Article

AI-Enhanced Decision-Making for Course Modality Preferences in Higher Engineering Education during the Post-COVID-19 Era

Amirreza Mehrabi [1,2,*], Jason Wade Morphew [1,*], Babak Nadjar Araabi [3], Negar Memarian [4] and Hossein Memarian [2]

1 Engineering Education Department, Purdue University, West Lafayette, IN 47907, USA
2 School of Engineering Science, University of Tehran, Tehran 1417466191, Iran; memarian@ut.ac.ir
3 School of Electrical and Computer Engineering, University of Tehran, Tehran 1417466191, Iran; araabi@ut.ac.ir
4 Department of Neurology, University of California, Los Angeles, CA 90095, USA; nmemarian@mednet.ucla.edu
* Correspondence: amehrabi@purdue.edu (A.M.); jmorphew@purdue.edu (J.W.M.); Tel.: +1-530-407-0036 (A.M.)

Abstract: The onset of the COVID-19 pandemic has compelled a swift transformation in higher-education methodologies, particularly in the domain of course modality. This study highlights the potential for artificial intelligence and machine learning to improve decision-making in advanced engineering education. We focus on the potential for large existing datasets to align institutional decisions with student and faculty preferences in the face of rapid changes in instructional approaches prompted by the COVID-19 pandemic. To ascertain the preferences of students and instructors regarding class modalities across various courses, we utilized the Cognitive Process-Embedded Systems and e-learning conceptual framework. This framework effectively delineates the task execution process within the scope of technology-enhanced learning environments for both students and instructors. This study was conducted in seven Iranian universities and their STEM departments, examining their preferences for different learning styles. After analyzing the variables by different feature selection methods, we used three ML methods—decision trees, support vector machines, and random forest—for comparative analysis. The results demonstrated the high performance of the RF model in predicting curriculum style preferences, making it a powerful decision-making tool in the evolving post-COVID-19 educational landscape. This study not only demonstrates the effectiveness of ML in predicting educational preferences but also contributes to understanding the role of self-regulated learning in educational policy and decision-making in higher education.

Keywords: online learning; technology-enhanced learning; educational machine learning; educational artificial intelligence; feature selection methods; random forest; SVM; decision tree

1. Introduction

The COVID-19 outbreak has revealed the complex decision-making processes universities must undertake when prompt and judicious actions are required, and decisions must be made under time pressure [1,2]. At the same time, the transition to remote instruction during the pandemic has revealed opportunities to expand the use of online learning, while also highlighting the drawbacks when the course modality does not align with instructor and student preferences. This decision-making process involves trade-offs between safeguarding students and enhancing access to education while maintaining high-quality, transformative learning experiences [3,4]. The swift transition from in-person to online classes during the COVID-19 pandemic has highlighted the importance of decision-making that considers the preferences of both instructors and students, given that the ability to incorporate students into the decision-making enhances the likelihood of successful school reform and student learning outcomes [5].

One benefit of the increase in online learning is the opportunity to leverage artificial intelligence (AI)'s ability to process large amounts of data to make timely decisions without the need to delay decision-making for human-controlled data analysis [6]. For instance, the decision to move instruction online or to being in person must be made quickly and should consider instructors' and students' preferences [7]; however, the time needed to survey both instructors and students may delay decision-making by several weeks. The affordances provided by AI can potentially enable administrators to make well-informed decisions that align with the needs of instructors and students while also optimizing the learning experience [8]. Through the utilization of AI, academic institutions can tailor course design and instructional modalities by using students' and faculty preferences in real-time, significantly enhancing educational equity within the university ecosystem [9,10].

Incorporating technological advancements within the educational milieu necessitates the establishment of robust and adaptive quality assurance protocols that are tailored to the needs of students and instructors within dynamically changing contexts [11–13]. Achieving equilibrium between individual preferences and optimal pedagogical outcomes demands the conceptualization and implementation of a theory-driven framework for decision-making and fortifying the quality and efficacy of technological integrations into the course fabric [11,14]. At the same time, the self-regulated learning (SRL) process, where students manage their learning through cognitive and motivational strategies, plays a key role in shaping course modality factors, further informing AI-driven decision-making models [15].

This paper aims to emphasize the potential of an AI tool that utilizes machine learning (ML) techniques to address challenges associated with decision-making involving large datasets. We ground this discussion by describing a study that developed an AI tool that predicts preferences for course modality (e.g., online or in-person) within seven Iranian Universities; however, this study aims to make a larger contribution to the potential for AI and ML in data-driven decision-making in higher education. The findings of this study can inform decision-making processes in higher-education institutions and contribute to understanding the factors influencing students' preferences for online learning. By identifying instructor and student preferences from existing large datasets available to institutions, decision-makers can maximize the benefits of technology-enhanced learning by making decisions about instructor and student preferences using a theory-driven data collection approach. This paper describes such an ML approach that addresses the following research questions:

RQ1: To what extent can machine learning (ML) predict student and instructor preferences for course modality in a post-Corona learning context?

RQ2: To what extent can psychological constructs associated with self-regulated learning (SRL) predict student and instructors' preferences for course modality?

2. Literature Review

2.1. Machine Learning in Education

Within education, ML methodologies play a pivotal role in the comprehensive examination of student performance [16], facilitation of learning processes, provision of nuanced feedback, and personalized recommendations [17,18], as well as educational administration and decision-making [19]. A wide variety of approaches have been proposed for how to employ ML in educational settings.

2.1.1. Machine Learning Methods

Within educational studies, most ML techniques applied for clustering and prediction purposes contain Artificial Neural Networks (ANN), decision trees (DTs), random forests (RFs), and support vector machines (SVMs) [20]. Most articles compare the accuracy of some of these methods to control the accuracy and interpretability of the model [18,21–23]. The appropriate model for improving prediction accuracy with categorical items depends on factors such as sample size, the number of variables, and theoretical context in handling

the data regarding the variables [24–27]. Studies have shown that SVM and RFs are accurate for clustering purposes with multidimensional and large data, such as predicting performance through activities and quarterly tests [18,28]. Conversely, methods such as regression tree, RF, and C4.5 (i.e., an extension of the Iterative Dichotomiser 3 algorithm) are highly effective for clustering datasets containing errors or random variations that obscure underlying patterns. These methods are particularly advantageous for datasets that are limited in size but require detailed, context-specific analysis [29]. For instance, ref. [30] used DTs to predict student performance to help teachers provide adapted instructional approaches and supplementary practice for students likely to have difficulty understanding the content.

2.1.2. Variable Selection in Machine Learning

In machine learning (ML), the learning model requires the inclusion of relevant variables and excluding the unhelpful variables to identify the most relevant variables that contribute to accurate predictions [31,32]. In this context, wrapper methods, such as Recursive Feature Elimination (RFE) and Boruta, have proven to be effective by considering the intricate interactions between variables and predictive models [33], while filter methods, such as chi-square and mutual information, focus on evaluating the individual relevance of predictors, shedding light on their independent contributions, as pointed out by [34]. This capability makes these methods ideal for capturing and simplifying complex relationships and interactions [35]. Employing a combination of filter and wrapper methods in variable selection is essential for leveraging diverse ML approaches [36], often using a strategic approach like majority voting identified by multiple selection methods [37].

2.2. Related Works

Faculty readiness for online teaching, particularly the challenges faced by those unfamiliar with the format, was extensively investigated by Singh et al. [12]. They notably focused on integrating laboratory components, distinct from other types of classes, within e-learning systems frameworks during the COVID-19 pandemic. In subsequent work, Singh et al. [38] investigated the impact of various course modalities on different aspects of learning, highlighting the importance of educational environment decisions. Carmona et al. [39] applied K-dependence Bayesian classifiers in machine learning to facilitate resource selection tailored to an individual student's prior knowledge, showcasing the integration of data-driven decision-making into teaching approaches and teacher preferences. Conversely, Kotsiantis et al. [40] pursued a methodological approach involving decision trees and neural networks. Their primary goal was to predict students' educational material preferences through designed surveys by considering student preferences. Hew et al. [41] used gradient-boosting tree methods with great precision to bridge the gap between expected and actual data about predicting students' success in Massive Open Online Courses (MOOCs), aiming for a cohesive approach to improve the reliability and accuracy of their research findings. Remarkably, to our knowledge, no existing literature has yet harnessed the potential of ML techniques to predict student and instructor preferences within the context of course modality using theory-driven approaches. Hebbecker et al. [42] delved into the theoretical foundations of classroom-level data-based decision-making while also examining the impact of teacher support on this process. Through latent mediation analyses on longitudinal data from teachers and students, the study explores the connections between instructional decision-making based on pedagogical approaches and students' reading progress.

2.3. Theory-Driven ML in Educational Research

Two main threats that reduce the repeatability and reliability of ML are the overfitting and underfitting of models due to limited sample sizes for training and largely defined variables [43]. Theory-driven ML approaches have been proposed to overcome the problem of inappropriate model fitting [9]. Theory-driven ML leverages existing theoretical frame-

works (e.g., self-regulated learning) to scaffold models by identifying relevant variables and collecting relevant data [9,44]. Theory-driven ML models use theories of learning to define variables (or features) to include in predictive models. Variable selection methods assist theory-driven ML in determining whether to reduce the problem's variables based on their importance for model predictability and their interdependence with other variables [35,44,45].

3. Theoretical Framework

Within a course, the conditions related to the student are overshadowed by the course-related conditions set by the instructor. E-learning systems have enhanced educational accessibility and introduced diverse avenues through which learners can interact with educational materials, thereby contributing substantively to the advancement of educational equity among a broad spectrum of students [46,47]. So, it is necessary to increase the scope of investigation in Technology-Enhanced Learning (TEL) as a requirement. TEL provides a framework for how an e-learning system meets the requirements of educational stakeholders [48]. People engage with e-learning systems, while e-learning technologies facilitate both direct and indirect interactions among various user groups that affect learning. E-learning systems encompass all activities aligned with pedagogical models and instructional strategies. Course conditions can be seen as part of the e-learning services component, which encompasses all activities aligned with pedagogical models and instructional strategies and controls the task and cognitive condition [46,49]. The external evaluation part of Cognitive Process-Embedded Systems (COPES) can be seen as part of the people component, as it involves learners engaging with the e-learning system, and the quality of the e-learning system affects their evaluation. By adding the conditions of the e-learning systems framework like pedagogy, instructional strategies, and the quality of Information Communication Technology (ICT), we can investigate most factors related to the course modality [46].

3.1. Self-Regulated Learning (SRL)

The concept of self-regulated learning involves perceiving learning as a dynamic process that entails the adaptive adjustment of cognition, behavior, and motivation to the content and the educational environment [15]. Refs. [50,51] proposed an SRL theory as a guiding framework in different learning environments for preferences of course modality, considering different conditions like students and instructors. One widely used theoretical model for investigating learning within self-regulated contexts is the Cognitive Process-Embedded Systems (COPES) model [46]. The COPES model investigates the process of learning during a task, considering the conditions that shape the satisfaction of tasks according to the learning goal and standards [46,52]. Conditions encompass both internal factors, such as the learner's characteristics and knowledge about the topic, and external factors, including environmental variables that are believed to impact the task-related internal conditions [46]. Conditions like the skill and knowledge of students, motivation models, and task content are inputs in information processing and decision-making, and the outcomes will be used to evaluate task success based on predefined standards and individual goals [15,46]. Given the success of the COPES model in explaining student behavior in self-regulated learning contexts, the conditions from the COPES model serve as useful variables to aid prediction in an ML model [15,46].

3.2. Course Modality as a Part of Technology-Enhanced Learning

Technology-enhanced learning (TEL) is the incorporation of technology in learning environments to promote the process of teaching and learning [53]. Any tool that aids in improving decision-making and learning experiences or adds value to educational environments by aligning the environment, tools, and content together can be classified as TEL [54]. Therefore, the preference for different course modalities is related to individual beliefs and experience with, and preferences for, the integration of technology within the

classroom [54,55]. As such, AI-based preference prediction tools should be investigated as a part of improving TEL [54,55].

3.3. E-Learning Systems

Within a course, conditions related to the student are often overshadowed by those set by the instructor. E-learning systems have enhanced educational accessibility, providing diverse avenues for interaction with materials, which advances educational equity among a broad spectrum of students [56]. These systems meet the needs of various stakeholders through a framework that integrates pedagogical models, instructional strategies, and Information Communication Technology (ICT) [48]. In the context of COPES, the course conditions align with the e-learning services component, which controls both task and cognitive conditions, while the external evaluation aspect aligns with the people component, as it involves learner engagement and system evaluation [48,52]. By incorporating the conditions of e-learning systems, such as pedagogy and ICT quality, into the COPES framework, we can investigate most factors related to course modality [48] (Figure 1).
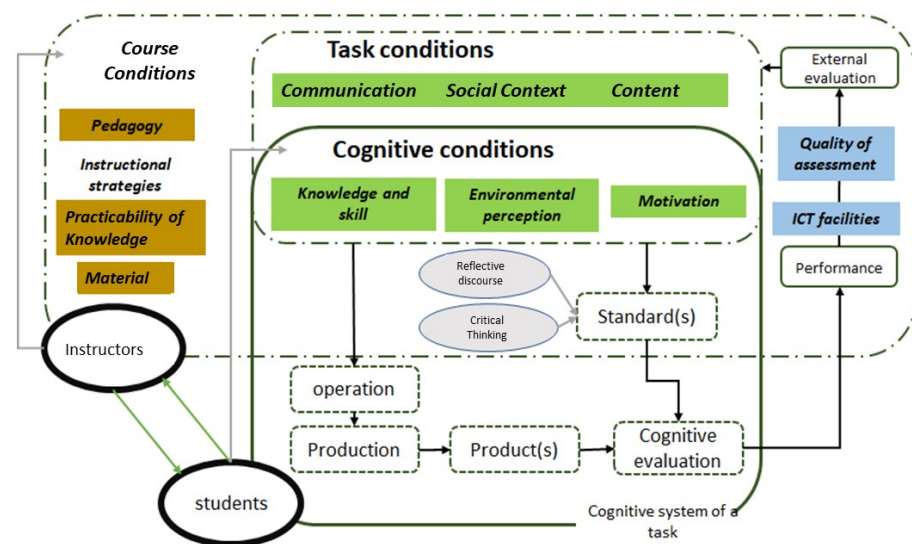

Figure 1. Framework from the combination of COPES and e-learning systems for course modality. Adapted from [46,48].

4. Research Design

4.1. Participants

The research was carried out among 140 instructors and 379 students from the engineering departments of seven Iranian universities, namely the University of Tehran, Sharif University of Technology, Isfahan University of Technology, Shiraz University, Sistan and Balouchestan University, Imam International University, and Ahvaz Shahid Chamran University. Participants were recruited via email from these seven universities. There were no restrictions based on age or other demographic factors, and participants did not receive any incentives or rewards for their involvement. For this study, two parallel surveys were developed and used to collect data from instructors and students. The surveys consisted of 50 and 49 questions, respectively, about the six dimensions of our theoretical framework of self-regulated learning and e-learning systems theory. Items were written to capture a variety of beliefs across six subscales, including theory and practice, motivation, pedagogy, knowledge, insight, and skills, working life orientation, quality of assessment, and information communication technology (Supplementary S17, Table S21). In addition, participants were asked to indicate the type of course in which they were enrolled or taught (i.e., theoretical–practical, theoretical, and practical) and assess their preferred situation in each type of course using Likert scales (Supplementary Table S1). The lack of an established

standardized survey called for the generation of pertinent survey items corresponding to distinct components of the framework. A pilot study of the surveys was conducted with ten students and five instructors to identify potential issues with item clarity.

Prior to analysis, data cleaning procedures were implemented to ensure the suitability and accuracy of the collected data [22,47,57,58]. This procedure consists of identifying and addressing any inconsistencies, errors, outliers, or missing values in the dataset. For the student survey, the items of F2 and F3 for fewer than 50 students were empty, and we put the value of the response variable instead of them. Only 160 students responded to the I4 response variable, as they did not have experience with practical types of classes (R code and libraries are in Supplementary S17 and Table S21).

4.2. Validity and Reliability Analysis

To examine the reliability of the survey items, Cronbach's alpha coefficient was used to assess the internal consistency of the items [59], while discriminant validity was assessed using the Heterotrait–Monotrait Ratio of Correlations (HTMT) method [60], which examines the correlations of indicators across constructs and within the same construct [61]. The HTMT correlation ratio is an approach to examine the extent to which latent constructs are independent of each other [62,63]. Acceptable values of composite reliability/Cronbach alpha range from 0.60 to 0.70, while the acceptable range of HTMT is less than 0.9 [62,63] (methodology in Supplementary S8; results in Tables S8 and S9).

Cronbach's alpha for instructors' surveys has a reliable value with a 0.95 confidence interval ($\alpha = 0.767$). However, for students' surveys, there are a couple of subscales with few items, which causes a reduction in Cronbach's alpha. For the students' survey, the 0.95 confidence interval has also a reliable value ($\alpha = 0.598$). This is consistent with the findings of Rempel et al. [62,64], who found that Cronbach's alpha can be reduced by subscales with few items. The result of Cronbach's alpha for each subscale for both surveys is presented in Table S22 of Supplementary S18. Tables S8 and S9 show the HTMT values for all subscales of both surveys. The highest HTMT value observed for instructors is 0.6 and for students, 0.532, which indicates that there is sufficient distinction between any two subscales in both surveys [62,63].

5. Methodology

5.1. Machine Learning Process

The first step in the machine learning (ML) process is variable selection, followed by modeling the problem using ML models as illustrated in Figure 2. In this study, both surveys feature a relatively large number of questions (variables) and a relatively limited number of responses. Given our interest in the most relevant variables and the varied accuracy of different variable selection methods, considering the contribution to accurate predictions and imbalanced data (where one class in some variables has significantly fewer instances compared to another class), we employed four interpretable variable selection methods with a significance level of $\alpha = 0.05$. To enhance the precision of variable selection and subsequent ML model implementation, a refined approach was adopted as outlined in [13,65,66]. This approach involved retaining questions that exhibited similarity or analogous distinctive response classes to those found in Likert scale questions, excluding binary (Yes/No) inquiries, and questions with unrestricted response options [67]. Additionally, questions related to a specific learning management system, socioeconomic indicators, or personal identifying details such as the name of the university and age, were excluded. Consequently, 38 questions were retained from the student survey, while 34 questions were preserved from the instructors' survey for variable selection [66].
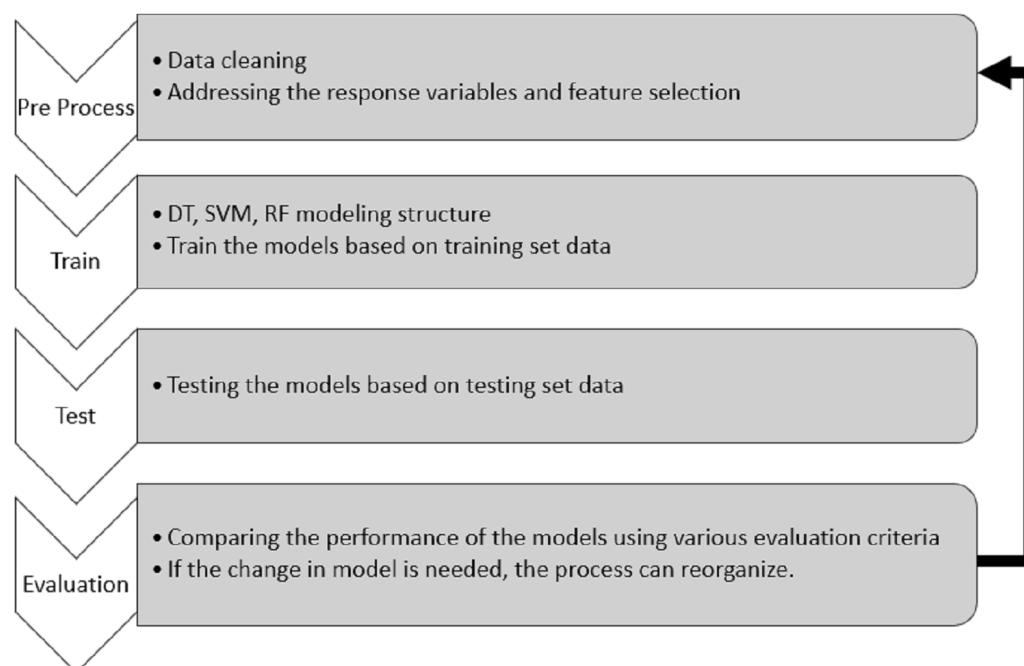
Figure 2. Process of ML.

The chi-square test, a filter method, computes the chi-square statistic for each variable about the target item and compares it to a critical value or *p*-value determined by the chosen significance level, indicating a significant association between categorical variables [67–69]. Variables with chi-square statistics exceeding the critical value or *p*-values below the significance level are selected [68,70]. The Maximum Relevance Minimum Redundancy (MRMR) method, another filtering method, aims to identify a subset of variables that have high relevance to the target variable while minimizing redundancy among the selected variables [71]. Wrapper methods employ a specific ML model to assess subsets of variables and identify the optimal subset that maximizes the model's performance [72–75]. Boruta, a wrapper method, employs a random forest model to evaluate variable importance, adding randomly generated variables and comparing their significance to original variables [76]. Recursive Feature Elimination (RFE), an embedded method, involves training a model on the full variable set, ranking variables based on their importance, eliminating the least important variables, and repeating this process [77,78].

Six response variables (Table S12) representing theoretical, practical, and theoretical–practical class types were analyzed using supervised ML models. The majority voting method was utilized to integrate outcomes and find the best sets of variables for each response variable [77,78]. This integration not only enhances the precision and reliability of the predictive model but also addresses challenges such as overfitting, noise, and the curse of dimensionality [77,78]. The REF method incorporates the population of the primary sample of variables entering the ML model. Moreover, the chi-square method lacks a definitive boundary for variable exclusion [36,65]. A threshold for the chi-square test was imposed, and the top variables according to REF criteria were selected from the chi-square test. The variable selection rule for majority voting to enter the ML stage is as follows: If an item has been eliminated by two or more variable selection methods, the item will be removed (Table S6). Tables S2–S5 in the Supplementary indicate the variables selected by each of the chi-square, MRMR, Boruta, and RFE methods, describing the situation of each variable concerning the response variables. For REF and Boruta, 70% of the data were used for training ML models and 30% for testing, with the DT (CRAT) model used for Boruta and RF for the REF with 500 trees [22,36,45,57,65]. Table S7 indicates the variables that remained for ML.

5.2. Machine Learning Models

After selecting the variables, the next step in the ML process is to model the problem using ML models. In this study, according to [9], three ML models were utilized, including DT, SVM, and RF)

The DT (i.e., the CRAN model) method constructs a tree-like model of decisions and their possible consequences. Each variable in each tree predicts the response variable, and the leaf of DT represents the prediction. The construction of the tree involves recursively partitioning the data based on the most informative variable, aiming to maximize the separation of the target variable. This process continues until a stopping criterion is met, such as reaching a maximum depth in variable searching or a minimum number of samples of data for each variable per leaf [79] (methodology in Supplementary S13; results in Tables S15 and S16).

The RF model constructs an ensemble of decision trees, where each tree is trained on a different subset of the data using a random selection of variables. During prediction, the model aggregates the predictions of individual trees to make the final prediction, resulting in a robust and reliable model. Additionally, RF identifies the key factors influencing the outcome. The cross-validation procedure for RF involved partitioning the dataset into 70% for training and 30% for testing, adhering to a significance level of $\alpha = 0.05$ [80] (methodology in Supplementary S12; results in Tables S13 and S14).

As the response variable has multiple dimensions, the multiclass SVM was used to find an optimal hyperplane that separates different classes. Train data lie closest to the hyperplane that separates the classes in a binary classification problem. This hyperplane is determined by identifying a subset of data according to each variable called support vectors. The SVM model identifies these support vectors during the training phase and uses them to compute the optimal hyperplane. After we selected the variables of the ML model, by separating the test and train starts, we needed to learn them.In this study, 70% of the data were used for training the ML models, and 30% of the data were used to test the models [81] (methodology in Supplementary S14; results in Tables S17 and S18).

5.3. Accuracy of ML Models

Accuracy in machine learning refers to the measure of how often a classification model correctly predicts the true class of an instance, typically expressed as a ratio of correct predictions to the total number of predictions:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

Equation (1) shows the formulation of accuracy in ML.

Precision is the ratio of correctly predicted positive observations to the total predicted positives. It indicates how precise the model's positive predictions are, providing a measure of the relevancy of its positive results:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

Equation (2) shows the formulation of precision in ML.

Recall (also known as sensitivity or the True Positive Rate) is the ratio of correctly predicted positive observations to all the actual positives. It indicates how well the model can capture all the positive instances:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

Equation (3) shows the formulation of recall in ML.

The F1-score is the harmonic mean of precision and recall. It provides a single metric that balances both concerns of false positives and false negatives, making it particularly useful when the class distribution is imbalanced:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

Equation (4) shows the formulation of the F1-score in ML.

The explication of True Positives (TP) as the quantification of instances accurately designated as positive, True Negatives (TN) as the quantification of instances accurately designated as negative, False Positives (FP) as the quantification of instances inaccurately designated as positive, and False Negatives (FN) as the quantification of instances inaccurately designated as negative, contained within the framework, provides a lucid exposition of the pivotal constituents that underlie the accuracy metric [22,57].

6. Results

6.1. Prediction Using Classification Techniques

The outcomes derived from the ML prediction methodologies concerning the application of five-point Likert scale responses to anticipate the six designated response categories reveal an elevated level of precision attributed to random forest (RF) models (Table 1: the results summary of Tables S13, S15 and S17; Table S22 indicates the F1-score, recall, and precision), which is consistent with prior scholarly investigations [82]. According to Table 1, SVM and DT have less accuracy in comparison to RF. However, in all of them, the accuracy is not at an acceptable level.

Table 1. The accuracy results of ML models based on test results for a five-point Likert spectrum for students and instructors (the results summary of Tables S14, S16 and S18). Table S22 indicates the F1-score, recall, and precision.

Models	Students			Instructors		
	I2 Theoretical	I4 Practical	I6 Theoretical–Practical	H2 Theoretical	H4 Practical	H6 Theoretical–Practical
SVM	0.45	0.43	0.56	0.48	0.53	0.42
DT	0.39	0.40	0.45	0.41	0.53	0.35
RF	0.45	0.43	0.58	0.49	0.62	0.37

The F1-scores for the "Strongly Agree" (SA) and "Strongly Disagree" (SD) classes exhibit significant variability across models and items, indicating inconsistent classification performance. For the SD class, certain models, such as RF and DT, achieve high recall (perfect in some cases) but low precision, resulting in moderate F1-scores, reflecting the models' tendency to over-predict this class. In contrast, the SA class often has very low or zero F1-scores, particularly for items like I4, H4, and H6, suggesting a severe difficulty in accurately identifying instances of "Strongly Agree" (Table S22).

We address class imbalance using the Random Over-Sampling Examples (ROSE) technique, which generates synthetic samples from minority classes to balance the dataset with 70% train and 30% test. ROSE applies a smoothed bootstrap resampling process, creating new data points that reflect the distributional characteristics of minority instances by drawing from a kernel density estimate, rather than simply duplicating existing samples. This method is particularly effective for datasets with ordinal data, as it enhances the diversity of training data and improves the generalization capabilities of machine learning models [83–85]. ROSE maintains the structural integrity of the data while improving the model's ability to differentiate between categories, especially in tasks involving ranking or classification [86,87]. The results demonstrate that while RF continues to perform well, showing improved accuracy in items such as I4 (practical) for both students and instructors, SVM and DT still face challenges. Instructors' practical responses (H4) particularly illustrate how ROSE reshapes the dataset, allowing SVM and RF to achieve slightly higher accuracy

compared to the non-ROSE sampling. However, the data size remains a limiting factor, as ROSE-driven improvements in accuracy, while notable, do not lead to a significant leap in predictive power across all items. This suggests that while ROSE helps mitigate class imbalance, further adjustments, such as merging Likert categories, might be needed to enhance model robustness.

The F1-score, recall, and precision metrics (Tables 2 and S24) further illustrate the impact of ROSE sampling on model performance across various categories. For students' theoretical and practical categories (I2 and I4), RF sees an improvement in precision and recall, particularly in categories with previously low representation, such as strongly disagree or agree, where synthetic samples boost prediction diversity. However, support for middle-range categories, such as neutral, remains less robust, highlighting the continued challenges faced by models like DT and SVM. Instructors' responses, particularly for theoretical–practical categories (H6), demonstrate a noticeable increase in recall for SVM after applying ROSE but at the cost of reduced precision, signaling a tendency for over-prediction in some categories. This trade-off is seen across models, where the addition of synthetic samples helps increase recall but introduces more false positives, reducing precision and yielding modest improvements in F1-scores. This suggests that while ROSE helps balance the dataset, its effectiveness in improving predictive performance is nuanced and dependent on the class distribution.

Table 2. The accuracy results of ML models based on test results for a five-point Likert spectrum for students and instructors via the ROSE method of sampling (F1-score, recall, and precision are in Table S24).

Models	Students			Instructors		
	I2 Theoretical	**I4 Practical**	**I6 Theoretical–Practical**	**H2 Theoretical**	**H4 Practical**	**H6 Theoretical–Practical**
SVM	0.44	0.66	0.47	0.64	0.46	0.54
DT	0.38	0.62	0.41	0.61	0.21	0.32
RF	0.47	0.74	0.58	0.64	0.36	0.43

Since the ROSE imbalance model did not sufficiently improve accuracy, a merging approach was employed to address the response imbalance between the "strongly agree" and "strongly disagree" categories. This approach consolidated these classes, reducing the number of categories in the Likert scale to enhance model performance. This approach reduces the students' and instructors' bias in responses, merges "strongly agree" and "agree" responses, and "strongly disagree" and "disagree" classes [88]. After merging the response classes, the implementation of ML changed to that of Table 3. And after merging these response classes, the performance of the machine learning models improved as shown in Table 3.

Table 3. Accuracy averages of models for students and instructors. Table S23 indicates the F1-score, recall, and precision.

Models	Students			Instructors		
	I2 Theoretical	**I4 Practical**	**I6 Theoretical–Practical**	**H2 Theoretical**	**H4 Practical**	**H6 Theoretical–Practical**
SVM	0.70	0.75	0.80	0.50	0.65	0.55
DT	0.71	0.72	0.71	0.53	0.65	0.48
RF	0.78	0.81	0.94	0.69	0.72	0.79

Results in Table 3 demonstrate a significant improvement in prediction accuracy (Table S11; Supplementary S10. Table S23 indicates the F1-score, recall, and precision), particularly for the RF model, which remained more accurate than the DT and SVM methods. The change from a 5-point Likert scale to a 3-point Likert scale was particularly beneficial for the analysis of instructors' survey responses due to an increase in the accuracy

of prediction for three response variables of the instructor. Unlike the RF model for the instructors, results indicated that DT and SVM could not make accurate predictions within the 3-point Likert scale for the instructors (specifically, the H4 and H6 response variables), which could be the result of the small sample, though this is not true for the students' variables.

Despite the observed accuracy levels within the five-point Likert scale, discernible discrepancies surfaced in the "strongly agree", "strongly disagree", and "Neutral" classifications (Tables 1 and 2). Regardless of the employed Likert scale, it is evident (Supplementary S17) that the "Neutral" class consistently manifested the highest error rates. This phenomenon can be attributed to the imbalance in the population of the "Neutral" class according to the population of this class compared to other classes. According to Tables S13–S18, the random forest method demonstrated superior accuracy levels concerning the "Neutral" category in comparison to the remaining classification methods, regardless of the employed Likert scale. While the DT model has acceptable accuracy in ranges of agree and disagree for 3-point Likert, their effectiveness decreases within the "Neutral" class of 3-point Likert in comparison with the 5-point Likert (Supplementary S17). Within the class reduction process from five-point Likert to three-point Likert, the "Neutral" class remained constant in quantity, unlike the other two categories that experienced augmentation which increased the imbalance of this class. This merging of imbalanced classes subsequently heightened the likelihood of errors within the Neutral class. Meanwhile, the SVM model demonstrated greater accuracy upon the transition from a five-point to a three-point Likert scale. This enhancement in accuracy can be the output of merging two imbalanced classes. Despite the smaller number of data records from instructors, their predictions demonstrated higher accuracy and a unanimous opinion in comparison to the students. This can be attributed to the defined theoretical framework and strong variable selection, which is aligned with instructors' preferences and lets their responses be more predictable.

The models exhibit strong performance in classifying the "Agree" (A) class, as evidenced by consistently high F1-scores of 0.809756, with both precision and recall exceeding 0.80. This suggests that the models accurately identify instances of agreement with minimal misclassification. For the "Disagree" (D) class, moderate F1-scores of 0.657143 and closely aligned precision and recall indicate a balanced but less robust performance. However, the models demonstrate lower performance in classifying the "Neutral" (N) class, reflected in significantly lower F1-scores (Table S23).

6.2. Subscales' Ranking and Framework

Given the results above, the determination of subscale rankings in predicting student and instructor preferences (RQ2) was undertaken by leveraging the advanced classification capabilities of the RF. The Mean Decrease in Impurity methodology, which assesses the significance of variables, was harnessed for this purpose (Supplementary S15; the most important variables of RF and DT are in Table S19). For both instructors and students, pedagogy and motivation are the most important subscales to increase the predictability of the response variables (course modality); however, the 'theory and practice' subscale indicated comparably diminished predictability. This diminished predictability might be related to the population of "Neutral" responses in related items in the survey as seen in the Supplementary Materials (S12–S14). The subscale of 'Knowledge, insight, and skill' also provides less information for predictability for the instructors when compared to the students. The subscale of 'working-life orientation' also is a prominent subscale for both students' and instructors' theoretical–practical response variables. Determinative subscales such as motivation, pedagogy, insight, and skills, working-life orientation, quality of assessment, and ICT assume pivotal roles in shaping preferences. It should be noted that the correlation of the theory and practice subscale with the responses is not high (Supplementary S12–S14).

7. Discussion

The results of the study indicate the superior predictive capabilities of the RF model in comparison to alternative models such as DT and SVM. Indeed, RF can handle high-dimensional response classes, imbalances, and multivariable situations. This study corroborates other studies where RF has been compared to other ML models, such as SVM and Logistic Regression [89,90]. A notable finding from this study is the predictive efficacy of RF models for "Neutral" responses, which is in an acceptable range of accuracy for both the 3-point and 5-point Likert scales. The SVM also provides accuracy in an acceptable range; however, DT has lower accuracy in comparison to SVM for most of the response variables. Consequently, the DT model tends towards divergent classifications with more clear distinct classes, thereby decreasing its suitability for high-level classification undertakings. This observation aligns with prior research [91,92]. However, the discriminatory boundaries between response classes like "agree" and "strongly agree" are not always an indicator of non-biased selection.

The COPES and e-learning systems theoretical frameworks were used to define the subscales and subsequently the variables. Utilizing four different variable selection methods helped to simplify the ML model and select the most relevant variables, which made the prediction more robust [91,92]. This means that some of the variables were fully aligned with the response variables and could help us to predict the course modality. Indeed, the defined subscales and variables from the theoretical framework with a high level of confidence can explain the course modality preferences for both instructors and students [91–94].

The subscales were defined according to three conditions in the theoretical framework, after which the ML ranked the subscales according to the responses (Table 4 and Figure 1). Motivation and pedagogy were categorized under course and cognition conditions. Alignment between pedagogy and motivation was the most prevalent amongst all other subscales. This result does not mean that other subscales like those related to task conditions are not important, but it means that for any decision-making about the course modality, the pedagogical approach of the instructor and how to motivate students should be prioritized and studied in any environment that is aligned with the result of studies like [91–96].

Table 4. Ranking of different dimensions based on the average reduction in impurity of each group of items in Supplementary S15 and Table S19.

	Groups					
	Students			Instructors		
	I2 Theo.	I4 Prac.	I6 T-P	H2 Theo.	H4 Prac.	H6 T-P
	Likert type (5-scale)					
1	Ped.	Ped.	Ped.	Ped.	Ped.	Work-life Ori.
2	Motiv.	Motiv.	Work-life Ori.	Motiv.	Motiv.	Ped.
3	Qual. of Assess. & ICT	Know., Ins. & Skill	Motiv.	Work-life Ori.	Work-life Ori.	Motiv.
4	Know., Ins. & Skill	Theory & Pract.	Qual. of Assess. & ICT	Qual. of Assess. & ICT	Qual. of Assess. & ICT	Theory & Pract.
5	Work-life Ori.	Work-life Ori.	Know., Ins. & Skill	Know., Ins. & Skill	Theory & Pract.	Qual. of Assess. & ICT
6	Theory & Pract.	Qual. of Assess. & ICT	Theory & Pract.	Theory & Pract.	Know., Ins. & Skill	Know., Ins. & Skill
	Likert type (3-scale)					
1	Ped.	Ped.	Ped.	Ped.	Ped.	Motiv.
2	Motiv.	Motiv.	Motiv.	Motiv.	Motiv.	Ped.
3	Qual. of Assess. & ICT	Know., Ins. & Skill	Qual. of Assess. & ICT	Know., Ins. & Skill	Work-life Ori.	Work-life Ori.
4	Know., Ins. & Skill	Theory & Pract.	Know., Ins. & Skill	Work-life Ori.	Qual. of Assess. & ICT	Theory & Pract.
5	Work-life Ori.	Work-life Ori.	Work-life Ori.	Theory & Pract.	Theory & Pract.	Qual. of Assess. & ICT
6	Theory & Pract.	Qual. of Assess. & ICT	Theory & Pract.	Qual. of Assess. & ICT	Know., Ins. & Skill	Know., Ins. & Skill

Note: Abbreviations represent the following—Ped. indicates Pedagogy (blue); Motiv. indicates Motivation (orange); Work-life Ori. indicates Work-life Orientation (green); Qual. of Assess. & ICT indicates Quality of Assessment and ICT (lighter gray); Theory & Pract. indicates Theory and Practice (light gray).

The findings from this study demonstrate the potential for ML to accurately classify student and instructor course modality preferences based on theoretical frameworks. So, the answer to the first research question is positive because the ML models with good accuracy can explain the multidimensional data and predict the preferences of students and instructors regarding the class modality.

The answer to the second research question is also affirmative. By considering the COPES and e-learning systems as the frameworks and studying the preferences about course modality in the context of course, task, and cognitive condition, we have a tool that predicts the preferences of students and instructors in higher education. Indeed, while these theories separately do not directly explain the course modality situation, the course modality is explainable by adding the course conditions that come from the e-learning framework to the COPES framework.

8. Conclusions

The necessity for adaptable instructional modalities is paramount to upholding educational standards and ensure accessibility in accurately predicting course modality preferences, which is critical for universities as they address the intricacies of education in the post-pandemic landscape. The alignment of pedagogical strategies with motivational factors has emerged as a significant predictor, emphasizing the importance of considering both instructional design and student engagement when selecting course modalities. This finding is instrumental for educational institutions aiming to optimize learning experiences, where the flexibility and adaptability of course delivery methods are crucial. Aligning educational strategies and course designs with the preferences revealed through ML analysis promises to create more stimulating and efficacious learning environments, thereby enhancing student engagement and achieving superior educational outcomes [91,92,94,95,97]. The results of this study are applicable to universities and educational institutions that have large amounts of high-dimensional data with lots of different subscales. As the type of data in access universities varies, in situations where survey data are not available, behavior data or historical preference records can be used to implement machine learning, specifically RF as the most robust model according to our result, to make decisions about the educational policies and activities like course modality based on students' and instructors' preferences [91,92,94,95]. The research further elucidates the significance of incorporating psychological constructs related to self-regulated learning within the decision-making paradigm. This underscores the merit of a data-informed approach in crafting educational experiences that are both personalized and attuned to the evolving dynamics of student and instructor needs. Such an approach is instrumental in fostering educational strategies that are responsive and tailored, reflecting the contemporary demands of the academic community. The investigation showcases the pivotal role of AI in transforming higher education by facilitating a harmonious balance between the quality of instruction and the multifaceted needs of the academic populace. This endeavor is crucial not only for augmenting the flexibility and robustness of educational frameworks but also for ensuring their continued relevance and efficacy in addressing the exigencies of forthcoming educational paradigms.

9. Future Research Directions

The development of predictive models for student and faculty preferences through this research marks a pivotal advancement in comprehending and accommodating the needs of the academic community, thereby providing a critical instrument for administrators striving to deliver educational experiences that are high in quality, equity, and effectiveness. Future research endeavors should aim to refine these predictive models, with a particular focus on elucidating the roles of motivation and pedagogy or exploring alternative models that more acutely incorporate these dimensions. Such efforts necessitate comprehensive data collection strategies within educational settings, encompassing a variety of sources to amass data conducive to the application of advanced AI Big Data methodologies. This approach

will facilitate educational institutions in making more informed, confident, and precise decisions regarding course modality selections, leveraging the rich tapestry of data at their disposal.

Author Contributions: Conceptualization, A.M., J.W.M., H.M. and N.M.; Methodology, A.M. and J.W.M.; Software, B.N.A. and A.M.; Validation, A.M., B.N.A. and J.W.M.; Formal Analysis, A.M.; Investigation, A.M., B.N.A. and J.W.M.; Resources, A.M.; Data Curation, B.N.A., A.M., H.M. and N.M.; Writing—Original Draft Preparation, A.M., J.W.M., B.N.A., H.M. and N.M.; Writing—Review and Editing, A.M., J.W.M., B.N.A., H.M. and N.M.; Visualization, A.M.; Supervision, J.W.M.; Project Administration, A.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The dataset used in this study is publicly accessible in the OSF repository at https://osf.io/v6tp3/?view_only=9af7602f772846ab866b54d6bbc2ed30, https://doi.org/10.17605/OSF.IO/V6TP3, accessed on 23 August 2024.

Conflicts of Interest: The authors declare no conflicts of interest. The funding sponsors had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Rauseo, S.; Rathnayake, D.; Marinciu, R. Decision-making under uncertainty: How university students navigate the academic implications of the COVID-19 pandemic challenges. In *Agile Learning Environments amid Disruption: Evaluating Academic Innovations in Higher Education during COVID-19*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 655–674.
2. Lekishvili, T.; Kikutadze, V. Decision-Making process transformation in post-COVID-19 world in higher educational Institutions. In *Digital Management in COVID-19 Pandemic and Post-Pandemic Times: Proceedings of the International Scientific-Practical Conference (ISPC 2021), Moscow, Russia, 2–4 November 2021*; Springer: Cham, Switzerland, 2023; pp. 169–177.
3. Krismanto, W.; Tahmidaten, L. Self-Regulated Learning in online-based teacher education and training programs. *Aksara J. Ilmu Pendidik. Nonform.* **2022**, *8*, 413. [CrossRef]
4. Skar, G.B.U.; Graham, S.; Huebner, A. Learning loss during the COVID-19 pandemic and the impact of emergency remote instruction on first grade students' writing: A natural experiment. *J. Educ. Psychol.* **2022**, *114*, 1553. [CrossRef]
5. Pekrul, S.; Levin, B. Building Student Voice for School Improvement. In *International Handbook of Student Experience in Elementary and Secondary School*; Thiessen, D., Cook-Sather, A., Eds.; Springer: Dordrecht, The Netherlands, 2007; Chapter 27, pp. 711–726. [CrossRef]
6. Winne, P.H. Cognition and Metacognition within Self-Regulated Learning. In *Handbook of Self-Regulation of Learning and Performance*, 2nd ed.; Shunk, D.S., Greene, J.A., Eds.; Taylor and Francis: Abingdon, UK, 2017; pp. 36–48. [CrossRef]
7. Lima, R.M.; Villas-Boas, V.; Soares, F.; Carneiro, O.S.; Ribeiro, P.; Mesquita, D. Mapping the implementation of active learning approaches in a school of engineering–the positive effect of teacher training. *Eur. J. Eng. Educ.* **2024**, 1–20. [CrossRef]
8. Alpaydin, E. *Machine Learning: The New AI*; The MIT Press: Cambridge, MA, USA, 2016.
9. Macarini, L.A.B.; Cechinel, C.; Machado, M.F.B.; Ramos, V.F.C.; Munoz, R. Predicting students success in blended learning—evaluating different interactions inside learning management systems. *Appl. Sci.* **2019**, *9*, 5523. [CrossRef]
10. Johri, A.; Katz, A.S.; Qadir, J.; Hingle, A. Generative artificial intelligence and engineering education. *J. Eng. Educ.* **2023**, *112*, 572–577. [CrossRef]
11. Hilbert, S.; Coors, S.; Kraus, E.; Bischl, B.; Lindl, A.; Frei, M.; Wild, J.; Krauss, S.; Goretzko, D.; Stachl, C. Machine learning for the educational sciences. *Rev. Educ.* **2021**, *9*, e3310. [CrossRef]
12. Singh, J.; Steele, K.; Singh, L. Combining the Best of Online and Face-to-Face Learning: Hybrid and Blended Learning Approach for COVID-19, Post Vaccine, & Post-Pandemic World. *J. Educ. Technol. Syst.* **2021**, *50*, 140–171. [CrossRef]
13. Singh, J.; Perera, V.; Magana, A.J.; Newell, B.; Wei-Kocsis, J.; Seah, Y.Y.; Strimel, G.J.; Xie, C. Using machine learning to predict engineering technology students' success with computer-aided design. *Comput. Appl. Eng. Educ.* **2022**, *30*, 852–862. [CrossRef]
14. Koretsky, M.D.; Nolen, S.B.; Galisky, J.; Auby, H.; Grundy, L.S. Progression from the mean: Cultivating instructors' unique trajectories of practice using educational technology. *J. Eng. Educ.* **2024**, *113*, 330–359. [CrossRef]
15. Zimmerman, B.J.; Campillo, M. Motivating Self-Regulated Problem Solvers. In *The Psychology of Problem Solving*; Cambridge University Press: Cambridge, MA, USA, 2003; pp. 233–262. [CrossRef]
16. Talib, N.I.M.; Majid, N.A.A.; Sahran, S. Identification of Student Behavioral Patterns in Higher Education Using K-Means Clustering and Support Vector Machine. *Appl. Sci.* **2023**, *13*, 3267. [CrossRef]
17. Martin, F.; Wang, C.; Sadaf, A. Student perception of helpfulness of facilitation strategies that enhance instructor presence, connectedness, engagement and learning in online courses. *Internet High. Educ.* **2018**, *37*, 52–65. [CrossRef]
18. Psyridou, M.; Koponen, T.; Tolvanen, A.; Aunola, K.; Lerkkanen, M.K.; Poikkeus, A.M.; Torppa, M. Early prediction of math difficulties with the use of a neural networks model. *J. Educ. Psychol.* **2023**, *116*, 212–232. [CrossRef]
19. Inusah, F.; Missah, Y.M.; Najim, U.; Twum, F. Data mining and visualisation of basic educational resources for quality education. *Int. J. Eng. Trends Technol.* **2022**, *70*, 296–307. [CrossRef]
20. Safaei, M.; Sundararajan, E.A.; Driss, M.; Boulila, W.; Shapi'i, A. A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. *Comput. Biol. Med.* **2021**, *136*, 104754. [CrossRef]

21. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* **2015**, *71*, 804–818. [CrossRef]

22. Bayirli, E.G.; Kaygun, A.; Öz, E. An analysis of PISA 2018 mathematics assessment for Asia-Pacific countries using educational data mining. *Mathematics* **2023**, *11*, 1318. [CrossRef]

23. Mehrabi, A.; Morphew, J. Board 73: AI Skills-based Assessment Tool for Identifying Partial and Full-Mastery within Large Engineering Classrooms. In Proceedings of the ASEE Annual Conference & Exposition, Portland, OR, USA, 23–26 June 2024; ASEE: Washington, DC, USA, 2024. [CrossRef]

24. Alghamdi, M.I. Assessing Factors Affecting Intention to Adopt AI and ML: The Case of the Jordanian Retail Industry. *MENDEL* **2020**, *26*, 39–44. [CrossRef]

25. Li, F.Q.; Wang, S.L.; Liew, A.W.C.; Ding, W.; Liu, G.S. Large-Scale Malicious Software Classification With Fuzzified Features and Boosted Fuzzy Random Forest. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 3205–3218. [CrossRef]

26. Liu-Yi, W.; Li-Gu, Z.H.U. Research and application of credit risk of small and medium-sized enterprises based on random forest model. In Proceedings of the 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 15–17 January 2021; pp. 371–374. [CrossRef]

27. Park, C.G. Implementing alternative estimation methods to test the construct validity of Likert-scale instruments. *Korean J. Women Health Nurs.* **2023**, *29*, 85–90. [CrossRef]

28. Abdelmagid, A.S.; Qahmash, A.I.M. Utilizing the Educational Data Mining Techniques 'Orange Technology' for Detecting Patterns and Predicting Academic Performance of University Students. *Inf. Sci. Lett.* **2023**, *12*, 1415–1431. [CrossRef]

29. Ahmad, I.; Basheri, M.; Iqbal, M.J.; Rahim, A. Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. *IEEE Access* **2018**, *6*, 33789–33795. [CrossRef]

30. Sengupta, S. Towards Finding a Minimal Set of Features for Predicting Students' Performance Using Educational Data Mining. *Int. J. Mod. Educ. Comput. Sci.* **2023**, *15*, 44–54. [CrossRef]

31. Osanaiye, O.; Cai, H.; Choo, K.K.R.; Dehghantanha, A.; Xu, Z.; Dlodlo, M. Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. *EURASIP J. Wirel. Commun. Netw.* **2016**, *2016*, 130. [CrossRef]

32. Yin, Y.; Jang-Jaccard, J.; Xu, W.; Singh, A.; Zhu, J.; Sabrina, F.; Kwak, J. IGRF-RFE: A hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset. *J. Big Data* **2023**, *10*, 15. [CrossRef]

33. Saeed, A.; Zaffar, M.; Abbas, M.A.; Quraishi, K.S.; Shahrose, A.; Irfan, M.; Huneif, M.A.; Abdulwahab, A.; Alduraibi, S.K.; Alshehri, F.; et al. A Turf-Based Feature Selection Technique for Predicting Factors Affecting Human Health during Pandemic. *Life* **2022**, *12*, 1367. [CrossRef]

34. Zaffar, M.; Hashmani, M.A.; Savita, K.; Rizvi, S.S.H.; Rehman, M. Role of FCBF Feature Selection in Educational Data Mining. *Mehran Univ. Res. J. Eng. Technol.* **2020**, *39*, 772–779. [CrossRef]

35. Tadist, K.; Najah, S.; Nikolov, N.S.; Mrabti, F.; Zahi, A. Feature selection methods and genomicbig data: A systematic review. *J. Big Data* **2019**, *6*, 79. [CrossRef]

36. Vommi, A.M.; Battula, T.K. A hybrid filter-wrapper feature selection using Fuzzy KNN based on Bonferroni mean for medical datasets classification: A COVID-19 case study. *Expert Syst. Appl.* **2023**, *218*, 119612. [CrossRef]

37. Zhou, X.; Li, Y.; Song, X.; Jin, L.; Wang, X. Thin Reservoir Identification Based on Logging Interpretation by Using the Support Vector Machine Method. *Energies* **2023**, *16*, 1638. [CrossRef]

38. Singh, J.; Evans, E.; Reed, A.; Karch, L.; Qualey, K.; Singh, L.; Wiersma, H. Online, Hybrid, and Face-to-Face Learning Through the Eyes of Faculty, Students, Administrators, and Instructional Designers: Lessons Learned and Directions for the Post-Vaccine and Post-Pandemic/COVID-19 World. *J. Educ. Technol. Syst.* **2022**, *50*, 301–326. [CrossRef]

39. Carmona, C.; Castillo, G.; Millán, E. Discovering Student Preferences in E-Learning. In Proceedings of the International Workshop on Applying Data Mining in E-Learning, Crete, Greece, 17–20 September 2007; pp. 33–42.

40. Kotsiantis, S.B.; Zaharakis, I.D.; Pintelas, P.E. Assessing Supervised Machine Learning Techniques for Predicting Student Learning Preferences. In Proceedings of the 3rd Congress on Information and Communication Technologies in Education, London, UK, 27–28 February 2018; Dimitracopoulou, A., Ed.; University of Aegean: Rhodes, Greece, 2019. [CrossRef]

41. Hew, K.F.; Hu, X.; Qiao, C.; Tang, Y. What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Comput. Educ.* **2020**, *145*, 103724. [CrossRef]

42. Hebbecker, K.; Förster, N.; Forthmann, B.; Souvignier, E. Data-based decision-making in schools: Examining the process and effects of teacher support. *J. Educ. Psychol.* **2022**, *114*, 1695. [CrossRef]

43. Turgut, Y.; Bozdag, C.E. A framework proposal for machine learning-driven agent-based models through a case study analysis. *Simul. Model. Pract. Theory* **2023**, *123*, 102707. [CrossRef]

44. Ouyang, F.; Wu, M.; Zheng, L.; Zhang, L.; Jiao, P. Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 1–23. [CrossRef]

45. Mehrabi, A.; Morphew, J. Investigating and predicting the Cognitive Fatigue Threshold as a Factor of Performance Reduction in Assessment. In Proceedings of the ASEE Annual Conference & Exposition, Portland, OR, USA, 23–26 June 2024; ASEE: Washington, DC, USA, 2024. [CrossRef]

46. Muis, K.R.; Chevrier, M.; Singh, C.A. The Role of Epistemic Emotions in Personal Epistemology and Self-Regulated Learning. *Educ. Psychol.* **2018**, *53*, 165–184. [CrossRef]

47. Kumar, R.; Sexena, A.; Gehlot, A. Artificial Intelligence in Smart Education and Futuristic Challenges. In Proceedings of the 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 11–12 May 2023; pp. 432–435.

48. Aparicio, M.; Bacao, F.; Oliveira, T. An e-Learning Theoretical Framework | Enhanced Reader. *Educ. Technol. Soc.* **2016**, *19*, 293–307.

49. Schrumpf, J. On the Effectiveness of an AI-Driven Educational Resource Recommendation System for Higher Education. *Int. Assoc. Dev. Inf. Soc.* **2022**, *1*, 883–901.

50. Whiteside, A.L.; Dikkers, A.G.; Lewis, S. More Confident Going into College': Lessons Learned from Multiple Stakeholders in a New Blended Learning Initiative. *Online Learn.* **2016**, *20*, 136–156. [CrossRef]

51. Sitzmann, T.; Ely, K. A Meta-Analysis of Self-Regulated Learning in Work-Related Training and Educational Attainment: What We Know and Where We Need to Go. *Psychol. Bull.* **2011**, *137*, 421–442. [CrossRef]

52. Balid, W.; Alrouh, I.; Hussian, A.; Abdulwahed, M. Systems engineering design of engineering education: A case of an embedded systems course. In Proceedings of the IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE) 2012, Hong Kong, China, 20–23 August 2012; p. W1D-7.

53. Passey, D. Technology-enhanced learning: Rethinking the term, the concept and its theoretical background. *Br. J. Educ. Technol.* **2019**, *50*, 972–986. [CrossRef]

54. Duval, E.; Sharples, M.; Sutherland, R. Research themes in technology enhanced learning. In *Technology Enhanced Learning: Research Themes*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1–10. [CrossRef]

55. Jackson, C.K. The full measure of a teacher: Using value-added to assess effects on student behavior. *Educ. Next* **2019**, *19*, 62–69. Available online: https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=15399664&v=2.1&it=r&id=GALE%7CA566264029&sid=googleScholar&linkaccess=fulltext (accessed on 10 August 2023).

56. Gunawardena, M.; Dhanapala, K.V. Barriers to Removing Barriers of Online Learning. *Commun. Assoc. Inf. Syst.* **2023**, *52*, 264–280. [CrossRef]

57. Rabbi, J.; Fuad, M.T.H.; Awal, M.A. Human Activity Analysis and Recognition from Smartphones using Machine Learning Techniques. *arXiv* **2021**. [CrossRef]

58. Sánchez-Ruiz, L.; López-Alfonso, S.; Moll-López, S.; Moraño-Fernández, J.; Vega-Fleitas, E. Educational Digital Escape Rooms Footprint on Students' Feelings: A Case Study within Aerospace Engineering. *Information* **2022**, *13*, 478. [CrossRef]

59. Bland, J.M.; Altman, D.G. Statistics notes: Cronbach's alpha. *BMJ* **1997**, *314*, 572. [CrossRef]

60. Afthanorhan, A.; Ghazali, P.L.; Rashid, N. Discriminant Validity: A Comparison of CBSEM and Consistent PLS using Fornell & Larcker and HTMT Approaches. *J. Phys. Conf. Ser.* **2021**, *1874*, 012085. [CrossRef]

61. Henseler, J.; Ringle, C.M.; Sarstedt, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **2015**, *43*, 115–135. [CrossRef]

62. Yusoff, A.S.M.; Peng, F.S.; Abd Razak, F.Z.; Mustafa, W.A. Discriminant validity assessment of religious teacher acceptance: The use of HTMT criterion. *J. Phys. Conf. Ser.* **2020**, *1529*, 042045. [CrossRef]

63. Ab Hamid, M.R.; Sami, W.; Sidek, M.M. Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. *J. Phys. Conf. Ser.* **2017**, *890*, 012163. [CrossRef]

64. Rempel, J.K.; Holmes, J.G.; Zanna, M.P. Trust in close relationships. *J. Personal. Soc. Psychol.* **1985**, *49*, 95–112. [CrossRef]

65. Remeseiro, B.; Bolon-Canedo, V. A review of feature selection methods in medical applications. *Comput. Biol. Med.* **2019**, *112*, 103375. [CrossRef] [PubMed]

66. Buenaño-Fernández, D.; Gil, D.; Luján-Mora, S. Application of machine learning in predicting performance for computer engineering students: A case study. *Sustainability* **2019**, *11*, 2833. [CrossRef]

67. Chopade, S.; Chopade, S.; Gawade, S. Multimedia teaching learning methodology and result prediction system using machine learning. *J. Eng. Educ. Transform.* **2022**, *35*, 135–142. [CrossRef]

68. Campbell, I. Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. *Stat. Med.* **2007**, *26*, 3661–3675. [CrossRef] [PubMed]

69. Borrego, M.; Froyd, J.E.; Hall, T.S. Diffusion of engineering education innovations: A survey of awareness and adoption rates in US engineering departments. *J. Eng. Educ.* **2010**, *99*, 185–207. [CrossRef]

70. Masood, H. Breast cancer detection using machine learning algorithm. *Int. Res. J. Eng. Technol. (IRJET)* **2021**, *8*, 1–5.

71. Rachburee, N.; Punlumjeak, W. A comparison of feature selection approach between greedy, IG-ratio, Chi-square, and mRMR in educational mining. In Proceedings of the 2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE), Chiang Mai, Thailand, 29–30 October 2015; pp. 420–424.

72. Chen, R.C.; Dewi, C.; Huang, S.W.; Caraka, R.E. Selecting critical features for data classification based on machine learning methods. *J. Big Data* **2020**, *7*, 52. [CrossRef]

73. Jia, W.; Sun, M.; Lian, J.; Hou, S. Feature dimensionality reduction: A review. *Complex Intell. Syst.* **2022**, *8*, 2663–2693. [CrossRef]

74. Kursa, M.B.; Jankowski, A.; Rudnicki, W.R. Boruta—A System for Feature Selection. *Fundam. Inform.* **2010**, *101*, 271–285. [CrossRef]

75. Saarela, M.; Jauhiainen, S. Comparison of feature importance measures as explanations for classification models. *SN Appl. Sci.* **2021**, *3*, 1–12. [CrossRef]

76. Yan, K.; Zhang, D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sens. Actuators B Chem.* **2015**, *212*, 353–363. [CrossRef]

77. Alotaibi, B.; Alotaibi, M. Consensus and majority vote feature selection methods and a detection technique for web phishing. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 717–727. [CrossRef]

78. Borandag, E.; Ozcift, A.; Kilinc, D.; Yucalar, F. Majority vote feature selection algorithm in software fault prediction. *Comput. Sci. Inf. Syst.* **2019**, *16*, 515–539. [CrossRef]

79. Jindal, A.; Dua, A.; Kaur, K.; Singh, M.; Kumar, N.; Mishra, S. Decision Tree and SVM-Based Data Analytics for Theft Detection in Smart Grid. *IEEE Trans. Ind. Inform.* **2016**, *12*, 1005–1016. [CrossRef]

80. Teo, S.G.; Han, S.; Lee, V.C.S. Privacy Preserving Support Vector Machine Using Non-linear Kernels on Hadoop Mahout. In Proceedings of the 2013 IEEE 16th International Conference on Computational Science and Engineering, Sydney, Australia, 3–5 December 2013; pp. 941–948. [CrossRef]

81. Sikder, J.; Datta, N.; Tripura, S.; Das, U.K. Emotion, Age and Gender Recognition using SURF, BRISK, M-SVM and Modified CNN. In Proceedings of the 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET), Prague, Czech Republic, 20–22 July 2022; pp. 1–6. [CrossRef]

82. Kuo, K.M.; Talley, P.C.; Chang, C.S. The accuracy of machine learning approaches using non-image data for the prediction of COVID-19: A meta-analysis. *Int. J. Med Inform.* **2022**, *164*, 104791. [CrossRef]

83. Krawczyk, B. Learning From Imbalanced Data: Open Challenges and Future Directions. *Prog. Artif. Intell.* **2016**, *5*, 221–232. [CrossRef]

84. Lunardon, N.; Menardi, G.; Torelli, N. ROSE: A Package for Binary Imbalanced Learning. *R J.* **2014**, *6*, 79–89. [CrossRef]

85. Buda, M.; Maki, A.; Mazurowski, M.A. A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks. *Neural Netw.* **2018**, *106*, 249–259. [CrossRef]

86. Sharma, A.; Verbeke, W. Improving Diagnosis of Depression With XGBOOST Machine Learning Model and a Large Biomarkers Dutch Dataset (N = 11,081). *Front. Big Data* **2020**, *3*, 15. [CrossRef] [PubMed]

87. Liu, X.Y. An Empirical Study of Boosting Methods on Severely Imbalanced Data. *Appl. Mech. Mater.* **2014**, *513–517*, 2510–2513. [CrossRef]

88. Beuthner, C.; Friedrich, M.; Herbes, C.; Ramme, I. Examining survey response styles in cross-cultural marketing research: A comparison between Mexican and South Korean respondents. *Int. J. Mark. Res.* **2018**, *60*, 257–267. [CrossRef]

89. Vora, D.R.; Iyer, K.R. Deep Learning in Engineering Education: Implementing a Deep Learning Approach for the Performance Prediction in Educational Information Systems. In *Deep Learning Applications and Intelligent Decision Making in Engineering*; IGI Global: Hershey, PA, USA, 2021; pp. 222–255.

90. Vora, D.R.; Iyer, K.R. Deep Learning in Engineering Education: Performance Prediction Using Cuckoo-Based Hybrid Classification. In *Machine Learning and Deep Learning in Real-Time Applications*; IGI Global: Hershey, PA, USA, 2020; pp. 187–218.

91. Saputra, N.A.; Hamidah, I.; Setiawan, A. A bibliometric analysis of deep learning for education research. *J. Eng. Sci. Technol.* **2023**, *18*, 1258–1276.

92. Davis, L.; Sun, Q.; Lone, T.; Levi, A.; Xu, P. In the Storm of COVID-19: College Students' Perceived Challenges with Virtual Learning. *J. High. Educ. Theory Pract.* **2022**, *22*, 66–82.

93. Li, H. The Influence of Online Learning Behavior on Learning Performance. *Appl. Sci. Innov. Res.* **2023**, *7*, 69. [CrossRef]

94. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Troussas, C.; Sgouropoulou, C. A hybrid machine learning model for grade prediction in online engineering education. *Int. J. Eng. Pedagog* **2022**, *12*, 4–23. [CrossRef]

95. Onan, A. Mining opinions from instructor evaluation reviews: A deep learning approach. *Comput. Appl. Eng. Educ.* **2020**, *28*, 117–138. [CrossRef]

96. Yogeshwaran, S.; Kaur, M.J.; Maheshwari, P. Project based learning: Predicting bitcoin prices using deep learning. In Proceedings of the 2019 IEEE global engineering education conference (EDUCON), Dubai, United Arab Emirates, 8–11 April 2019; pp. 1449–1454.

97. Lameras, P.; Arnab, S. Power to the Teachers: An Exploratory Review on Artificial Intelligence in Education. *Information* **2022**, *13*, 14. [CrossRef]