black-box decisions hard to explain to educators.
19	(Peng, 2017)	Research on Model of Student Engagement in Online Learning	Prediction of student engagement in online learning environments with behavioural data, Interaction patterns, performance metrics	SVMwith RBF kernel achieved 87.6 % accuracy in engagement classification	SVM struggled with imbalanced data where low engagement cases were underrepresented
20	(Salem and Shaalan, 2025)	Unlocking the power of machine learning in E-learning: A comprehensive review of predictive models for student performance and engagement	Evaluate ML models for predicting student performance and engagement	XGBoost yield the highest accuracy (avg. 92% AUC) for performance prediction	Weak on unstructured data (text/video)

Source: Developed for this research.

2.1 Summary of Literature Review

The transformative shift in mobile learning, accelerated by the COVID-19 pandemic, has made video conferencing platforms essential, heightening the need to understand User Experience (UX) (Miranda et al., 2024; Singh & Renuga Devi, 2022). While machine learning (ML) algorithms like XGBoost, Random Forest, and LSTM have excelled in predicting student performance and engagement, a unified approach to UX—encompassing usability, affect, and value—remains under-explored (Al Noman Bhuyan et al., 2025; Ayouni et al., 2021).

Existing research often isolates these dimensions, focusing solely on usability (Luo et al., 2020), affect (Yu et al., 2022), or value (Alsayat & Ahmadi, 2022). Although behavioral data (e.g., clickstreams, camera usage) are key predictors, they frequently overlook subjective aspects like frustration or perceived usefulness (Khoudi et al., 2025; Tan et al., 2022). Hybrid models integrating facial recognition attempt to bridge this gap but face hardware and complexity constraints (Yu et al., 2022; Alencar & Netto, 2020).

While ensemble and hybrid methods (e.g., CNN-LSTM) improve accuracy, they often suffer from interpretability and computational issues (Alsayat & Ahmadi, 2022; Ayouni et al., 2021). A significant gap is the underutilization of direct, ordinal feedback (e.g., Likert-scale ratings), which offers clearer insights than ambiguous behavioral or text-based proxies (Singh & Renuga Devi, 2022; Rawat et al., 2021).

In summary, the literature reveals a need for an integrated ML framework that concurrently predicts usability, affect, and value. This study addresses the fragmentation in current models and the neglect of ordinal data to provide actionable insights for optimizing mobile learning platforms.

3.Research methods

3.1 Data Collection and Preprocessing

3.1.1 Dataset Description

The foundation of this study's analytical phase is a meticulously curated dataset focused on user experience with mobile learning video conferencing platforms. The data was collected via a structured survey administered to students, capturing their perceptions across three critical dimensions of UX, each measured by seven questions rated on a 5-point Likert scale (1: Strongly Disagree to 5: Strongly Agree).

Usability (7 Questions):

This dimension assessed the functional and operational aspects of the platforms.

1. I often face technological glitches
2. Teachers use break room for group work
3. Teachers create forum for discussion
4. Teacher asked us to submit assignments
5. Interactive features (rose/hat/gifts) enhance engagement
6. Screen sharing & voice features are useful
7. Attendance tracking is important

User Affect (7 Questions):
This dimension captured the emotional and psychological responses of users.
1. Looking at screen all the time is tiring
2. I sometimes log into class and then do not attend
3. There are distractions at home
4. I get bored and rather want to do things I like
5. I am exhausted with video conferencing learning
6. I maintain concentration during online learning
7. I feel comfortable with online learning

User Value (7 Questions):
This dimension evaluated the perceived benefit and effectiveness of the platform for learning.
1. I actively interact with teachers
2. I actively interact with friends
3. I work effectively in groups online
4. My learning grows thanks to online learning
5. Reflection prompts help me process info
6. Interactive quizzes maintain attention
7. Real-time translation improves understanding

The dataset, comprising 517 responses, was found to be complete with no missing values in the critical response fields. Non-essential columns like timestamps were removed to streamline analysis.

3.1.2 Data Preprocess

The dataset, collected through structured surveys targeting users of mobile learning video conferencing platforms, was found to be well-organized and largely complete. No missing values were detected in the critical response fields associated with usability, affect, and user value. However, non-essential columns such as timestamps and respondent identifiers were excluded to optimize memory usage and reduce computational overhead.

Given that all survey responses were recorded using ordinal scales from "5- strongly agree" to " 1- strongly disagree", a one-hot encoding strategy was applied to convert these string-based categories into numerical representations. This transformation preserved the semantic integrity of the data while aligning it with the input requirements of various machine learning algorithms. Additionally, the dataset was segmented into three distinct groups usability, user affect, and user value to facilitate targeted analysis and model training. This structured organization enabled a more coherent and interpretable dataset, laying the groundwork for subsequent predictive modelling.

3.1.3 Data Exploration

The exploratory phase involved a detailed examination of each response column, both individually and within their respective groups. Special attention was given to the distribution of feedback across the usability, affect, and value dimensions. Visualization techniques, primarily bar charts, were employed to illustrate the frequency of each ordinal response.

The distribution of responses for usability, user affect, and user value is presented in Figures 3.1 through 3.3. These visualizations serve as a foundational reference for the predictive modeling phase, offering clarity on the data's structure and guiding the interpretation of machine learning outcomes.

Figure 3.1: Bar chart showing the distribution of responses for the Usability dimension across the three prototypes. The X-axis represents the Likert scale ratings (1-Strongly Disagree to 5-Strongly Agree), and the Y-axis shows the frequency of each rating. Each bar represents the number of teachers who selected a particular rating for the usability-related questions. This visualization provides an overview of the perceived ease of use and functionality of each prototype.

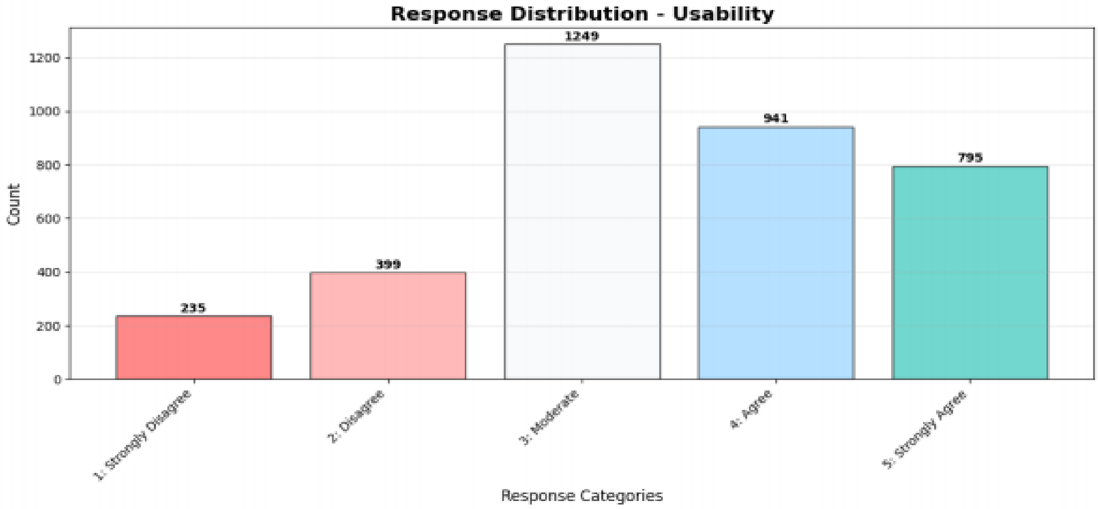

Figure 3.1: Bar Chart Visualization - Response Distribution for Usability

Source: Developed for this research.

Figure 3.2: Bar chart showing the distribution of responses for the User Affect dimension across the three prototypes. The X-axis represents the Likert scale ratings (1-Strongly Disagree to 5-Strongly Agree), and the Y-axis shows the frequency of each rating. Each bar represents the number of teachers who selected a particular rating for the affect-related questions. This visualization provides an overview of the emotional and psychological responses of teachers to each prototype.

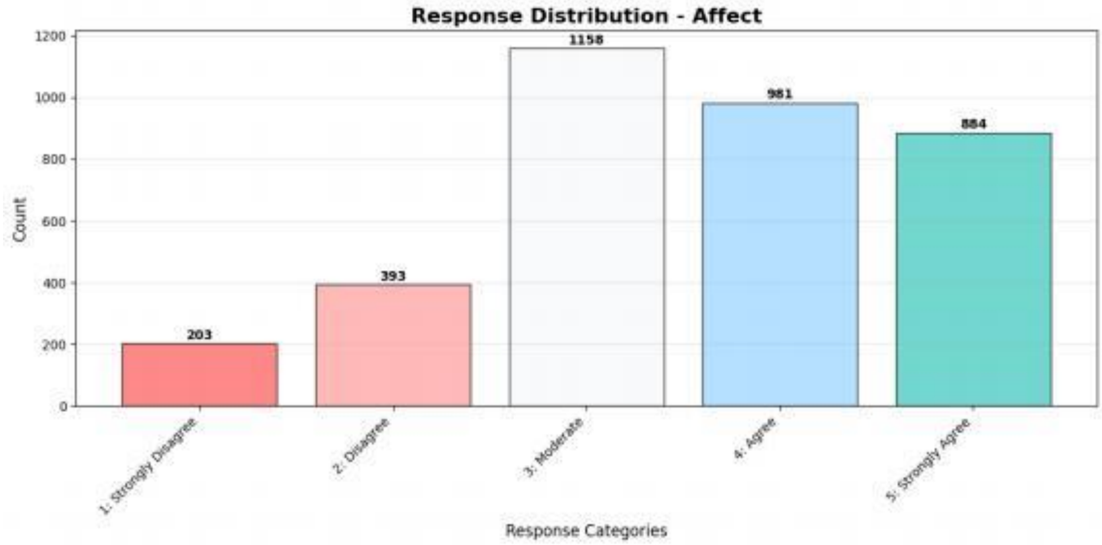

Figure 3.2: Bar Chart Visualization - Response Distribution for User Affect
Source: Developed for this research.

Revised Caption: "Figure 3.3: Bar chart showing the distribution of responses for the User Value dimension across the three prototypes. The X-axis represents the Likert scale ratings (1-Strongly Disagree to 5-Strongly Agree), and the Y-axis shows the frequency of each rating. Each bar represents the number of teachers who selected a particular rating for the value-related questions. This visualization provides an overview of the perceived benefit and effectiveness of each prototype in supporting teaching and learning outcomes.

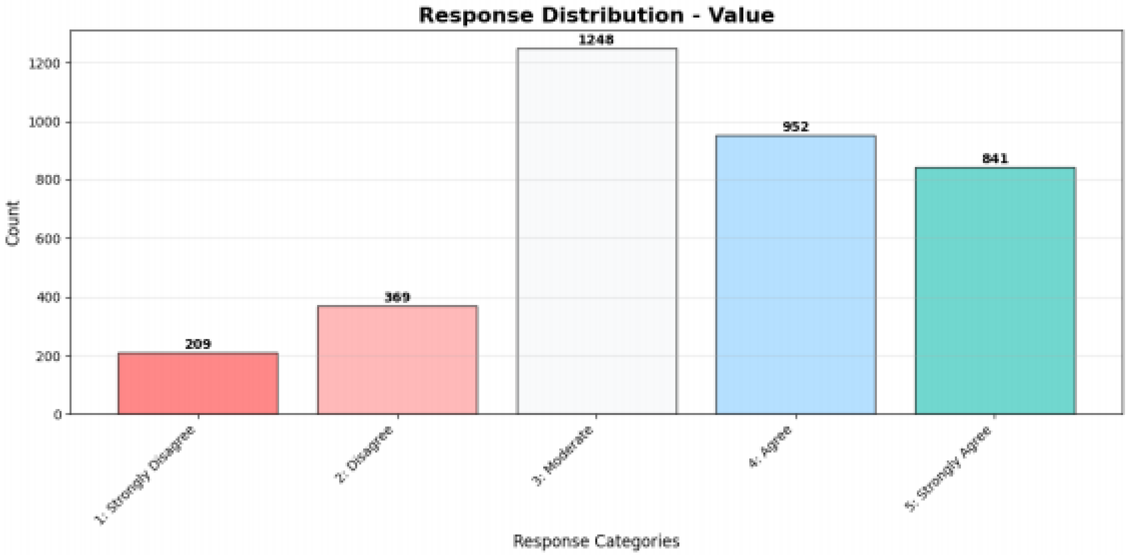

Figure 3.3: Bar Chart Visualization - Response Distribution for User Value
Source: Developed for this research.

City University eJournal of Academic Research (CUeJAR)
e-ISSN : 2682-910X
CUeJAR Homepage: https://www.city.edu.my/CUJAR

OPEN ACCESS

3.2 Machine Learning Models

3.2.1 Machine Learning Algorithms

The questionnaire used in this study was thoughtfully designed to capture the multifaceted nature of user experience in mobile learning video conferencing platforms. It is divided into three distinct sections, each representing a core dimension of UX:

1. Usability: Measures the ease of use, interface clarity, and functional accessibility of the platform.
2. User Affect: Captures emotional responses, engagement levels, and perceived comfort during usage.
3. User Value: Assesses the perceived benefit, relevance, and strategic value of the platform in supporting learning outcomes.

Each section comprises multiple ordinal-response items, typically rated on a Likert scale. These responses provide structured, interpretable data ideal for classification tasks. To extract meaningful insights, the dataset was analyzed through three strategic machine learning scenarios, each designed to explore different facets of the data:

- Scenario 1 - Usability

Focusing solely on usability items, this scenario aims to isolate the functional component of UX. Desired correlations include the relationship between usability and perceived learning effectiveness and the predictive power of interface design on overall satisfaction.

- Scenario 2 - User Affect

This scenario explores the emotional and perceptual layer of UX. Key correlations include the impact of emotional engagement on continued platform use and the linkage between affective responses and perceived value.

- Scenario 3 - User Value

By isolating user value items, this scenario investigates the strategic and outcome- oriented aspects of UX. Correlations of interest include the influence of perceived value on platform recommendation likelihood and the relationship between value perception and affective engagement.

The strategic analysis of the three focused scenarios usability, user affect, and user value offers targeted insights into the distinct contributions each dimension makes to the overall user experience in mobile learning video conferencing platforms. By isolating these groups, the study reveals how functional ease (usability), emotional engagement (affect), and perceived benefit (value) independently influence user satisfaction and platform effectiveness. This approach enables a clearer understanding of which UX factors are most predictive and actionable, guiding developers and educators toward more precise enhancements. The findings underscore the importance of designing learning tools that not only function smoothly but also resonate emotionally and deliver meaningful value to users.

The strategic analysis of the three focused scenarios usability, user affect, and user value offers targeted insights into the distinct contributions each dimension makes to the overall user experience in mobile learning video conferencing platforms. By isolating these groups, the study reveals how functional ease (usability), emotional engagement (affect), and perceived

CUeJAR
Received: 17th October 2025
Revised: 14th November 2025
Accepted: 5th December 2025

benefit (value) independently influence user satisfaction and platform effectiveness. This approach enables a clearer understanding of which UX factors are most predictive and actionable, guiding developers and educators toward more precise enhancements. The findings underscore the importance of designing learning tools that not only function smoothly but also resonate emotionally and deliver meaningful value to users.

NO	Author, Year	Title	Machine Learning Algorithms								
			LR	NNW	SVM	L GBM	XGB	RF	LSTM	NB	KNN
1	(Kwok & Ng, 2023)	Looking through the fog of remote Zoom teaching: a case study of at-risk student prediction	1					1	1		
2	(Khoudi et al., 2025)	Leveraging machine learning and clickstream data to improve student performance prediction in virtual learning environments	1				1	1	1	1	1
3	(Al Noman Bhuyan et al., 2025)	A Machine Learning Approach for Predicting Student Engagement with Online Learning Platform	1	1	1			1			
	(Althibyani, 2024)	Predicting student success in MOOCs: a comprehensive analysis using machine learning models	1	1	1		1				1
5	(Nguyen et al., 2025)	Early prediction of student performance based on behavioral data in blended learning	1	1	1		1	1			1
6	(Tong and Zhan, 2023)	An evaluation model based on procedural behaviors for predicting MOOC learning performance: students' online learning behavior analytics and algorithms construction	1		1		1	1	1		
7	(Miranda et al., 2024)	Machine learning's model-agnostic interpretability on the prediction of students' academic performance in video-conference-assisted online learning during the covid-19 pandemic	1		1	1		1			
8	(J.-E. Lee et al., 2023)	A comparison of machine learning algorithms for predicting student performance in an online mathematics game	1		1		1	1			1
9	(Alsayat and Ahmadi, 2022)	A Hybrid Method Using Ensembles of Neural Network and Text Mining for Learner Satisfaction Analysis from Big Datasets in Online Learning Platform			1		1		1		
10	(Alhothali et al., 2022)	Predicting Student Outcomes in Online Courses Using Machine Learning Techniques: A Review	1		1		1	1	1		

Figure 3.4: Contingency Table for Machine Learning Algorith

Source: Developed for this research.

NO	Author, Year	Title	Machine Learning Algorithms								
			LR	NNW	SVM	L GBM	XGB	RF	LSTM	NB	KNN
11	(Tan et al., 2022)	An intelligent tool for early drop-out prediction of distance learning students	1		1		1	1			1
12	(Nguyen, 2024)	Applying Learning Analytics to Predict the Student's Learning Outcome Based on Online Learning Activities	1		1		1	1	1		
13	(Singh and Renuga Devi, 2022)	Analysis of Student Sentiment Level using Perceptual Neural Boltzmann Machine Learning Approach for E-learning Applications			1				1		
14	(Yu et al., 2022)	Analysis of student e-learning engagement using learning affect: Hybrid of facial emotions and domain model						1	1		
15	(Ayouni et al., 2021)	A new ML-based approach to enhance student engagement in online environment			1		1		1		
16	(Luo et al., 2020)	Students' Online Behavior Patterns Impact on Final Grades Prediction in Blended Courses	1		1		1	1			
17	(Rawat et al., 2021)	A systematic analysis using classification machine learning algorithms to understand why learners drop out of MOOCs	1		1		1	1		1	1
18	(Alencar and Netto, 2020)	Measuring Student Emotions in an Online Learning Environment			1			1	1		
19	(Peng, 2017)	Research on Model of Student Engagement in Online Learning	1					1		1	
20	(Salem and Shaalan, 2025)	Unlocking the power of machine learning in E-learning: A comprehensive review of predictive models for student performance and engagement				1	1		1		
	Total		14	3	15	2	13	15	11	3	6

Figure 3.5: Contingency Table for Machine Learning Algorithms (Cont'd)

Source: Developed for this research.

According to Figure 3.4 and 3.5, the contingency table for machine learning algorithms summaries the ML algorithms applied by different researcher based on the literature review studies. The top 5 ML algorithms (LR, SVM, XGB, RF, LSTM) is then applied in this research for further exploration.

1. Logistic Regression (LR)

LR is a linear classification algorithm used for binary or multiclass classification. It models the probability of a class using the logistic (sigmoid) function.

Strengths:

• Simple, interpretable, and computationally efficient.

•Works well with linearly separable data.

• Provides probabilities for class membership.

•Less prone to overfitting with regularization (L1/L2).

Weaknesses:

•Assumes a linear relationship between features and log-odds.

•Struggles with non-linear decision boundaries.

•Sensitive to outliers and multicollinearity.

• Requires feature scaling for good performance.

2. Support Vector Machine (SVM)

SVM is a supervised learning model used for classification and regression. It finds the optimal hyperplane that maximizes the margin between classes

Strengths:

•Effective in high-dimensional spaces.

•Works well with clear margin separation.

•Kernel trick allows handling non-linear data (RBF, polynomial).

•Robust to overfitting in high dimensions.

Weaknesses:

• Computationally expensive for large datasets.

•Requires careful tuning of hyperparameters.

•Poor interpretability compared to linear models.

•Performs poorly with overlapping or noisy data.

3. XGBoost

XGBoost is a gradient-boosted decision tree algorithm that optimizes performance using sequential tree building with gradient descent.

Strengths:

• Highly accurate, often state-of-the-art for structured data.

•Handles missing values and feature scaling automatically.

•Regularization prevents overfitting.

•Parallel processing makes it faster than traditional Gradient Boosting Decision Trees.

Weaknesses:

• More complex and harder to interpret than simpler models.

•Requires hyperparameter tuning (learning rate, max depth).

•Can overfit if not properly regularized.

• Slower training than RF due to sequential boosting.

4. Random Forest (RF)
RF is an ensemble learning method that builds multiple decision trees and combines their predictions (bagging).
Strengths:
• Handles non-linear data well.
•Robust to overfitting (due to averaging multiple trees).
•Works with missing values and outliers.
•Provides feature importance scores.
Weaknesses:
• Less interpretable than single decision trees.
• Can be computationally expensive for large datasets.
•May overfit noisy data if trees are too deep.
• Bias towards features with more categories.
5. Long Short-Term Memory (LSTM)

LSTM is a type of recurrent neural network (RNN) designed for sequential data such as time series and NLP. It uses memory cells to retain long-term dependencies.
Strengths:
•Excellent for sequential/temporal data.
•Handles long-term dependencies better than vanilla RNNs.
•Can process variable-length sequences.
• Stateful architecture for context-aware predictions.
Weaknesses:
• Computationally intensive and slow to train.
•Requires large amounts of data to avoid overfitting.
•Hyperparameter tuning is complex (layers, units, dropout).
•Black-box nature reduces interpretability.

3.2.2 Data Partition

Splitting a dataset into training and testing subsets is essential for evaluating a machine learning model's ability to generalize to unseen data and to prevent overfitting. In this project, the initial 70:30 train-test split resulted in poor accuracy across models. To improve performance, the split ratio was adjusted to 80:20 and finally to 90:10.
The choice of a 90:10 partition for a limited dataset (517 rows) is justified for several reasons. With small datasets, maximizing the training data (90%) is crucial for the model to learn diverse patterns effectively. A larger test set (10%) in such cases might not provide enough data for robust evaluation. Furthermore, forgoing a separate validation set is a practical decision to avoid having insufficient data in any single subset. Thus, the 90:10 split balances the need for ample training data with a test set adequate for a reliable performance assessment.

3.2.3 Performance Evaluation Metrics

• Accuracy

Accuracy is a basic yet essential metric that indicates how correct a model's predictions are across all classes. It is determined by dividing the number of correct predictions by the total number of predictions made. Although accuracy gives an overall view of performance, it can be misleading for datasets with an uneven class distribution where one class is much more common than others.

• Precision

Precision measures how reliable positive predictions are. It represents the proportion of true positives among all instances predicted to be positive. Precision is vital in situations where false positives are costly because it shows how well the model avoids incorrectly labeling negatives as positives.

• Recall

Recall also known as sensitivity or the true positive rate, recall measures how well the model identifies all actual positive cases. It is calculated by dividing the number of true positives by the total of true positives and false negatives. Recall is especially important when failing to detect positive cases has serious consequences.

• F1 score

The F1 Score combines precision and recall into a single value using their harmonic mean, balancing both metrics. It is valuable when dealing with imbalanced datasets or when it is important to consider both false positives and false negatives together.

4.Results and Discussion

4.1 Result Comparison

Table 4.1: Performance Evaluation Comparison Matrix for Scenario 1 -Usability

Data Partition	Model	Accuracy	Precision	Recall	F1-Score
70:30	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	0.9487	0.9213	0.9880	0.9535
	Random Forest	0.9295	0.9091	0.9639	0.9357
	XGBoost	0.9359	0.9195	0.9639	0.9412
	LSTM	0.9551	0.9419	0.9759	0.9586
80:20	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	0.9712	0.9492	1.000	0.9739
	Random Forest	0.8942	0.8947	0.910	0.9027
	XGBoost	0.9327	0.9153	0.964	0.9391
	LSTM	0.9327	0.9455	0.928	0.9369
90:10	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	1.0000	1.0000	1.000	1.0000
	Random Forest	0.9423	0.9310	0.964	0.9474
	XGBoost	0.9423	0.9310	0.9643	0.9474
	LSTM	0.9423	0.9630	0.9286	0.9455

Source: Developed for this research.

Table 4.2: Performance Evaluation Comparison Matrix for Scenario 2- User Affect

Data Partition	Model	Accuracy	Precision	Recall	F1-Score
70:30	Logistic Regression	0.9872	0.9798	1.0000	0.9898
	Support Vector Machine	0.9615	0.9505	0.9897	0.9697
	Random Forest	0.9551	0.9500	0.9794	0.9645
	XGBoost	0.9231	0.9048	0.9794	0.9406
	LSTM	0.9615	0.9417	1.0000	0.9700
80:20	Logistic Regression	0.9904	0.9848	1.000	0.9924
	Support Vector Machine	0.9808	0.9846	0.9846	0.9846
	Random Forest	0.9519	0.9412	0.9846	0.9624
	XGBoost	0.9231	0.9130	0.9692	0.9403
	LSTM	0.9615	0.9420	1.0000	0.9701
90:10	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	0.9808	0.9697	1.0000	0.9846
	Random Forest	0.9038	0.8649	1.0000	0.9275
	XGBoost	0.9423	0.9394	0.9688	0.9538
	LSTM	0.9423	0.9143	1.0000	0.9552

Source: Developed for this research

Table 4.3: Performance Evaluation Comparison Matrix for Scenario 3- User Value

Data Partition	Model	Accuracy	Precision	Recall	F1-Score
70:30	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	0.9615	0.9663	0.966	0.9663
	Random Forest	0.9615	0.9882	0.943	0.9655
	XGBoost	0.9359	0.9759	0.910	0.9419
	LSTM	0.9808	0.9886	0.977	0.9831
80:20	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	0.9712	0.9831	0.966	0.9748
	Random Forest	0.9808	1.0000	0.966	0.9831
	XGBoost	0.9712	1.0000	0.950	0.9744
	LSTM	0.9808	1.0000	0.966	0.9831
90:10	Logistic Regression	1.0000	1.0000	1.000	1.0000
	Support Vector Machine	0.9808	1.0000	0.9667	0.9831
	Random Forest	0.9615	0.9667	0.9667	0.9667
	XGBoost	0.9423	0.9655	0.9333	0.9492
	LSTM	0.9423	0.9355	0.9667	0.9508

Source: Developed for this research.

According to Table 4.1, in terms of usability the result highlighted that LR achieved a perfect score with accuracy, precision, recall, and F1-score = 1.0000 across all partition ratios (70:30, 80:20, 90:10), indicating flawless classification. This suggests that usability-related responses were highly structured and predictable. Besides, SVM and LSTM also performed exceptionally well, with F1-scores consistently above 0.95, especially under the 90:10 split where SVM reached perfect scores. Moreover, RF and XGBoost showed slightly lower but still strong performance, with F1-scores ranging from 0.93 to 0.95.

According to Table 4.2, in terms of user affect the result highlighted that LR again led with near-perfect scores, reaching 1.0000 across all metrics in the 90:10 partition. LSTM consistently achieved high recall of 1.0000, indicating its strength in identifying all relevant affective responses. SVM and RF maintained strong performance, though Random Forest showed a slight dip in precision under the 90:10 split, suggesting some misclassification.

According to Table 4.3, in terms of user value the result highlighted that LR maintained perfect scores across all partitions, reinforcing its dominance in structured classification tasks. LSTM, RF, and SVM all achieved F1-scores above 0.96, with RF and LSTM peaking at 0.9831 under the 80:20 split. XGB showed slightly lower performance under the 90:10 split with F1-score = 0.9492, though still within a high-performing range.

In short, for usability and user value, the LR algorithm consistently delivered perfect scores across all data partitions 70:30, 80:20, and 90:10 making it the most reliable and accurate model in terms of user interface effectiveness and task completion. In the dimension of user affect, the same prototype and algorithm combination again stood out, particularly at the 90:10 partition, where LR achieved flawless performance across all metrics. Overall, the Mobile Learning Video Conferencing prototype combined with Logistic Regression, especially at the 90:10 partition, emerges as the most effective configuration across all three dimensions.

LR achieved perfect scores (1.0000) across all metrics and partitions for all three UX dimensions. This suggests that the model was able to classify the ordinal responses with complete accuracy. While impressive, such results are typically indicative of a small and highly structured dataset, which may lack the complexity and variability and well-separated classes, likely contributed to this flawless performance. However, this also raises concerns about overfitting, as the model may have memorized patterns rather than learned generalizable insights.

SVM also demonstrated near-perfect performance, with accuracy and F1-scores consistently above 0.95 across all partitions. Its strength lies in its ability to handle high- dimensional data and find optimal decision boundaries, which is particularly effective when the dataset is clean and the classes are well-defined. The slight variations in precision and recall across partitions suggest that SVM is sensitive to training data volume, but overall, it maintained robust predictive power. Like LR, SVM's high performance may be partially attributed to the limited dataset size, which simplifies the classification task.

RF showed strong but slightly lower performance compared to LR and SVM, with accuracy and F1-scores ranging from 0.90 to 0.95. This model benefits from ensemble learning, reducing variance and improving generalization. However, its performance may have been affected by the small dataset, which limits the diversity of decision trees and reduces the effectiveness of bootstrapping. RF's ability to handle non-linear relationships and feature interactions makes it a valuable model, but its results suggest that it may require larger datasets to fully leverage its strengths.

XGB performed comparably to Random Forest, with metrics consistently above 0.90, though not reaching the perfection of Logistic Regression or SVM. As a gradient boosting algorithm, XGB excels in handling complex patterns and minimizing error iteratively. Its slightly lower scores may reflect the limited depth of patterns in the dataset, which restricts the model's ability to refine predictions. Nonetheless, XGB's stability across partitions indicates that it is a reliable choice for ordinal classification tasks, especially when scalability and performance tuning are required.

LSTM, a deep learning model designed for sequential data, achieved high accuracy and F1-scores, often exceeding 0.95. While the dataset is not inherently time-series, the model's ability to capture contextual dependencies may have contributed to its success. The use of ordinal responses with consistent patterns likely allowed LSTM to learn meaningful representations despite the absence of temporal features. However, LSTM models are computationally intensive and require careful tuning, which may not be justified given the dataset's simplicity. Its strong performance here suggests potential for future applications involving more complex or dynamic UX data.

5. Summary

This study addresses literature gaps by using machine learning to predict UX dimensions—usability, affect, and user value—in mobile video conferencing. While prior work has focused on student performance and engagement in online learning (Kwok & Ng, 2023; Khoudi et al., 2025), emotional and perceptual aspects of UX remain understudied. The use of ordinal survey data provides a structured alternative to text-based sentiment analysis (Alsayat & Ahmadi, 2022; Singh & Renuga Devi, 2022), reducing the ambiguity common in NLP models.

Performance was exceptionally high across all models, with Logistic Regression and SVM achieving near-perfect scores across UX dimensions, aligning with Peng (2017) and Salem & Shaalan (2025). However, these results may surpass literature due to the small, clean dataset, raising overfitting concerns similar to Luo et al. (2020) and Althibyani (2024).

LSTM performed strongly in predicting affect and user value, matching deep learning's recognized capacity to capture emotional subtleties (Yu et al., 2022; Alencar & Netto, 2020). Unlike previous multimodal approaches, this study used only ordinal data, showing well-encoded survey responses can be equally effective.

While Random Forest and XGBoost perform well in other studies (Tan et al., 2022; Rawat et al., 2021), they showed slightly lower precision and F1-scores here, possibly due to sensitivity to dataset size and diversity—reinforcing that ensemble methods may underperform in data-constrained settings.

Testing three data splits (70:30, 80:20, 90:10) revealed model stability trends. As in Nguyen et al. (2025), higher training ratios inflated scores with less reliable testing. The 70:30 split was most balanced, supporting realistic validation in small-scale studies.

6. References

Al Noman Bhuyan, Abdullah, et al. “A Machine Learning Approach for Predicting Student Engagement with Online Learning Platform.” 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), 13 Feb. 2025, pp. 1–6, https://doi.org/10.1109/ecce64574.2025.11013151.

Alencar, Márcio, and José Netto. “Measuring Student Emotions in an Online Learning Environment.” Proceedings of the 12th International Conference on Agents and Artificial Intelligence,2020,pp.563–569, https://doi.org/10.5220/0008956505630569.

Alhothali, Areej, et al. “Predicting Student Outcomes in Online Courses Using Machine Learning Techniques: A Review.” Sustainability, vol. 14, no. 10, 19 May 2022, p. 6199, https://doi.org/10.3390/su14106199.

Alodwan, T. (2021). Educational Research and Reviews Online learning during the COVID-19 pandemic from the perspectives of English as foreign language students. 16(7), 279–288. https://doi.org/10.5897/ERR2021.4169

Alsayat, Ahmed, and Hossein Ahmadi. “A Hybrid Method Using Ensembles of Neural Network and Text Mining for Learner Satisfaction Analysis from Big Datasets in Online Learning Platform.” Neural Processing Letters, 18 Aug. 2022, https://doi.org/10.1007/s11063- 022-11009-y.

Althibyani, Hosam A. “Predicting Student Success in MOOCs: A Comprehensive Analysis Using Machine Learning Models.” PeerJ Computer Science, vol. 10, 23 Aug. 2024, pp. e2221– e2221, https://doi.org/10.7717/peerj-cs.2221.

Ayouni, Sarra, et al. “A New ML-Based Approach to Enhance Student Engagement in Online Environment.” PLOS ONE, vol. 16, no. 11, 10 Nov. 2021, p. e0258788, https://doi.org/10.1371/journal.pone.0258788.

Bryson, J. R., & Andres, L. (2020). Covid-19 and rapid adoption and improvisation of online teaching: Curating resources for extensive versus intensive online learning experiences. Journal of Geography in Higher Education, 44(4), 1–16.

Demuyakor, J. (2020). Coronavirus (COVID-19) and online learning in higher institutions of education: A survey of the perceptions of ghanaian international students in china. Online Journal of Communication and Media Technologies, 10(3), e202018. https://doi.org/10.29333/ojcmt/8286

Dhawan, & Almaiah. (2020). (PDF) A Literature Review of E-Learning and E-Teaching in the Era of Covid-19 Pandemic. ResearchGate.

https://www.researchgate.net/publication/344927168_A_Literature_Review_of_E-Learning_and_E-Teaching_in_the_Era_of_Covid-19_Pandemic

Khoudi, Zakaria, et al. "Leveraging Machine Learning and Clickstream Data to Improve Student Performance Prediction in Virtual Learning Environments." Information Discovery and Delivery, 14 Mar. 2025, https://doi.org/10.1108/idd-08-2024-0120.

Lee, Ji-Eun, et al. "A Comparison of Machine Learning Algorithms for Predicting Student Performance in an Online Mathematics Game." Interactive Learning Environments, 16 May 2023, pp. 1–15, files.eric.ed.gov/fulltext/ED629740.pdf, https://doi.org/10.1080/10494820.2023.2212726.

Lee, Jungwon, et al. "A Comparison and Interpretation of Machine Learning Algorithm for the Prediction of Online Purchase Conversion." Journal of Theoretical and Applied Electronic Commerce Research, vol. 16, no. 5, 5 May 2021, pp. 1472–1491, https://doi.org/10.3390/jtaer16050083.

Luo, Yangyang, et al. Students' Online Behavior Patterns Impact on Final Grades Prediction in Blended Courses. 1 Dec. 2020, https://doi.org/10.1109/eitt50754.2020.00034.

Miranda, E., Aryuni, M., Rahmawati, M. I., Hiererra, S. E., & Dian Sano, A. V. (2024). Machine Learning's model-agnostic interpretability on the prediction of students' academic performance in video-conference-assisted online learning during the COVID-19 pandemic. Computers and Education: Artificial Intelligence, 7, 100312. https://doi.org/10.1016/j.caeai.2024.100312

Nguyen, Hong Thi, et al. "Early Prediction of Student Performance Based on Behavioral Data in

Blended Learning." The International Journal of Information and Learning Technology, vol. 42, no. 3, 7 May 2025, pp. 311–328, https://doi.org/10.1108/ijilt-04-2024-0069.

Nguyen, Viet Anh. "Applying Learning Analytics to Predict the Student's Learning

Outcome Based on Online Learning Activities.” Proceedings of the 2024 10th International Conference on Frontiers of Educational Technologies, 14 June 2024, pp. 140–146, https://doi.org/10.1145/3678392.3678401.

Peng, Wang. “Research on Model of Student Engagement in Online Learning.” EURASIA Journal of Mathematics, Science and Technology Education, vol. 13, no. 7, 18 June 2017, https://doi.org/10.12973/eurasia.2017.00723a.

Rawat, Seema, et al. “A Systematic Analysis Using Classification Machine Learning Algorithms to Understand Why Learners Drop out of MOOCs.” Neural Computing and Applications, 31 May 2021, https://doi.org/10.1007/s00521-021-06122-3.

Safat, W., Asghar, S., & Gillani, S. A. (2021). Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques. IEEE Access, 9, 1–1. https://doi.org/10.1109/access.2021.3078117

Salem, Maha, and Khaled Shaalan. “Unlocking the Power of Machine Learning in E-Learning: A Comprehensive Review of Predictive Models for Student Performance and Engagement.” Education and Information Technologies, 5 Apr. 2025, https://doi.org/10.1007/s10639- 025-13526-4.

Singh, Laishram Kirtibas, and R. Renuga Devi. “Analysis of Student Sentiment Level Using Perceptual Neural Boltzmann Machine Learning Approach for E-Learning Applications.” IEEE Xplore, 1 July 2022, ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9850860.

Tan, Choo Jun, et al. “An Intelligent Tool for Early Drop-out Prediction of Distance Learning Students.” Soft Computing, 19 Jan. 2022, https://doi.org/10.1007/s00500-021-06604-5.

Tong, Yao, and Zehui Zhan. “An Evaluation Model Based on Procedural Behaviors for Predicting MOOC Learning Performance: Students’ Online Learning Behavior Analytics and Algorithms Construction.” Interactive Technology and Smart Education, 6 Feb. 2023, https://doi.org/10.1108/itse-10-2022-0133.

UNESCO. (2021). UNESCO. Www.unesco.org. https://www.unesco.org/en

Yu, Weiwei, et al. “Analysis of Student E-Learning Engagement Using Learning Affect: Hybrid of Facial Emotions and Domain Model.” 2022 IEEE 25th International Conference on Computational Science and Engineering (CSE), Dec. 2022, pp. 98–105, https://doi.org/10.1109/cse57773.2022.00023.