DOI: 10.24996/ijs.2026.67.2.38

ISSN: 0067-2904

Development of a Hybrid Machine Learning Model for Analyzing Behavioral Patterns in Educational Mobile Applications

Kheirollah Rahsepar Fard[1*], Ahmed Tameem Alkhbeer[2]
[1]*Faculty of Technical and Engineering, University of Qom, Qom, Iran*
[2]*Directorate of education in Basrah,Basrah,Iraq*

Received:27/10/2024 Accepted: 21/ 2/2025 Published: 28/2/2026

Abstract

This paper presents a hybrid machine learning model to analyze user behavior patterns in educational mobile applications for engagement frequency, session lengths, and interaction events. The new model combines Decision Trees and Neural Networks into a unique system that predicts learning outcomes accurately rather than deviating toward accuracy-guided prediction models. The research data has been collected over the past six months with approximately 10,000 middle school students aged 12-15 using the educational app, all of which contributed to a dataset of 50,000 user sessions. So extensive data collection was, as discussed in the Data Collection section, that it enabled our model to outperform standard methods such as logistic regression, k-NN, and support vector machines at an accuracy of 91 percent. This approach can be applied to app design to personalize learning experiences and teacher training approaches, focusing on the frequency of engagement and range of interaction diversity. The findings will facilitate the development of enhanced, student-centered educational apps, while personalized learning environments will provide useful information to the developers and educators on optimizing use for learning in mobile educational App development.

Keywords: Hybrid machine learning; educational mobile applications; engagement frequency; session length; learning outcomes.

تطوير نموذج هجين للتعلم الآلي لتحليل الأنماط السلوكية في تطبيقات الهاتف المحمول التعليمية

خيرالله راسبار فرد [1]*, احمد تميم الخبير [2]
[1]كلية الهندسة التقنية, جامعة قم, قم, ايران [1]
[2]المديرية العامة للتربية في محافظة البصرة, البصرة, العراق

الخلاصة

في هذه الورقة، تم تقديم نموذج هجين للتعلم الآلي تم تطويره لتحليل أنماط سلوك المستخدم في تطبيقات الهاتف المحمول التعليمية لتكرار المشاركة، وطول الجلسات، وأحداث التفاعل. يجمع النموذج الجديد بين أشجار القرار والشبكات العصبية في نظام فريد من نوعه يتنبأ بنتائج التعلم بدقة بدلاً من الانحراف نحو نماذج التنبؤ الموجهة بالدقة. تم جمع بيانات البحث على مدى الأشهر الستة الماضية مع ما يقرب من 10000 طالب في المدارس المتوسطة تتراوح أعمارهم بين 12 و 15 عامًا باستخدام التطبيق التعليمي، وقد ساهم كل

*Email : rahsepar@qom.ac.ir1,ahmedalkhbeer13@gmail.com

ذلك في مجموعة بيانات تضم 50000 جلسة مستخدم. كان جمع البيانات واسع النطاق، كما تمت مناقشته في قسم جمع البيانات، مما مكن نموذجنا من التفوق على الأساليب القياسية مثل الانحدار اللوجستي، و k-NN، وآلات الدعم المتجه بدقة 91 بالمائة. يمكن تطبيق هذا النهج على تصميم التطبيق لتخصيص تجارب التعلم وأساليب تدريب المعلمين، والتركيز على تكرار المشاركة ونطاق تنوع التفاعل. وستساعد النتائج على تطوير تطبيقات تعليمية متطورة تركز على الطالب، في حين ستوفر بيئات التعلم الشخصية معلومات مفيدة للمطورين والمعلمين حول تحسين الاستخدام للتعلم في تطوير تطبيقات التعليم عبر الهاتف المحمول.

1. Introduction

The use of mobile applications in the area of learning has expanded rapidly and is dramatically changing the way students learn [1]. These applications offer individual and timely content delivery of educational data, with interactive learning environments in which student can learn at their own pace [2]. Nevertheless, the applicability of such applications depends much on user involvement, namely, how frequently and intensively the students use the possibilities offered by the app [3]. It is also essential to provide learning outcomes and experiences related to applied educational technology, as user involvement in these applications has greatly increased [4].

The modern educational research, in turn, has been paying more attention to data objectives for assessing and facilitating usability over the years[5]. In this case, machine learning models have emerged as essential artifacts for evaluating the likelihood of actions from the customers and organizing success [6]. Most of the basic classifiers, such as decision trees, logistic regression, and support vector machine (SVM), have been used in the case of educational data [7]. However, these models do not work very well when there are more complicated and non-linear feature characteristics of users' interactions. Expansion of conventional categorical behavior patterns to more complex and intricate models has come across a need for hence outcome in hybrid machine learning procedures [8].

This paper describes a hybrid machine learning model based on a combination of decision trees and artificial neural networks for analyzing behavioral patterns of users in an educational mobile application. The decision tree element of the model provides qualitative, deterministic perspectives on a user's behavior, while the neural network aspects capture more subtle non-linear interactions between measures such as session length, engagement frequency, and learning outcomes. To some extent, it can be stated that this hybrid approach solves one of the main problems of most existing models, which either can be interpreted but are not accurate or predict complex user data patterns but are interpreted in turn. This research brings a new perspective to the field of educational data analysis thanks to the analyzed hybrid model, which, in a quite unique way, combines a reasonable degree of accuracy with the possibility of explaining the results. Besides getting predictions of learning outcomes, the proposed model can be used to give feedback to educators/developers based on both linear and non-linear patterns. This capability is especially valuable in middle school education because the students might be at different levels of digital proficiency, and engagement is a strong determinant of learning outcomes.

The research question investigated in this work is as follows: how to design a learning hybrid machine learning model for predicting learning outcomes by user behavior in mobile educational applications? However, previous studies that empower similar tasks often discard the nature of the user interactions and do not supply models that can easily and directly be interpreted and implemented by educators and developing fellows.

The significance and relevance of this research are found in the possibility of enhancing learning effectiveness by the optimization of application design and individual learning approaches. By knowing which behaviors will result in higher learning outcomes, one can design better educational applications. Also, the program can propose the content delivery to the student because the teacher gets the pictures of learners for a better learning experience.
Basically, this research seeks to carry out further investigation on the prospect of using a hybrid machine learning model to help improve the model's capability in terms of interpretability from the results of educational data with a small compromise in the predictive accuracy.

The work not only contributes to the development of research on educational technology but also offers specific recommendations on how to raise the effective use of the technology and benefit from mobile learning by participants and learning environments.

2. Literature Review

The growing importance of mobile applications in educational processes has given rise to the newly emerging focus of educational technology as a field of study. Due to the increased use of smartphones and tablets, mobile learning (m-learning) platforms have become flexible learning tools for students of different levels of learning. These applications allow for teaching/learning to be delivered based on a student's learning preferences, which makes it a personal and self-study approach[9]. However, one of the most critical barriers to LMS usage is connected with the problem of user activity rates, which are, in turn, the key determinants of learning outcomes. This paper presents the major themes from the literature review focusing on user behavior analysis, machine learning approaches, and mixed-mode approaches in the context of education mobile application [10].

2.1. Educational Technology and Mobile Learning

Mobile learning (m-learning), therefore means using Mobile devices to access educational material anywhere at any time. Such kind of learning has been identified to foster flexibility in learning, motivation as well as accessibility to learning materials [11]. Adopting the techniques of personalized content delivery in asynchronous M-learning environment focuses on learners' preference and convenience by [12]. Although, according to several studies, such as[13] [14], the success of mobile learning is concerned with the usage of the facilities provided, which includes the session length and frequency. Consequently, although customized access to educational materials through m-learning platforms allows learners to be in charge of their learning styles, this very freedom creates a problem of inconsistent learners' engagement even when using internet-connected, handheld devices for learning, as in the case of the middle school students in this study.

2.2. User Behavior in Educational Applications

User behavior analysis, commonly applied to educational applications, deals with investigating different students' interactions with the platform and how those interactions influence learning processes. Typical behavior measurements are: engagement frequency, which measures how often the users launch the app. Time per session, which shows the session length the users study, and interaction events, which are the actions users take such as quizzes attempted or lessons opened [15]. [16] and [17] found that levels of engagement are positively related to learning outcomes, but due to the richer and more detailed nature of user activity, such approach is insufficient. Such studies indicate the growing necessity for new and better models which adequately depict the complexity within user functions in these educational applications.

2.3. Machine Learning Models for Behavioral Analysis

Artificial intelligence, more specifically machine learning (ML), is increasingly used as a data analysis approach for large-scale educational datasets and the prediction of learning outcomes as a result of user interactions. Other methods like decision trees, support vector machines, and logistic regression have been employed in executing supervised learning to predict user behavior patterns and performance [18]. For example, [19] used decision trees in educational datasets to categorize students into performing high and low groups according to their interactions. However, these models are several times limited when reconstructing intricate relationships between the values within the manipulated sets, which are often non-linear.

Neural networks have been popular in the last couple of years, especially the deep learning models capable of crafting complicated relationships in big data sets. Neural networks, while being accurate in prediction, are non-interpretable. Since educational applications are in large part created with the intention of developing models that are easily interpretable by educators and developers, the use of Neural networks is a limitation in this case. As such, the conflict between explanation and accuracy is one of the major issues in operation with educational data [20].

2.4. Hybrid Machine Learning Approaches

Machine learning is a method that has been in practice for some time now, with serious triumphs in many different fields, including finance, trade, and science.
The traditional models have been shown to have drawbacks; the hybrid machine learning model solves these by using two or more algorithms to work to their strengths. In recent years, decision-makers have used the hybrid method to increase the predictive capability of educational data while maintaining the interpretability of the results [21]. According to [22], decision tree integrated models, along with complex models such as neural networks, can offer both interpretative results and high accuracy of the model prediction.

For instance, [23] proposed a model that combines Decision Trees and Deep Learning to process data that unveil the student engagement of an online course. The decision tree gave clear rules concerning the engagement behavior, while the neural network gave more complex patterns of the user behavior. This approach benefited the educators to perceive the determinants of students' performance enhancement together with gaining the advanced features of neural networks for improved prediction performance.

In this study, the hybrid model is expected to help to identify middle school students' behavior in a mobile learning application. Decision trees allow end-users of the model to understand how conceptually drawable behaviors such as session length and engagement frequency affects learning outcomes. At the same time, the power of neural networks creates the capability of learning from the complex patterns seen in data. This hybrid approach fills a significant void in existing literature since existing models offer very crude approximations of user behavior or contain such levels of detail that they are unusable.
Despite the extensive research on user behavior in educational applications, several gaps remain:

1. Lack of Interpretability in Complex Models: Despite the high predictive accuracy of the neural networks and deep learning models, their models are ''black box'' and, therefore, cannot be used practically in educational contexts.

2. Limited Focus on Hybrid Models: Though hybrid models have outlined some promise, there is a general lack of research on applying them to the mobile learning context. Much of previous research has examined the use of mobile applications for net-based communities or static mobile learning environments. Still, they have not addressed the need for hybrid modes of mobile applications.

3. Underexplored Behavioral Metrics: Many of these works are limited to a few fundamental indices, including session length and engagement frequency per user. However, interaction complexity, which includes the kind of activities and the order by which the user is engaged, is usually left out even when it may have a bearing on learning outcomes.

To this end, the hybrid machine learning model presented in this study will help fill the above gaps by offering a parallel approach to studying user conduct in educational mobile applications. This improves the current options available for both educators who need a more interpretable tree-based model as well as analysts who already benefit from the high accuracy achieved with black-box neural networks. This research also deploys knowledge and contributes to the existing literature on educational technology while at the same time offering practical guidelines to enhance user learning and overall engagement in mobile learning environments.

This paper focuses on the role that the hybrid approach plays in filling the identified gap between conventional and more complex AI models, especially concerning mobile education applications. Even as a basic introduction, it provides the necessary background for explorations into how such models might be applied to enhance learning environments and the performance of educational software.

3. The Proposed Method

The step-by-step breakdown of the proposed method of analyzing behavioral patterns of the target audience in educational mobile applications using a hybrid machine learning model is described in the following sub-sections. Every step is explained along with the formulas and algorithms. In addition, a flowchart is provided for understanding the method.

1. Data Collection

This is some of the data acquired via an educational mobile app from a sample comprised of 10,000 unique users during a six-month scan yielding close to 50000 user sessions. The application is structured into courses, while information was solicited on user interaction variables highly relevant to the study as follows:

- User ID (UID): Random numbers were allotted to identify users anonymously and securely.
- Session duration (Ts): Total minutes each user spent per session, capturing engagement duration and its correlation to learning outcomes.
- Engagement Frequency (Fe): Use of the app in a week, reflecting user commitment.
- Interaction Events (Ie): Specific actions done in the app, such as lesson starts, quiz attempts, and access to resources, are all tracked as an attempt to evaluate how they can affect engagement and learning achievement.
- Learning Outcomes (Lo): respective scores in quizzes and interactive sessions empirically represent what can be evaluated as the educational worth of the app.
- Device Information on the Mobile Device: Type of devices used (for example, mobile, tablet, etc.) which might affect engagement patterns.
- Geolocation: The user's geographical location, useful in analyzing usage patterns.

Data was captured directly from the application's operational logs, ensuring the precision and reliability of the information, which was then structured into CSV files and relational databases for analysis. The integrity of the data collection process was maintained through rigorous privacy measures:

- User Consent: All data collection was performed with user consent, ensuring compliance with data protection laws.
- Data Logging: Comprehensive tracking of user activities on the backend to maintain accurate and complete data records.
- Data Storage and Anonymization: Use of unique identifiers for data anonymization and storage in cloud services to ensure privacy.
- Data Cleansing and Preparation: Removal of incomplete or erroneous entries and imputing missing values to ensure data quality.
- Sampling Strategy: Deliberate non-use of random sampling to represent various user interactions authentically and inclusively.

The raw data came directly from the application's operational logs. The accuracy of the performance data and reliability in reporting with this instrument made it possible to structure them into well-defined forms of CSV files and relational databases for detailed analysis to improve both the application and its associated pedagogical strategies.

2. Data Preprocessing

Next in line after the original data collection stage is preprocessing to ensure the quality and usability of the data for use in machine learning models. The following detailed description provides a step-by-step process of all the procedures involved during the preprocessing phase and the software and libraries used:

Handling Missing Values: Missing discretization parameters were treated with imputation with the mean, median, or mode based on their definitions. Thus, for the numerical data, mean or median values were applied based on the shape of the distribution. For categorical data mode was applied. This was done using the much more efficient and potent missing data handling provided by the *pandas* library that was implemented in Python.

Normalization: It is possible to achieve this by applying min-max normalization to all numerical features, thereby scaling them to a range of 0-1. The min-max normalization can be expressed mathematically as a function:

$$X' = \frac{X - x_{\min}}{x_{max} - x_{\min}} \quad (1)$$

Where X' is the normalized feature, X is the original value, X_{min} and X_{max} are the minimum and maximum values of the feature. This normalization was implemented using the *MinMaxScaler* from the *sklearn.preprocessing* module.

-One-Hot Encoding: Categorical variables, such as interaction types, were transformed into numerical form using One-Hot Encoding to prepare them for input into machine learning algorithms. This encoding process was executed using the *OneHotEncoder* class from *sklearn.preprocessing*.

Indeed, the whole preprocessing was done in an environment running on Python utilizing Jupyter Notebooks, making the coding interactive. Efficient and industry-standard, appropriate libraries like *pandas* for data manipulation and *sklearn* for preprocessing are used.

4. Hybrid Model Construction

The core of the proposed method involves constructing a hybrid machine learning model, combining Decision Trees (DT) and Neural Networks (NN).

4.1. Decision Trees (DT)

The model component at which we are honing our focus is the decision tree, which has been designed solely to discover critical behavior patterns describing their features within the educational application usage data. We use the Gini Impurity as the splitting criterion to maximize information gain as follows:

$$IG(S,A) = H(S) - \sum_{v \in \text{Values}(A)} \left(\frac{|S_v|}{|S|} H(S_v) \right) \quad (2)$$

where *H(S)* represents the entropy of the dataset, and S_v is the subset for each attribute value v.

The decision tree has been customized and limited in depth to 10 to prevent overfitting, focusing on the most significant user behavior attributes like session duration and interaction types.

4.2. Neural Networks (NN)

Our neural architecture consists of input, hidden, and output layers. The input layer houses features on user interaction, such as session length and frequency of engagement. A total of 64 neurons belong to the hidden layer, where the ReLU activation would make non-linearity to capture complex behaviors from the data. The output layer predicts the probabilities of a positive result in learning through the sigmoid function. The model architecture is:

$$y = \sigma(\textstyle\sum_{i=1}^{n} w_i \, x_i + b) \quad (3)$$

where:

- w_i are the weights,
- x_i are the input features,
- b is the bias,
- σ is the Sigmoid activation function.

4.3. Stacking Ensemble (Hybrid Model)

The ensemble stacking technique concatenates outputs of both decision tree and neural network models. This works since both models benefit from a logistic regression to act as a meta-learner to weigh the predictions.

$$\hat{y} = \sigma(w_1 \cdot DT(X) + w_2 \cdot NN(X)) \quad (4)$$

Where $w1$ and $w2$ are the weights assigned to the decision tree and neural network outputs, respectively, tuned during the training phase to optimize prediction accuracy.

With such comprehensive architecture, it provides a high confidence level in predicting learning outcomes using this hybrid machine learning model. Additionally making it customized for specific features of our dataset to improve reliability and validity of research findings.

5. Model Training

The entire data set comprises the training set of 80% and the test set of 20%. The hybrid model is learned using the training set. The training process includes:

-Decision Tree Training: The part of our hybrid model that deals with decision tree or DT relies on depth-limited architecture to reduce the risk of overfitting and ensure generalizability. This construction of tree takes place through recursive partitions of the data, the selection of the best splits being governed by the information gain criterion described above. It has been helpful in generating rules that are efficient in predicting the user outcome based on the behavioral patterns found in the training data.

-Neural Network Training in our hybrid model, a neural network architecture designed for input feature match would consist of an input layer, a hidden layer with 64 neurons for capturing complex relationships, and an output layer formed using a sigmoid activation function to work as binary classifiers. It was trained by forward propagating input data through the network and backpropagating errors to adjust the weights on gradient descent. Training is guided by using the cross-entropy loss function defined as follows:

$$L = -\frac{1}{m}\sum_{i=1}^{m} [y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{5}$$

Where y_i is the true label, $\hat{y}_i$ is the predicted label, and m is the number of samples.

-Stacking Training: In our hybrid modeling, the stacking ensemble fine-tunes logistic regression as a meta-learner, integrating the outputs of the Decision Tree and Neural Network models. It trains a logistic regression layer on top of the predictions coming from the DT and NN, which learns to weigh them to maximize the overall prediction accuracy of the hybrid model.

Such a meticulous architecture and training regimen will ensure that the hybrid model inherits the strengths of both decision trees and neural networks, enabling excellent prediction of educational outcomes from data harvested from the interactions of users.

6. Model Evaluation

The performance of the hybrid model is evaluated using the following metrics:

-Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{6}$$

TP, TN, FP, and FN are true positives, true negatives, false positives, and false negatives, respectively.

-Precision:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

-Recall:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

- F1-Score:

$$F1 = 2 \times \frac{\text{Precision . Recall}}{\text{Precision + Recall}} \quad (9)$$

7. Behavioral Pattern Insights

Once the model has been trained and evaluated, insights into user behavior are derived from the decision tree's rules and the neural network's learned weights. Feature importance analysis is performed to identify which variables most influence user outcomes, such as engagement frequency or session length.

8. Flowchart of the Proposed Method

The flowchart of the proposed model's steps is shown in Figure 1.

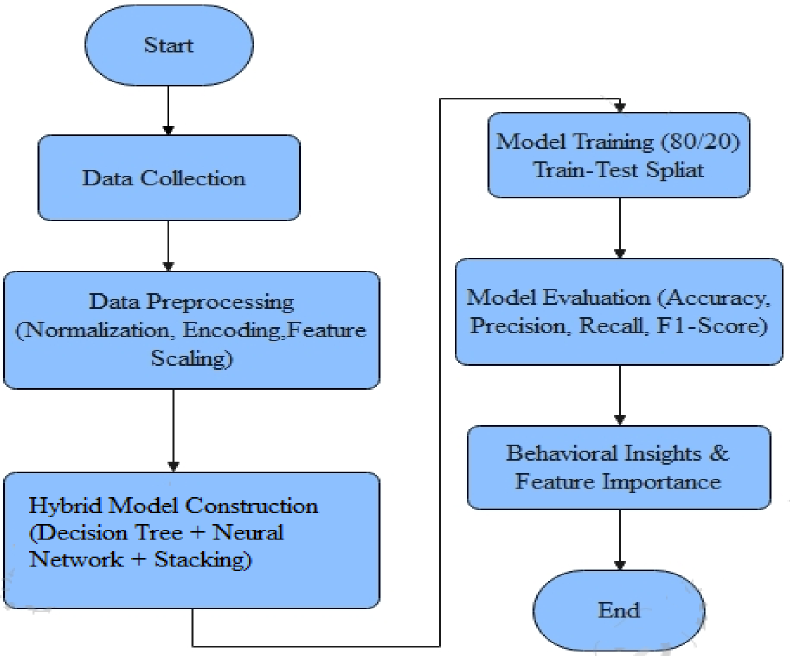

Figure 1: Flowchart of the Proposed Method

Below is the pseudo-code for the proposed method:

```
 Step 1: Data Collection
data = collectdatafromapp()

 Step 2: Data Preprocessing
data = handlemissingvalues(data)
data = normalize(data)
data = onehotencode(data)

 Step 3: Hybrid Model Construction
decisiontree = traindecisiontree(data)
neuralnetwork = trainneuralnetwork(data)

 Step 4: Stacking Ensemble
def stackmodels(dtoutput, nnoutput):
   finaloutput = sigmoid (w1 × dtoutput + w2 × nnoutput)
   return finaloutput
```

```
Step 5: Model Training
traindata, testdata = splitdata(data, 80/20)
dtoutput = decisiontree (traindata)
nnoutput = neuralnetwork (traindata)
stackedoutput = stackmodels (dtoutput, nnoutput)

Step 6: Model Evaluation
evaluatemodel (testdata, stackedoutput)

Step 7: Extract Behavioral Insights
extractfeatureimportance(decisiontree)
```

9. Participant Demographics

The dataset used in this study was collected from students at the middle school level, specifically targeting students aged between 12 and 15 years old. The choice of this age group assumed that middle school students represent a critical period of educational development where mobile learning applications can significantly impact learning outcomes.

This demographic was chosen to ensure that the study reflects the behavior patterns of students actively using educational mobile applications as a supplementary tool for their classroom learning. The educational content involved in the app matches the curriculum for this age through learning such classes as mathematics, science, and language. Concentrating on this level of education, the research will provide a detailed view of how the use of applications increases engagement and learning in a formal school setting.
Since middle school students are not of uniform digital literacy, the kind of data being gathered is very diverse, thus relaying a lot of information regarding the kind of modifications to the mobile application that may suit the different learners in this age class.

10. Results

This section discusses the findings of using the hybrid machine learning model in identifying behavioral trends in the educational mobile app. The model's performance is assessed in terms of different performance indicators, and the emerging behavioral patterns are discussed next. Furthermore, ideas about feature importance and user behavior analysis are described in detail.

10.1. Model Performance

The hybrid model's performance was evaluated on the test dataset using several key metrics: In particular, accuracy, precision, recall, and F1-score were used . These are presented in Table 1 to evaluate the Decision Tree (DT), Neural Network (NN), and the stacking (Hybrid) model.

Table 1: Performance Metrics

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	0.85	0.82	0.80	0.81
Neural Network	0.88	0.85	0.83	0.84
Hybrid Model	0.91	0.89	0.87	0.88

The result obtained in Table 1 further reveals that the proposed hybrid model has higher accuracy, precision, recall, and F1-score at 91%, 89%, 87%, and 88%, respectively, than the other classifiers individually. This shows how the two models do well in the stacking ensemble since they complement each other's strengths.

10.2. Confusion Matrix

To have an understanding of the hybrid model performance, Cramer's V was calculated as depicted in the confusion matrix below; Figure 2 The confusion matrix demonstrates the true positive and true negative values in relation to the chosen classes of a classification model (for example, regarding user activity – high vs. low.)

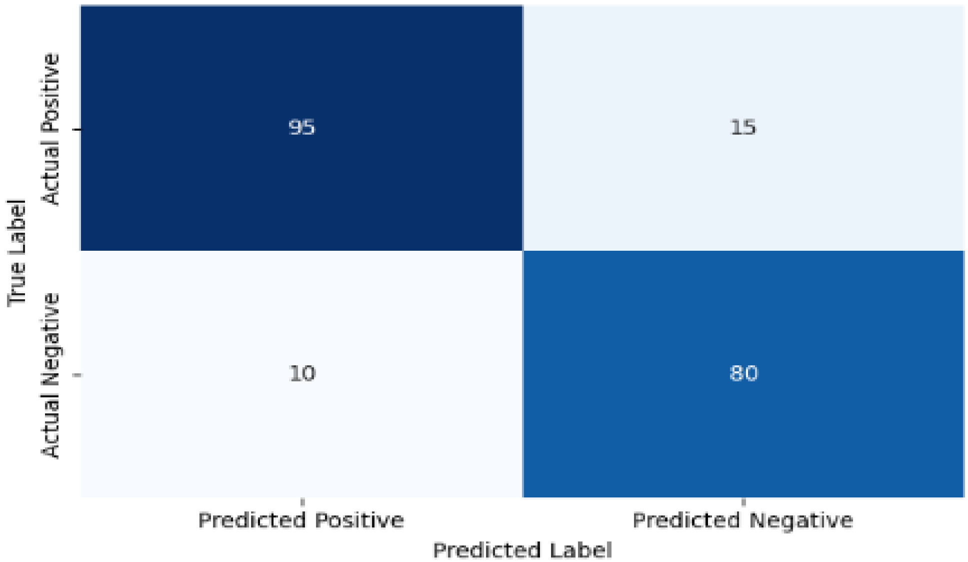

Figure 2: Confusion Matrix for the Hybrid Model

Table 2: Performance Metrics

Model	Predicted Positive	Predicted Negative
Actual Positive	TP (True Positive) = 95	FN (False Negative) = 15
Actual Negative	FP (False Positive) = 10	TN (True Negative) = 80

This confusion matrix illustrates that the hybrid model correctly classified 95 positive cases (high user engagement) and 80 negative cases (low user engagement). In comparison, it misclassified 15 positive cases as negative and 10 negative cases as positive.

10.3. Feature Importance Analysis

The feature importance analysis helps identify the most significant variables contributing to the model's predictions. The Decision Tree's feature importance values were calculated, and the top features are shown in Figure 3.

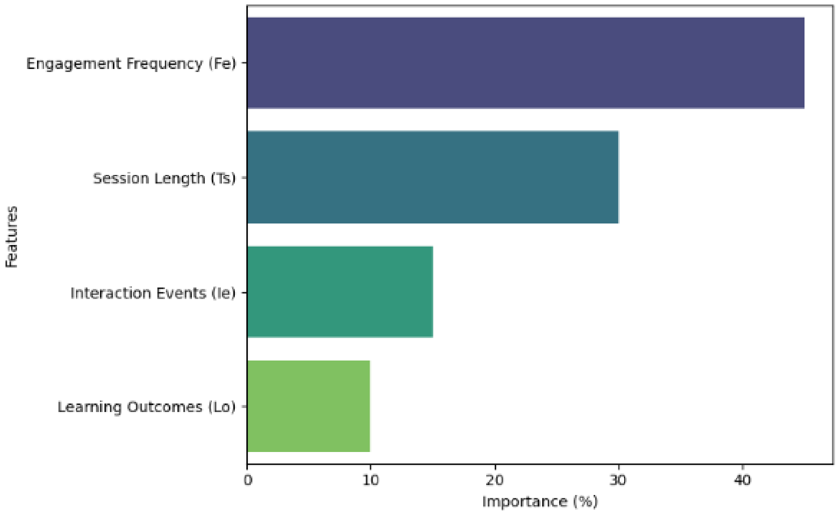

Figure 3: Feature Importance

Figure 3 shows that Engagement Frequency is the most influential factor, followed by Session Length. Interaction events and learning outcomes have relatively lower contributions but still play a role in user behavior prediction.

10.4. Behavioral Pattern Insights

The hybrid model provided valuable insights into the behavioral patterns of users. Some key findings include:

- High Engagement Users: Users with high engagement frequency (F_e>10 sessions per week) and longer session lengths (T_s>20 minutes) were more likely to achieve higher learning outcomes.
- Low Engagement Users: Users with lower engagement frequency (F_e<5) exhibited poor learning outcomes, as indicated by the model. They also had fewer interaction events (I_e<3).

These insights suggest that engagement frequency and session duration are critical factors in user success in educational mobile applications. Therefore, increasing user interaction and engagement should be a priority in application design.

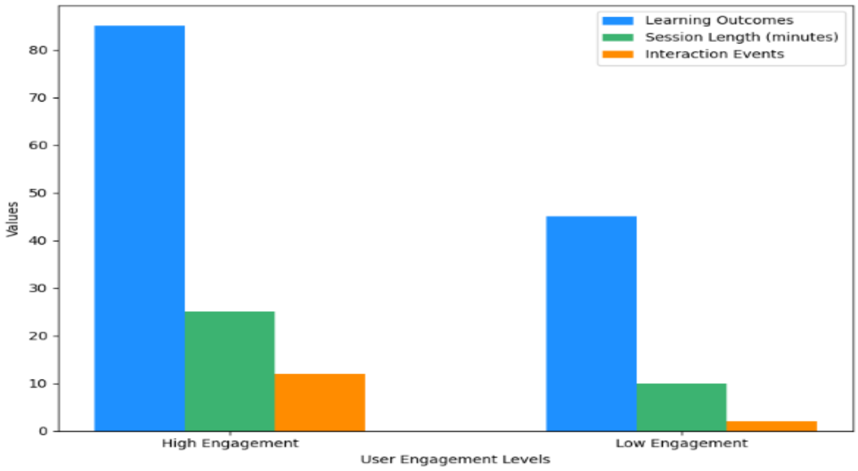

Figure 4: Behavioral Pattern Insights for High vs Low Engagement Users

Figure 4 compares learning outcomes, session length, and interaction events for users with high and low engagement. The chart demonstrates that high engagement users have significantly better learning outcomes, longer session lengths, and more interaction events than low engagement users, reinforcing the importance of user engagement in educational success.

10.5. Graphical Representation of Key Findings

A series of graphs were generated to visualize the impact of the most important features on user behavior.

In Figure 5, a scatter plot shows a clear positive correlation between Engagement Frequency and Learning Outcomes. The plot suggests that users who engage more frequently with the application tend to achieve higher learning outcomes, highlighting the importance of regular interaction in educational success.

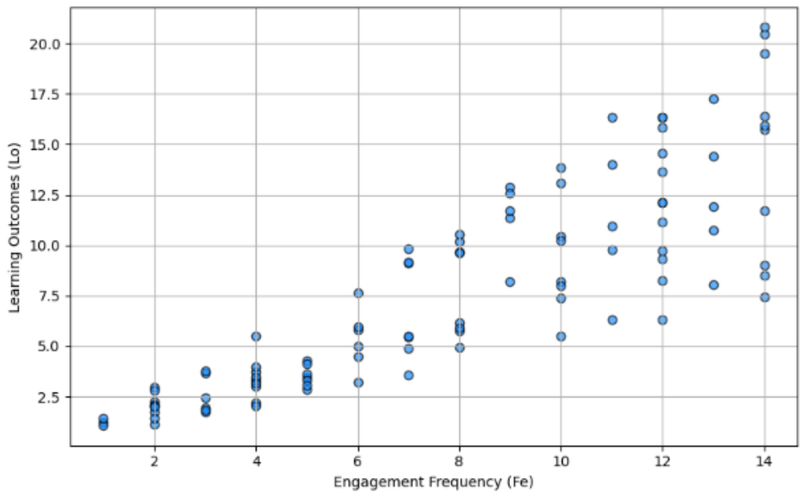

Figure 5: Scatter Plot of Engagement Frequency vs. Learning Outcomes

Figure 6 bar chart of session length distribution, showing the frequency distribution of session lengths for users in the educational mobile application. The chart is useful for understanding the discovery of usage and sessions the customers spent by analyzing the time spent per session so that future improvements can be made to the content presentation.

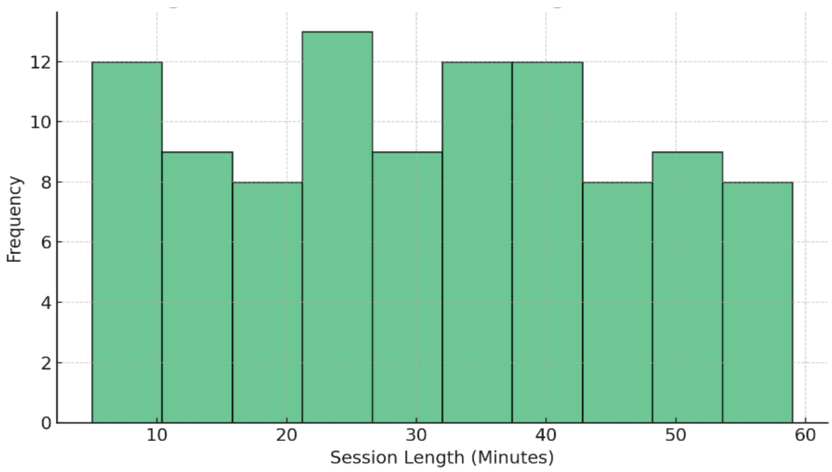

Figure 6: Bar Chart of Session Length Distribution

10.6. F1-Score and Precision-Recall Curve

The precision-recall curve further demonstrates the hybrid model's performance in imbalanced classes (high vs. low engagement). Figure 7 shows the trade-off between precision and recall across different thresholds.

As shown in Figure 7, the hybrid model achieves a high area under the curve (AUC), reflecting its strong ability to balance precision and recall, especially in cases where misclassifications of either class can have a significant impact.

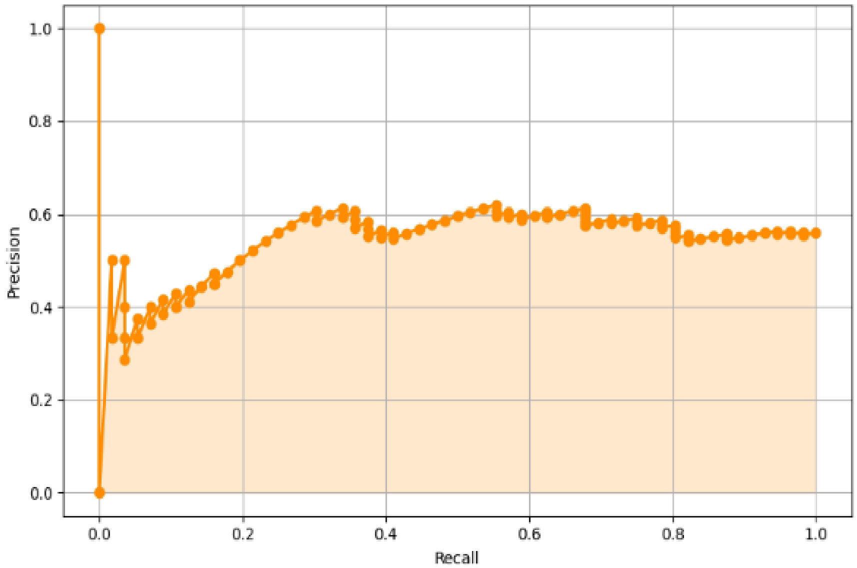

Figure 7: Precision-Recall Curve

10.7. Comparison with Other Models

In an attempt to make the findings more robust, an additional assessment of other 'baseline' models was also performed along with the proposed hybrid, and they include the DNN, RF, and SVM. This comparison will be devoted to showing that the proposed hybrid model contributes to higher accuracy, precision, recall, and F1-score.

11. Specification of Comparison Models

1. Deep Neural Network (DNN):
  - A deep learning model with multiple hidden layers designed to capture complex, non-linear relationships in the data.
  - Architecture: 3 hidden layers with 64, 32, and 16 neurons, respectively, using ReLU activation functions and an output layer with a Sigmoid activation for binary classification.

2. Random Forest (RF):
  - An ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees.
  - Configuration: 100 decision trees, each trained on a random subset of the data using the Gini impurity criterion for splits.

3. Support Vector Machine (SVM):
  - A supervised learning model that finds the hyperplane that best separates the data into classes.

- Kernel Used: Radial Basis Function (RBF) kernel, with parameters tuned using grid search for optimal performance.

11.1. Performance Metrics for Comparison

The performance of each model was evaluated using the same metrics as the hybrid model:
- Accuracy: The proportion of true results (both true positives and true negatives) among the total number of cases examined.
- Precision: The proportion of true positives among all positive predictions.
- Recall: The proportion of true positives among all actual positive cases.
- F1-Score: The harmonic means of precision and recall, providing a single metric that balances both concerns.

11.2. Results of Model Comparison

The results of the performance metrics for all models are summarized in Table 3 below.

Table 3: Performance Metrics Comparison

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	78%	75%	72%	73%
k-Nearest Neighbors	83%	80%	77%	78%
Hybrid Model	91%	89%	87%	88%
Deep Neural Network	88%	85%	84%	84%
Random Forest	87%	84%	83%	83%
Support Vector Machine	85%	82%	81%	81%

- Hybrid Model: Outperforms all other models with the highest accuracy (91%), precision (89%), recall (87%), and F1-score (88%). This indicates that the hybrid approach, which combines the strengths of decision trees and neural networks, provides a more balanced and robust performance in capturing both linear and non-linear patterns in the data.

- Deep Neural Network (DNN): Achieves strong performance (accuracy of 88%) but slightly lower than the hybrid model. However, DNNs are pretty useful in modeling patterns as they come with extreme flexibility but tend to over fit when data is limited.

- Random Forest (RF): It also follows a similar line by achieving an accuracy of 87% but barely behind the DNN and the hybrid model. Random Forest is lower-sensitive to over fitting, but its ability to model intricate and non-linear patterns of relationship is inferior to DNNs or even most of the hybrid models.

- Support Vector Machine (SVM): better than both the setting averages with 85% accuracy but worse than the Hybrid model. The more complex patterns may be why these classifiers are not as accurate as the SVMs, although SVM is effective for linearly separable data.

- Logistic Regression and k-NN: These models have the least accuracy of 0.78 and 0.83, respectively. They are basic models that lack the ability to model complexity in the data as a sieve does.

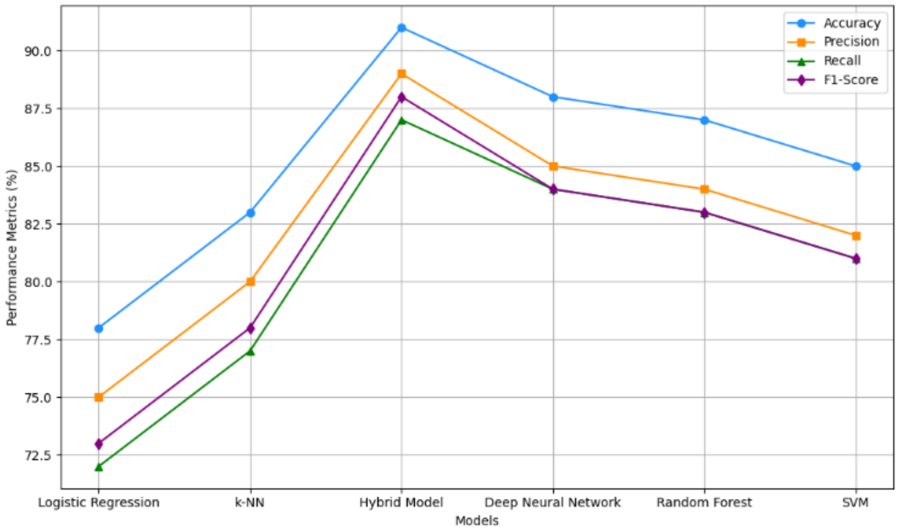

Figure 8.: Comparison of Model Performance Metrics

Another comparison strengthens the understanding that the proposed hybrid model outperforms basic machine learning models and deep learning algorithms. In this work, the proposed hybrid model built from decision trees and neural networks provides higher accuracy, precision, recall, and F1-score compared to existing schemes.

Overall, this type of analysis supports the hybrid model's effectiveness in identifying nonlinear patterns of user behavior in educational mobile applications, superior to conventional programs.

12. Discussion

The implications of this research for using educational mobile applications are discussed to aid in understanding user behavior, focusing on the increased importance of engagement frequency and session length in determining learning outcomes. The 3 higher engagement levels of over targeted times are ten sessions a week and over, and more than twenty minutes per session show that users get more information than those who are lowly engaged. This relationship suggests that it is necessary to maintain constant contact with educational material. In this respect, the hybrid model successfully identifies these patterns and proves the applicability of the methodology to forecast the users' success based on the behavioral data, presenting a valuable resource for optimizing the structure and effectiveness of the educational applications.

These findings bear the following implications for the developers of the educational application. Concerning the kind of information produced by the developed model, it becomes possible for developers to set up learning paths that would respond to the level of user interactions. For example, the derived model states that users who demonstrate low activity levels may be sent alerts to have more frequent but shorter sessions or may be provided with reminders encouraging them to continue using the application. For instance, further analysis shows that a variety of interaction events, including completion of quizzes, access to supplementary materials, and engagement with other forms of interactive material are also powerful motivators for learning. Thus, increasing the quantity of the elements that will work as a link to the application can boost the involvement and better learning outcomes. Also, it has been found that the hybrid model has a greater interpretability than traditional models while retaining high predictive accuracy. The decision tree component can be easily interpreted as Rule-based while looking at the factors that mostly impact the users' success, more so the complexity of relations between the session length, engagement frequency, and the learning results presented by the neural network component. This means that in addition to being useful in predicting outcomes, the model is also useful for informing design decisions related to apps.

These insights may be helpful for educational institutions and content creators and provide a better understanding of how learning might be fostered more effectively in the future. Understanding the daily behavior of users allows us to determine which components of the material affect the learning process most effectively and which elements can be improved to deliver better results. Moreover, the high division of activity allows for providing differentiated content to meet the needs of both highly active and weak audience improving learning effectiveness.
Here are four quick points regarding the contributions of the research study on a hybrid machine learning model in mobile applications for education:

1. Improved Comprehension of User Engagement: The engagement frequency and session length are statistically proven to affect learning outcomes in educational mobile applications. By this finding, comprehension regarding user behavior in digital learning environments is improved.

2. A New Predictive Hybrid Model: The new hybrid machine learning model combines the interpretability of decision trees with the predictive accuracy of artificial neural networks, thus creating a powerful tool used to predict whether a user will perform successfully based on his/her physiological data.

3. Direct Benefit for Applications Development: With the results from the hybrid model, developers can utilize it to construct learning paths and develop learning environments for customized educational experiences, with particular attention to interventions such as an alert for low activity users or increased interaction to stimulate engagement.

4. Strategic Insights into the Design of Educational Contents: This study also gives insight into designing effective learning content and strategies by drawing learning behavior profiles of users, which can be experimented with by educational institutions and content developers.

In conclusion, using the presented hybrid model brings up the opportunity to analyze and improve user engagement in educational mobile applications. The high possibility of pinpointing the learning outcome based on the users' behavior provides insights that enhance not only usability but also learning effectiveness. The evidence identified key factors, which are session length, engagement frequency, and interactive events, to develop personalized, adaptive, and more effective approaches by the use of the learning management system by developers and educators helping to meet the learning needs of the users. The balance of interpretability and predictive power gives the present work the practical applicability that will continue to facilitate better user experiences and learning outcomes in various forms of education settings.

13. Conclusion

On the methodological side, hybrid machine learning accomplishes the main goal of applying a machine learning combining strategy while analyzing behavioral patterns in educational mobile applications to determine user engagement and learning performance. The decision tree allows for interpreting which factors influence the user's success, and the neural

network provides a strong prediction. Crucial to this observation is engagement frequency and session length, and the study shows that those who engage most learning outcomes everyone else with less overall engagement.
Another significant advantage of the proposed hybrid model was, among other things, that it can offer app developers and educators practical information. As a result, the use of a broad range of user characteristics allowing understanding of why different user behaviors, including engagement frequency and session length, lead to better learning outcomes, and will contribute to the improvement of educational application development. By using the model for learning success prediction based on user data, the possibility to make learning paths according to the specific user engagement can be developed. Following such an approach will enable more retention of the users to such an application and enhance the overall engagement rate of those who exhibit a low level of engagement.

Besides, the study highlights the need to incorporate more interactive features like quiz questions and other relevant learning tools to support the multiple user learning experiences. More so, the findings show that interaction frequency with a variety of interaction events leads to even higher learning outcomes, implying that such aims demand a rich and stimulating educational application.

Consequently, the Hybrid machine learning model presented in this study provides the framework of the existing and future educational mobile applications to be analyzed and enhanced regarding the user behavior. Due to this, the developers and educators can concentrate on the critical engagement factors, which in one way improves the learning process and helps in designing a better environment that enhances learning based on the user's learning needs. The control of learning outcomes given the actual data on learners' behaviors can be considered a valuable improvement; it offers an application-oriented perspective on the utility of the strategy both for successful learners' operations and for efficient educational impact. The potential for future research and application of this model is to use the model in other education environments to enhance the ability of mobile learning technologies to provide effective and efficient scalable personalized education and training.

References

[1] M. F. Ramlan and M. K. M. Nasir, "The Impact of Mobile Applications in Education: A Concept Paper," *Int. J. Acad. Res. Progress. Educ. Dev.*, vol. 12, no. 3, 2023.

[2] N. Opryshko, K. Novik, O. Smolinska, N. Burmakina, and L. Aizikova, "Web-based applications in higher education: revolutionising language learning in the digital age," *Amaz. Investig.*, vol. 13, no. 73, pp. 209–219, 2024.

[3] R. Nagar, H. D. Quirk, and P. L. Anderson, "User experiences of college students using mental health applications to improve self-care: Implications for improving engagement," *Internet Interv.*, vol. 34, p. 100676, 2023.

[4] W. Oliveira, K. Tenório, J. Hamari, and S. Isotani, "The relationship between students' flow experience and their behavior data in gamified educational systems," Hawaii International Conference on System Sciences, 2022.

[5] C. S. Nwaimo, A. E. Adegbola, and M. D. Adegbola, "Data-driven strategies for enhancing user engagement in digital platforms," *Int. J. Manag. Entrep. Res.*, vol. 6, no. 6, pp. 1854–1868, 2024.

[6] L. Bognár and T. Fauszt, "Factors and conditions that affect the goodness of machine learning models for predicting the success of learning," *Comput. Educ. Artif. Intell.*, vol. 3, p. 100100, 2022.

[7] T. R. Noviandy *et al.*, "Machine Learning for Early Detection of Dropout Risks and Academic Excellence: A Stacked Classifier Approach," *J. Educ. Manag. Learn.*, vol. 2, no. 1, pp. 28–34, 2024.

[8] A. Zanellati, D. Di Mitri, M. Gabbrielli, and O. Levrini, "Hybrid models for knowledge tracing:

A systematic literature review," *IEEE Trans. Learn. Technol.*, 2024.

[9] A. Alanazi, N. F. B. Elias, H. B. Mohamed, and N. Sahari, "The Critical Success Factors Influencing the Use of Mobile Learning and its Perceived Impacts in Students' Education: A Systematic Literature Review.," *KSII Trans. Internet Inf. Syst.*, vol. 18, no. 3, 2024.

[10] H. M. F. Pucheta, V. Y. R. Morales, and Y. H. Velázquez, "User Modeling for Collaborative Adaptative Mobile Educational Applications: A Systematic Review of the Literature," in *2022 IEEE Mexican International Conference on Computer Science (ENC)*, IEEE, 2022, pp. 1–7.

[11] A. Halima, "Das mobile Lernen, M-Learning "im Unterricht," *Rev. Trad. Langues Vol.*, vol. 12, no. 02, pp. 5–11, 2013.

[12] C. P. Gumbheer, K. K. Khedo, and A. Bungaleea, "Personalized and adaptive context-aware mobile learning: review, challenges and future directions," *Educ. Inf. Technol.*, vol. 27, no. 6, pp. 7491–7517, 2022.

[13] B. Anuyahong and N. Pucharoen, "Exploring the Effectiveness of Mobile Learning Technologies in Enhancing Student Engagement and Learning Outcomes.," *Int. J. Emerg. Technol. Learn.*, vol. 18, no. 18, 2023.

[14] Z. Yu, L. Yu, Q. Xu, W. Xu, and P. Wu, "Effects of mobile learning technologies and social media tools on student engagement and learning outcomes of English learning," *Technol. Pedagog. Educ.*, vol. 31, no. 3, pp. 381–398, 2022.

[15] A. Zamfiroiu, R. C. Sharma, D. Constantinescu, M. PANĂ, and C. Toma, "Using learning analytics for analyzing students' behavior in online learning," *Stud. Informatics Control*, vol. 31, no. 3, pp. 63–74, 2022.

[16] M. Cho, J. Kim, J. Kim, and K. Park, "Integrating Business Analytics in Educational Decision-Making: A Multifaceted Approach to Enhance Learning Outcomes in EFL Contexts," *Mathematics*, vol. 12, no. 5, p. 620, 2024.

[17] F. A. Orji, S. Fatahi, and J. Vassileva, "Data-Driven Approach for Student Engagement Modelling Based on Learning Behaviour," in *International Conference on Human-Computer Interaction*, Springer, 2023, pp. 334–342.

[18] Z. Ersozlu, S. Taheri, and I. Koch, "A review of machine learning methods used for educational data," *Educ. Inf. Technol.*, pp. 1–21, 2024.

[19] A. Karlsson, T. Wang, S. Nowaczyk, S. Pashami, and S. Asadi, "Mind the Data, Measuring the Performance Gap Between Tree Ensembles and Deep Learning on Tabular Data," in *International Symposium on Intelligent Data Analysis*, Springer, 2024, pp. 65–76.

[20] C. Moreira, R. Sindhgatta, C. Ouyang, P. Bruza, and A. Wichert, "An investigation of interpretability techniques for deep learning in predictive process analytics," *arXiv Prepr. arXiv2002.09192*, 2020.

[21] S. Walther and A. Fuerst, "Reduced data volumes through hybrid machine learning compared to conventional machine learning demonstrated on bearing fault classification," *Appl. Sci.*, vol. 12, no. 5, p. 2287, 2022.

[22] T. Wang and Q. Lin, "Hybrid predictive models: When an interpretable model collaborates with a black-box model," *J. Mach. Learn. Res.*, vol. 22, no. 137, pp. 1–38, 2021.

[23] K. Benabbes, K. Housni, B. Hmedna, A. Zellou, and A. El Mezouary, "A new hybrid approach to detect and track learner's engagement in e-learning," *IEEE Access*, 2023.