International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I1P125
Volume 73 Issue 1, 305-312, January 2025

Original Article

Determining the Level of Adaptability of Students in Online Education Using Machine Learning Algorithms

Lida Violeta Asencios-Trujillo[1], Lucía Asencios-Trujillo[2], Carlos Jacinto La-Rosa-Longobardi[3], Djamila Gallegos-Espinoza[4], Cristina Piñas-Livia[5]

[1,2,3,4,5] *Graduate School, Enrique Guzmán y Valle National University of Education, Lima-Peru*

[1]*Corresponding Author : asenciostrujillolida@gmail.com*

Received: 23 July 2024 Revised: 14 October 2024 Accepted: 13 November 2024 Published: 31 January 2025

Abstract *- In the context of online education, student adaptability is a critical factor for their success. This study aims to predict the level of adaptability of students in online education environments using Machine Learning models. A dataset of 1205 records was used, which includes several demographic and contextual characteristics, such as age, gender, educational level, and type of institution, among others. Data preprocessing included the transformation of categorical features using one-hot encoding. The dataset was then divided into training and test sets to evaluate the model's performance. The Random Forest algorithm was selected for the classification task due to its ability to handle data with multiple characteristics and its robustness against overfitting. The results show that the Random Forest model achieved an accuracy of 91.29% in predicting the level of adaptability. The recall and f1-score values for the different categories ("Low", "Moderate", "High") indicated good performance, especially for the "Low" and "Moderate" categories. All information collected for this study is anonymous, ensuring data privacy. The dataset includes data at the national and international levels, providing a broad and generalizable perspective.*

Keywords **-** *Adaptability, Online education, Machine learning, Classification, Random forest.*

1. Introduction

Global education has undergone a significant transition to online learning, driven in large part by the COVID-19 pandemic. This situation forced academic institutions around the world to adopt virtual teaching to ensure continuity of learning [1]. However, this transition was not without its challenges, and one of the most pressing issues is the level of adaptability of students to these digital environments. The concept of adaptability in online education is multifaceted, as it encompasses the ability of students to adjust to different virtual platforms, handle various technological tools, and maintain motivation and engagement in an environment where human interaction is limited [2].

Previous studies have shown that adaptability is crucial for academic success and student retention in online programs [3]. Students with high adaptability tend to perform better academically and experience a more positive learning experience[4]. However, assessing student adaptability presents several challenges. Factors such as access to technological resources, level of family support, previous technological skills, knowledge and preventive measures of diseases, and the type of educational institution can influence adaptability[5-7]. Geographic and demographic diversity adds another layer of complexity to the problem [8, 9]. Using machine learning techniques to address this challenge has shown promise. Machine learning models can identify complex patterns and make predictions based on large volumes of data[10]. This allows educational institutions to gain valuable insights to design personalized support strategies that foster adaptability and improve student academic achievement [11, 12].

In this context, the present study seeks to develop a Machine Learning model to predict the level of adaptability of students in online education environments. By identifying the key factors influencing adaptability, this study aims to contribute to the development of more effective educational strategies and a better understanding of students' needs in an increasingly digitized world.

2. Literature Review

The rise of online education has led to a significant change in the way teaching is delivered and learning is assessed. Prior to the COVID-19 pandemic, online education was already an emerging field, but the closure of schools and universities globally forced most institutions to migrate to virtual platforms[13]. This massive transition posed challenges, one of the most important of which was the level of adaptability of the students.

2.1. Adaptability in Online Education

Adaptability in online education refers to the ability of students to adjust to new technologies, virtual platforms, and teaching methods. Previous studies suggest that adaptability is related to academic achievement and student satisfaction [14]. Students with higher levels of adaptability tend to be more successful in virtual settings, while those with lower adaptability may struggle with isolation and lack of technological support [15]. Assessing student adaptability can be tricky due to the diversity of factors involved. Access to technology, type of educational institution, and prior skills can all influence adaptability. One study suggested that students with limited access to technological resources have a harder time adapting to online education [16]. Machine Learning has been widely used to predict educational outcomes, such as academic performance and student retention[17].

These models can identify patterns in large volumes of data and help educational institutions make informed decisions. A recent study showed that the use of machine learning models can improve the accuracy of predictions in educational settings [18]. In the context of adaptability, machine learning models offer a valuable tool for predicting the level of adaptation of students[19]. Some studies have explored the use of classification models to identify factors that contributeto adaptability and to predict student behavior in virtual settings [20-23]. These approaches can help educational institutions identify students who need additional support and design strategies to improve their online experience.

3. Methodology

To develop a Machine Learning model that predicts the level of adaptability of students in online education, we follow a process that encompasses data preprocessing, feature selection, dataset splitting, and model training.

3.1. Data collection

The dataset used in this study was collected from educational sources and contains detailed information about students in online education environments. The dataset consists of 1205 records, each representing a unique student. 14 characteristics were collected for each student, covering demographic, academic, and contextual aspects.

The features included in the dataset are:

- Gender: Category that indicates whether the student is “Boy” or “Girl”.
- Age: Age range of the student, with values such as “11-15”, “16-20”, “21-25”, among others.
- Educational level: Type of institution where the student studies, such as “School”, “College”, or “University”.
- Type of institution: Distinguishes between “Government” and “Non-Government” institutions.
- IT Student: Indicates whether the student is in an information technology program.
- Location: Reflects whether the student is in an urban or rural location.
- Other contextual factors: Characteristics related to the study environment, such as “Load-shedding”, “Internet Type”, “Network Type”, “Class Duration”, and “Device”.

The target variable in the dataset is the “Level of Adaptability,” which indicates the student’s ability to adapt to online education. This variable has three possible values: “Low”, “Moderate”, and “High”, which allows students to be categorized according to their level of adaptability.

3.1.1. Data Source and Anonymization

The dataset was collected anonymously to ensure the privacy of the students. The data comes from national and international educational sources, providing a broad perspective of adaptability in online education. Anonymization ensures that individual students cannot be identified, which is essential to protect their privacy. To provide geographic context and demonstrate the diversity of sources, geospatial information about where the data is coming from can be displayed. Figure 1 shows a map with the location of the main data sources, indicating that most of it comes from Peru, with additional points in Latin America and some in other regions of the world. The use of geospatial maps provides a visual method to illustrate the geographic distribution of data without compromising student anonymity.

Fig. 1 Location of data collected for the study (South America)

3.2. Data Preprocessing

Data preprocessing is a critical step in ensuring that the dataset is in a format suitable for machine learning analysis and models. Since all the characteristics of the dataset are categorical, it is necessary to convert them into numerical variables so that the algorithms can process them.

3.2.1. One-Hot Encoding

The one-hot encoding method was used to convert the categorical features into a numerical format. This method creates additional columns, where each column represents a unique category and is assigned a value of 0 or 1, indicating the absence or presence of that category. This eliminates the implicit hierarchy that can arise when using direct numerical representations, such as ordinal encoding, that could induce bias in the model.

The implementation of one-hot encoding was done using the Scikit-learn library, specifically the OneHotEncoder method, which allows you to convert multiple categorical columns into an array of numerical features. By applying this technique, 22 new features were obtained from the original 13 (excluding the target field). To avoid multicollinearity, the drop='first' parameter was used, which removes the first category of each feature after encoding to avoid redundancies.

3.2.2. Processing Missing Values

Before applying one-hot encoding, missing values were checked in the dataset. Incomplete records or records with inconsistent data were removed to maintain the quality of the dataset. Using techniques such as Pandas’ dropna() ensures that the dataset is clean and free of faulty records.

3.2.3. Normalization and Scaling

Although normalization or scaling techniques were not applied in this study because all characteristics were categorical, it is important to consider these techniques when working with numerical data of different scales. Scaling may be required for some machine learning models that are sensitive to the magnitude of the data.

3.2.4. Implications of Preprocessing

Preprocessing data using one-hot encoding increases the number of features, which can have implications for the machine learning model. More features mean more complexity, but also the ability to capture more detailed information. This requires a robust model that can handle a high number of features without falling into overfitting.

3.3. Splitting the Dataset

Dividing the dataset into training and test sets is a critical step in the machine learning process. It ensures that the model is trained with sufficient data and that there is a separate set to validate its performance. The goal is to evaluate how the model performs with data not seen during training to simulate real-world situations.

3.3.1. Split Process

To split the dataset, we use the train_test_split function of the Scikit-learn library. This method allows you to specify the proportion of data to be used for training and testing, ensuring that the process is random and reproducible.

In this study, the dataset was divided into two sets: 80% was used for training and the remaining 20% for testing. This approach is commonly used because it provides a balance between training set size and model validation.

3.3.2. Considerations for Splitting the Dataset

Randomness: To avoid bias, the division process was performed randomly, ensuring that each record has an equal probability of being included in any set. A seed (random_state) was established to ensure the experiment's reproducibility.

Training/Test Ratio

The use of an 80/20 is common in machine learning, but it can vary depending on the size of the dataset and the complexity of the problem. A higher ratio for training may be necessary if the dataset is small, while a higher ratio for testing may be useful if more robust validation is required.

Class Balance

Since the target field has three values (“Low”, “Moderate”, “High”), it is important to ensure that both partitions maintain a balanced distribution of these classes. The process of random splitting tends to maintain this balance, but it’s always good to check it after partitioning.

3.3.3. Implications for the Model

Proper dataset splitting helps prevent overfitting by ensuring that the model is trained with sufficient data, but does not memorize patterns specific to the training set. It also allows for more accurate evaluation of the model, as the test suite provides a production-like scenario, with data not seen during training.

3.4. Machine Learning Model

To predict the level of adaptability of students in online education, the Random Forest algorithm was selected, a model based on decision trees that is characterized by its ability to handle data with multiple characteristics and its robustness in the face of overfitting.

3.4.1. Reasons to Choose Random Forest

Random Forest is an ensemble learning algorithm that creates multiple decision trees during the training process and then combines their results to get a final prediction. This ensemble approach offers several advantages.

- Robustness vs. Overfitting: By generating multiple trees, Random Forest is less prone to overfitting, as it averages the results and reduces the variability of predictions.

- Categorical Feature Management: Random Forest can handle categorical and numerical features without the need for extensive transformation.
- Importance of Characteristics: The model provides a measure of importance for each characteristic, which helps identify the most influential ones for prediction.

3.4.2. Random Model Parameters

The model was configured with the following parameters:

- Number of Trees: 100 decision trees were trained, which is common to obtain a balance between accuracy and computational efficiency.
- Division Criterion: The entropy criterion was used to decide where to divide each node of the tree. This criterion measures the impurity of the data and allows you to choose the best division to maximize information gain.
- Tree Depth: Although no maximum depth was specified, Random Forest tends to limit depth naturally to avoid oversetting.

3.4.3. Training Process

The training process involved the use of the training dataset (80% of the total). The model was allowed to create multiple decision trees, each trained on a random subset of data and features, to increase the diversity of decisions. This helps reduce variance and improves model stability.

3.5. Model Evaluation

Once trained, the model was evaluated using the test set (20% of the total). Metrics such as accuracy, recall, and f1-score were measured for each category of the target variable ("Low", "Moderate", "High"). Overall accuracy was also calculated to determine the effectiveness of the model in predicting the level of adaptability.

3.5.1. Future Implications and Improvements

Random Forest is a versatile and robust model, but it can be optimized by adjusting hyperparameters, such as the number of trees, the maximum depth, and the division criterion. Future improvements could include exploring other machine learning algorithms and applying advanced techniques such as hyperparameter tuning to improve performance.

3.6. Model Evaluation

A critical part of the machine learning process is evaluating the model to determine its performance and reliability. In this study, after training the Random Forest model, the test set (20% of the dataset) was used to evaluate its effectiveness in predicting the level of adaptability of students in online education.

3.6.1. Evaluation Metrics

To obtain a complete evaluation of the model, the following metrics were used:

- Accuracy: Measures the percentage of correct predictions in relation to the total predictions made. It is a general metric that provides an idea of the accuracy of the model. In this study, an accuracy of 91.29% was achieved, indicating a high level of accuracy.
- Recall: Evaluates the proportion of true positives correctly identified in each category. A high recall indicates that the model is able to correctly identify most true cases in a specific class.
- F1-Score: It is the harmonic average between accuracy and recall. This metric is useful for balancing both aspects, especially when classes are out of balance. A high f1-score suggests that the model is able to maintain a balance between accuracy and recall.

3.6.2. Evaluation by Category

To understand how the model performed in each category of the target variable ("Low", "Moderate", "High"), the recall and the f1-score were calculated for each.

- High: 88% accuracy, 65% recall, and 75% f1-score.
- Low: 93% accuracy, 94% recall, and 94% f1-score.
- Moderate: 90% accuracy, 94% recall, and 92% f1-score.

3.6.3. Overall Accuracy and Considerations

In addition to the metrics by category, the overall accuracy was calculated to have an overview of the model's performance. The accuracy of 91.29% indicates that the model can correctly predict the level of adaptability with a high degree of accuracy. However, some categories may have lower accuracy or recall, suggesting potential areas for improvement.

3.6.4. Possible Improvements

Techniques such as adjusting hyperparameters, using different machine learning algorithms, and adding more training data can be explored to improve model performance. Hyperparameter tuning can help optimize the model for better results in categories with lower recall.

3.7. Tool and Libraries Used

The use of specialized tools and libraries is essential for the analysis and development of Machine Learning models. In this study, we used several Python libraries for data preprocessing, model training, and performance evaluation. The key tools used and their purpose are described below.

3.7.1. Pandas

Pandas is a widely used Python library for data manipulation and analysis. It allows you to work with flexible data structures such as DataFrames, which are similar to tables in databases. In this study, Pandas was used to.

- Data Reading: Pandas made it easy to read the dataset from a CSV file and convert the data to a structured format.

- Preprocessing: Pandas functions were used for initial preprocessing, such as removing duplicate values, treating missing values, and manipulating columns.
- Data Exploration: Pandas provided functions to explore the structure of the dataset, calculate descriptive statistics, and transform data for analysis.

3.7.2. Scikit-Learn

Scikit-learn is a Python library for machine learning, offering a wide range of algorithms and tools for data preprocessing, model training, and performance evaluation. In this study, Scikit-learn was used to.

- One-Hot Encoding: Scikit-learn provided the OneHotEncoder method for converting categorical features into numerical format.
- Splitting the Dataset: The train_test_split function divided the dataset into training and test sets.
- Model Training: The Random Forest algorithm was selected for training, and Scikit-learn offered an interface to configure parameters such as the number of trees and the division criterion.
- Model Evaluation: Scikit-learn provided metrics to evaluate model performance, such as accuracy, recall, and f1-score.

3.7.3. Seaborn and Matplotlib

Seaborn and Matplotlib are Python libraries used for visualizations. Matplotlib is a general-purpose graphics library, while Seaborn focuses on statistical visualizations. In this study, Seaborn and Matplotlib were used to:

- Exploratory Data Analysis: Seaborn and Matplotlib allowed for the creation of graphs to visualize the distribution of the target field, explore relationships between features, and generate heat maps to show correlations.
- Visualization of Results: They were used to create graphs that show the performance of the model, such as the distribution of predictions and evaluation metrics.

4. Result

After the training and evaluation of the Machine Learning model, significant results were obtained in terms of accuracy and ability to predict the level of adaptability of students in online education. A Random Forest model with 100 trees was used to make the prediction, and the results were evaluated using metrics such as accuracy, recall, and f1-score.

4.1. Classifier Comparison

Figure 2 shows the comparison between different classifiers (Random Forest, Decision Tree, and KNN) in terms of accuracy and mislabeled points. The graph represents accuracy in a thin red bar overlaid with thicker bars indicating the number of mislabeled points for each classifier.

- Random Forest: With a high level of accuracy, this sorter had a total of 119 mislabeled points.
- Decision Tree: More accurately, it had 39 mislabeled points, indicating fewer errors compared to Random Forest.
- KNN: Showed lower accuracy than Decision Tree, with a total of 76 mislabeled points, suggesting an intermediate performance.

This analysis helps to compare the performance between different ranking models, showing that Decision Tree has the fewest errors, while Random Forest has the highest number of mislabeled points.

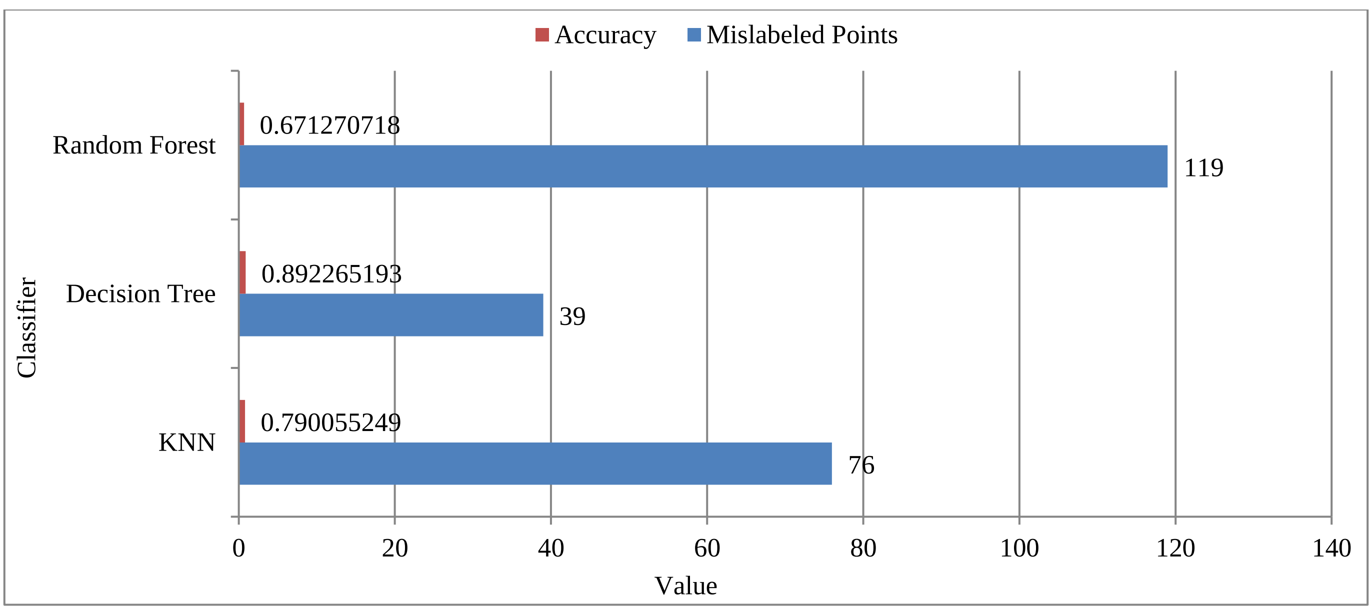

Fig. 2 Comparison of mislabeled points and accuracy

4.2. Model Performance Analysis

The model achieved an overall accuracy of 91.29%, indicating a high accuracy level in predicting adaptability. This accuracy suggests that the model is reliable in identifying patterns and making accurate predictions in the dataset. To better understand the model's performance, we analyzed the accuracy, recall, and f1-score metrics for each category:

- High: 88% accuracy, 65% recall, and 75% f1-score.
- Low: 93% accuracy, 94% recall, and 94% f1-score.
- Moderate: 90% accuracy, 94% recall, and 92% f1-score.

The graphs below show the recall and f1-score for each category:

- Recall Chart by Category:
- F1-Score Chart by Category:

These charts illustrate the distribution of model performance in each category.

4.3. Machine Learning Model Output

The Random Forest model was trained to predict the level of adaptability of students in online education. The results shown below include the learning curve and the confounding matrix, which help evaluate the model's performance and identify potential areas for improvement. The learning curve in Figure 3 shows how the model's accuracy changes as the number of training examples increases. The X-axis represents the number of examples used for training, while the Y-axis indicates the model's accuracy. The red curve shows the accuracy of the model in the training set. This curve stabilizes at around 95%, which could indicate a possible overadjustment. The green curve shows the accuracy of the model in cross-validation. This curve has an ascending pattern, suggesting that the model improves as the number of examples increases. This difference between the training curve and the cross-validation curve suggests that the model could benefit from more training data and adjustments to reduce overfitting.

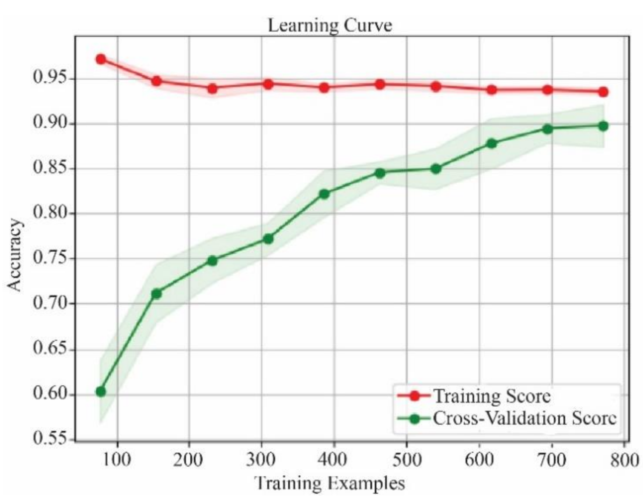

Fig. 3 Learning curve

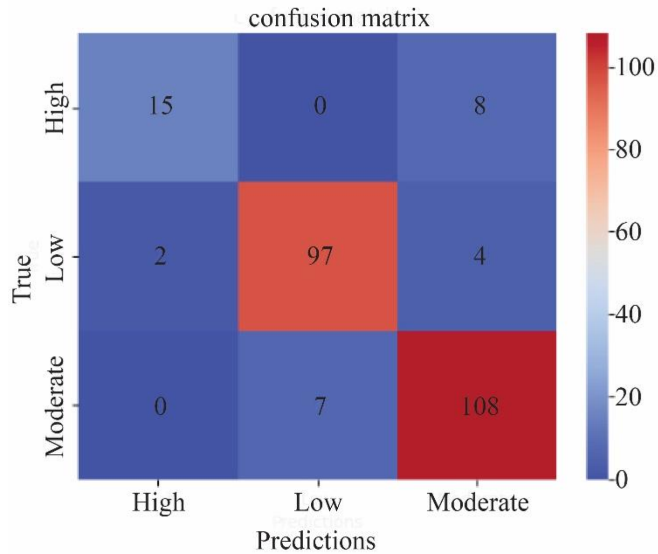

Fig. 4 Confusion matrix

4.4. Model Analysis

The overall accuracy of 91.29% indicates a good performance of the Random Forest model in predicting the level of adaptability. However, recall is lower in the "High" category, suggesting potential areas for improvement. Figure 4 shows the model's confusion matrix, which can help visualize how the model predicts each category. The confounding matrix shows how the model predicted the "High", "Low", and "Moderate" categories compared to the actual values of the test set. Each row represents the actual values, and each column represents the model's predictions. Ideally, the correct predictions are located on the main diagonal.

4.4.1. Correct Predictions (main diagonal)

- For the "High" category, the model correctly predicted 15 cases.
- For the "Low" category, the model had 97 correct predictions.
- For the "Moderate" category, the model got it right 108 times.

4.4.2. Errors (outside the main diagonal):

- The model mistakenly predicted 2 cases as "Low" instead of "High".
- There were 8 incorrect predictions of "High" as "Moderate".
- In the "Moderate" category, the model had 7 incorrect predictions towards "Low".

5. Discussion

The results obtained from the Random Forest model to predict the level of adaptability of students in online education suggest several interesting points to discuss. The model achieved an overall accuracy of 91.29%, indicating high accuracy. However, the confounding matrix showed that the model struggles to accurately predict the “High” category, suggesting areas for improvement. Compared to previous research on prediction in online education, these results align with the trend of using machine learning models to improve accuracy and obtain reliable predictions. The high global accuracy level indicates that the Random Forest model can be a valuable tool for this type of study. Although the model performed strongly, the lower accuracy in the “High” category suggests that more data or adjustments to the model are needed. This could involve more data collection, focus on class balance, and hyperparameter tuning to improve the model’s ability to accurately distinguish between categories. A possible improvement would be to implement oversampling techniques to balance classes or test other machine learning models that are less likely to be confused between adjacent categories. Using hyperparameter tuning techniques, such as grid search, could also help find optimal settings for the model.

These results have practical implications for the educational field, especially in online education environments. The Random Forest model can be used by educational institutions to identify students who need additional support in terms of adaptability. This could lead to implementing customized strategies to enhance the learning experience and increase student retention in virtual environments.

6. Conclusion

In conclusion, the results obtained from the Random Forest model indicate that, although it has a good overall performance, there are areas for future improvements, especially in accuracy for the “High” category. Analysis of the correlation matrix suggests that certain features could be optimized to improve the model. These results are a step forward in using Machine Learning to predict the level of adaptability in online education and offer a solid basis for future research.

References

[1] Wiston Forero-Corba, and Francisca Negre Bennasar, “Machine Learning and Artificial Intelligence Techniques and Applications in Education: A Systematic Review,” *Trends in Digital Education*, vol. 27, no. 1, pp. 209-253, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Catherine M. Pank, “Online Education,” *American Journal of Nursing*, vol. 107, no. 5, pp. 74-76, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[3] Junqi Zhu et al., “Effects of Online Learning on College Students in Eastern China: A Structural Equation Model,” *Frontiers in Public Health*, vol. 10, pp. 1-13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Audrey Perry, and Tia Andrighetti, “Evaluation of Online Distant Synchronous Interprofessional Simulations,” *Journal of Perinatal and Neonatal Nursing*, vol. 37, no. 2, pp. 123-130, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Felipe Maia de Toledo Piza et al., “Assessing Team Effectiveness and Affective Learning in a Datathon,” *International Journal of Medical Informatics*, vol. 112, pp. 40-44, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[6] Gina Camayo et al., “Emotional Intelligence, Depression, and Risk Factors in Pregnant Peruvian Andean Adolescents: A Multivariate Logistic Regression Study,” *International Journal of Preventive Medicine*, vol. 13, no. 1, pp. 1-5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Fernan Cosme Chaname Zapata et al., “Knowledge and Prevention Measures Against Sexually Transmitted Infections in Peruvians Andean Adolescents,” *Journal of Public Health*, vol. 23, no. 1, pp. 1-7, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Yuling Chen, Ningyu Fan, and Hongyang Wu, “AI-Enabled Metaverse for Education: Challenges and Opportunities,” *20th International Conference, Held as Part of the Services Conference Federation, SCF 2024*, Bangkok, Thailand, vol. 15429, pp. 103-113, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Natalia I. Vargas-Cuentas et al., “Chagas Disease Study Using Satellite Image Processing: A Bolivian Case,” *Acta Astronautica*, vol. 144, pp. 216-224, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Catherine Nabiem Akpen et al., “Impact of Online Learning on Student’s Performance and Engagement: A Systematic Review,” *Discover Education*, vol. 3, no. 1, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Xiaoshan Huang et al., “Social Presence in Technology-Rich Learning Environments: How Real we are Feeling Connected and How Does it Matter for Learning?,” *Information and Learning Sciences*, vol. 124, no. 11-12, pp. 396-424, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Enrique Lee Huamaní et al., “Prototype of a Mobile Application for the Detection of Car Accidents on the Roads of Peru,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 3, pp. 37-42, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Tzu-Shen Lin, and Yuh-Huey Jou, “Adolescents’ Perceptions of Parental Guan-Chiao and Parent-Child Relationships: Mediating Role of Parent-Child Digital Interaction,” *Bulletin of Educational Psychology*, vol. 56, no. 2, pp. 281-304, 2024. [CrossRef] [Publisher Link]

[14] Jelena Jardas Antonić, Antonija Srok, and Nenad Vretenar, "Blended Learning in the Spotlight of Educational Management-A Scientometric Analysis," *Journal of Information and Organizational Sciences*, vol. 48, no. 2, pp. 279-294, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[15] Larry Johnson et al., "*NMC Horizon Report: 2016 Higher Education Edition*," Austin, Texas: The New Media Consortium, pp. 1-56, 2016. [Google Scholar] [Publisher Link]
[16] S. Adams Becker et al., *NMC Horizon Report: 2017 Higher Education Edition*, New Media Consortium, 2017. [Google Scholar] [Publisher Link]
[17] Said A. Salloum et al., "Predicting Student Adaptability to Online Education Using Machine Learning," *Communications in Computer and Information Science*, Thessaloniki, Greece, vol. 2162, pp. 187-196, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[18] Daniel Bailey, and Norah Almusharraf, "Facebook in Class: The Instructor's Influence on Engagement and Language Play in Online Social Media Forums," *Computer-Assisted Language Learning Electronic Journal*, vol. 22, no. 3, pp. 66-85, 2021. [Google Scholar] [Publisher Link]
[19] R. Rivas et al., "Biomedical Engineering Education in Peru in 2015: A Unique and Innovative Collaboration in Latin America," *World Congress on Medical Physics and Biomedical Engineering*, Toronto, Canada, pp. 1612-1615, 2015. [CrossRef] [Google Scholar] [Publisher Link]
[20] Daojun Wang, Meishu Wang, and Xinli Xing, "Research on Integrating Blockchain and Machine Learning LPP Algorithm in Online Education Platform Under Covid-19 Environment," *Scalable Computing: Practice and Experience*, vol. 25, no. 3, pp. 1442-1454, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[21] Wil Van der Aalst, *Data Science in Action*, Process Mining, pp. 3-23, 2016. [CrossRef] [Google Scholar] [Publisher Link]
[22] Randall S. Davies, and Richard E. West, *Technology Integration in Schools*, Handbook of Research on Educational Communications and Technology, 4th ed., pp. 841-853, 2013. [CrossRef] [Google Scholar] [Publisher Link]
[23] Larry Cuban, *Oversold and Underused: Computers in the Classroom*, Harvard University Press, 2009. [CrossRef] [Google Scholar] [Publisher Link]