Array 19 (2023) 100303

Contents lists available at ScienceDirect

Array

journal homepage: www.sciencedirect.com/journal/array

Forecasting students' adaptability in online entrepreneurship education using modified ensemble machine learning model

Amit Malik [a], Edeh Michael Onyema [b,c], Surjeet Dalal [d,*], Umesh Kumar Lilhore [e], Darpan Anand [f], Ashish Sharma [g], Sarita Simaiya [h]

[a] Department of Computer Science and Engineering, SRM University, Delhi-NCR, Sonipat, Haryana, India
[b] Department of Vocational and Technical Education, Faculty of Education, Alex Ekwueme Federal University, Ndufu-Alike, Abakaliki, Nigeria
[c] Adjunct Faculty, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India
[d] Department of Computer Science and Engineering, Amity University Haryana, Gurugram, Haryana, India
[e] Department of Computer Science and Engineering, Chandigarh University, Gharuan, Punjab, 140413, India
[f] Department of Computer Engineering, Sir Padampat Singhania University, Udaipur, India
[g] Department of Computer Engineering and Applications, GLA University, Mathura, UP, 281406, India
[h] Apex Institute of Technology (CSE), Chandigarh University Gharuan, Mohali, Punjab, India

ARTICLE INFO

Keywords:
Entrepreneurship
Online education
Adaptability
Academic forecasting
Machine learning

ABSTRACT

Entrepreneurship education has become essential in recent years. This education system may not be unconnected with the global agitation for value creation, employability skills and job creation. Engaging in entrepreneurial training provides students with the skills needed to enhance their ability to create marketable and profitable solutions to emerging problems. To do this, many emerging entrepreneurs rely on technology to engage in entrepreneurship education. This study presents a machine learning technique to predict the adaptability level of students in online entrepreneurship education. The suitability of different algorithms like Random Forest, C5.0, CART and Artificial Neural Network was examined using the Kaggle Educational dataset. The algorithms recorded a high accuracy rate and affirmed machine learning techniques' ability to forecast students' adaptation to online entrepreneurship training. The findings of this research contribute to the field of online entrepreneurship education by providing a reliable and efficient approach for predicting students' adaptability. The proposed modified ensemble machine learning model can assist educators and administrators in identifying students who may require additional support, tailoring instructional strategies, and designing targeted interventions to enhance their adaptability and overall learning experience in online entrepreneurship education.

1. Introduction

The need for entrepreneurship training is growing globally [1]. This need is because government alone can no longer provide jobs for everyone, and people must find ways to create jobs and wealth for themselves. Entrepreneurship education has become necessary as unemployment bites harder and poverty increases [2]. Many educational institutions and governments are refocusing their policies and curriculum on education on problem-solving [3]. This factor is necessary to bridge the employment gap and encourage critical thinking. Mobile phones and laptops increase the chances of students' involvement in entrepreneurship learning. These technologies aid access to entrepreneurship mentors/coaches, colleagues and training. Through e-learning, young entrepreneurs can access multiple resources to enhance their entrepreneurial mindsets and skills. Online systems allow students to develop more interest in entrepreneurship training and communicate with their mentors. Technology plays a massive role in spreading entrepreneurship education, but there are concerns about students' adaptability to online education. These could be from different causative factors, including the digital divide and network issues. However, technology is believed to be very much needed to promote entrepreneurship education, learning and practice. This study attempted to predict the adaptability of students in online entrepreneurship education courses to recommend possible ways to improve their adaptation.

* Corresponding author.
E-mail addresses: amit.m@srmuniversity.ac.in (A. Malik), mikedreamcometrue@gmail.com (E.M. Onyema), sdalal@ggn.amity.edu (S. Dalal), umeshlilhore@gmail.com (U.K. Lilhore), darpan.anand@spsu.ac.in (D. Anand), ashish.sharma@gla.ac.in (A. Sharma), saritasimaiya@gmail.com (S. Simaiya).

https://doi.org/10.1016/j.array.2023.100303
Received 20 March 2023; Received in revised form 20 June 2023; Accepted 24 June 2023
Available online 3 July 2023

1.1. Related concepts

1.1.1. Adaptability

There is an ongoing effort to integrate entrepreneurship studies into educational curriculums. While this is a welcome idea [4], it requires a change of mindsets and methodology by the teachers and students. The need for higher-education entrepreneurial education has grown significantly in the recent decade. U.S. colleges and universities provided more than twice as many entrepreneurship courses as they did ten years earlier in 2004 [5]. The distribution of these courses has a great deal of shared information. Identifying opportunities, creating company plans, and implementing those plans are all part of the Entrepreneurship course's fundamental threshold ideas. The first two sections cover marketing, operations, finance, team development, legal challenges, and innovation. The third section looks at how a business may get started and expand.

The substance of a course is inextricably linked to the goals it seeks to achieve with its students. Entrepreneurship courses at business schools, for example, emphasize marketing and financial aspects, while those at engineering schools focus on innovation and operations [6]. The ability to adapt one's behaviour, ideas and feelings to new, unpredictable settings and circumstances is called adaptability [7]. Student's academic and personal growth, as well as their motivation, engagement, accomplishment, and psychological well-being, depend heavily on their ability to adapt. COVID-19 forced many people to transit online. The transition to online education came with many challenges, including sustainability issues. However, e-learning platforms were very helpful in ensuring continued education during that period. This study attempted to examine students' adaptability level in learning entrepreneurship online during that lockdown. Hopefully, the investigation will highlight the growing impact of technology on entrepreneurship training and education.

1.1.2. Entrepreneurship

Innovation and entrepreneurship education are becoming increasingly important in many nations worldwide, while certain industrialized countries have created a relatively adult entrepreneurship education and support system. At least 400 institutions in the United States offer entrepreneurship courses, and several of the country's most prestigious universities now offer entrepreneurship courses and degrees. 'The Graduate Enterprise Program,' a government initiative begun in 1998, is now in its third year. A law promoting university technology transfer was enacted in Japan in 1998 to encourage students to pursue entrepreneurial endeavours while still in school. We systematically evaluated the technologies used from three perspectives: pedagogy, usability, and technology. A total of five specific cases have been chosen for in-depth investigation and comparison [8].

There are following research contributions of paper given below:

- This paper makes an original contribution by zeroing in on the topic of teaching entrepreneurship over the internet. Research gives insights into forecasting students' adaptation in a specialized field of education by taking into account the specific problems and dynamics of this setting.
- In order to predict students' adaptability, this research suggests a modified ensemble machine learning approach. To address the complexities of adaptability prediction in the context of online entrepreneurship education, the model is likely an original extension or adaption of current ensemble methodologies.
- This paper explores the possibility of combining data from several sources in order to improve forecast accuracy. This finding highlights the significance of including a wide range of input factors into a prediction model, including demographic data, previous academic achievement, learning practices, and interaction patterns.
- The results of this work may be useful for those working in the fields of online entrepreneurship education and policymaking. The research can guide the creation of individualized treatments, instructional techniques, and support systems to boost students' learning results and overall educational experience by identifying characteristics that impact adaptability and anticipating their degrees of adaptability.

This paper is designed with aiming of prediction of adaptability of online education. section 2 defines the various existing works carried out in this problem domain. section 3 explains the about datasets used for experimental purpose. In this paper, educational dataset from Kaggle is being taken. section 4 demonstrates proposed Stacking Ensemble Learning. Next section 5 demonstrate the result followed by conclusion of the paper in section 6.

2. Review of literature

Various creators have introduced research concentrates on business expectation and their suggestions; nonetheless, research concentrates on instructive advancements in E.E. have been restricted notwithstanding the functional turn of events and shown off on the web and hybrid courses utilizing the web and instructive apparatuses since most recent twenty years.

By leading both on the web and actual overviews, Liu et al. [9] gathered understudies' data on the three individual levels (School, College, and University). The review structure comprises a person's socio-segment factors. To find out about the viability of online training, we have applied a few A.I. algorithms to anticipate the flexibility level of the understudies to online schooling. Among utilized algorithms, the Random Forest classifier accomplished the best precision of 89.63% and berated different algorithms.

Meifeng et al. [10] presented an independent web-based schooling system in light of wise suggestions. This framework can change the review strategies per the student's review circumstance and hence works on the viability during the time spent internet considering. The interest investigation is completed, the framework's structure is presented, and the key innovation is proposed in this paper. Examining the independent web-based schooling system in light of the astute proposal gives a new strategy for online instruction, which can be a significant point to be discussed.

Wang et al. [11] explored the job of flexibility in assisting secondary school understudies with exploring their web-based getting the hang of during a time of COVID-19 that involved completely or somewhat far off internet learning. Drawing on the Job Demands-Resources hypothesis and information from an example of 1548 Australian secondary school understudies in nine schools, we inspected the job of flexibility in foreseeing understudies' web-based learning self-viability in science and their fin-ish-of-year arithmetic accomplishment. It was found that past the impacts of web-based learning re-quests, on the web and parental learning backing, and foundation ascribes, versatility was fundamentally connected with more elevated levels of internet learning self-viability and gains in later accomplishment; internet learning self-adequacy was additionally altogether combined with improvements in achievement.

Rubia et al. [12] explore the significance of the evaluation framework in the Web-based distance learning schooling climate and lays out an astute appraisal framework model, which is portrayed for its positive presentation and versatility. In addition, the execution of the evaluation framework is given in light of this model. In an incorporated instruction climate, this evaluation framework helps out different subsystems during educating and growing experiences and assumes a significant part in working on the presentation of the entire schooling system.

Kandakatla et al. [13] build a transformative game model with the support of centre ventures, fire up endeavours and the public authority, examines the balance point and relating solidness conditions for the developmental game model, and do a recreation investigation on various soundness conditions from the parts of the underlying readiness of the members, the power of the sponsorship and discipline, the expense and advantage of the methodology choice. The initial expectation will undoubtedly influence the subject's positive system choice and adversely affect the subject's adverse technique determination.

Lu et al. [14] inspect formal versus casual, youthful high-innovation adventures, evaluating the degree to which strong familiarity directs the connection between innovation speculations and firm execution in an arising economy. We observe that innovation speculations are decidedly connected with forceful execution. However, firm casualness decidedly directs the connection between assets and athletic execution. The impact of investments will generally depend upon the degree of firm casualness in a rising economy. More unique asset requirements and more elevated levels of strong familiarity address essential variables in arising economies, yet in addition, are fundamental for contextualizing asset-based hypothesis.

Suet al. [15] presents the pioneering profile of STEM (Science, Technology, Engineering and Mathematics) understudies at the University of Jaén (Spain). The Kaiser-Meyer-Olkin (KMO) test and Pearson's connection test were utilized for factor examination and relationships. The outcomes demonstrate that understudies have highly innovative mentalities, despite no tremendous contrasts in age and orientation. Then again, measurably, contrasts have been tracked in light of the designing mastery region. Understudies of the speciality in Industrial Electronics Engineering have higher mentalities concerning innovativeness and development, which are vital for business ventures.

Xu et al. [16] utilize the standards of helpful arrangement to dissect and upgrade a few parts of the Entrepreneurship course. The attention is on surveying and adjusting the appraisal assignments to guarantee a successful assessment and the accomplishment of understudy learning results. The course update process and the beneficial arrangement and creative appraisal can be applied to different courses in the field and more extensively to educational planning, instructing, and learning in advanced education.

Wang et al. [17] utilized A.I. (ML) algorithms to distinguish low-commitment understudies in a sociology course at the Open University (O.U.) to evaluate the impact of commitment on understudy execution. The info factors of the review included the most extraordinary schooling level, eventual outcomes, score on the appraisal, and the number of snaps on virtual learning climate (VLE) exercises. The results showed that the J48, choice tree, JRIP, and angle helped classifiers display better execution regarding the exactness, kappa worth, and review contrasted with the other tried models.

Xueli et al. [18] propose the conduct order-based e-learning execution (BCEP) expectation system, which chooses the elements of e-learning ways of behaving, utilizes highlight combination with conduct information as indicated by the conduct characterization model to acquire the classification including upsides of each sort of conduct, lastly assembles a learning execution indicator in light of A.I. Further-more, because current e-learning conduct order techniques don't wholly consider the method involved with learning, we likewise propose a web-based conduct arrangement model in light of the e-educational experience called the cycle conduct grouping (PBC) model. Trial results with the Open University Learning Analytics Dataset (OULAD) show that the learning execution indicator in light of the BCEP forecast outline work has a decent expectation impact, and the presentation of the PBC model in learning execution expectation is superior to conventional grouping strategies. We develop an e-gaining execution indicator according to another viewpoint and give another answer for the quantitative assessment of e-learning grouping techniques.

Pardede et al. [19] also endeavour to blueprint and think to gauge the accompanying feeling and social elements of commitment (abilities, cooperation/collaboration, and execution, close to home). The outcomes uncovered that the exploratory gathering is measurably fundamentally higher than those in the benchmark group. These experimental outcomes infer the capability of a versatile e-learning climate to connect with understudies towards learning.

3. Materials and methods

This section specifies the dataset taken for research work and the methods involved in this problem domain.

3.1. Dataset

This dataset has been taken from Kaggle to evaluate the effectiveness of online education. The target feature is adaptively level and the various feature sets. The link to this dataset is given below:

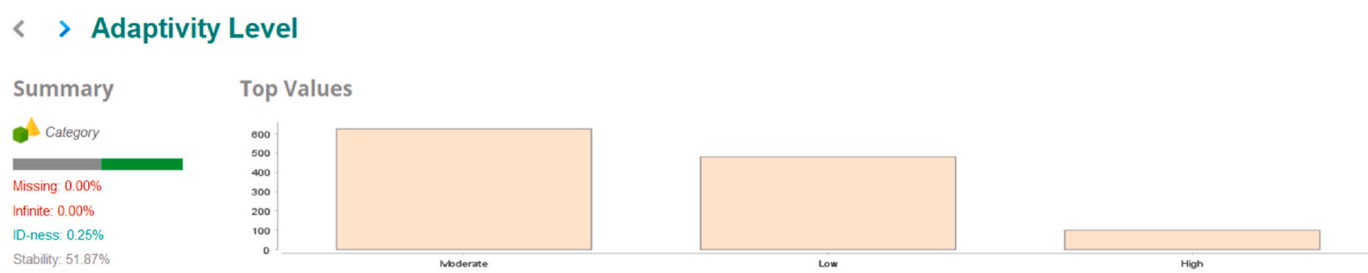
http://www.kaggle.com/datasets/mdmahmudulhasansuzan/students-adaptabilitylevel-in-online-education.

The statistical analysis of the target variable 'Adaptivity Level' is shown in Fig. 1. The target values contain three different types of values, i.e. Moderate, Low and High. It also shows their counts and percentages.

3.1.1. Missing values

Human mistakes in data processing, faulty equipment, respondents' unwillingness to answer particular questions, drop-out in studies, and combining unrelated data are the most common causes of missing numbers. Performance deterioration, data analysis challenges, and skewed outputs can all be caused by missing values in any data-related domain [20]. Lost data can have varying degrees of significance depending on the amount, the pattern, and the cause behind its absence. A technique known as imputation can be used to address missing data, which involves deleting instances and replacing them with possible or estimated values [21]. In this paper, the authors applied Hot-deck imputation approach for handling missing values in the context of forecasting students' adaptability in online entrepreneurship education. First authors define the criteria to identify similar students based on which missing values will be imputed. The criteria can include attributes such as age, gender, prior educational background, online learning experience, or any other relevant factors that may influence students' adaptability. Then they find students in the dataset who have complete information for the similarity criteria. These students will serve as the "donor" or reference group for imputing missing values. For each student with missing values, the authors identify the most similar students (donors) from the reference group based on the defined similarity criteria. Similarity can be measured using distance metrics like Euclidean distance, cosine similarity, or other suitable measures. Once the similar students (donors) are identified, authors impute the missing values in the target student's adaptability based on the values of the corresponding features in the donor students. This can be done by directly assigning the values of the donors or using statistical methods like taking the mean or median value of the donor values. It is important to be cautious of potential donor selection bias. The authors ensure that the selected donors represent a diverse range of characteristics and that the imputed values do not introduce bias into the dataset. Bias may arise if there are systematic differences between the students with missing values and the selected donors.

Finally they assess the impact of hot-deck imputation on the forecasting model's performance and compare the results obtained with and without imputed values to evaluate the effectiveness of the imputation technique. It is crucial to examine if the imputed values provide reasonable estimates and do not distort the relationship between predictors and the target variable. The authors conduct sensitivity analysis to examine the robustness of the imputed values and their impact on the forecasting results. Hot-deck imputation provides a way to fill in missing values based on similar students' characteristics. However, it is important to acknowledge the limitations of this approach. Hot-deck imputation assumes that similar students will have similar values for missing features, which may not always hold true. Careful consideration should be given to the suitability of this method for the specific dataset and the potential biases that may be introduced during imputation.



3 Distinct Values:

Value	Count	Percentage
Moderate	625	51.87%
Low	480	39.83%
High	100	8.30%

Fig. 1. Details of Adaptivity level.

3.1.2. *Outlier data*

Data that is markedly different from the rest of the data is known as an outlier. An outlier might be caused by measurement variability or an experimental mistake; the latter are often eliminated from the data set. Analyzing an outlier might complicate data. Detecting outliers in Python uses used following methods such as:

- Rescaling the data
- Marking the outliers
- Dropping outliers

3.2. *Methods*

3.2.1. *Bayesian networks*

Probabilistic graphical models such as Bayesian networks are frequently employed. The structure and parameters make up this system. Random variables are associated with nodes in a directed acyclic graph (DAG) representing conditional independencies and dependencies. Each node has its own set of conditional probability distributions. A Bayesian network is a compact, adaptable, and easily understandable representation of a joint distribution of probabilities [22]. As directed acyclic networks enable the modelling of causal linkages between variables, it is also a valuable resource in searching for new information.

In most cases, a Bayesian network is learned by analyzing relevant data. In this case, we have used the Tree Augmented Naïve Bayes model, which considers the probability of fields used in prediction with the dependent target field and other independent areas, thereby increasing the model's accuracy. The fields' dependency is tested using a likelihood ratio with a 0.01 significance level [23].

Fig. 2 demonstrates the target and predictor variables, classified by two different colours. It identifies the relevant variables that can affect students' adaptability in online education. These variables may include demographic information, prior academic performance, learning behaviors, self-efficacy, motivation, social support, and any other factors deemed important. Each variable will be represented as a node in the Bayesian network. It determine the relationships and dependencies between the variables. This step involves analyzing how the variables interact with each other and influence adaptability.

Also, the predictor variables are assigned different values of "importance" using the variance-based method to decide the importance of a field concerning our dependent area, as shown in Fig. 3. Bayesian networks allow for personalized predictions by considering individual student attributes as input variables. By incorporating specific information about each student, the predictor can generate tailored predictions, enabling personalized interventions and support. The sensitivity value is calculated using the predictor variables and the target variables' variance [24] to rank the predictor variables.

According to the values calculated and shown in the above chart, the most influencing variable is the 'financial condition' with an importance value of 0.21, and the least influencing variable is the 'education level'. However, the variable rank does not affect the accuracy and only

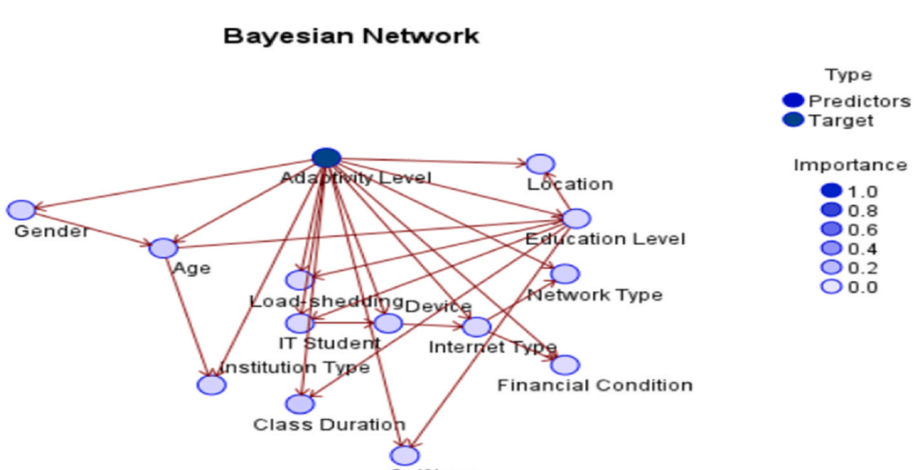


Fig. 2. Bayesian Network for prediction adaptively of online education.

provides insight into the model. The Bayesian Network's accuracy is 76% in the current setup.

3.2.2. Random forest model

As the name suggests, many individual decision trees are in a random forest, yet they all work together to make decisions. The class forecast of each tree in the random forest is based on the number of votes it receives. The wisdom of crowds is a simple yet powerful principle at the heart of random forest [25]. In this model, the predictor importance is as below in Fig. 4. Random Forest provides a measure of feature importance, which indicates the relative contribution of each predictor variable in making accurate predictions. This information can help identify the most influential predictors and understand their impact on the outcome variable. Feature importance can be calculated based on metrics such as Gini impurity or information gain.

Features names for short as given below:

1. Class Duration - F1
2. Financial Condition - F2
3. Age - F3
4. Gender - F4
5. Location - F5
6. Institution Type - F6
7. Internet Type - F7
8. Load-shedding - F8
9. Self Lms - F9
10. I.T. Student - F10
11. Network Type_4G - F11
12. Network Type_3G - F12
13. Education Level_University - F13
14. Education Level_School - F14
15. Device_Mobile - F15
16. Education Level_College - F16
17. Device_Computer - F17
18. Device_Tab - F18
19. Network Type_2G - F19

The bootstrap technique is implemented for building the trees and

forest measured the accuracy of more than 93%, providing the highest significance to the variable' class duration' [27].

3.2.3. Neural network

The input, hidden, and output layers are all possible layers in neural networks. Many neurons are interconnected in the hidden layer or layers in between. These neurons' weights (or strength) are "fine-tuned" as the network "learns" from data, allowing accurate predictions to be made [28]. The Neural Network has been designed for the current problem, consisting of one hidden layer and seven neurons, as shown in Fig. 5.

The Neural Network has been designed for the current problem, consisting of one hidden layer with six neurons, as shown in Fig. 4. The activation function is 'hyperbolic tangent' for the hidden layer and 'softmax' for the output layer because the target variable has more than two classes, which is a classification problem. The model predicted the variable' class duration' as the most important one with a value of 0.21 and showed an accuracy of up to 74% in the standard neural Network.

3.2.4. CART

In artificial intelligence, machine learning has been the most quickly developing issue. In machine learning, several algorithms have garnered a lot of traction because of their openness. CART (Classification and Regression Trees) is a well-known name for the Decision Tree method [29]. There are several applications for this algorithm, which is a straightforward machine learning algorithm. In the current problem, the predictor variables have been assigned to the following level, as shown in Fig. 6.

CART's most important predictor fields in the current setup are 'financial condition' with a value of 0.36 and 'class duration' with a value of 0.35. The model measured an accuracy of more than 72%. As we are dealing with the categorical data, our target variable is absolute, so we have used the 'Gini' method to measure the impurity, a probabilistic approach whose value ranges between 0 and 1. The variable that recorded the maximum impurity decrease is used for splitting the tree further [30].

```
Class Duration in [ "0" ] [ Mode: Low ] => Low (111; 0.928)
Class Duration in [ "1-3" "3-6" ] [ Mode: Moderate ] (757)
    Financial Condition in [ "Mid" ] [ Mode: Moderate ] (550)
        Institution Type in [ "Government" ] [ Mode: Low ] (154)
            Education Level in [ "College" "School" ] [ Mode: Moderate ] (94)
                Location in [ "No" ] [ Mode: Low ] => Low (20; 0.7)
                Place in [ "Yes" ] [ Mode: Moderate ] => Moderate (74; 0.662)
            Education Level in [ "University" ] [ Mode: Low ] => Low (60; 0.683)
        Institution Type in [ "Non-Government" ] [ Mode: Moderate ] (396)
            Age in [ "16-20" ] [ Mode: Low ] (75)
                Class Duration in [ "1-3" ] [ Mode: Low ] => Low (48; 0.729)
                Class Duration in [ "3-6" ] [ Mode: Moderate ] => Moderate (27; 0.778)
            Age in [ "1-5" "11-15" "21-25" "26-30" "6-10" ] [ Mode: Moderate ] => Moderate (321;
0.791)
    Financial Condition in [ "Poor" "Rich" ] [ Mode: Low ] (207)
        A device in [ "Computer" ] [ Mode: High ] => High (23; 0.783)
        A device in [ "Mobile" "Tab" ] [ Mode: Low ] (184)
            Financial Condition in [ "Poor" ] [ Mode: Low ] (144)
                Age in [ "1-5" ] [ Mode: Moderate ] => Moderate (26; 0.808)
                Age in [ "11-15" "16-20" "21-25" "26-30" "6-10" ] [ Mode: Low ] => Low (118;
0.602)
            Financial Condition in [ "Rich" ] [ Mode: Moderate ] => Moderate (40; 0.425)
```

calculates the estimated mean through random sampling. The 'random subspace method' [26] implements this technique. A total of 20 numbers of trees are built with value one as the least possible minimum leaf size. The tree is split by choosing the best split among the group of random dependent variables, which usually decreases the overall variance of the model, giving a better model. In this case, also the random

3.3. Proposed Stacking Ensemble Learning

It is a machine learning algorithm that combines predictions of machine learning models, like bagging and boosting. It involves two base models, level-0 and level-1 models [31]. The other is commonly known as the meta-model or level-1. This model is used in predicting the base model.

- Level-0 Models: Fitting the model after training the model on the data.
- Level-1 Models: We use this model to learn and then combine the forecast made by the level-0 model.

This model quickly automates most of the procedures below and generates ensembles as supplied instructions.

parameters and evaluating its performance on different subsets of the training data to find the best configuration. The Label role-labeled input training data is available via the inner training set of the port in this model. A Prediction Model may be educated with this data to predict the labeled data better. Specifically, the model port of the inner subprocess requires the Prediction Model. The compiled Prediction Models are made available as a Multi-Label Model at the operator's model output port. The input and output port extender permits more items to be sent to and from the subprocess.

4. Results

Algorithm 1:
Input:
M number of algorithms;
N number of rows in the training set
Output:
Prediction of Adaptability level

Step 1. Set up the ensemble.
a) Initially, N-base algorithms with specific parameters can be specified. (Random Forest, Neural Network & CART)
b) Next, define a method for meta-learning. (NaiveBayes with default parameters).
Step 2. Train the ensemble.
a) On the training set, run each of the M basic algorithms.
b) The user may do k-fold cross-validation and gather each learner's projected values from M algorithms.
c) A new N x M matrix may be created by combining the N predicted values from each M method that has undergone cross-validation. The "level-one" data includes this matrix and the original response vector.
d) The meta-learning algorithm should be trained on data from level one. It is possible to use the "ensemble model" to make predictions on a test set using the L base learning and meta-learning models.
Step 3. Predict new data.
a) To produce ensemble predictions, first, generate predictions from the primary learners.
b) Feed those predictions into the meta-learner to construct the ensemble forecast.

Output received from the level-0 model is used in the training level-1 model [35].

In Fig. 7, the proposed ensemble model is based on multi-level modelling with Feature weight by Tree Importance. Once the ensemble model is trained and fine-tuned, it can be used to make predictions on new, unseen data. The model has been applied to new student data, and it will output predictions or probabilities indicating the students' adaptability in online entrepreneurship education. It perform hyperparameter tuning and cross-validation to optimize the ensemble model's performance. This step involves adjusting the model's

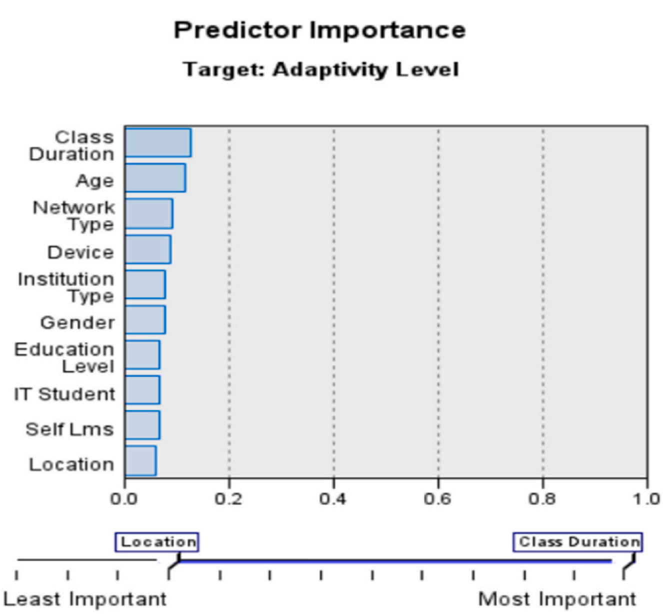


Fig. 3. Predictor Importance in Bayesian Network for the current problem.

The results are presented according to the performance metrics applied in each algorithm. Figs. 8–12 show the metrics of each model studied after being tested individually ten times and the standard deviation information.

Accuracy represents the proportion of correctly predicted instances over the total number of instances in the dataset. It gives a clear indication of how well the model is performing in terms of making correct predictions. High accuracy implies that the model is successfully capturing the underlying patterns in the data and producing accurate results. It enables comparison between different algorithms or variations of the same model. By assessing accuracy, practitioners can identify which models perform better and select the most suitable one for a given task or problem. In many real-world applications, the ultimate goal is to maximize accuracy. For example, in binary classification tasks,

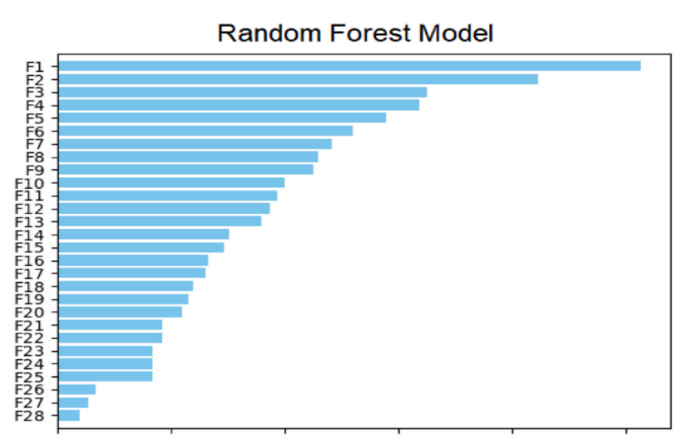


Fig. 4. Predictor Importance in Random Forest for the current problem.

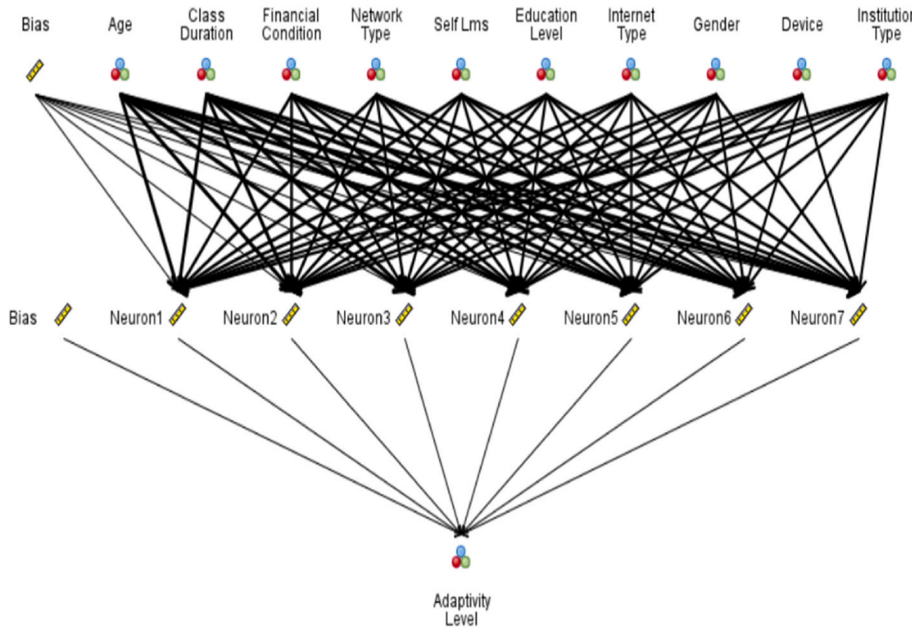


Fig. 5. Neural Network for prediction adaptively of online education.

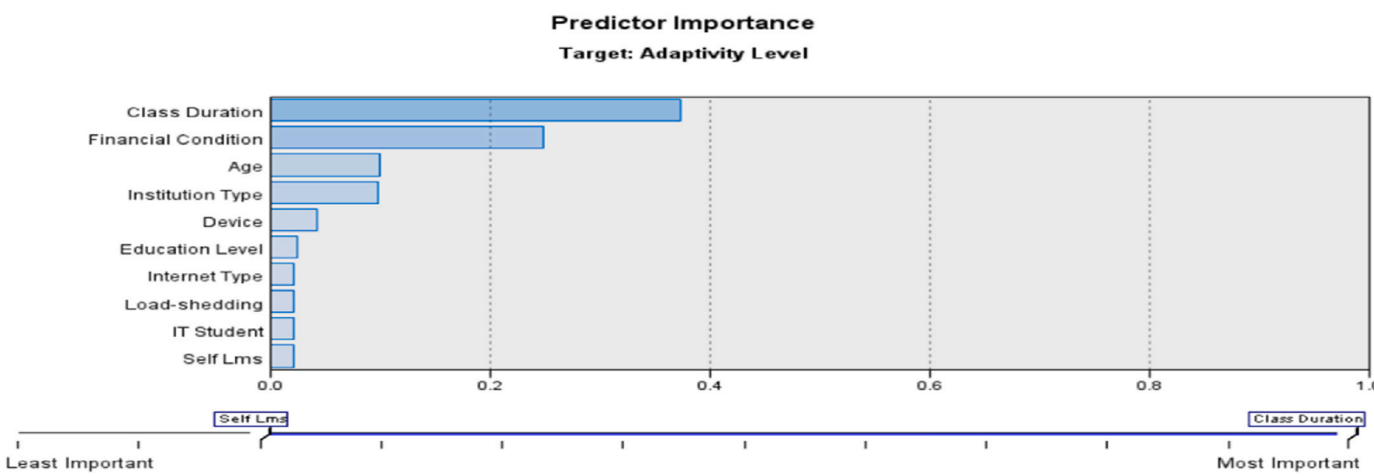


Fig. 6. Predictor Importance in CART Model for the current problem.

accurately identifying positive and negative instances is often the primary objective. Accuracy directly aligns with this objective, making it a key metric to focus on.

Table 1 depicts the performance analysis of the existing machine learning models with the proposed model. As the res.

Gains analysis provides additional insights beyond accuracy and helps in assessing the effectiveness of a model. Gains analysis is particularly useful when dealing with imbalanced datasets, where the distribution of positive and negative instances is uneven. Accuracy alone may not accurately reflect the model's performance in such cases. Gains analysis provides a clearer understanding of how well the model performs in predicting the positive instances compared to a random or baseline prediction. Gains analysis is commonly used in ranking tasks, such as search engine result pages or recommender systems. It helps evaluate the ability of the model to correctly order or rank the instances based on their relevance or importance. Gains curves allow for a visual representation of the model's performance in different sections of the ranking, providing insights into the effectiveness of the model across different positions.Fig. 13 demonstrates the graphically respresentation of the results achieved by various machine learning models.

The use of a modified ensemble machine learning model in forecasting students' adaptability in online entrepreneurship education offers several advantages:

4.1. Improved prediction accuracy

Ensemble models, such as Random Forest, Gradient Boosting, or Stacking, combine multiple base models to make predictions. This aggregation of diverse models helps to reduce bias and variance, leading to improved prediction accuracy compared to individual models. By leveraging the collective intelligence of the ensemble, the modified ensemble model can provide more robust and reliable forecasts of students' adaptability.

4.2. Handling complex relationships

Online entrepreneurship education involves various factors that influence students' adaptability, including demographics, academic background, online learning activities, and engagement metrics. The modified ensemble model can capture complex relationships and interactions among these factors. It can automatically learn and adapt to nonlinear patterns in the data, enabling more accurate predictions

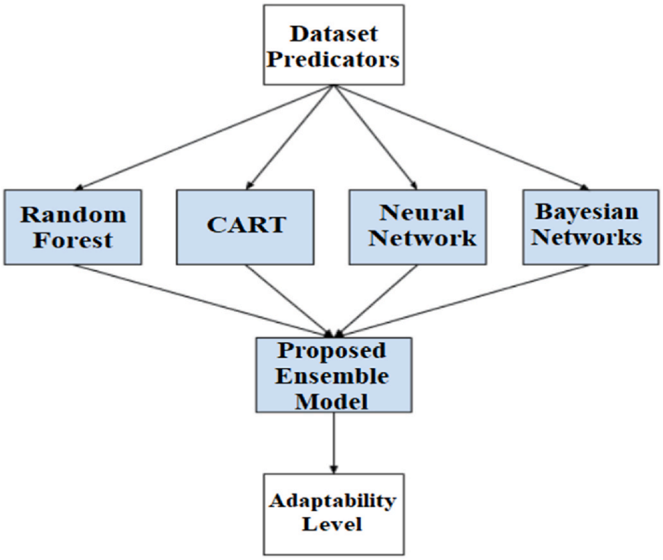


Fig. 7. Proposed Ensemble Model for a current problem.

Results for output field Adaptivity Level

Comparing $B-Adaptivity Level with Adaptivity Level

Correct	901	74.77%
Wrong	304	25.23%
Total	1,205	

Coincidence Matrix for $B-Adaptivity Level (rows show actuals)

	High	Low	Moderate
High	45	11	44
Low	7	324	149
Moderate	18	75	532

Performance Evaluation

High	2.047
Low	0.685
Moderate	0.347

User Defined Score for $B-Adaptivity Level

Mean	74.772
Sum	90100.0
Minimum	0.0
Maximum	100.0
Standard Deviation	43.45

Fig. 8. Metrics of Bayes Network for the current problem.

compared to traditional linear models.

4.3. Feature importance assessment

Ensemble models provide a measure of feature importance, indicating which factors have the most significant impact on predicting students' adaptability. This information can help educators and institutions understand the key determinants of adaptability and tailor their interventions accordingly. By identifying the most influential factors, they can focus on providing targeted support and resources to enhance students' adaptability in online entrepreneurship education.

4.4. Robustness to noisy data and outliers

Ensemble models are inherently robust to noisy data and outliers. By combining predictions from multiple models, the modified ensemble model can mitigate the influence of individual outliers or noisy instances, reducing their impact on the overall prediction. This robustness allows for more reliable adaptability forecasts, even in the presence of data imperfections.

Results for output field Adaptivity Level

Comparing $N-Adaptivity Level with Adaptivity Level

Correct	1,086	90.12%
Wrong	119	9.88%
Total	1,205	

Coincidence Matrix for $N-Adaptivity Level (rows show actuals)

	High	Low	Moderate
High	63	11	26
Low	0	452	28
Moderate	8	46	571

Performance Evaluation

High	2.37
Low	0.802
Moderate	0.566

User Defined Score for $N-Adaptivity Level

Mean	90.124
Sum	108600.0
Minimum	0.0
Maximum	100.0
Standard Deviation	29.846

Fig. 9. Metrics of Neural Network for the current problem.

Results for output field Adaptivity Level

Comparing $R-Adaptivity Level with Adaptivity Level

Correct	876	72.7%
Wrong	329	27.3%
Total	1,205	

Coincidence Matrix for $R-Adaptivity Level (rows show actuals)

	High	Low	Moderate
High	27	15	58
Low	0	357	123
Moderate	10	123	492

Performance Evaluation

High	2.174
Low	0.594
Moderate	0.343

User Defined Score for $R-Adaptivity Level

Mean	72.697
Sum	87600.0
Minimum	0.0
Maximum	100.0
Standard Deviation	44.57

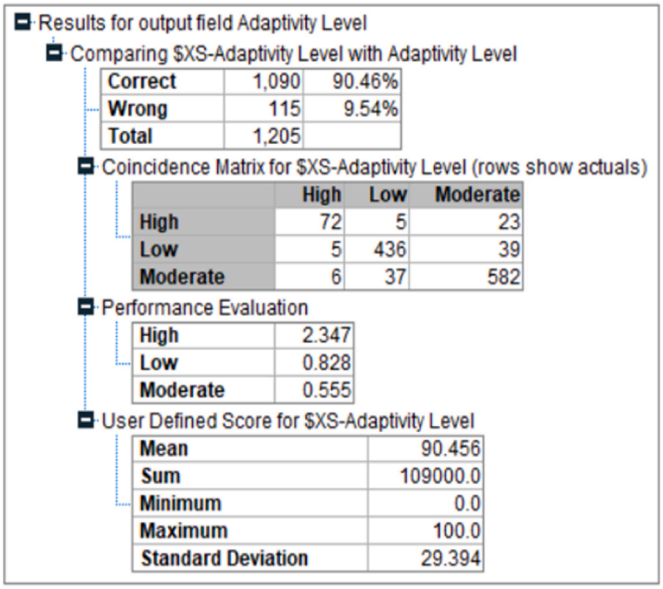
Fig. 10. Metrics of CART for the current problem.

4.5. Generalization capability

Ensemble models are known for their ability to generalize well to unseen data. The modified ensemble model, trained on a diverse set of base models, can effectively capture the underlying patterns and trends in the data. This generalization capability enables the model to make accurate predictions for new students entering online entrepreneurship education, ensuring the practical applicability of the forecasts.

4.6. Flexibility and adaptability

The modified ensemble model can accommodate various types of data, including categorical and numerical variables, as well as handle missing values and outliers. It is a versatile approach that can be customized to the specific requirements of the problem at hand. The flexibility and adaptability of the model make it suitable for forecasting



Results for output field Adaptivity Level

Comparing $XS-Adaptivity Level with Adaptivity Level

Correct	1,090	90.46%
Wrong	115	9.54%
Total	1,205	

Coincidence Matrix for $XS-Adaptivity Level (rows show actuals)

	High	Low	Moderate
High	72	5	23
Low	5	436	39
Moderate	6	37	582

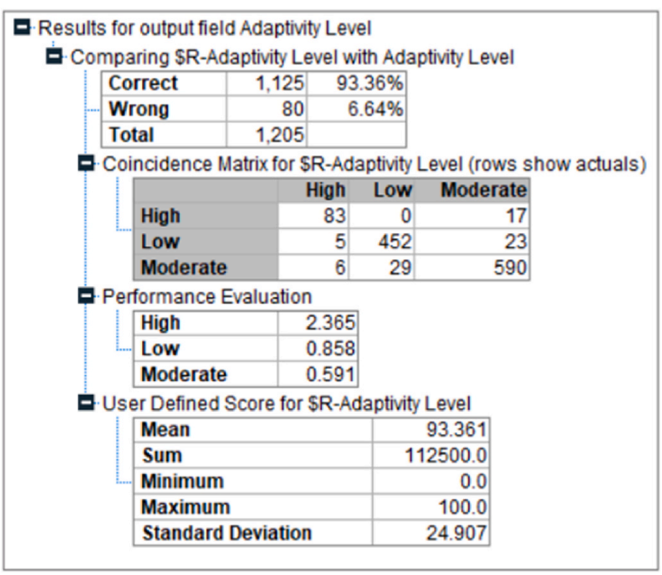
Performance Evaluation

High	2.347
Low	0.828
Moderate	0.555

User Defined Score for $XS-Adaptivity Level

Mean	90.456
Sum	109000.0
Minimum	0.0
Maximum	100.0
Standard Deviation	29.394

Fig. 11. Metrics of Random Forest for the current problem.



Results for output field Adaptivity Level

Comparing $R-Adaptivity Level with Adaptivity Level

Correct	1,125	93.36%
Wrong	80	6.64%
Total	1,205	

Coincidence Matrix for $R-Adaptivity Level (rows show actuals)

	High	Low	Moderate
High	83	0	17
Low	5	452	23
Moderate	6	29	590

Performance Evaluation

High	2.365
Low	0.858
Moderate	0.591

User Defined Score for $R-Adaptivity Level

Mean	93.361
Sum	112500.0
Minimum	0.0
Maximum	100.0
Standard Deviation	24.907

Fig. 12. Metrics of Proposed Ensemble Learning for the current problem.

adaptability in online entrepreneurship education, where the characteristics of the data can vary.

4.7. Scalability

Ensemble models can efficiently handle large datasets, making them suitable for forecasting adaptability for a large number of students in online entrepreneurship education. The model training and prediction processes can be parallelized, enabling scalability and faster processing times, even with extensive data.

By leveraging the advantages of a modified ensemble machine learning model, forecasting students' adaptability in online entrepreneurship education becomes more accurate, robust, and adaptable. These benefits can empower educators and institutions to better understand students' needs, design targeted interventions, and enhance the overall educational experience for aspiring entrepreneurs in the online realm.

Table 1
Performance comparison.

Model	Accuracy	Gains
Naive Bayes	74.71%	106.0
Random Forest	90.12%	190.0
CART	72.70%	218.0
Neural Network	90.46%	262.0
Proposed Model	93.96%	312.0

5. Discussion

In this work, five machine learning classifier models (Bayesian Network, Random Forest, CART, Neural Network. and their ensembles) were tested to predict online education adaptively. These techniques used the public database from Kaggle as an evaluation set, which contains multiple predictors like state, gender, and internet connection, as well as the financial status for classification. The experimental results obtained in the performance metrics among the proposed techniques were analyzed according to their means and standard deviation. The best results were obtained with a proposed ensemble learning method combining B.N., N.N., CART and Neural Network, which achieved slightly greater accuracy than the current literature—the proposed model is among the most accurately tested models in this research. The CART model was the one with the lowest performance. However, it is close to the one already found in scientific articles, and it was the method with the lowest standard deviation. The study's outcome consistently affirmed students' willingness to engage in online learning and entrepreneurship. The study also supports the assertion of existing studies on supportive technologies' role in enhancing entrepreneurship education and storing education data.

The forecasting of students' adaptability in online entrepreneurship education using a modified ensemble machine learning model has several practical implications:

5.1. Personalized support

The model can identify students who may struggle with online entrepreneurship education and require additional support. By forecasting adaptability, educators and institutions can provide personalized interventions and resources to help students overcome challenges and enhance their learning experience.

5.2. Targeted interventions

The model's predictions can guide the development of targeted interventions to improve students' adaptability. Educational institutions can design specific programs, workshops, or online resources to address the identified areas of weakness or low adaptability, enabling students to develop the necessary skills and mindset for success in online entrepreneurship education.

5.3. Early warning system

The model can serve as an early warning system, identifying students at risk of low adaptability before they experience significant difficulties. This allows for timely interventions and support systems to be put in place to prevent academic struggles, disengagement, or dropout.

5.4. Curriculum design

Insights from the model can inform the design and refinement of online entrepreneurship education curricula. The predictions can highlight areas where the curriculum may need adjustments to better align with students' adaptability needs and improve their overall learning outcomes.

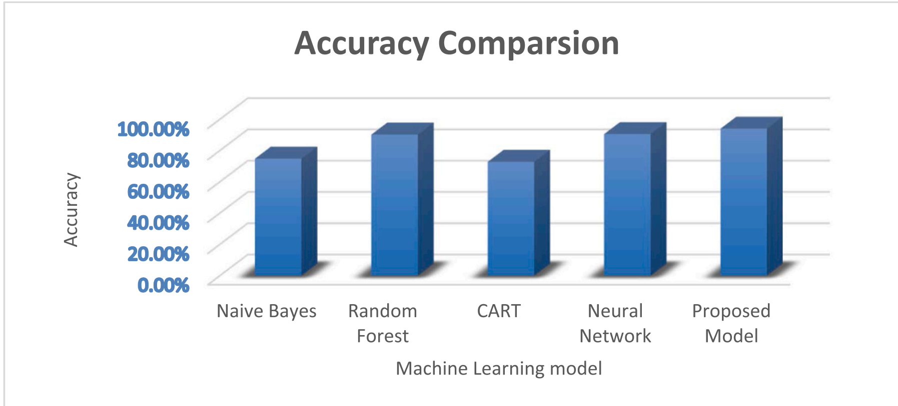


Fig. 13. Performance Comparison of various models.

5.5. Resource allocation

By identifying students with higher adaptability, institutions can allocate resources, such as mentorship programs or advanced coursework, to further enhance their learning experience and challenge them appropriately.

5.6. Program evaluation

The model's predictions can be used to evaluate the effectiveness of online entrepreneurship education programs. By comparing the predicted adaptability with actual outcomes and performance, institutions can assess the program's impact and make data-driven decisions for program improvement and resource allocation.

5.7. Enhancing student success

By understanding students' adaptability levels, educators and institutions can provide tailored support and guidance to enhance student success in online entrepreneurship education. This can contribute to higher retention rates, improved academic performance, and better overall student satisfaction.

5.8. Scalability and efficiency

The use of a modified ensemble machine learning model allows for scalability and efficiency in predicting students' adaptability. Once the model is trained and fine-tuned, it can be applied to new student data in real-time, providing quick and automated adaptability assessments for a large number of students.

It is important to note that while the modified ensemble machine learning model offers valuable insights and predictions, it should be used as a tool to support decision-making rather than the sole determinant of students' outcomes. Human expertise, contextual understanding, and a holistic approach to education remain crucial in providing comprehensive support to students in online entrepreneurship education.

6. Conclusions and future Scope

The results show that all students met the course goals to a high degree of satisfaction but that students who took the course in a classroom felt they had done a better job than those who took it online. Students were able to adapt online based on the results. The discrepancies can be traced partly to the educational techniques utilized in the online course. That is to say, it is not the delivery modality to blame for the difficulties students reported in the online study. Students' learning outcomes would likely be enhanced if the online course included pedagogical activities that required them to learn in action. It was discovered that students who elected to take the course online had an even greater desire to pursue a career as an entrepreneur due to their participation in the system overall.

Given that entrepreneurship education aims to help students find and realize their entrepreneurial potential, this point should not be neglected. The study would contribute to the growing knowledge of the use of technology in entrepreneurship. The future work will focus on developing a mobile platform to assist startups in marketing their products and connecting and interacting with global audiences.

CRediT authorship contribution statement

Amit Malik: Conceptualization, Methodology, Software. **Edeh Michael Onyema:** Data curation, Writing – original draft. **Surjeet Dalal:** Visualization, Investigation. **Umesh Kumar:** Supervision. **Darpan Anand:** Software. **Ashish Sharma:** Validation. **Sarita Simaiya:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Hasan Suzan M, Samrin NA, Biswas AA, Pramanik A. Students' adaptability level prediction in online education using machine learning approaches. 2021 12th International Conference on Computing communication and networking technologies (ICCCNT) 2021:1–7. https://doi.org/10.1109/ICCCNT51525.2021.9579741.

[2] QingsongTu, Liu Jian. Research on autonomous online education system based on intelligent recommenda-tion. 2011 IEEE International Symposium on I.T. In Medicine and education 2011:410–3. https://doi.org/10.1109/ITiME.2011.6132136.

[3] Martin Jour, Collie Andrew J, Rebecca J, Nagy, Robin P. Adaptability and high school students' online learn-ing during COVID-19: a job demands-resources perspective. Front Psychol 2021. https://doi.org/10.3389/fpsyg.2021.702163.

[4] El-Sabagh HA. Adaptive e-learning environment based on learning styles and its impact on development stu-dents' engagement. Int J EducTechnol High Educ 2021;18:53. https://doi.org/10.1186/s41239-021-00289-4.

[5] Onyema EM, Elhaj MAE, Bashir SG, et al. Evaluation of the performance of K-nearest neighbor algorithm in determining student learning styles. International Journal of Innovative Sci., Eng. &Techn. 2020;7(1):91–102.

[6] Igwe Paul Agu, Madichie Nnamdi O, Newbery Robert. Determinants of livelihood choices and arti-sanal entrepreneurship in Nigeria. Int J Entrepreneurial Behav Res 2019;25(4):674–97. ISSN 1355-2554.

[7] Hixson C, Paretti M, Lesko JJ. Work in progress: integrating entrepreneurship into undergraduate engi-neering education. Proc. - Front. Educ. Conf. FIE 2012:12–3. https://doi.org/10.1109/FIE.2012.6462399.

[8] Huang-Saad A, Morton C, Libarkin J. Unpacking the impact of engineering entrepreneurship education that leverages the Lean LaunchPad Curriculum. Proc. - Front. Educ. Conf. FIE 2016. https://doi.org/10.1109/FIE.2016.7757373. vol. 2016-November, no. June 2015.

[9] Liu T. Thinking of college students' innovation and entrepreneurship education under the background of big data. Proc. - 2021 6th Int. Conf. Smart Grid Electr. Autom. ICSGEA 2021 2021:546–50. https://doi.org/10.1109/ICSGEA53208.2021.00129.

[10] Meifeng H, Lin S. The feasibility and implementation strategy of integrating sustainable development con-cept into innovation and entrepreneurship education. Proc. - 2021 2nd Int. Conf. Educ. Knowl. Inf. Manag. ICEKIM 2021 2021:301–4. https://doi.org/10.1109/ICEKIM52309.2021.00073.

[11] Wang J, Chen C. The construction of 'The whole-course and personalized' comprehensive entrepreneurship education model. Proc. 2013 6th Int. Conf. Inf. Manag. Innov. Manag. Ind. Eng. ICIII 2013 2013;vol. 3:45–50. https://doi.org/10.1109/ICIII.2013.6703616.

[12] La Rubia MD, Rus-Casas C, Bueno-Rodriguez S, Aguilar-Pena JD, Eliche-Quesada D. Study of the en-trepreneurial attitudes of STEM students. IEEE Access 2021;9: 112005–18. https://doi.org/10.1109/ACCESS.2021.3103118.

[13] Kandakatla R, Aluvalu R, Devireddy S, Kulkarni N, Joshi G. Role of Indian higher education institutions towards aatmanirbhar India: government policies and initiatives to promote entrepreneurship and innovation. Proc. 2021 world Eng. Educ. Forum/global Eng. Deans Counc. WEEF/GEDC 2021 2021:8–14. https://doi.org/10.1109/WEEF/GEDC53299.2021.9657261. November.

[14] Lu L. Research on the optimization of the innovation and entrepreneurship education ecosystem of private uni-versities from the perspective of big data. Iccse 2021 - IEEE 16th Int. Conf. Comput. Sci. Educ. 2021:706–9. https://doi.org/10.1109/ICCSE51940.2021.9569497. no. Iccse.

[15] Su F, Chang J, Li X. Research on the evolution path and influence factors of core Enterprise oriented entrepreneurship ecosystem under the government regulation. IEEE Access 2021;9:90863–80. https://doi.org/10.1109/ACCESS.2021.3084342.

[16] Xu DY. Research on problems and countermeasures of college entrepreneurship education. 2011 IEEE 18th Int. Conf. Ind. Eng. Eng. Manag. IE EM 2011 2011: 191–4. https://doi.org/10.1109/IEEM.2011.6035137. no. PART 1.

[17] Wang F. Research on innovation and entrepreneurship education for college students from the perspective of big data. J. Phys. Conf. Ser. 2020;1693(1):701–4. https://doi.org/10.1088/1742-6596/1693/1/012014.

[18] Xueli H. Reform and practice on teaching method of introduction to innovation and entrepreneurship course in higher vocational education based on information technology. Proc. - 2021 Int. Conf. Educ. Inf. Manag. Serv. Sci. EIMSS 2021 2021: 574–7. https://doi.org/10.1109/EIMSS53851.2021.00129.

[19] Pardede E, Lyons J. Redesigning the assessment of an entrepreneurship course in an information technology degree program: embedding assessment for learning practices. IEEE Trans Educ 2012;55(4):566–72. https://doi.org/10.1109/TE.2012.2199757.

[20] Manapat JZ, De La Pena FB. Redefining philippine higher education through humanitarian engineer-ing, entrepreneurship, and design (HEED). GHTC 2018 - IEEE Glob. Humanit. Technol. Conf. Proc. 2019. https://doi.org/10.1109/GHTC.2018.8601919.

[21] Chen S, Zhang Y, Huang Q, Li X. Practical platform construction of computer professionals students training for innovation and Entrepreneurship. Proc. - 2020 Int. Conf. Big data Informatiz. Educ. ICBDIE 2020 2020:421–5. https://doi.org/10.1109/ICBDIE50010.2020.00105.

[22] Baruah B, Ward A, Jackson N. Online business simulation platforms for teaching entrepreneurship to en-gineering students in Higher Education. 29th Annu. Conf. Eur. Assoc. Educ. Electr. Inf. Eng. EAEEIE 2019 - Proc. 2019. https://doi.org/10.1109/EAEEIE46886.2019.9000424.

[23] Xu Y. On construction of entrepreneurship education community: the view from student satisfaction. 2011 IEEE 3rd Int. Conf. Commun. Softw. Networks, ICCSN 2011 2011:45–8. https://doi.org/10.1109/ICCSN.2011.6013772.

[24] Li Y. Multi-level of innovation and entrepreneurship educational system in China in case study of KTU 'three-stage creative and entrepreneurial educational system' and 'four innovations park' in the Yango Universi-ty. Proc. - 2020 Int. Symp. Adv. Informatics, Electron. Educ. ISAIEE 2020 2020:125–8. https://doi.org/10.1109/ISAIEE51769.2020.00035.

[25] Chen L, Ifenthaler D. Mobile learning and social media applied in entrepreneurship education courses in a higher vocational college. TALE 2021 - IEEE Int. Conf. Eng. Technol. Educ. Proc. 2021:1023–7. https://doi.org/10.1109/TALE52509.2021.9678928.

[26] Jin Z, Wu Y. Investigation and research on the present situation of innovation and entrepreneurship educa-tion for engineering postgraduates. Proc. - 2020 Int. Symp. Adv. Informatics, Electron. Educ. ISAIEE 2020 2020:110–4. https://doi.org/10.1109/ISAIEE51769.2020.00032.

[27] Bosman LB, Phillips M. Integrating the entrepreneurial mindset into the engineering classroom. IEEE Trans Educ 2021;65(2):1–6. https://doi.org/10.1109/te.2021.3101360.

[28] Siqueira ACO, Bruton GD. High-technology entrepreneurship in emerging economies: firm informality and contextualization of resource-based theory. IEEE Trans Eng Manag 2010;57(1):39–50. https://doi.org/10.1109/TEM.2009.2028323.

[29] Sarnin SS, Naim NF, Idris A. Entrepreneurship education at engineering faculty in Malaysia. Proc. 2019 IEEE 11th Int. Conf. Eng. Educ. ICEED 2019 2019;vol. 2019: 187–91. https://doi.org/10.1109/ICEED47294.2019.8994954.

[30] Lim YM, Lee WP, Lim TM. Entrepreneurial and commercialization pathway through project-based learning in higher-education. Tale 2019 - 2019 IEEE Int. Conf. Eng. Technol. Educ. 2019. https://doi.org/10.1109/TALE48000.2019.9225964.

[31] Kunicina N, Bilic I, Zabasta A, Caiko J, Ribickis L. Development of entrepreneurship skills for students creative thinking support in higher education. 2019 Int. Conf. Eng. Appl. ICEA 2019 - Proc. 2019. https://doi.org/10.1109/CEAP.2019.8883450.