A

Project Report on

PREDICTING A STUDENT'S ADAPTABILITY LEVEL IN ONLINE EDUCATION

Submitted to the partial fulfillment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE & ENGINEERING

Submitted by:
Rishika Sharma (RA1911026030046)
Ayush Rawat (RA1911026030045)
Shubhaangi Verma (RA1911026030026)
Tanishk Arora (RA1911026030019)

Supervised by:
Dr. Anand Pandey
Assistant Professor

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)
DELHI-NCR CAMPUS, GHAZIABAD (U.P)

SRM Institute of Science and Technology
Delhi NCR Campus, Modinagar,
Ghaziabad (UP)-201204

APRIL 2023

BONAFIDE CERTIFICATE

This is to certify that project Report entitled "Predicting a Student's Adaptability in Online Education", which is submitted by Rishika Sharma (RA1911026030046), Ayush Rawat (RA1911026030045), Shubhaangi Verma(RA1911026030026) and Tanishk Arora (RA1911026030019) in the partial fulfillment of the requirement for the award of degree B.Tech(CSE with specialization in Artificial Intelligence and Machine Learning) of SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Ghaziabad is a record of the candidate own work carried out by them under my own supervision.

…..
(Signature)
Dr. Anand Pandey
Project Supervisor
Assistant Professor (Sr.G)
Department of CSE

…………………………........
(Signature)
Dr. Ashwani Kumar
HOD (CSE AIML)
Associate Professor

INTERNAL EXAMINER

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

We would like to express our heartfelt appreciation to Dr. Anand Pandey, Assistant Professor and Project Supervisor at SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, for his invaluable insights and expertise in the subject matter, which motivated us to work diligently.

Our profound gratitude goes out to Dr. Jitendra Singh and Dr. Rakesh Kumar Yadav, Project Coordinators at SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, for their enlightening guidance and skillful coordination, which served as a perpetual source of inspiration.

We would also like to extend our sincere thanks to Dr. S. Vishwanathan, Director of SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, for his unwavering support that enabled us to undertake and complete our project work.

Our special thanks go to Dr. D. K. Sharma, Dean (Academics), and Dr. R. P. Mahapatra, Dean (E&T) at SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, for their valuable guidance and unconditional support.

We would like to express our gratitude to Dr. Ashwani Kumar, Head of the Department of Computer Science and Engineering with specialization in Artificial Intelligence and Machine learning at SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, for his suggestions and encouragement in completing this project.

We also owe our thanks to all the teaching and non-teaching staff members of our college who provided us with direct or indirect help throughout our studies and project work.

Finally, we would like to express our sincere appreciation to our parents, family members, and friends for their unwavering support and encouragement, and to all our well-wishers.

Rishika Sharma (R1911026030046)

Ayush Rawat (RA1911026030045)

Shubhaangi Verma (RA1911026030026)

Tanishk Arora (RA1911026030019)

DECLARATION

We, Rishika Sharma (RA1911026030046), Ayush Rawat (RA1911026030045), Shubhaangi Verma (RA1911026030026) and Tanishk Arora (RA1911026030019) hereby declare that the work which is being presented in the project report "Predicting a Student's Adaptability in Online Education" is the record of authentic work carried out by us during the period from January '23 to May '23 and submitted by us in partial fulfillment for the award of the degree "Bachelor of Technology in Computer Science and Engineering" to SRM IST, NCR Campus, Ghaziabad (U.P.). This work has not been submitted to any other University or Institute for the award of any Degree/Diploma.

Rishika Sharma (RA1911026030046)

Ayush Rawat (RA1911026030045)

Shubhaangi Verma (RA1911026030026)

Tanishk Arora (RA1911026030019)

ABSTRACT

The Covid-19 pandemic has compelled a significant move towards online learning, increasing the focus on students' capacity to adapt to this new approach to education. It is crucial to comprehend the elements that influence a student's capacity to adjust to online learning. In light of this, our team has been working to create a classifier that can assess a student's level of adaptability to online learning by looking at a variety of indicators relating to their personal, academic, and learning habits. Members of our staff have firsthand experience with the difficulties of online learning. For the past two years, we have all participated in online learning, and we have all experienced varied degrees of difficulties adjusting to this form of instruction. We are exploring how the various criteria linked with an individual affect their level of adaptability to online learning in light of their personal experiences. We intend to create a classifier that can precisely predict a person's level of adaptation to online learning by analyzing these aspects. We intend to achieve this by developing an ideal model using state-of-the-art machine learning methods and visualization tools. In order to do this, it will be necessary to analyze enormous quantities of student data and use sophisticated algorithms to look for trends and connections between different characteristics and adaptability levels. Once a reliable classifier has been created, it will be a useful tool for educators and decision-makers trying to enhance online education and make it more available to all students. We intend to support continuous efforts to enhance online learning and make sure that all students have the resources they need to be successful in an educational environment that is continually changing by utilizing the power of machine learning and data analytics.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1

INTRODUCTION

1.1 EFFECT OF COVID ON EDUCATION SYSTEM

The COVID-19 pandemic has significantly impacted the educational industry on a local and international scale. Numerous difficulties have arisen as a result of schools and universities' forced adaptation to a new manner of teaching and learning.

One of the most significant effects of the epidemic on education is the shift to online learning. To maintain social distance and stop the virus from spreading, many colleges and universities have switched to remote instruction. While flexible and convenient, online learning has also brought attention to the differences in access to technology and the internet. It may be challenging for students from low-income families to participate in online learning because they lack access to dependable computers or internet connections.

Due to closures and quarantines, the pandemic has also disturbed the regular academic calendar, forcing schools and institutions to change their schedules. This has led to lost instructional time, which may have an effect on the academic development and long-term success of the pupils. The pandemic's negative social and emotional effects on kids have had a substantial influence on schooling as well. The mental health of students has declined as a result of social isolation, a lack of routine, and anxiety about the future. Some students have also had to take on bigger duties at home, such as looking after younger siblings or older relatives, which may make it harder for them to concentrate in class.

Additionally, the epidemic has had an economic impact on colleges and universities. To enable remote instruction, many institutions have been forced to make investments in cutting-edge technology like video conferencing and learning management systems. At the same time, a lot of families have been struggling financially, which may affect their capacity to cover educational costs like tuition and materials.

Finally, the pandemic has brought to light the value of resilience and adaptability in schooling. Schools and institutions had to immediately change their focus to remote instruction and come up with fresh approaches to student engagement. This has caused a renewed emphasis on creative and adaptable teaching methods that can assist students in succeeding in any learning setting.

The trend to online learning is only one example of how the COVID-19 epidemic has significantly impacted the educational sector, changes to the academic calendar, difficulties for students' social and emotional development, financial ramifications for institutions and families, and the demand for educational resilience and adaptability. Although it is yet unknown how long these effects will last, it is obvious that the epidemic has increased the demand for creative and adaptable methods of teaching and learning

1.2 MACHINE LEARNING INTRODUCTION

"Machine learning" is a subfield of artificial intelligence that focuses on developing statistical models and algorithms that enable computer systems to automatically improve at a job by taking in new information from the input**.** The explosion of data has made machine learning more and more significant in recent years since it gives machines the chance to learn and make predictions or judgements based on relationships and patterns in the data. The machine learning process, which has three main stages—data preparation, model training, and prediction or inference—is depicted in the diagram below.

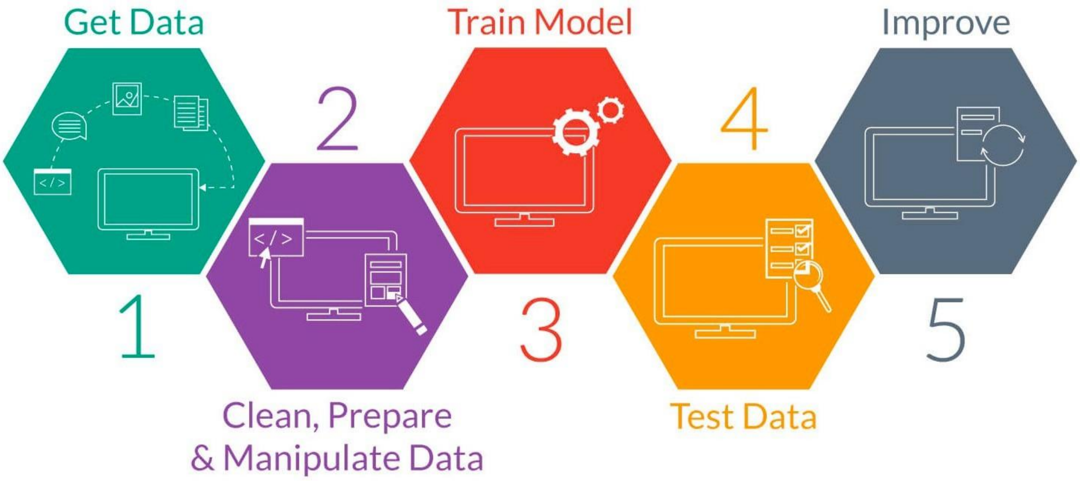

Fig. 1.1 Machine Learning Process

Data is gathered and preprocessed during the data preparation stage to make sure it is acceptable for machine learning algorithms. This could entail cleaning, normalizing, or otherwise changing the data. Machine learning algorithms are applied to the prepared data in the model training stage in order to learn from it and produce a predictive model. As shown in Fig. 1.1, the trained model is then utilized to make predictions or choices based on fresh, unobserved data during the prediction or inference stage.

Machine learning has several applications, such as fraud detection, natural language processing, image recognition, and recommendation systems. Machine learning is becoming a more crucial technology as data volume increases for gaining insights from and making predictions from this data.

1.3 OBJECTIVES

The COVID-19 epidemic has forced several educational institutions to adopt online instruction. The sudden shift to remote learning has left many students and teachers unable to adapt to the new format as a result of the abrupt move to remote learning. In this situation, educators and administrators may find it to be a crucial tool to be able to forecast how well students would adapt to online learning using machine learning models. The following are the goals of utilizing machine learning models to forecast students' capacity for online learning:

1. Identify at-risk students: Machine learning models can locate students who may be in danger of falling behind or quitting classes by examining the data gathered by online learning platforms. This can increase students' chances of success by enabling educators to intervene and give those who need it most specialized attention.

2. Personalize learning: Machine learning models can also be utilized to tailor each student's educational experience. Models can offer specialized recommendations on resources and activities that are most likely to interest and encourage each individual student by analyzing data on their learning patterns and preferences. This may boost student engagement and enhance learning outcomes.

3. Optimize course design: By examining data on student engagement and performance, machine learning models may also be utilized to improve course design. Education professionals can modify their course designs to better meet the requirements of their students and enhance learning outcomes by determining which resources and activities are most effective.

4. Improve student retention: Another way to increase student retention is to predict how well students will adapt to online learning. The likelihood that students will finish their courses and continue their education can be increased by educators by identifying at-risk students early on and offering focused support.

5. Enhance student success: Improving student success is the ultimate objective of forecasting students' adaptability to online learning. Teachers can assist students in achieving their academic objectives and setting themselves up for future success by offering focused support, individualized instruction, boosting student retention, and optimizing course design.

Machine learning models must be trained on massive datasets of student performance data, including details on their use of online learning platforms, their learning preferences and patterns, and their demographic and socioeconomic background, in order to accomplish these goals. This information can be used to spot trends and forecast behavior, enabling educators to give students who need it the most specialized attention. Additionally, machine learning models can be incorporated into online learning platforms to give teachers real-time information about their students' engagement and performance.

In conclusion, predicting how well students will adapt to online learning using machine learning models can give educators useful information on student involvement and performance. Teachers can assist students in overcoming the difficulties of online learning and achieving their academic objectives by identifying at-risk students, personalizing learning, optimizing course design, promoting student retention, and enhancing student achievement.

CHAPTER-2

LITERATURE REVIEW

The literature review of all the research papers that we have read for this project is contained in this chapter. First off, the research article we're going to talk about is titled Barriers to Online Learning in the Time of COVID-19: A National Survey of Medical Students in the Philippines. 584 medical students from 36 different medical schools nationwide participated in the study. Demographics, technological access, online learning experiences, and perceived impediments to online learning were all included in the poll. According to 96.6% and 96.7% of respondents, respectively, most medical students had access to a computer or laptop and a steady internet connection, according to the survey. However, 28.1% of respondents said they didn't have enough bandwidth for online learning, and 21.1% said they had trouble connecting to the internet while taking online courses. In addition, the study found that medical students encountered a number of challenges when attempting to learn online, including technical difficulties, such as slow internet and computer malfunctions (47.6%), domestic distractions (41.6%), and difficulties comprehending course material in the absence of instructor interaction (39.2%).

Lack of desire (35.6%), trouble managing time (33.0%), and trouble connecting with peers and instructors (27.6%) were among the other obstacles. The survey also discovered that the hurdles experienced by medical students varied significantly depending on their year level, gender, and region. While male students reported higher levels of enthusiasm and time management than female students, third-year students reported the largest hurdles to online learning. Students from metropolitan locations reported more technical difficulties and difficulty comprehending course topics, whereas students from rural areas reported more distractions at home. Overall, the study emphasizes the necessity for educational institutions to address the challenges that Philippine medical students encounter when trying to learn online. According to the survey, educational institutions should assist students in resolving technical challenges and enhancing their access to reliable internet connections.

Institutions should also think about giving students the chance to interact with teachers and classmates in order to help them understand the course materials.

The report recommends that educational institutions create plans to deal with domestic distractions and enhance students' time management abilities. This can entail delivering time management classes or giving students the means to set up dedicated study places at home. The report also emphasizes the necessity for schools to be aware of the many hurdles that students experience depending on their year level, gender, and region. Institutions should think about creating specific interventions to deal with these particular obstacles and offer assistance to students based on their individual requirements. In conclusion, the study discovered that during the COVID-19 epidemic, medical students in the Philippines encountered a number of obstacles to online learning, including technological difficulties, domestic distractions, and content comprehension challenges.

The second research article we've looked at is titled Gender disparities in Digital Learning During COVID-19: Proficiency Beliefs, Intrinsic Value, Learning Engagement, and Perceived Teacher Support. With a mean age of 36.79 years, the study used a cross-sectional survey of 899 individuals in Spain. Demographics, psychological discomfort, personality qualities, coping mechanisms, and social support were all included in the survey.

According to the study, a number of distinct factors were significantly related to psychological discomfort. One of the Big Five personality traits, neuroticism, was a significant predictor of psychological discomfort; the level of distress was inversely correlated with the level of neuroticism. Avoidance coping was positively correlated with psychological discomfort, but coping techniques including problem-solving and positive reappraisal were negatively correlated. A significant predictor of psychological discomfort was also social support. Lower levels of distress are associated with higher levels of social support.

The study also discovered that psychological anguish varied significantly depending on demographic characteristics. Psychological discomfort was reported to be more prevalent in women than in males, particularly in those with children as compared to those without. Additionally, those with lower income levels than those with higher income levels reported higher degrees of distress. The study has significant ramifications for people's psychological health during the COVID-19 pandemic.

According to the study, those who have greater levels of neuroticism may be more prone to psychological distress and may profit from specialized interventions that enhance their coping mechanisms and social networks. According to the study, those who use avoidance coping mechanisms may be more likely to experience psychological discomfort and may benefit from interventions that enhance their capacity for problem-solving and positive self-evaluation.

The report also emphasizes the need for measures that address how different demographic groups—such as women, those with children, and those with lower income levels—have been affected by the pandemic differently. For these populations to strengthen their coping mechanisms and social support throughout the pandemic, specific treatments may be beneficial.

The last study we looked at was titled The impact of gender, educational achievement, and personality on online learning outcomes during the COVID-19 outbreak. A mixed-methods technique was employed for the study, which included interviews with 12 students and 10 faculty members in addition to a survey of 463 students and 72 faculty members. Participants' experiences with online learning, opinions of its benefits and drawbacks, and suggestions for enhancing the online learning environment were all included in the survey and interviews.

According to the survey, online learning was not always a positive experience for both students and academic staff. While some participants claimed that online learning offered more flexibility and convenience, others complained about technological issues, a lack of interaction, and a decline in motivation and interest. Faculty members also mentioned difficulties with the lack of guidance and assistance for online instruction.

The survey also discovered that students' experiences varied significantly depending on their socioeconomic position and access to technology. Online learning was more challenging for students from low-income families and those without access to dependable internet and equipment.

The study makes the case that a number of elements, including access to technology, motivation, engagement, and faculty support, might influence the success of online learning. The report makes the recommendation that institutions give faculty members enough assistance and training to ease their transition to online instruction. Universities should support students who may be having difficulties with online learning due to socioeconomic difficulties or a lack of access to technology, according to the study's other recommendation.

Table 2.1- Tabular Summary of Literature Survey

S No.	AUTHOR	TITLE	PUBLICATION	RESULT
1.	Ronnie Baticulon, Jinno Jenkin Sy, Nicole Rose I Alberto, Maria Beatriz C. Baron	Barriers to Online Learning in the Time of COVID-19: A National Survey of Medical Students in the Philippines [5]	Research Gate	According to a poll conducted in the Philippines in May 2020, obstacles to online learning include school information, access to technology resources, study habits, living circumstances, and self-evaluation of ability.
2.	Selma Korlat, Marlene Kollmayer, Julia Holzer, Marko Lüftenegger, Elisabeth	Gender Differences in Digital Learning During COVID-19: Competence Beliefs, Intrinsic Value,	Frontiers in psychology	According to the study, older people who are more conscientious, amiable, and open perform better than younger people who are highly extraverted and neuroticism.

	Rosa, Pelikan Barbara Schober, Christiane Spiel	Learning Engagement, and Perceived Teacher Support [6]		
3.	Zhonggen Yu	The effects of gender, educational level, and personality on online learning outcomes during the COVID-19 pandemic [7]	Springer Open	According to the study, a student's level of adaptability is influenced by their gender. The fact that boys are viewed as having a higher level of aptitude, comfort, and engagement with computers may give them an advantage over girls.
4.	Di Xu, Shanna Smith Jaggars	Adaptability to Online Learning: Differences Across Types of Students and	Teachers College , Columbia University	In order to assess how successfully students adjust to the online environment, this study contrasts their

		Academic Subject Areas [2]		capability to persevere and achieve good marks in face-to-face courses with their ability to do so in online courses. Academic performance in online courses fell for all research participants, but some had a harder time transitioning than others, particularly men, younger students, Black students, and students with lower grade point averages. Students in these online classrooms struggled in areas like English and social science, in part due to unfavorable peer pressure.
5.	Anne-Mette Nortvig,	A Literature Review of the	University College Absalon, Denmark	According to the findings from the

	Anne Kristine Petersen and Søren Hattesen Balle	Factors Influencing E-Learning and Blended Learning in Relation to Learning Outcome, Student Satisfaction and Engagement [3]		research papers included in the review, some of the factors that seem to predominate more are educator presence in online settings, interactions between students, teachers, and content, and purposefully designed connections between online and offline activities as well as between campus-related and practice-related activities. In addition to exploring and challenging the applicability of research focusing on comparisons between various forms of online learning, blended learning, or

				"traditional" face-to-face teaching and learning, the article makes a few significant suggestions. The teaching approach alone does not fully account for the complexity of learning and teaching.
6.	Paola Xhelili, Eliana Ibrahimi, Erinda Rruci, Kristina Sheme	Adaptation and Perception of Online Learning during COVID-19 Pandemic [4]	Albanian University	According to the report, student's biggest problems were the lack of an internet connection and technological equipment. High academic performers and students enrolled in technology-based programs felt more at ease and satisfied taking classes online.

CHAPTER-3

SYSTEM ANALYSIS

3.1 EXISTING PROBLEM

Worldwide educational institutions have been forced to switch from traditional in-person training to remote and online instruction as a result of the COVID-19 epidemic. Even though online education has grown in popularity recently, the abrupt transition to remote learning has presented a number of difficulties for students, particularly when it comes to adjusting to this new method of learning. We will examine the current issues that students encounter as they adjust to online education in this article.

Lack of in-person interaction with peers and teachers is one of the main issues that students have with online education. Students can connect with their instructors and fellow students in real-time, ask questions, and receive immediate feedback when learning in person. Online learning, on the other hand, restricts students to virtual interactions, which might be less interesting and productive than in-person interactions. Additionally, students who are uncomfortable using technology may find it difficult to interact with their lecturers and peers during online lessons.

The absence of structure and routine in online learning is another problem for students. In conventional in-person instruction, students follow a fixed plan and routine that aids in time management and helps them stay on task. Although some students may find it difficult, online learning frequently requires students to plan their own schedules and manage their time autonomously. It may be challenging for students who have trouble with self-control and time management to remain motivated and focused in an online learning environment. Additionally, online learning could be less participatory and engaging than conventional in-person learning, which makes it challenging for students to stay interested and focused. When taking classes online, many students find it difficult to stay motivated and interested, especially if they are not accustomed to this style of instruction. Additionally, interruptions and distractions from family members, pets, or household tasks may prevent students from concentrating and focusing on their academics.

Finally, a key issue that students with online education experience is the absence of social interaction and support. Students who are accustomed to socializing with their classmates and creating a sense of community at school could find it difficult to do so in an online learning setting. Due to this lack of social engagement, students may experience emotions of loneliness and isolation, which can be harmful to their mental health and general wellbeing.

In conclusion, online learning has taken over as the standard in the modern educational system. One of the biggest problems that students have with online education is the absence of in-person interaction. Other problems include the lack of structure, regularity, engagement, and social support. To support students in their online learning journey, schools and institutions need to address these issues and put in place efficient ways.

3.2 EXISTING SOLUTIONS

With the aid of machine learning, there has been an increase in interest in predicting student’s' ability to adapt to online learning. Examples of the work done in this field are as follows:

1. To predict student persistence in online courses, researchers at the University of Memphis used machine learning. The researchers trained a prediction model that could recognize students at risk of dropping out using demographic information, academic history, and data on online learning behavior. The model's forecast precision was 86%.

2. Machine learning was utilized in a different study that was published in the Journal of Educational Data Mining to forecast student success in an online course. The researchers trained a prediction model that could recognize students who were likely to succeed in the course using demographic data, academic background, and online learning behavior data. The model's forecast precision was 80%.

3. Machine learning was employed in a study by academics at the University of Southern California to forecast students' participation in online debates. The researchers created a prediction model that could recognize students who were likely to be very active in online discussions after using natural language processing techniques to analyze the content of students' online postings and comments. The model's forecast precision was 82%.

4. Machine learning was utilized in a study that appeared in the International Journal

of Distance Education Technologies to forecast student satisfaction with an online course. The researchers trained a prediction model that could identify students who were likely to be satisfied with the course using demographic data, academic background, and online learning behavior data. The prediction accuracy of the model was 85%. These findings show that machine learning has the potential to predict student adaptability in online learning. Educational institutions can offer tailored support to students to help them succeed in online courses by identifying students who are at risk of dropping out and forecasting success, engagement, or satisfaction.

3.3 PROPOSED SOLUTION

In online learning, student adaptation can be predicted using a variety of different machine learning models. The particular problem, the data, and the resources at hand all influence the model selection. In order to predict student adaptation in online learning, we will employ the following machine learning models:

1. Decision Tree: This model can be used to forecast if a student will finish an online course or drop out. It is frequently used for classification problems. Decision trees can handle categorical and continuous data and can be simple to read.

2. Random Forest: A decision tree extension that can be applied to classification and regression issues. Random forests are capable of coping with missing data, noisy data, and nonlinear relationships.

3. Logistic Regression: This model can be used to forecast whether or not a student would drop out of an online course and is frequently employed for binary classification issues. Implementing logistic regression is straightforward, it can handle both continuous and categorical data, and it produces coefficients that can be understood.

4. Support Vector Machine: SVM is a strong classifier that can handle high-dimensional data and non-linear correlations. SVM can be used to forecast online course participation or student achievement.

5. Artificial Neural Networks: This model is incredibly adaptable and can be applied to both regression and classification issues. Large datasets can be handled using neural networks, which can also capture complicated correlations between input properties and output variables.

6. K Nearest Neighbors: KNN is a non-parametric technique used in machine learning for both classification and regression tasks. It is a straightforward and comprehensible algorithm that locates the "nearest" data points to a given input based on a distance metric, and then predicts based on the class or value of those neighbors.

7. Gaussian Naive Bayes: This probabilistic approach is utilized in machine learning for classification applications. It is a simple and efficient strategy based on the Bayes theorem and assumes that the features are conditionally independent given the class variable.

CHAPTER 4

METHODOLOGIES

Preparing unprocessed data for a machine learning model is a technique known as data preparation. It is a crucial stage in the development of a machine learning model. The complexity of real-world data, together with the possibility of numerous forms of noise, missing values, and inconsistent formatting, make it unsuitable for inclusion in machine learning models. Therefore, it is essential to perform data preprocessing tasks to clean, transform, and prepare the data for analysis. By doing so, the accuracy and efficiency of machine learning models can be significantly improved, leading to more reliable results and better decision-making.

4.1 INITIAL DATASET

A machine learning model is trained with an initial dataset, which is a set of labelled or unlabeled data. The dataset is often made up of two types of information: input data, which includes properties of the data such as features or traits, and output data, which includes the outcome or target variable that is anticipated for each input data point. The performance of the machine learning model depends heavily on the size and quality of the initial dataset. While a smaller dataset with less diversity can result in overfitting and subpar generalization, a larger dataset with a varied range of input data can enhance the model's accuracy and robustness. The below figure [Fig.4.1] shows the head of the dataframe.

	Gender	Age	Education Level	Institution Type	IT Student	Location	Load-shedding	Financial Condition	Internet Type	Network Type	Class Duration	Self Lms	Device	Adaptivity Level
0	Boy	21-25	University	Non Government	No	Yes	Low	Mid	Wifi	4G	3-6	No	Tab	Moderate
1	Girl	21-25	University	Non Government	No	Yes	High	Mid	Mobile Data	4G	1-3	Yes	Mobile	Moderate
2	Girl	16-20	College	Government	No	Yes	Low	Mid	Wifi	4G	1-3	No	Mobile	Moderate
3	Girl	11-15	School	Non Government	No	Yes	Low	Mid	Mobile Data	4G	1-3	No	Mobile	Moderate
4	Girl	16-20	School	Non Government	No	Yes	Low	Poor	Mobile Data	3G	0	No	Mobile	Low
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1200	Girl	16-20	College	Non Government	No	Yes	Low	Mid	Wifi	4G	1-3	No	Mobile	Low
1201	Girl	16-20	College	Non Government	No	No	High	Mid	Wifi	4G	3-6	No	Mobile	Moderate
1202	Boy	11-15	School	Non Government	No	Yes	Low	Mid	Mobile Data	3G	1-3	No	Mobile	Moderate
1203	Girl	16-20	College	Non Government	No	No	Low	Mid	Wifi	4G	1-3	No	Mobile	Low
1204	Girl	11-15	School	Non Government	No	Yes	Low	Poor	Mobile Data	3G	1-3	No	Mobile	Moderate

1205 rows × 14 columns

Fig 4.1: Initial Dataset

4.2 DATA EXPLORATION

The practice of analyzing and visualizing a dataset to better comprehend its characteristics and trends is known as data exploration. Gaining insights into the data, spotting outliers and anomalies, and figuring out which characteristics would be most crucial for predicting the target variable are the aims of data exploration. The below figure [Fig. 4.2] shows all the features of the data frame.

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1205 entries, 0 to 1204
Data columns (total 14 columns):
 #   Column               Non-Null Count  Dtype 
---  ------               --------------  ----- 
 0   Gender               1205 non-null   object
 1   Age                  1205 non-null   object
 2   Education Level      1205 non-null   object
 3   Institution Type     1205 non-null   object
 4   IT Student           1205 non-null   object
 5   Location             1205 non-null   object
 6   Load-shedding        1205 non-null   object
 7   Financial Condition  1205 non-null   object
 8   Internet Type        1205 non-null   object
 9   Network Type         1205 non-null   object
 10  Class Duration       1205 non-null   object
 11  Self Lms             1205 non-null   object
 12  Device               1205 non-null   object
 13  Adaptivity Level     1205 non-null   object
dtypes: object(14)
memory usage: 131.9+ KB
```

Fig. 4.2: Data Frame Information

The association between various elements in a dataset is visualized using a heatmap, a common data visualization approach. With the use of a heatmap, it is possible to spot patterns and connections in the data that might not be immediately obvious from a straightforward review of the raw data. Model training is the process of using a labelled dataset to train a machine learning algorithm or model. Model training aims to train the model to spot patterns in the data and utilize those patterns to make precise predictions about brand-new, untainted data.

The below figure [Fig. 4.3] represents the relation between different features available in the dataset using a heatmap.

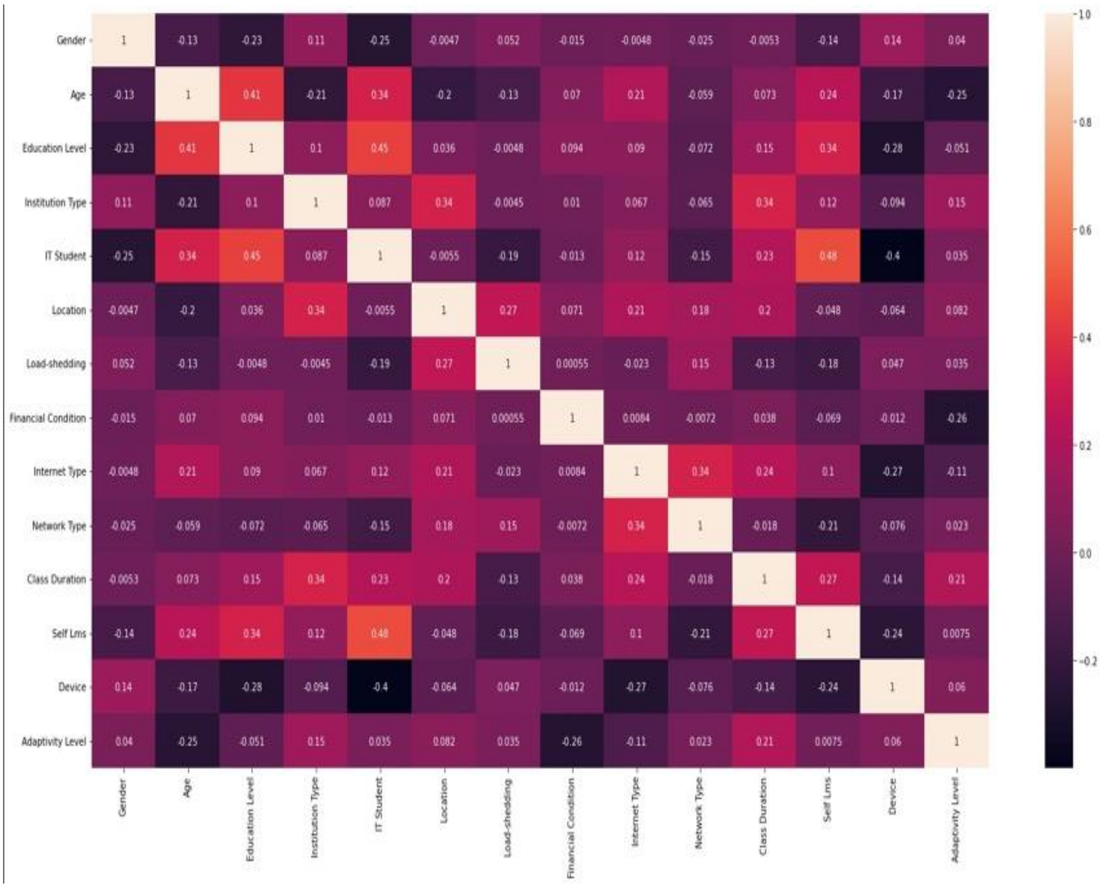

Fig. 4.3 Heatmap

The below figures [Fig. 4.4] and [Fig. 4.5] shows the distribution of data across the features Education Level and IT Student.

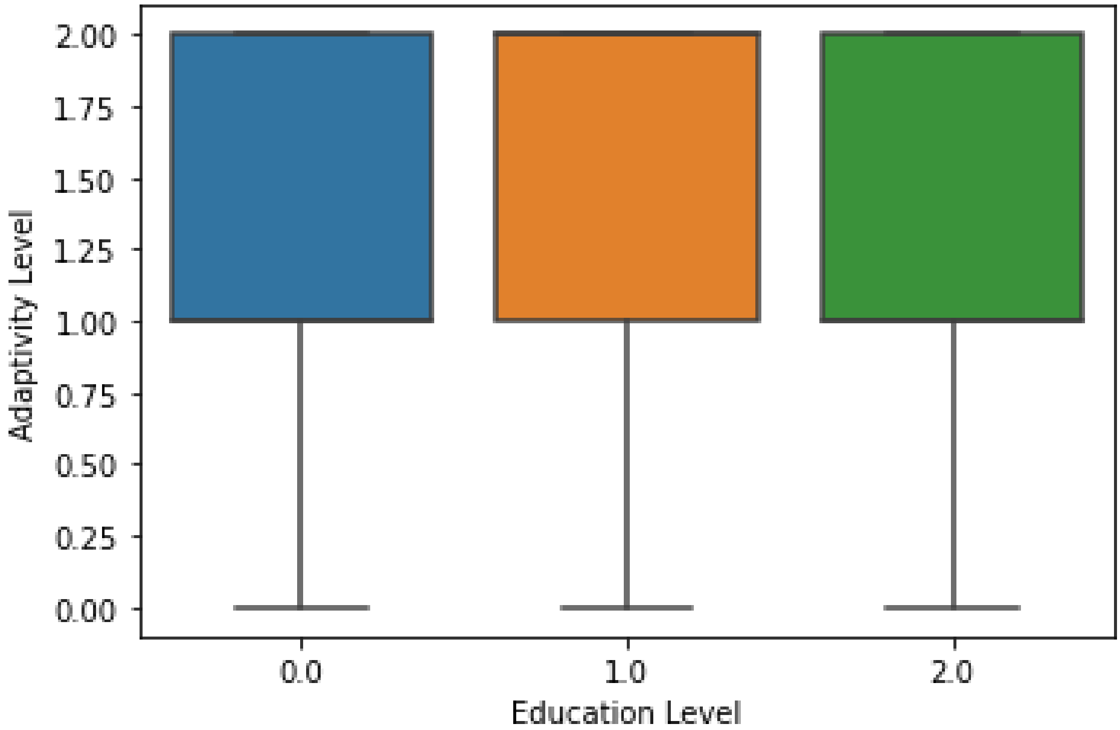

Fig. 4.4 Boxplot of Education Level

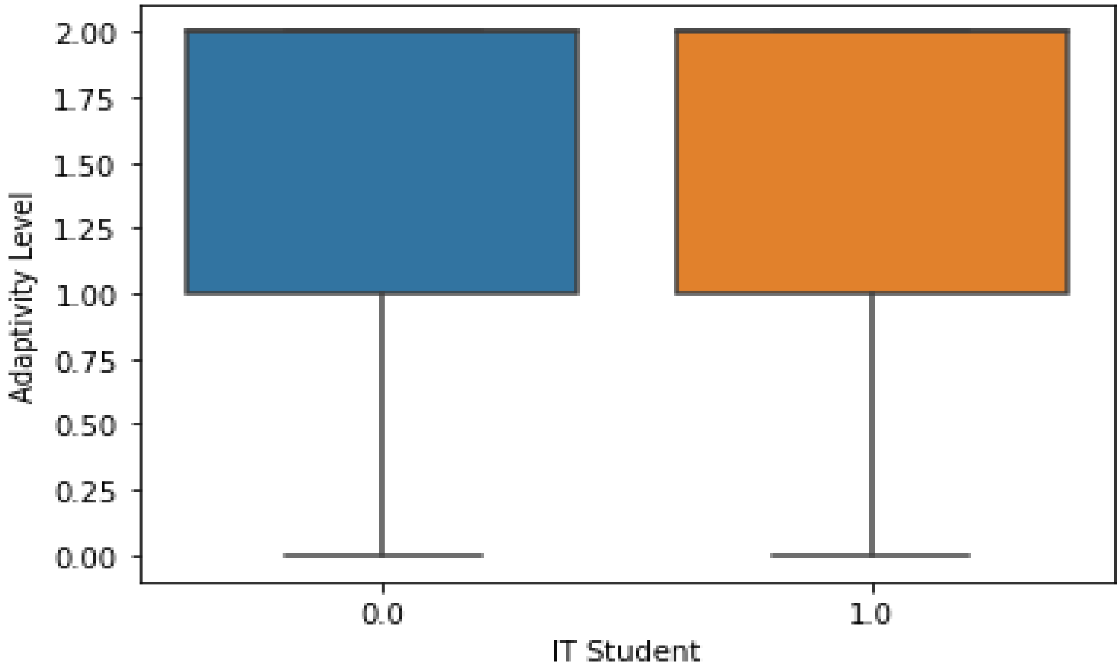

Fig. 4.5 Boxplot of IT Student

The below bar plots show the distribution of the number of students across various features [Fig. 4.6].

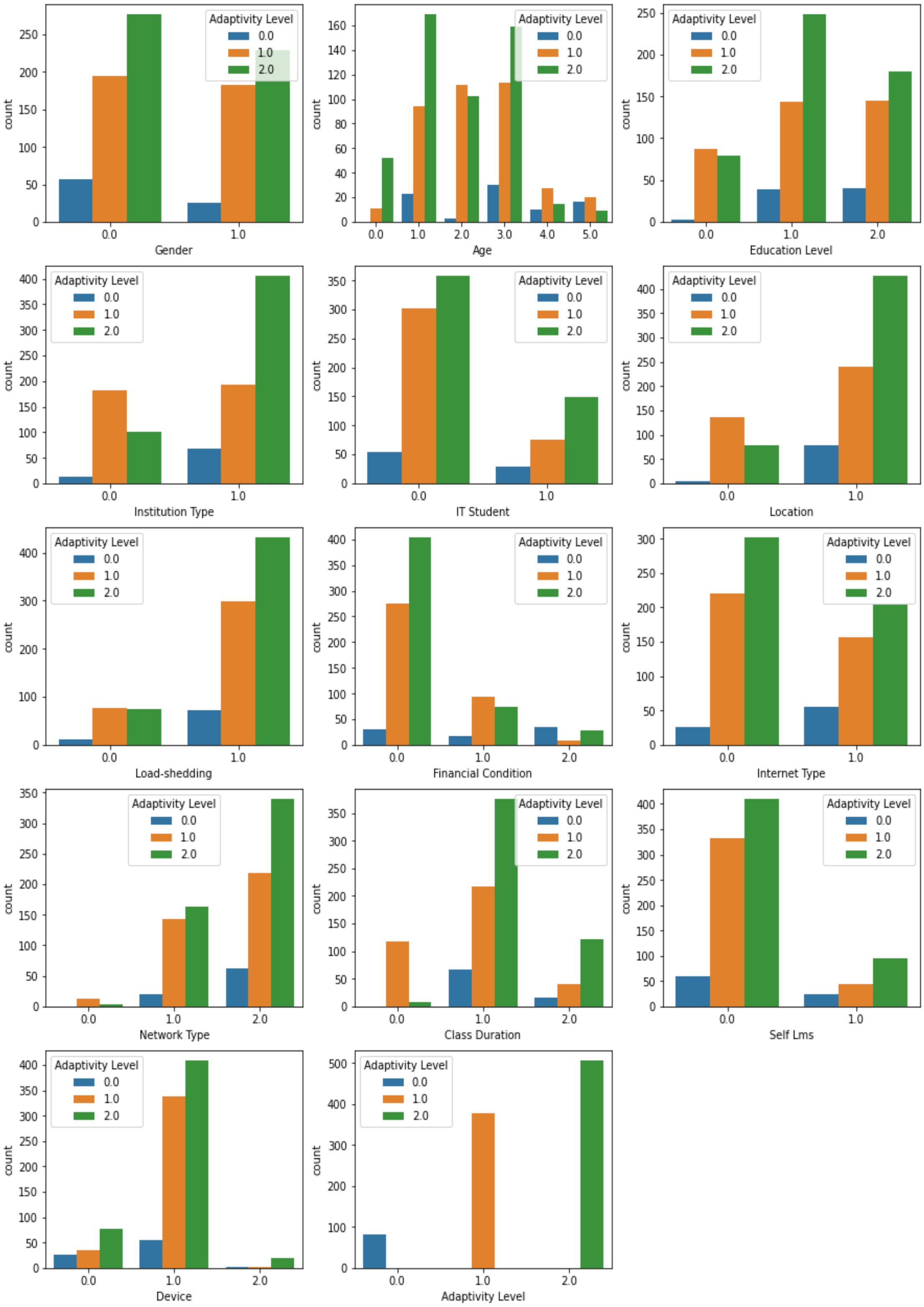

Fig 4.6 Bar Plots

The below figure [Fig. 4.7] represents the distribution of the number of students with the network type.

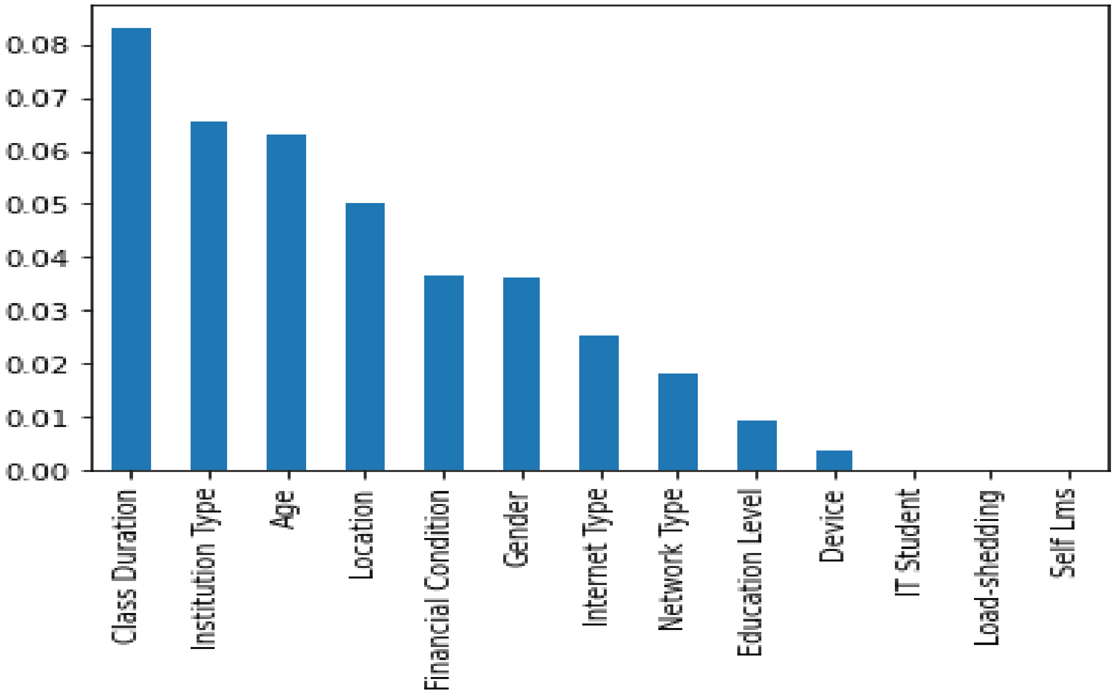

Fig. 4.7 Histogram

The below figure [Fig. 4.8] represents a scatter plot between age and gender.

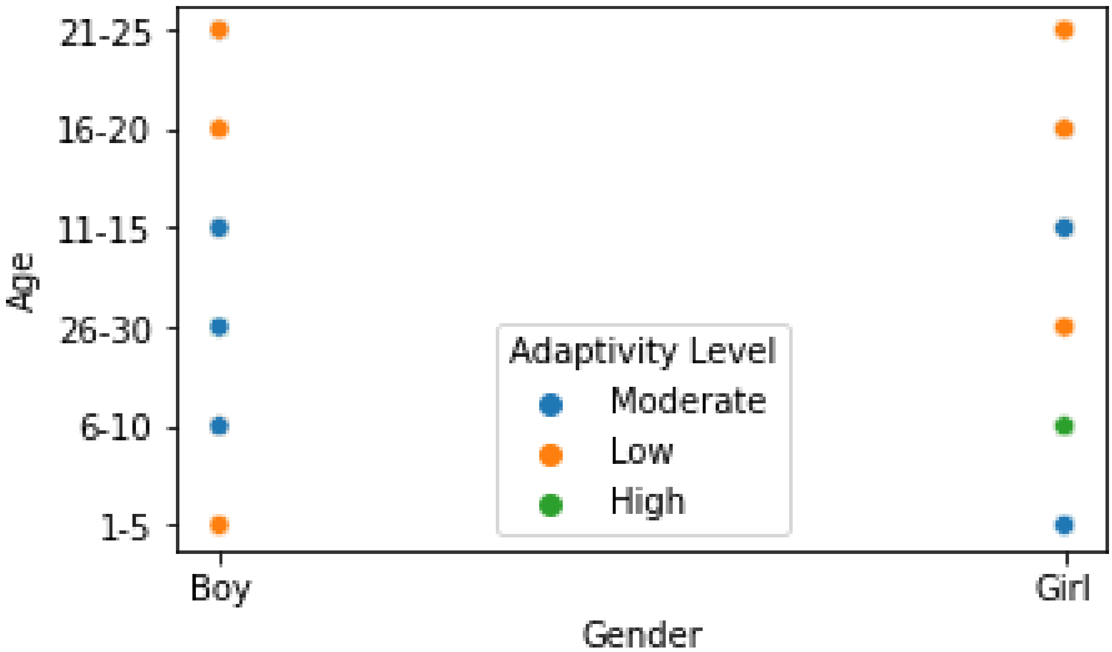

Fig. 4.8 Scatter Plot of Age and Gender

The below figure [Fig. 4.9] represents a scatter plot between education level and gender.

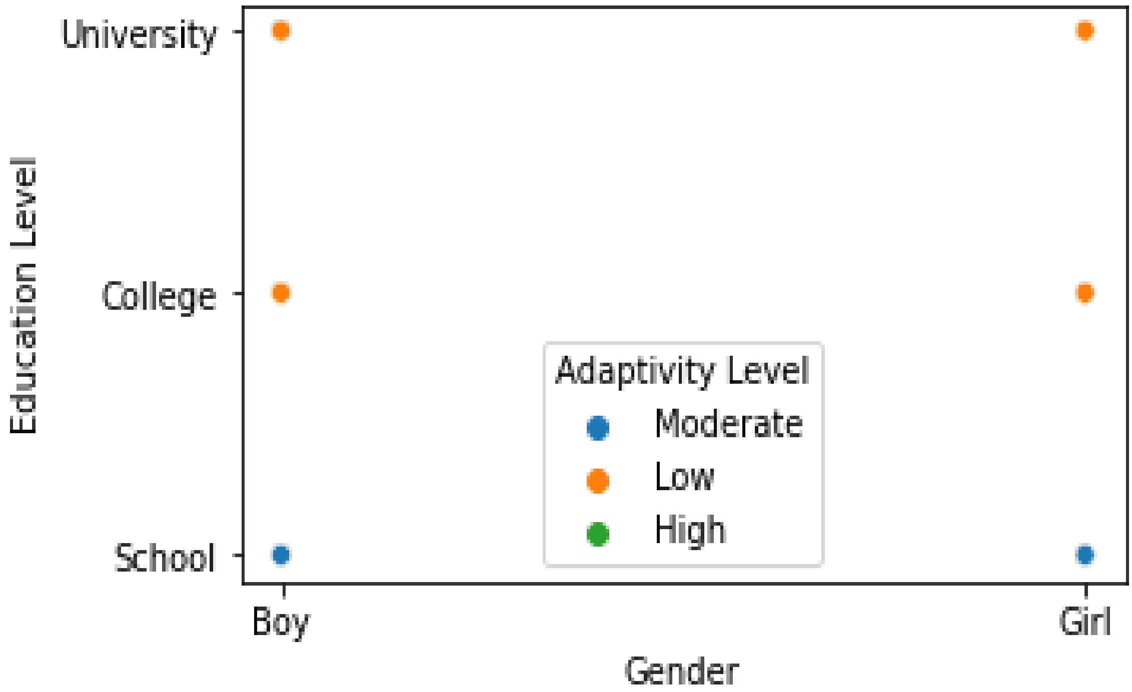

Fig. 4.9 Scatter Plot of Education Level and Gender

4.3 MODEL TRAINING/ IMPLEMENTATION

Model training is the process of using a labelled dataset to train a machine learning algorithm or model. Model training aims to train the model to spot patterns in the data and utilize those patterns to make precise predictions about brand-new, untainted data.

```
#Implementing Logistic Regression

model = lm.LogisticRegression(max_iter=100)
y_pred = model.fit(X_train, y_train).predict(X_test) #training the model

accuracy = accuracy_score(y_pred, y_test)
target_names = ['low adaptivity', 'moderate adaptivity', 'high adaptivity']

print("accuracy is: ", accuracy)
print(classification_report(y_test, y_pred, target_names=target_names))
```

```
accuracy is:  0.6473029045643154
                     precision    recall  f1-score   support

     low adaptivity       0.80      0.22      0.35        18
moderate adaptivity       0.66      0.56      0.60       104
    high adaptivity       0.64      0.79      0.70       119

           accuracy                           0.65       241
          macro avg       0.70      0.52      0.55       241
       weighted avg       0.66      0.65      0.63       241
```

Fig. 4.10: Logistic Regression

```
#Implementing Gaussian Naive Bayes
from sklearn.naive_bayes import GaussianNB
model = GaussianNB()
y_pred = model.fit(X_train, y_train).predict(X_test) #training the model

accuracy = accuracy_score(y_pred, y_test)
print("accuracy is: ", accuracy)
print(classification_report(y_test, y_pred, target_names=target_names))
```

```
accuracy is:  0.6307053941908713
                     precision    recall  f1-score   support

      low adaptivity       0.32      0.39      0.35        18
 moderate adaptivity       0.65      0.62      0.64       104
     high adaptivity       0.67      0.67      0.67       119

            accuracy                           0.63       241
           macro avg       0.55      0.56      0.55       241
        weighted avg       0.64      0.63      0.63       241
```

Fig. 4.11: Gaussian Naïve Bayes

```
#Implementing Decision Tree
from sklearn import tree
clf = tree.DecisionTreeClassifier(criterion="entropy",max_depth=10)
clf = clf.fit(X_train, y_train)
y_pred=clf.predict(X_test)
accuracy = accuracy_score(y_pred, y_test)
print("accuracy is: ", accuracy)
print(classification_report(y_test, y_pred, target_names=target_names))
```

```
accuracy is:  0.8298755186721992
                     precision    recall  f1-score   support

      low adaptivity       0.72      0.72      0.72        18
 moderate adaptivity       0.86      0.84      0.85       104
     high adaptivity       0.82      0.84      0.83       119

            accuracy                           0.83       241
           macro avg       0.80      0.80      0.80       241
        weighted avg       0.83      0.83      0.83       241
```

Fig. 4.12: Decision Tree

```
#Implementing Random Forest classifier
from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier(criterion="entropy",n_estimators = 100, max_depth =
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

accuracy = accuracy_score(y_pred, y_test)
print("accuracy is: ", accuracy)
print(classification_report(y_test, y_pred, target_names=target_names))
```

```
accuracy is:  0.8672199170124482
                     precision    recall  f1-score   support

     low adaptivity       1.00      0.61      0.76        18
moderate adaptivity       0.88      0.85      0.86       104
    high adaptivity       0.85      0.92      0.88       119

           accuracy                           0.87       241
          macro avg       0.91      0.79      0.83       241
       weighted avg       0.87      0.87      0.87       241
```

Fig. 4.13: Random Forest

```
## Implementing SVM
from sklearn.svm import SVC
clf = SVC(kernel='poly',degree=7)
clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)

accuracy = accuracy_score(y_pred, y_test)
print("accuracy is: ", accuracy)
print(classification_report(y_test, y_pred, target_names=target_names))
```

```
accuracy is:  0.8589211618257261
                     precision    recall  f1-score   support

     low adaptivity       0.62      0.72      0.67        18
moderate adaptivity       0.87      0.93      0.90       104
    high adaptivity       0.90      0.82      0.85       119

           accuracy                           0.86       241
          macro avg       0.79      0.82      0.81       241
       weighted avg       0.86      0.86      0.86       241
```

Fig. 4.14: Support Vector Machine

```
#Implementing Artificial Neural Network
from sklearn.neural_network import MLPClassifier
clf = MLPClassifier(hidden_layer_sizes=(512,256,128,64), learning_rate='adaptive',random_state=1, max_iter=1000)
```

```
from sklearn.metrics import log_loss
```

```
classes=np.unique(y_train)
train_loss=[]
test_loss=[]
for epoch in range(15):
    clf.partial_fit(X_train, y_train,classes)
    train_loss.append(log_loss(y_train,clf.predict_proba(X_train)))
    test_loss.append(log_loss(y_test,clf.predict_proba(X_test)))

epoch=range(15)
plt.plot(epoch, train_loss, label = 'train loss')
plt.plot(epoch, test_loss, label = 'test loss')
plt.xlabel("Epoch")
plt.ylabel("Loss")
plt.legend()
plt.show()
```

Fig. 4.15: Multi-Layer Perceptron Training

The below figure [Fig. 4.16] shows the relationship between training and testing loss.

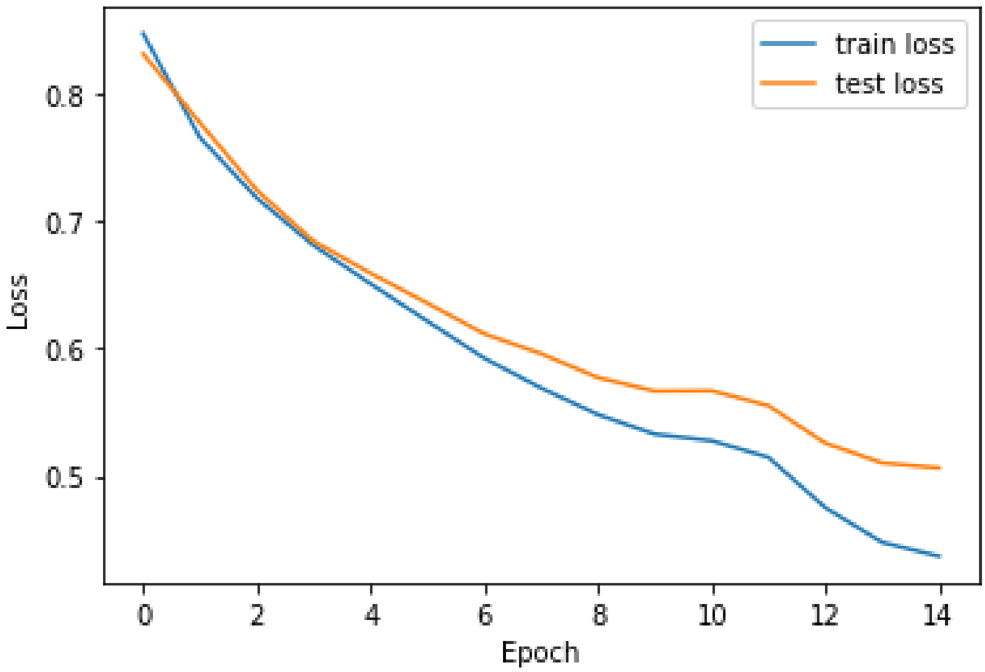

Fig. 4.16: Loss plot for MLP

```
[ ] print("Training accuracy: ", clf.score(X_train, y_train))
    print("Testing accuracy: ",clf.score(X_test, y_test))

    Training accuracy:  0.8215767634854771
    Testing accuracy:  0.8257261410788381
```

Fig. 4.17: Accuracy for MLP

```
#Implementing K-Nearest Neighbours
from sklearn.neighbors import KNeighborsClassifier
clf = KNeighborsClassifier(n_neighbors=3)
clf.fit(X_train, y_train)
y_pred=clf.predict(X_test)
accuracy = accuracy_score(y_pred, y_test)
print("accuracy is: ", accuracy)
print(classification_report(y_test, y_pred, target_names=target_names))

accuracy is:  0.8174273858921162
                     precision    recall  f1-score   support

      low adaptivity      0.87      0.72      0.79        18
 moderate adaptivity      0.85      0.77      0.81       104
     high adaptivity      0.79      0.87      0.83       119

            accuracy                          0.82       241
           macro avg      0.84      0.79      0.81       241
        weighted avg      0.82      0.82      0.82       241
```

Fig. 4.18: K-Nearest Neighbors

CHAPTER 5

RESULTS

By training the above models on the given dataset, we can say that Random Forests give the best performance amongst all the classifiers used for prediction as it has more accuracy, precision, recall, F1-score than other classifiers. Below is a table which can be used for comparison between different models.

Table 5.1- Tabular Summary of Results

Models	Class Name	Accuracy	Precision	Recall	F-1
Linear Regression	Low	69.71%	0.88	0.31	0.54
	Moderate		0.74	0.56	0.64
	High		0.66	0.87	0.75
Naive Bayes	Low	70.12%	0.58	0.61	0.59
	Moderate		0.72	0.61	0.66
	High		0.70	0.80	0.75
Random Forest	Low	86.72%	1.00	0.67	0.80
	Moderate		0.88	0.85	0.86
	High		0.84	0.92	0.88

5.1 GUI OUTPUT

The following figures in this section shows the GUI we have created with the help of ReactJs in the frontend and Flask in the backend using libraries such as Numpy, Picle, Flask, Flask_cors, React_router-dom, Bootstrap etc. After providing proper inputs in gender, age, education level etc. sections we click on “Submit" button to predict.

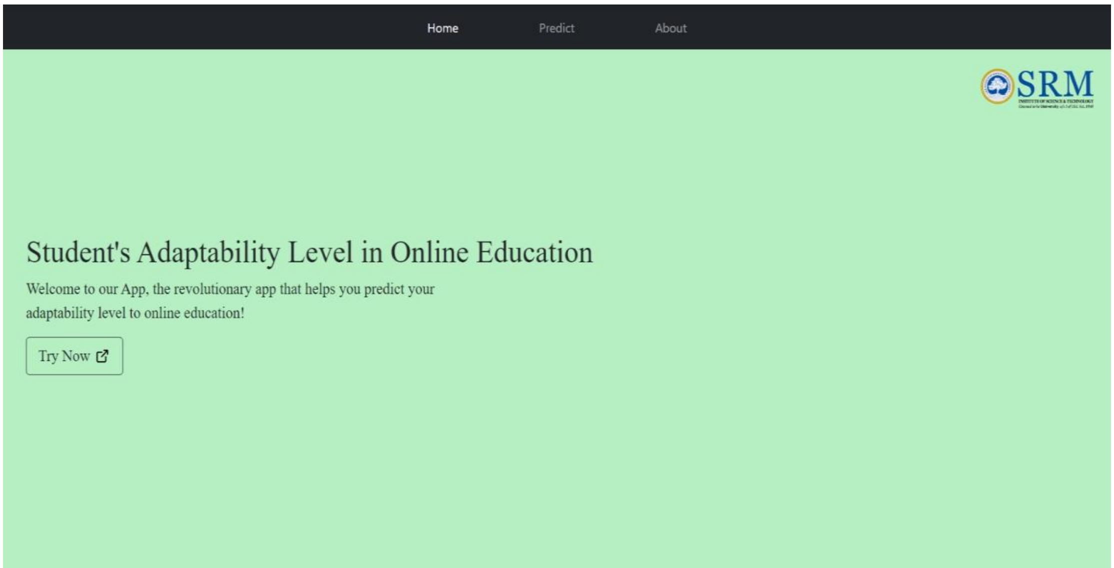

Fig. 5.1: Home page in GUI Interface

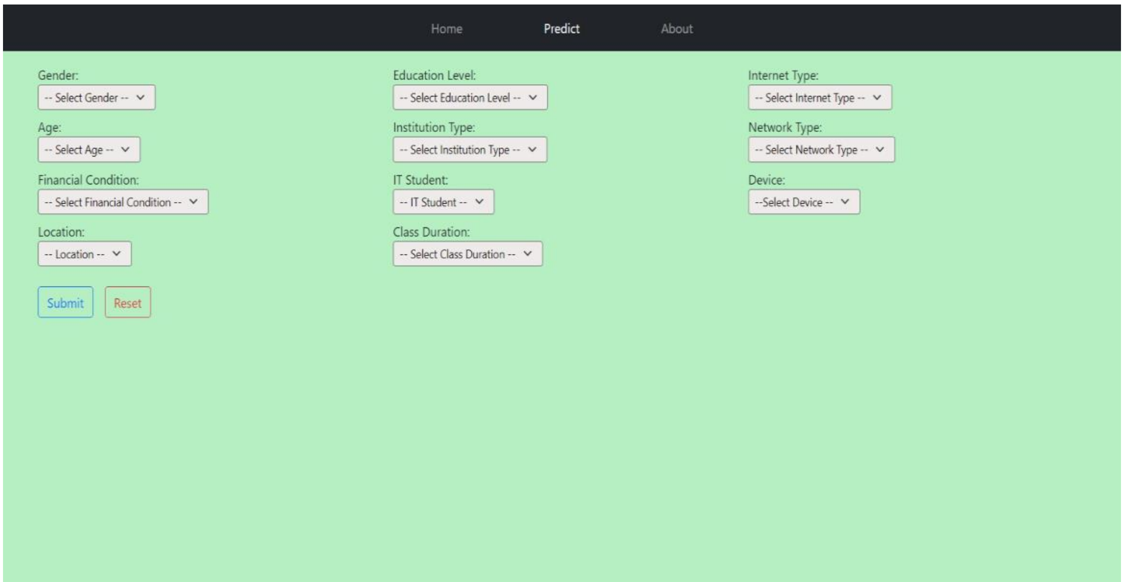

Fig. 5.2 Prediction Interface

The prediction from the model are shown below-

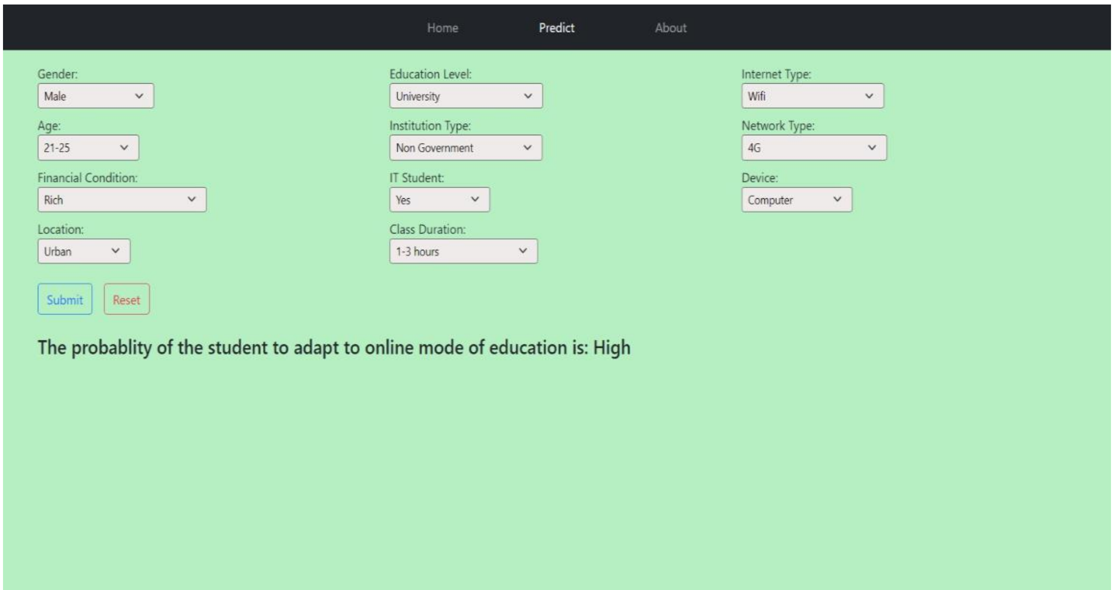

Fig. 5.3: A test case where the model predicts high level of adaptability

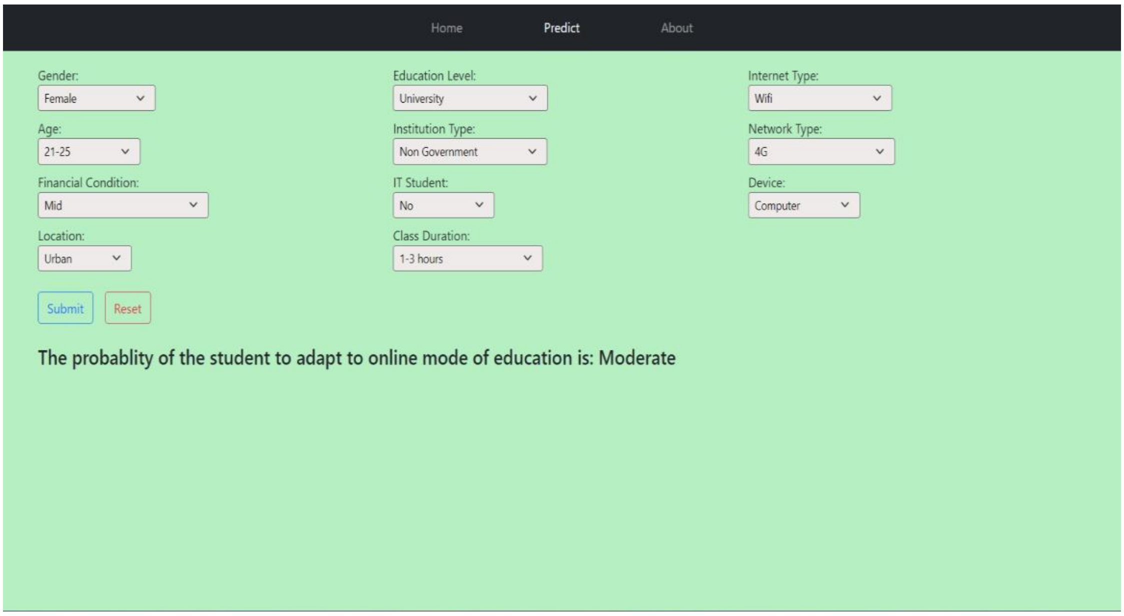

Fig. 5.4: A test case where the model predicts moderate level of adaptability

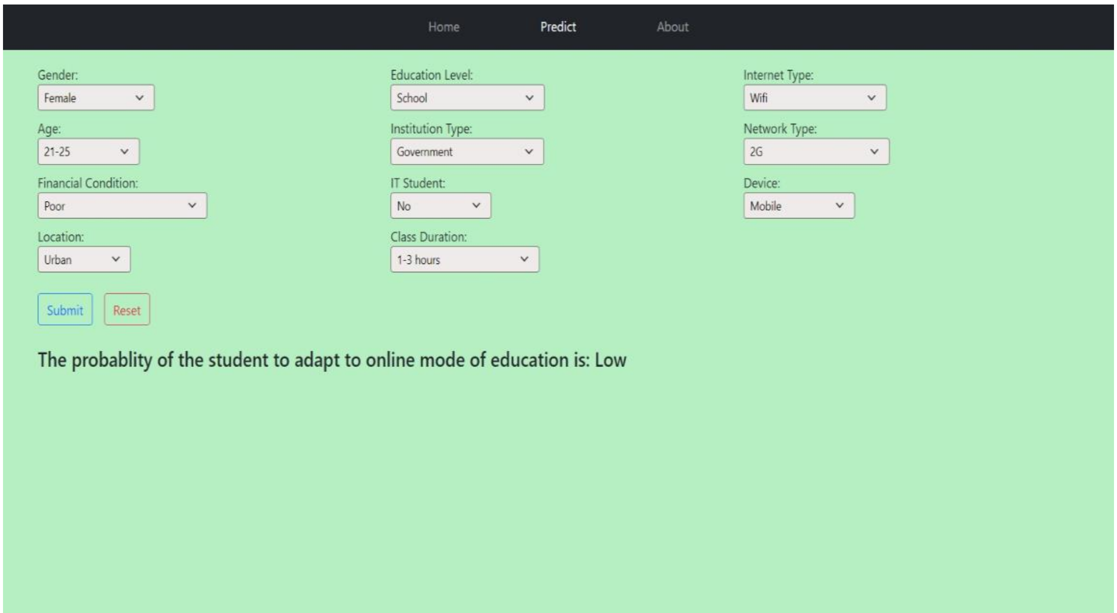

Fig. 5.5: A test case where the model predicts low level of adaptability

CHAPTER 6

Conclusion and Future Scope

The goal of this study is to predict how adaptive students will be to learn via the internet using a variety of machine learning models. This project will continue to employ models like Logistic Regression, Naive Bayes, Random Forest, KNN, SVM, and ANN. The results show that Random Forest outperforms the other models

The significance of this research resides in its ability to assist those in charge of making educational decisions in raising the standard of instruction for pupils. The capacity to forecast a student's level of adaptation in an online learning environment allows decision-makers to create individualized learning plans that are tailored to the needs of each individual learner. For instance, if a student is anticipated to have a low degree of adaptability, the decision-makers can provide them with more resources, such individualized coaching or study schedules, to assist them succeed in online learning.

The project also emphasizes the value of adopting ensemble learning methods for categorical data, such as Random Forest. Random Forest is a great option for this project since it employs many decision trees to produce precise predictions and decision trees function well with categorical data. The application of several machine learning models also highlights the significance of considering various techniques and choosing the most successful one.

In conclusion, this project offers a method to forecast students' level of adaptation in online education using machine learning models, which makes a significant addition to the field of education. Decision-makers can help students succeed and raise the standard of education by using predictions to personalize the learning experience.

REFERENCES

[1] **Dataset**: https://www.kaggle.com/datasets/mdmahmudulhasansuzan/students-adaptability-level-in-online-education

[2] D. Xu and S. S. Jaggars, "Adaptability to online learning: Differences across types of students and academic subject areas," 2013.

[3] A.-M. Nortvig, A. K. Petersen, and S. H. Balle, "A literature review of the factors influencing E-learning and blended learning in relation to learning outcome, student satisfaction and engagement," Electron. J. E-Learn., vol. 16, no. 1, pp. 46–55, 2018.

[4] P. Xhelili, E. Ibrahimi, E. Rruci, and K. Sheme, "Adaptation and perception of online learning during COVID-19 pandemic by Albanian university students," IJonSE, vol. 3, no. 2, pp. 103–111, 2021.

[5] R. E. Baticulon et al., "Barriers to online learning in the time of COVID-19: A national survey of medical students in the Philippines," Med. Sci. Educ., vol. 31, no. 2, pp. 615–626, 2021.

[6] S. Korlat et al., "Gender differences in digital learning during COVID-19: Competence beliefs, intrinsic value, learning engagement, and perceived teacher support," Front. Psychol., vol. 12, p. 637776, 2021.

[7] Z. Yu, "The effects of gender, educational level, and personality on online learning outcomes during the COVID-19 pandemic," Int. J. Educ. Technol. High. Educ., vol. 18, no. 1, p. 14, 2021.