




# Impact of Covid-19 on Bangladeshi university students mental health: ML and DL analysis

Md Monir Ahammod Bin Atique<sup>1</sup>, Md Ilias Bappi<sup>1</sup>, Kyungbeak Kim<sup>1</sup>, Kwanghoon Choi<sup>1</sup>, Md Martuza Ahamad<sup>2</sup>, Khondaker Masfiq Reza<sup>3</sup>

**1** Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju, South Korea.

**2** Department of Computer Science and Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj-8100, Bangladesh.

**3** Department of Computer Science and Engineering, Brac University, Dhaka, Bangladesh.

 These authors contributed equally to this work.

\* Corresponding author:

E-mail: monir024@jnu.ac.kr

E-mail: kyungbaekkim@jnu.ac.kr

E-mail: i\_bappi@jnu.ac.kr

## Abstract

The Covid-19 outbreak has adversely influenced university students across the world both physically and psychologically. The psychological struggle faced by students, is effected by various factors, including disruptions to daily routines and academic activities, increased reliance on smartphones and the internet, limited social interaction, and confinement to their homes. These impediments reflect a broader issue of imbalance in cognitive health status among them during the pandemic. In Bangladesh, despite having the necessary population to study, understanding the impact of Covid-19 on the mental health status of university students has been limited. Hence, it is imperative to diagnose mental health issues and deal with the underlying reasons in order to enhance students' psychological well-being, which leads to academic achievement. Nowadays, Artificial Intelligence (AI) based prediction models have the potential to play a crucial role in predicting mental state early. The purpose of the study is to explore the following effects of the pandemic on the mental health of Bangladeshi university students using Machine Learning (ML) and Deep Learning (DL) techniques. A reliable AI prediction system requires real-world data, that was collected by a survey through a Google form (online questionnaires) among 400 students of 16 universities, and the respondents were 253. In this paper, after data preprocessing, ten widely known ML and four DL models were developed to automatically and accurately predict mental well-being during or after the Covid-19 circumstance. According to our findings, the Random Forest (RF) algorithm and Siamese Neural Networks (SNNs) outperformed other models in terms of accuracy (86% and 75%). Additionally, Chi-Square test was conducted, which revealed the five most common and significant predictors ("Stable family income", "Disruption of daily life", "Own income", "Sleep status", and "Fear of getting infected with Covid-19") of psychological health conditions. Overall, this work could assist university administrations, government agencies, and health specialists in taking appropriate measures to understand and maintain students' mental health. This research also suggests proper monitoring, government support, and social awareness during and after the worldwide epidemic to keep an excellent mental health state of university students.

## Introduction

Mental health refers to general psychological condition of any individual. It is also defined as general ability to deal with people, stresses as well as one's contribution in the community [1]. Several reasons have been identified through numerous studies [2–4] regarding the importance of maintaining good mental health conditions in daily life. One of the most significant reasons is that every person has a direct link between his or her physical and mental health. According to medical science, a patient with a chronic disease is said to have a close connection to the possibilities of developing depression [5]. Studies also show correlations among mental health well-being and social relationships within families and communities [6]. Therefore, it is obvious that mental health status has a direct impact on individual personnel's and social development. A survey shows that individuals with balanced mental health are more keen on personal advancement through appropriate planning and intentional engagement [7]. However, healthy psychological state is affected when a global pandemic arises.

In 2019, people were first introduced to the Novel Coronavirus (Covid-19) disease, which was derived from SARS-Cov-2 (Severe Acute Respiratory Syndrome Coronavirus-2). Before December 31, 2019, the globe had never faced such a disease. It contributed to lung injury, cardiac injury, and even death. The World Health Organization (WHO) declared Covid-19 as a global pandemic on March 11, 2020. During the mass contagion, social isolation was implemented by the governments of

many countries. This outbreak had reached Bangladesh as well. As of 2020, Bangladesh was the ninth most densely populated country and territory in the world [8]. Considering this huge population, the Bangladeshi government also implemented lockdowns all over the country to protect against the devastating scenario. As a result, many young adults have had to face a new kind of environment where going out of the home and maintaining social relationships have become challenging or impossible, leading to a new circumstance called “social isolation”, as mentioned earlier. This situation had a great impact on people’s mental health, especially on those who already had a mental illness before the health disaster. Social isolation may lead to loneliness which can make a person more vulnerable to ill mental conditions like depressive symptomatically, higher suicide attempts, and completed suicide [9]. This may also lead to dementia - which refer to not being able to think, remember, make decisions for everyday activity [9]. As a strong linkage between mental health and social isolation, it is found that the longer the isolation may take place, people become more mentally sick and more vulnerable to depression and anxiety [10]. Young people are the ones who suffer the most from this isolation.

University students are similar to other young adults in acknowledging the importance of addressing mental health needs. In other words, they are not unique in this regard; they share the same awareness of the importance of mental health as other young generation. This infectious disease crisis severely affected their academic and social activities. During this time, students stayed home and faced several problems, such as loneliness, career concerns, and financial issues. They can’t meet with their friends and relatives physically, which affects many aspects of their mental health. Most of the time, they are isolated, and sometimes they feel negative thoughts, which sometimes cause depression and anxiety. Many university students had medium to high levels of stress and suicidal thoughts at that time. A research found that students with poor mental conditions have anxiety and depression, which in some cases went as high as clinical cut-offs [11, 12]. This type of psychological degradation is less visible except during the viral outbreak.

Over the decades, there has been increasing worldwide recognition of the adverse effects of stress, anxiety, and depression on individuals, institutions, and communities [13]. In that ongoing context in recent years, Covid-19 has imposed a negative impact on mental well-being, especially on university students. Therefore, a group of studies has been conducted globally to figure out the efficient method that can aid the respected authority in identifying whose mental state is poor or better including severe levels of anxiety and depression symptoms. Many previous works have been taken into account before moving forward with this study. Most of the studies [14–22] focused on physiological and statistical techniques for exploring the factors such as gender, education status, etc which have a relation of being mentally sick. However, most of the above studies did not concentrate on AI-driven methods such as ML, DL, and Data Mining. Similar to the other fields, AI technologies can be used in the field of mental health for quite some time. Various ML and DL techniques [23–37] have been utilized to analyze the mental health condition of any individual with descent accuracy. Different mainstream algorithms such as Decision Tree (DT), RF, Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), Extreme Gradient Boosting (XGB), Gradient Boosting Classifier (GBC) and Artificial Neural Networks (ANNs) including Convolutional Neural Network (CNNs) and Recurrent Neural Networks (RNNs) have been developed several times for the prediction. Nevertheless, very few studies have been conducted exploring the university (undergraduate, graduate, and Ph.D.) students data. Although Covid-19 has a considerable influence on the psychological well-being of students at higher education institutions, there is comparatively less research on predicting mental health states. Particularly in highly

populated countries like Bangladesh, there exists a research limitation in developing practical and reliable studies on university students. While there were several studies [13,38] in Bangladesh, there's still plenty of room to improve both the statistical analysis of features and the application of DL techniques. Generally, research on mental health status is not given the same level of attention as their physical health in Bangladesh. Another limitation of the existing research is not using real-world datasets, and that is why there is enough scope for improvement in the data collection methods.

The main goal of our study is to build an AI-driven system including proper analysis of features where we utilized our primary dataset which is based on real feedback on Covid-19's impact from Bangladeshi university students. During the experiment, we collected data through an online questionnaire and trained not only ten well-known ML models such as RF, CatBoost, XGB, GBC, SVM, DT, K-Nearest Neighbors (K-NN), Light Gradient Boosting Machine (LGBM), AdaBoost, and Multinomial Naive Bayes (MNB) but also four DL techniques including CNNs, RNNs, Hybrid Neural Network, and SNNs. These models were then assessed on testing subset how well they predicted the students' well-being status (Poor or Better). Each model was evaluated against performance metrics. Finally based on this evaluation, best-performing model (RF and SNN) was determined. Additionally, we also performed Chi-Square test as a statistical associativity of features which contributed to figuring out the most associative attributes influenced on mental well-being. Through this test, we found five highly significant factors ("Stable family income", "Disruption of daily life", "Own income", "Sleep status", and "Fear of getting infected") that greatly influenced psychological health conditions during this epidemic of Bangladeshi higher education students.

The following are primary contributions of our proposed research work in the context of mental health prediction:

- A real-world dataset was collected through an online survey that have a huge impact on the student's mental health status.
- The descriptive data analysis was applied to obtain valuable insights from the dataset. The factors that affect mental health was examined thoroughly.
- Performed a statistical test of the features to understand how these features contribute to the level of mental health status.
- To accurately predict student's mental health labels (Poor or Better) using AI-driven classification models.
- To determine the adverse impact of social isolation on university students mental and psychological well-being during the Covid-19 outbreak.
- Our findings can assist government agencies, university authorities, and healthcare practitioners to ensure university students' mental fitness and provide them with practical solutions to maintain a healthy educational experience.

The rest of the paper is organized as follows: Related Works describes the similar works. Then, Research Methodology discusses the methodology and presents a brief description of the dataset, descriptive data analysis of the features, preprocessing tasks, feature selection and ranking, proposed models, and evaluation criteria of the applied models. Experimental Results Analysis shows Chi-Square test, ML, and DL result analysis. Discussion presents the general discussions of the whole study, limitations and strengths. At last, Conclusion demonstrates the concluding statements of this study, and mentions the direction of future research.

## Related Works

119

Mental health issues encompass a broad range of conditions that affect thoughts, emotions, and behaviors, thereby impacting an individual's overall well-being. These conditions may create anxiety, depression, stress, or other forms of psychological imbalance that affect daily life. Covid-19 significantly impacted mental health during the pandemic, prompting many studies to be conducted both during and after this period. Below, we discuss some relevant research works pertaining to our study.

120  
121  
122  
123  
124  
125

## Mental Health Status, Anxiety, and Depression Levels

126

Previously related works [14–16] found many factors such as age, gender, family type, marital status, education, and employment status that directly contribute to mental health issues among the students doing higher studies. Many studies conducted investigation ground on statistical surveys in the United States and China [17–20] reveals the adverse impact of Covid-19 prevalence on mental health especially amongst college or university students. Their findings found the existence of poor mental health among most university students during the pandemic period. In China, 24.9% Chinese university students faced anxiety situations in that time [17]. Another research conducted in the United States [18] colleges during the Covid-19 pandemic revealed that 24.5% of 1881 students experienced depression (PHQ-9), while 20.7% reported anxiety (GAD-7). Additionally, the experiment highlighted the adverse effects of remote learning on personal relationships and identified poor sleep quality as a contributing factor to higher risks of depression and anxiety. Similar to other countries around the world, Covid-19 has had a devastating effect on Bangladesh's education sector. On March 17, 2020, the government of Bangladesh closed all schools, colleges, and institutions in the country due to the presence of eight verified cases to avoid these devastating diseases. According to statistics [39]. in Bangladesh, approximately 1 million university students are doing their bachelor (undergraduate), graduate, and PhD studies at 53 public, 107 private, and 3 international universities. Some researchers in Bangladesh did a few experiments on the impact of Covid-19 on university students' mental well-being and found mild to severe-level depression symptoms. One study [21] found severe anxiety symptoms amongst students who are living with families in urban areas. Moreover, the author reported that nearly 88% of students faced mild to severe anxiety symptoms throughout the pandemic. Another survey [22] showed that approximately 69% and 47% of students reported mild to severe levels of psychological impact and mild to extremely severe levels of depression, respectively due to the pandemic.

127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152

## Using ML and DL Classifier on Mental Health Prediction

153

Since the outbreak, epidemiological data in Bangladesh has revealed that mental health issues due to the Covid-19 and mass isolation are prominent [24]. Early and after the pandemic, one contributing factor to these mental health implications for Bangladeshis was found to be fear [25]. Diverse studies have reiterated that Covid-19 related worries and fears for Bangladeshi Diverse studies have reiterated that Covid-19-related worries and fears among Bangladeshi individuals are associated with increased mental health issues. Thus, collectively, these findings shed light on how the pandemic can affect the mental and psychological well-being of university students in Bangladesh.

154  
155  
156  
157  
158  
159  
160  
161

In parallel, many researchers have investigated different factors affecting students' psychological well-being during the pandemic. Cross-national research [26] and surveys [38] conducted among university students in Bangladesh. Negative perceptions and poor mental health were found to be linked during the Covid-19 outbreak in that region [40]. During the outbreak in April of 2020, a study examined how Bangladeshi

162  
163  
164  
165  
166

167 university students' mental health, anxiety, and depression levels changed as the  
168 eruption advanced [27]. University students, according to their findings, had a  
169 significant proportion of anxiety and depression symptoms and moderate to poor  
170 mental health ratings, based on the cut scores for their individual measurements. Early  
171 in the prevalence, 65.9% of Bangladeshi medical students reported low levels of worry or  
172 depression [28], and 3.3% of individuals exhibited significant symptoms of depression.  
173 Furthermore, the use of social media as a coping mechanism during this worldwide  
174 pandemic has been observed [37], with distinct gender differences in coping strategies  
175 noted.

176 In the realm of applying ML on mental health studies, Reshmy Krishnan et al.  
177 worked on engineering and medical students data during Covid-19 using the Hospital  
178 Anxiety and Depression Scale (HADS) assessment [29]. They conducted an online  
179 survey using a Google form which contains questionnaire, reaching 120 medical and 152  
180 engineering students respondents. They applied different ML algorithms like SVM,  
181 which is a robust handling method of health related data, and J48 DT. The SVM  
182 performed exceptionally well in their investigation. However, the authors only used an  
183 online survey through social media, and they couldn't take face-to-face interviews with  
184 the students. Similarly, Hanif Abdul Rahman et al. [30] explored a dataset of 17  
185 Southeast Asian university students. Highlighting RF and adaptive boosting as optimal  
186 algorithms, the research identifies key predictors such as "sports activities", "body mass  
187 index", "GPA", "sedentary hours", and "age". The dataset was gathered through a  
188 cross-sectional survey methodology, and it does not allow for the establishment of causal  
189 relationships. Even though the accuracy is satisfactory but still room for improvement  
190 in data collection. Mostafa Rezapour and Lucas Hansen [31] applied several ML models  
191 and statistical techniques such as Synthetic Minority Oversampling, and Chi-Square test  
192 to analyze the mental health condition based on Covid-19 mental health data. They  
193 focused on a dataset from a mental health survey conducted during the Covid-19  
194 isolation period among front-line workers. Roles such as nurses, emergency room staff,  
195 and surgeons had to undergo unique situations like high pressure, long hours of duty,  
196 the possibility of trauma, and emotional strain, which contributed to mental health  
197 challenges. However, this type of research can be applied to the mental health  
198 conditions of university students as well.

199 Every day, around 2.5 quintillion bytes of data are created on social media.  
200 Examining this vast amount of data can be a valuable way to assess the impact of  
201 Covid-19 on mental health well-being [32]. The authors, Choudhury et al. [33] collected  
202 a data set of Twitter users who are diagnosed with depression by crowd-sourcing and  
203 developed ML models for predicting the risk of depression. They also used "Social  
204 engagement", "Prescribed medicines", "Ego network", and "Linguistic style" of the user  
205 as features in their paper. To avoid overfitting, they used Principal Component  
206 Analysis (PCA) and used a SVM classifier with the radial basis function (rbf) kernel for  
207 depression prediction. Tianhua Chen [34] investigated the mental health status of UK  
208 university students during the plague. This analysis aims to create a framework using a  
209 method called feature permutation importance. The goal was to make different ML  
210 methods work together, even those that can't directly tell how important each factor is.  
211 This way, it can look at how different factors affect students' mental well-being from  
212 various angles, going beyond the usual focus on regression. The dataset used this  
213 approach in an online survey at a UK university and found that factors like "Past  
214 medical history", "Well-being history", and "Facing adversity" are significant for mental  
215 well-being. Additionally, keeping good relationships through regular communication  
216 with family and friends, along with regular exercise, generally contributes to better  
217 mental well-being.

218 Few studies closely support our work. Among them, Madhuri Mahalingam et al. [35]

presented a investigation to forecast anxiety levels of the university students campuses in Lebanon. Using ML based techniques they experimented on student survey items, such as demographics and self-rated health, to predict anxiety symptoms among the sample of 329 respondents. The best-performing model, Multi-Layer Perceptron (MLP) obtained the highest AUC score, and the most highly ranked feature to predict anxiety was determined to be self-rated health. They collected the dataset from one public and one private Lebanese university. However, this research could be improved by adding a wide range of university students' mental well-being data. Besides, Ishrak Jahan Ratul et al. [13] proposed several supervised ML algorithms to examine university students' perceived social and psychological stress among 444 university students real-world data which is also a closely related study to our work. They used some tuning and statistical techniques to get better results such as PCA, Chi-Square test, Grid Search Cross-Validation (GSCV), and Genetic Algorithm. Results showed that 11.26% of individuals had high social stress, while 24.10% had extremely high psychological stress. The MLP model showed the highest accuracy when combined with PCA and GSCV. The accuracy could be improved considering large data samples and this experiment also could be extended for predicting the social and psychological stress of university students after the pandemic. In additional research, El-Sayed Atlam et al. [36] utilized statistical and ML methods to investigate the impact of the Covid-19 pandemic on the psychological well-being of university students in three Arab countries: Egypt, Jordan, and Saudi Arabia. A questionnaire comprising demographic details and five primary dimensions - digital device usage, sleep patterns, social interactions, emotional state, and academic performance was distributed among students. Analysis of 1766 responses revealed a clear link between students' psychological health and online education during the pandemic. However, due to the small sample size, the dataset may be biased, limiting the effectiveness of the ML model's performance.

Shifting focus from the board of ML methodologies, the domain of DL also has a profound footprint on mental health prediction. The researchers [32] used tweet data from Twitter to determine the impact of Covid-19 on people's psychological well-being, with a particular emphasis on depression. A novel pipeline was proposed based on recurrent neural networks (in the form of long short-term memory) and a CNN, which is capable of identifying depressive tweets. Additionally, another observation [41] demonstrated the design of a DL-based mental health monitoring scheme (DL-MHMS) for students. In this survey, they used the most efficient CNN technique to classify mental health status as positive, negative, and normal based on the EEG signals collected from college students. In another examination, the authors [42] introduced a model focused on predicting anxiety levels during an epidemic. They utilized a Gated Recurrent Unit (GRU), a variant of RNN, to develop a model capable of accurately forecasting anxiety levels. The dataset they utilized was created by collecting relevant tweets, from which they extracted the top hashtags mentioning Covid-19 and quarantine in Arabic. Furthermore, recent exploration has shown that there has been limited exploration of social media data employing DL techniques to predict well-being status among college students and others. However, there is a noticeable gap in research focusing on mental health problems among university students using DL methods. Therefore, we employed this in our research experiment, as it tends to excel when dealing with categorical data exhibiting similarities, mirroring the characteristics of our dataset.

Overall, the importance of understanding the psychological well-being of students during the Covid-19 pandemic is underscored by a variety of studies employing diverse methodologies. Studies conducted in the United States, China, Lebanon, and Bangladesh have revealed the significant adverse impact of the prevalence on mental health among college or university students. These studies collectively contribute to our

understanding of the psychological well-being of students. While ML and DL approaches offer powerful tools for analysis, other methodologies also provide valuable insights into the challenges faced by students in adapting to the new norms imposed by the pandemic. Further research in this area is essential to developing comprehensive prediction tools for supporting students mental health in times of crisis.

## Research Methodology

This section covered a detailed explanation of how the proposed methodology has been carried out. Firstly, raw data is collected from university students through online survey and then analyzed the data descriptively to produce meaningful insights. After analyzing the raw data, we moved on to the next process, which was the data-preprocessing step. Following data pre-processing, we performed "Feature Selection and Ranking" process. Then we developed ML and DL models to predict the mental health status of the students. Finally, all the models are assessed using performance evaluation metrics. The whole research methodology is depicted in Fig 1.

## Dataset Description

This research focuses on university students between the ages of 19 and 30, who are the most important contributors to this primary dataset. From December 4, 2021, to December 13, 2021, the survey was conducted in 16 renowned universities of Bangladesh through a Google form among 400 students. There are 15 distinct questionnaires in that form and the response was 253 (Records) which refers to 63% response rate. There were 213 male students and 40 female students among those who responded. Dataset descriptions are given in Table 1. This dataset has the following characteristics:

- This primary dataset is also a standard dataset. It is utilized as a binary classification problem.
- In the output column, there are two levels: "Better" and "Poor". "Better" refers to any student who has a stable mental health state. Contrarily, students having low mental well-being state included as "Poor" level.
- Some features of the dataset are "Gender", "Stable family income", "Worry about academic delay", and so on.
- Dataset has 253 students entries with 15 attributes.
- In this primary dataset, 213 are male and 40 female students.

## Descriptive Data Analysis

The research revealed that fourteen features in total influence mental health status significantly. To present a proper idea of those features and how these features correlate with the label "Mental Health Status," the statistical graph is depicted in Fig 2A, Fig 2B, and Fig 2C. Besides, a brief description with intuitive analysis is enlisted below for better understating. In this comparison, we examine the "Poor" mental health status between the "Yes" and "No" categories of features, including "Disruption of daily life", "Worry about academic delay", "Exercise during Covid-19", "Worry about social support during Covid-19", "Stable family income", "Fear of getting infected with Covid-19", and "Own income".

**Sleep status hours during Covid-19:** In our analysis, we discovered that if any university student had a sleep status (hour) in the period of Covid-19 between 1.5 and



**Table 1.** Dataset features with description

Feature	Data Types	Description
Gender	Categorical	Gender of the students (1 = Male, 0 = Female)
Worry about social support during Covid-19	Categorical	(1 = Yes, 0 = No)
Residence	Categorical	(1 = Urban, 0 = Rural)
Stable family income	Categorical	(1 = Yes, 0 = No)
Living with parents	Categorical	(1 = Yes, 0 = No)
Relative or acquaintance infected with Covid-19	Categorical	(1 = Yes, 0 = No)
Worry about academic delay	Categorical	(1 = Yes, 0 = No)
Disruption of daily life	Categorical	(1 = Yes, 0 = No)
Exercise during Covid-19	Categorical	(1 = Yes, 0 = No)
Own income	Categorical	(1 = Yes, 0 = No)
Quarrel between parents	Categorical	(1 = No, 0 = Absent - did not live with parents, 2 = Yes)
Sleep status hour during Covid-19	Numeric	1 = Less than normal, 2 = Normal ( 7-9 hours ), 3 = More than normal
Internet browsing hours during Covid-19	Numeric	1 = less than 2 hours, 2 = 2-4 hours, 3 = 5-6 hours , 4 = greater than 6 hours
Fear of getting infected with Covid-19	Categorical	(1 = Yes, 0= No)
Mental Health Status	Categorical	(0 = Better , 1=Poor)

2.5, then his or her mental health status was better. The value range indicates normal (7-9 hours) or slightly less than normal (5-7 hours) sleeping hours in a day. 314 315

**Internet browsing hours during Covid-19:** Looking at the internet browsing hours of a university student, we found that if it is greater than 3 (means greater than 6 hours), then mental health status is inferior. In our exploration, we identified certain circumstances in which a student’s mental health state improves through the relaxation of their mind using the internet. 316 317 318 319 320

**Disruption of daily life:** Mental health status is considerably poor when disruption of life is rated “Yes”. More number of students estimated “Better” mental health state, despite the “Disruption of daily life,” features marked as “Yes”. There could be other reasons for this different scenario. 321 322 323 324

**Worry about the academic delay:** There are a considerable number of students whose mental health status is “Poor”. The reason behind this is that they are worried about academic delays. Even though a high number of students rated their mental health status as “Yes” while considering the feature “Worry about the academic delay”. There might be other features that are responsible for rating their mental health status as “Better”. 325 326 327 328 329 330

**Quarrel between parents:** If there are no quarrels between parents, then the 331

mental health status of a student is better compared to other situations. Here, "Absent" refers to the parents' status being late. 332  
333  
**Exercise during Covid-19:** Students' mental health is relatively poor if they do not exercise during the Covid-19 epidemic. 334  
335  
**Residence:** Students who are in rural areas their mental health status is much better than students in urban areas. 336  
337  
**Relative or acquaintance infected with Covid-19:** If a student's family or friend is not infected with Covid-19, his or her mental health will be slightly better. 338  
339  
**Gender:** There is a significant improvement in mental health status among males compared to females. 340  
341  
**Worry about social support during Covid-19:** In our research, we discovered that more students experience poor mental health conditions when they are concerned about social support. 342  
343  
**Stable family income:** We have seen that, when there is no stable family income, students' mental health status tends to be poorer. 344  
345  
**Living with parents:** In our research, it shows that people who live with their parents act to have better mental health status than those who don't. 346  
347  
**Fear of getting infected with Covid-19 :** If there is a history of Covid-19 infection in the student's family or network of friends, and there's a fear of getting infected, then his or her mental health tends to be poorer due to the fear of infection. 348  
349  
**Own income:** The mental health condition of those who do not have their own income is not better compared to those who do. 350  
351  
352  
353

## Data Preprocessing 354

Data preprocessing is one of the most important phases in ML and DL task, but it is difficult and time-consuming. In our experiment, we followed the preprocessing steps which is described below: 355  
356  
357

- Handling missing values: The dataset utilized in this investigation contains no missing values. Therefore, we proceed without focusing on this step. 358  
359
- Data type conversion: There are 13 categorical features among the total 15 features in our primary dataset. For categorical attributes, we used label encoding technique for data conversion from categorical to numerical ones. For example, the three values in the "Quarrel between parents" feature were changed as follows: the value allocated to "Absent" was given to 0, the value assigned to "No" to 1, and the value given to "Yes" to 2. 360  
361  
362  
363  
364  
365

Then, we focused on feature selection and ranking to understand which feature is mostly responsible for poor mental well-being and which does not. 366  
367

## Feature Selection and Ranking 368

Feature Selection and Ranking is a process that enable us to identify the valuable features in any dataset. It is used in our experiment to readily differentiate the features that play a crucial role in predicting mental health state. A univariate feature selection approach is applied for the feature selection process. The best features are chosen by utilizing the "SelectKBest" function to extract the top best features which are shown in Fig 3. Each feature is compared to the target variable to check for a statistically significant association between them. It is obvious from the figure that, "Quarrel between parents" is the prime reason leading to the poor mental health status of a student. Secondly, "Worry about social support during Covid-19" is almost the same preference for low mental well-being. Thirdly, "Disruption of daily life" has an impact 369  
370  
371  
372  
373  
374  
375  
376  
377  
378

on student’s mental health status. “Stable family income”, “Fear of getting infected with Covid-19”, and “Own income” are also a consequence of the impact on mental state. We have noticed that “Worry about academic delay” is also a reason for poor mental well-being. Other features also had minimal influence on mental well-being.

## Analysis of Prediction Technologies

The process of analyzing prediction technologies involves selecting from a variety of ML and DL techniques that can be utilized in the experimentation. In this case, several predictive models were used, such as those based on classification and neural network algorithms. The goal is to identify the best classifier for analyzing the problem. Each classifier must therefore be trained on the featured set, and the classifier with the best classification results is used for prediction. The classification and neural network algorithms taken into consideration are:

### Machine Learning Models

To determine the level of mental wellness, ten ML classifiers were used based on their performance. The following is a list of all the supervised ML techniques with theoretical explanation we applied in our experiment.

**Random Forest:** There are a number of decision trees that make up the RF. Input data is used to train a set of decision trees. This algorithm generates a large number of decision trees from random subsets of a dataset, and then uses the predictions from each tree to forecast the final result based on the predictions that received the most votes. With an ensemble technique that outperforms the basic tree structure [43].

**CatBoost:** Categorical Boosting is referred to as CatBoost, a new gradient-boosting method that does a better job of dealing with categorical characteristics [44]. It also uses a novel method for calculating leaf values while selecting the tree structure, which helps reduce overfitting. Consequently, this new algorithm outperforms the previous methods, such as Gradient Boosted Decision Trees (GBDT), XGB, and LGBM.

**Extreme Gradient Boosting:** XGB is a scalable ML method that applies a gradient boosting architecture for tree boosting. It is an end-to-end tree-boosting method that is commonly used in data research [45]. It can address real-world size issues with far fewer resources.

**Support Vector Machine:** The SVM is a supervised ML technique used for classification and regression tasks [46]. It’s primary goal is to find the optimal hyperplane that directly separates the information focused on two-part by augmenting the edge.

**K-Nearest Neighbour:** The KNN algorithm is a popular classification approach in ML and pattern recognition. The approach works on the idea that if the majority of the k most comparable samples to a query point q in the feature space belong to a specific category [47], then the query point q must also belong to that category.

**Decision Tree:** It is also a supervised ML strategy that constructs a classification model in a tree and produces classification rules in a top-down manner [48]. It also generates rule sequences used to determine the class of a new observation. Sometimes post-pruning approaches are employed to dealing with overfitting issues.

**Gradient Boosting Machine:** The learning mechanism in gradient boosting machines, or simply GBMs, fits new models sequentially to estimate the response variable more accurately. The basic concept behind this technique is to build new base-learners that are maximally correlated with the negative gradient of the loss function, which is connected with the entire ensemble [49]. The loss functions used can be arbitrary, however for a better understanding, if the error function is the conventional squared-error loss, the learning approach will result in successive error-fitting.

**AdaBoost:** AdaBoost is a common approach for building a robust classifier from a linear combination of member classifiers. During the training process, the member classifiers are chosen to reduce mistakes in each iteration step [50]. AdaBoost provides a straightforward and practical approach for generating ensemble classifiers.

**Multinomial Naive Bayes:** The MNB classifier is one of the NB classifier variants frequently used as a baseline in text categorization. However, this classifier is not entirely Bayesian [51].

**Light Gradient Boosting Machine:** LGBM is a free and open source distributed gradient boosting framework for ML that was created by Microsoft. It uses decision tree algorithms to do ranking, classification, and other ML tasks.

## Deep Learning Models

Despite using ML, to assess more accurately the level of mental wellness in different structures, four DL techniques were utilized, considering their performance in this experiment. The following is a list of the neural network techniques we applied in our experiment.

**Convolutional Neural Networks:** CNNs have been successfully used in many applications where important information about data is embedded in the form of features, such as speech and image [52]. However, CNNs offer noticeable advantages for tabular data analysis, especially in scenarios where data proves to exhibit relationship or sequential patterns in the tabular data, likewise appropriate in our dataset.

**Recurrent Neural Networks:** The RNNs are efficient in learning from past experiments and can simultaneously process a sequence of inputs and outputs, which is a kind of sequence-to-sequence network that is exceptionally efficient in handling genomic data [53]. In our cases, features in tabular data have an implicit sequence or hierarchical structure that an RNN can exploit.

**Hybrid Neural Networks:** A Hybrid Neural Network is a kind of network that combines elements from several architectures rather than being solely based on one specific architecture [54].

**Siamese Neural Networks:** SNNs, where similarity has always been a key aspect, are typically used for verification and identification tasks [55]. In this network, it first examines the features to determine their similarity. Instead of learning to classify inputs, it learns to differentiate between two inputs.

## Evaluation Metrics

To select the best performing model for the prediction of mental health status, the performance of all the classifiers was evaluated using metrics such as Accuracy, Precision, Recall, F1-score, and Area under the curve (AUC).

**Accuracy:** The accuracy metric, in general, measures the proportion of right predictions to the total number of instances examined [56]. Here, ‘TP’ stands for true positive, ‘TN’ stands for true negative, ‘FP’ stands for false positive, and ‘FN’ stands for false negative.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

**Precision:** Precision is used to calculate the percentage of positive patterns that are accurately predicted out of all predicted patterns in a positive class [56].

$$Precision = TP/(TP + FP) \quad (2)$$

**Recall:** When we need to identify the number of positives that may fairly be predicted, recall is another viable selection of evaluation metric [57]. The recall is used

to measure the fraction of positive patterns that are correctly classified.

$$Recall = TP / (TP + FN) \quad (3)$$

**F1 Score:** The F1 score maintains a balance between accuracy and recalls for any classifier [57]. An F1 score is a number between 0 and 1, representing the concordant accuracy and recall measures.

$$F_1 = \left( \frac{2 * Precision * Recall}{Precision + Recall} \right) \quad (4)$$

**Area Under the Curve (AUC):** The Area under the ROC curve (AUC) indicates how well a model can discriminate across classes (the positive and negative classes). The higher the AUC, the better the model predicts 0's as 0 and 1's as 1 [58].

## Experimental Results Analysis

In the dataset, there are 14 features and one label of which 13 features are categorical types and two are numerical ones. Before delving into the result analysis of ML and DL techniques, we conducted Chi-Square test to find the associativity of the features with respect to levels. Then we analyzed our ML and DL models' performance in terms of Accuracy, Precision, Recall, F1-score, and AUC.

### Finding Associative Features using Chi-Square Test

Chi-Square test technique works with the target column and features by considering the important variables. Mostly it is used where the target variable is categorical, and the likelihood aspect is maintained while evaluating the statistical significance. It examines the degree of any inconsistencies between the actual and expected results based on the variables involved in the relation of the sample size and number. Degrees of Freedom (DOF) are being applied to determine whether a null hypothesis can be rejected or not based on the total number of experimental variables and samples. We have applied the Chi-Square test and found the five most associative features ("Stable family income", "Disruption of daily life", "Own income", "Sleep status", and "Fear of getting infected with Covid-19") which was presented in Table 2. Following that, we applied ML models considering these five features only and found 72.54 accuracy while applying RF. Here, with a 95% confidence interval (CI), and a p-value of less than (or equal to) 0.05 was considered to be significant. The DOF and critical value were also mentioned in the Table 2.

### Prediction Analysis using ML and DL

The proposed methodology of our experiment, which was already depicted earlier. The whole methodology is divided into different steps to do our research efficiently. The first step is the data collection and it's description, which was also previously mentioned. Following data preprocessing step, we split our dataset into two subsets using the "train\_test\_split" function: training (80%) and testing (20%). We used the "stratify" parameter when splitting the dataset to keep the class levels balanced in both the training and test subsets, matching the distribution of the collected dataset [59]. Then, we trained ten ML techniques using our processed dataset. The Python programming language (Python 3) has been used to fit the dataset to ML techniques. This technique included RF, CatBoost, XGB, GBM, SVM, DT, K-NN, LGBM, AdaBoost, and MNB. After learning the insight and figuring out the pattern of the training subset, then we

**Table 2.** Analysis of associations between different features

Feature	Chi2	P-value	DOF	Critical-value
Gender	0.073	0.787	1	3.841
Worry about social support during Covid19	15.156	9.897	1	3.841
Residence	0.17	0.68	1	3.841
Stable family income	6.66	<b>0.009</b>	1	3.841
Living with parents	0.516	0.472	1	3.841
Relative or acquaintance infected with Covid19	0.443	0.505	1	3.841
Worry about academic delay	2.632	0.104	1	3.841
Disruption of daily life	11.886	<b>0.00056</b>	1	3.841
Exercise during Covid19	1.301	0.253	1	3.841
Own income	4.275	<b>0.038</b>	1	3.841
Quarrel between parents	21.516	2.126	2	5.991
Sleep status hour during Covid19	10.492	<b>0.005</b>	2	5.991
Internet browsing hours during Covid19	0.94	0.815	3	7.814
Fear of getting infected with Covid19	5.419	<b>0.019</b>	1	3.841

provided a testing subset to evaluate the model how accurately predicted on new data samples. Lastly, the performance of the classifiers was compared using performance metrics such as Confusion Matrix, Precision, Recall, F1-score, and AUC Score. The performance metrics for all the applied ML models are presented in Table 3. It is evident from this table that “accuracy” is the best evaluation metric for predicting mental health status.

In comparison, the RF model yielded the highest accuracy of 86 percent outperformed other models. It also had decent precision (87), recall (86), and f1-score (86) scores compared to other classifiers. The next best classifier is CatBoost which achieved performance over 80 in terms of evaluation metrics. A bar chart depicted in Fig 4 illustrates the performance of all the classifiers in which the RF surpasses other techniques. Considering AUC, again RF achieved the highest score compared to the other models, which is 81 as shown in Fig 5.

In DL result analysis, we followed the same data pre-processing technique as mentioned in ML prediction analysis. Afterward, we used our processed dataset to apply four different DL methods. These methods included CNNs, RNNs, Hybrid Neural Networks, and SNNs. Once the model gained insights and identified patterns within the training subset, it then moved on to evaluate the accuracy of predicting new data samples by utilizing the testing subset. Additionally, we compared the performance of the classifier using the confusion matrix similar to what we devoted in ML. The performance metrics for all the DL classifiers are detailed in Table 4.

From the table, it reveals that the SNN demonstrates the highest accuracy of 75 percent. While the CNN model also performs well, its accuracy (72%) is slightly less

**Table 3.** Result analysis over implemented ML algorithms

Model	Accuracy	Precision	Recall	F1-score	AUC
RF	<b>0.86</b>	<b>0.87</b>	<b>0.86</b>	<b>0.86</b>	<b>0.81</b>
CatBoost	0.82	0.83	0.82	0.82	0.77
XGBoost	0.81	0.80	0.80	0.80	0.76
GBM	0.78	0.81	0.78	0.76	0.70
SVM	0.76	0.76	0.76	0.76	0.71
LGBM	0.76	0.76	0.76	0.76	0.71
KNN	0.74	0.77	0.75	0.71	0.65
AdaBoost	0.72	0.72	0.73	0.70	0.64
DT	0.72	0.72	0.73	0.72	0.68
MNB	0.70	0.73	0.71	0.65	0.59

than SNN. SNN excels due to its ability to effectively handle common class imbalance issues and identify similar features within the dataset, requiring fewer samples per class during training. Additionally, it had the best precision (81), recall score (70), and F1-score (67) compared to other classifiers, which is presented in Table 4. In most cases, an AUC of 0.5 typically indicates no discrimination, meaning the model’s ability to differentiate between classes based on the test dataset. An AUC range between 0.7 to 0.8 is generally considered acceptable [60]. In this study, the SNN demonstrated an AUC of (72%), outperforming other models which are depicted in Fig 6.

**Table 4.** Result analysis over implemented DL techniques

Model	Accuracy	Precision	Recall	F1-score	AUC
SNN	<b>0.75</b>	<b>0.81</b>	<b>0.70</b>	<b>0.67</b>	<b>0.72</b>
CNN	0.72	0.60	0.57	0.58	0.70
Hybrid	0.69	0.56	0.53	0.54	0.71
RNN	0.70	0.61	0.63	0.64	0.69

## Discussion

The analysis of [36] pointed out that Covid-19 lockdowns had a variety of consequences, including increased psychological pressure on the daily lives of students and parents. The authors also indicated the effects of the epidemic in many areas of life, particularly in the education sector. Lack of digital skills, electricity, accessibility, facilities, and network challenges are the major obstacles that students face when enrolling in online education during this epidemic that contributed to students’ poor mental status [61]. The researchers also discovered that over 97 percent of university students were experiencing mental health issues due to the arising of the current epidemic. The primary goal of our study is to examine the psychological status of Bangladeshi university students during and after the current pandemic.

The emergence of the Covid-19 prevalence is presently posing a severe threat to

world health. Bangladeshi students are not safe from this threat also. To improve university students' mental health, it is essential to conduct a mental health assessment as soon as possible in Bangladesh. Manually doing the examination can be difficult; however, ML and DL provide a promising technique for performing the tasks in this pandemic situation. Our study conducted a Chi-Square statistical test (at  $p$  less than 0.05) to identify the five most significant features: "Stable family income", "Disruption of daily life", "Own income", "Sleep status (hour) during Covid-19", and "Fear of getting infected with Covid-19". It shows that students from rural areas have a better mental health status than students from urban areas. This might be because rural regions have a lower Covid-19's impact rate. We also discovered that students who are concerned about academic delays tend to have poorer mental health statuses. In our experiment, RF, CatBoost, and XGB performed better than the other classifiers during testing. Algorithmically, the RF model tackles the overfitting problem of decision trees by implementing randomization to achieve better generalization and outperform the other classifier. Because of these reasons, in this paper, it (RF) achieved the highest accuracy 86% than other classifiers. Our dataset has many categorical features, and CatBoost generally works well for categorical types of features. Therefore, in this research, we have applied this gradient boosting algorithm. On the other hand, XGB attempts to create new trees that complement the existing ones and boost training for difficult-to-classify data points. At last, to evaluate the performance of the classifiers, we use the measured area under the ROC curve, also known as AUC. In comparison, the larger the AUC, the higher the model's performance. The RF classifier surpasses the other models by having the greatest AUC and the lowest False positive rate.

Elsewhere, we evaluated four DL algorithms where SNN shows the highest accuracy compare to the others. It outperforms when the features are alike to each other. Additionally, this method is good at handling common class imbalance issues and finding similar features in the dataset. It also worked well in our study while needing fewer samples for training each class. Our second best-performing DL model was CNN and it performs well where data contains categorical relationships.

## Strengths and Limitations

Strengths in this research, include 1) this research was conducted during the pandemic, focusing particularly on university students in Bangladesh, where mental health research is limited; 2) we collected a dataset from diverse dimensions such as 16 universities; 3) we also achieved decent enough accuracy by applying well-known classification models to predict accurately whether a student's mental health status is poor or better. Along with some substantial advantages, there are some limitations in this experiment. Limitations are 1) we cannot include all the university students' data, hence the findings of our analysis may not be generalizable to all Bangladeshi students; 2) as the data is self-reported via an online questionnaire, some of it may be biased; 3) the dataset was limited to build more robust prediction; additional data could be added to improve forecast accuracy.

## Conclusion

Globally, Covid-19 has impacted negatively on people's lives, and it has hampered individuals' physical and mental well-being, especially university students in Bangladesh. In this paper, we employed AI experiments and statistical test on our primary dataset where data was collected through an online questionnaire. Then we trained the dataset with ML and DL techniques where RF and SNNs revealed the best accuracy in predicting mental health state. On the other hand, in the Chi-Square feature test, we



found the five most associative and significant predictors that greatly influence the psychological health conditions of students during the pandemic in Bangladesh. We hope that our findings will assist university administrators and healthcare providers in maintaining the mental health of students during and aftermath the epidemic. Based on our findings, further research is needed to analyze the long-term impact of the Covid-19 pandemic on students' mental health performance. In the future, we would add more features to the dataset to track the development of student's mental and physical status over a more extended period, such as several years. As our dataset size is limited, more students' records from educational institutions could be included in future studies to get more reliable and robust predictions.

## Acknowledgement

This work was supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(IITP-2024-00156287, 50%). This work was partially supported by the Korea Internet & Security Agency (KISA) - Information Security College Support Project (50%)

## References

1. WHO;. Available from: <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response/>.
2. Connell J, Brazier J, O'Cathain A, Lloyd-Jones M, Paisley S. Quality of life of people with mental health problems: a synthesis of qualitative research. *Health and quality of life outcomes*. 2012;10:1–16.
3. Keyes CL. Promoting and protecting positive mental health: Early and often throughout the lifespan. In: *Mental well-being: International contributions to the study of positive mental health*. Springer; 2012. p. 3–28.
4. Borg M. The nature of recovery as lived in everyday life: Perspectives of individuals recovering from severe mental health problems. 2007;.
5. Galson SK. Mental health matters. *Public Health Reports*. 2009;124(2):189–191.
6. Navaneetham P, Kanth B. Effects of Personal Relationships on Physical and Mental Health among Young Adults-A Scoping Review. *The Open Psychology Journal*. 2022;15(1).
7. Sharma HL, Rani R. Impact of mental health on Personal Growth Initiative (PGI) among university postgraduates. *new science*. 2014;4(3).
8. World Population;. Available from: <https://worldpopulationreview.com/country-rankings/countries-by-density/>.
9. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. *International psychogeriatrics*. 2020;32(10):1217–1220.
10. Pancani L, Marinucci M, Aureli N, Riva P. Forced social isolation and mental health: a study on 1,006 Italians under COVID-19 lockdown. *Frontiers in psychology*. 2021;12:663799.

11. Wang X, Hegde S, Son C, Keller B, Smith A, Sasangohar F. Investigating mental health of US college students during the COVID-19 pandemic: Cross-sectional survey study. *Journal of medical Internet research*. 2020;22(9):e22817. 645-647
12. Chen T, Lucock M. The mental health of university students during the COVID-19 pandemic: An online survey in the UK. *PloS one*. 2022;17(1):e0262562. 648-650
13. Ratul IJ, Nishat MM, Faisal F, Sultana S, Ahmed A, Al Mamun MA. Analyzing perceived psychological and social stress of university students: A machine learning approach. *Heliyon*. 2023;9(6). 651-653
14. Conversano C, Di Giuseppe M, Miccoli M, Ciacchini R, Orrù G. Mindfulness, age and gender as protective factors against psychological distress during COVID-19 pandemic. *Frontiers in psychology*. 2020;11:562965. 654-656
15. Shyadligeri AA, Vaz FS, Lokapure SR. Psychological distress among first-year medical students amidst COVID-19-related uncertainty at a medical college in Western India: A cross-sectional study. *Education for Health*. 2022;35(1):20–25. 657-659
16. Kecojevic A, Basch CH, Sullivan M, Davi NK. The impact of the COVID-19 epidemic on mental health of undergraduate students in New Jersey, cross-sectional study. *PloS one*. 2020;15(9):e0239696. 660-662
17. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry research*. 2020;287:112934. 663-665
18. Lin C, Tong Y, Bai Y, Zhao Z, Quan W, Liu Z, et al. Prevalence and correlates of depression and anxiety among Chinese international students in US colleges during the COVID-19 pandemic: A cross-sectional study. *Plos one*. 2022;17(4):e0267081. 666-669
19. Son C, Hegde S, Smith A, Wang X, Sasangohar F, et al. Effects of COVID-19 on college students' mental health in the United States: Interview survey study. *Journal of medical internet research*. 2020;22(9):e21279. 670-672
20. Wang X, Hegde S, Son C, Keller B, Smith A, Sasangohar F. Investigating mental health of US college students during the COVID-19 pandemic: Cross-sectional survey study. *Journal of medical Internet research*. 2020;22(9):e22817. 673-675
21. Islam MA, Barna SD, Raihan H, Khan MNA, Hossain MT. Depression and anxiety among university students during the COVID-19 pandemic in Bangladesh: A web-based cross-sectional survey. *PloS one*. 2020;15(8):e0238162. 676-678
22. Khan AH, AH SM, Hossain S, Hasan M, Ahmed H, Sikder MT. The Impact of COVID-19 Pandemic on Mental Health and Wellbeing among Home-quarantined Bangladeshi Students. A cross-sectional pilot study *J Affect Disord*. 2020;277:121–128. 679-682
23. Shafiee NSM, Mutalib S. Prediction of mental health problems among higher education student using machine learning. *International Journal of Education and Management Engineering (IJEME)*. 2020;10(6):1–9. 683-685
24. Ahmed MZ, Ahmed O, Aibao Z, Hanbin S, Siyu L, Ahmad A. Epidemic of COVID-19 in China and associated psychological problems. *Asian journal of psychiatry*. 2020;51:102092. 686-688

25. Griffiths MD, Pakpour AH, Mamun MA. Correction to: Psychometric validation of the Bangla Fear of COVID-19 Scale: Confirmatory factor analysis and rasch analysis. *International journal of mental health and addiction*. 2022;20(4):2520–2522. 689  
690  
691  
692
26. Ochnik D, Rogowska AM, Kuśnierz C, Jakubiak M, Schütz A, Held MJ, et al. A comparison of depression and anxiety among university students in nine countries during the COVID-19 pandemic. *Journal of clinical medicine*. 2021;10(13):2882. 693  
694  
695
27. Faisal RA, Jobe MC, Ahmed O, Sharker T. Mental health status, anxiety, and depression levels of Bangladeshi university students during the COVID-19 pandemic. *International Journal of Mental Health and Addiction*. 2021; p. 1–16. 696  
697  
698
28. Safa F, Anjum A, Hossain S, Trisa TI, Alam SF, Rafi MA, et al. Immediate psychological responses during the initial period of the COVID-19 pandemic among Bangladeshi medical students. *Children and Youth Services Review*. 2021;122:105912. 699  
700  
701  
702
29. Krishnan R, Kumari S, Badi AA, Jeba S, James M. Predictive machine learning model for mental health issues in higher education students due to COVID-19 using HADS assessment. *Arab Gulf Journal of Scientific Research*. 2023;. 703  
704  
705
30. Abdul Rahman H, Kwicklis M, Ottom M, Amornsriwatanakul A, H Abdul-Mumin K, Rosenberg M, et al. Machine learning-based prediction of mental well-being using health behavior data from university students. *Bioengineering*. 2023;10(5):575. 706  
707  
708  
709
31. Rezapour M, Hansen L. A machine learning analysis of COVID-19 mental health data. *Scientific reports*. 2022;12(1):14965. 710  
711
32. Banna MHA, Ghosh T, Nahian MJA, Kaiser MS, Mahmud M, Taher KA, et al. A hybrid deep learning model to predict the impact of COVID-19 on mental health from social media big data. *IEEE Access*. 2023;11:77009–77022. 712  
713  
714
33. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. In: *Proceedings of the international AAAI conference on web and social media*. vol. 7; 2013. p. 128–137. 715  
716  
717
34. Chen T. Investigating the mental health of university students during the COVID-19 pandemic in a UK university: a machine learning approach using feature permutation importance. *Brain Informatics*. 2023;10(1):27. 718  
719  
720
35. Mantas J, et al. A machine learning study to predict anxiety on campuses in Lebanon. *Healthcare Transformation with Informatics and Artificial Intelligence*. 2023;305:85. 721  
722  
723
36. Atlam ES, Ewis A, Abd El-Raouf M, Ghoneim O, Gad I. A new approach in identifying the psychological impact of COVID-19 on university student's academic performance. *Alexandria Engineering Journal*. 2022;61(7):5223–5233. 724  
725  
726
37. Prowse R, Sherratt F, Abizaid A, Hellemans KG, Patterson ZR, McQuaid RJ. Coping with the COVID-19 pandemic: examining gender differences in stress and mental health among university students. *Frontiers in psychiatry*. 2021;12:650759. 727  
728  
729
38. Faisal RA, Jobe MC, Ahmed O, Sharker T. Mental health status, anxiety, and depression levels of Bangladeshi university students during the COVID-19 pandemic. *International journal of mental health and addiction*. 2022;20(3):1500–1515. 730  
731  
732  
733

39. Study in Bangladesh;. Available from: 734  
[https://www.studyinbangladesh.com.bd/about\\_us](https://www.studyinbangladesh.com.bd/about_us). 735
40. Sayeed A, Kundu S, Al Banna MH, Hasan MT, Begum MR, Khan MSI. Mental health outcomes during the COVID-19 and perceptions towards the pandemic: Findings from a cross sectional study among Bangladeshi students. *Children and youth services review*. 2020;119:105658. 736
41. Du C, Liu C, Balamurugan P, Selvaraj P. Deep learning-based mental health monitoring scheme for college students using convolutional neural network. *International Journal on Artificial Intelligence Tools*. 2021;30(06n08):2140014. 740
42. Almqren MA, Almqren L, Alhayan F, Cristea AI, Pennington D. Using deep learning to analyze the psychological effects of COVID-19. *Frontiers in Psychology*. 2023;14:962854. 743
43. Ho TK. Random decision forests. In: *Proceedings of 3rd international conference on document analysis and recognition*. vol. 1. IEEE; 1995. p. 278–282. 746
44. Dorogush AV, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:181011363*. 2018;. 748
45. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*; 2016. p. 785–794. 750
46. Hui-mei D, Ying M, Wei W, Hui D, Feng W. An Auto-flag Method of Radio Visibility Data Based on Support Vector Machine. *Chinese Astronomy and Astrophysics*. 2017;41(1):125–135. 753
47. Kuang Q, Zhao L. A practical GPU based kNN algorithm. In: *Proceedings. The 2009 International Symposium on Computer Science and Computational Technology (ISCSCI 2009)*. Citeseer; 2009. p. 151. 756
48. Morgan JN, Sonquist JA. Problems in the analysis of survey data, and a proposal. *Journal of the American statistical association*. 1963;58(302):415–434. 759
49. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*. 2013;7:21. 761
50. An TK, Kim MH. A new diverse AdaBoost classifier. In: *2010 International conference on artificial intelligence and computational intelligence*. vol. 1. IEEE; 2010. p. 359–363. 763
51. Xu S, Li Y, Wang Z. Bayesian multinomial Naïve Bayes classifier to text classification. In: *Advanced multimedia and ubiquitous engineering*. Springer; 2017. p. 347–352. 766
52. Zhu Y, Brettin T, Xia F, Partin A, Shukla M, Yoo H, et al. Converting tabular data into images for deep learning with convolutional neural networks. *Scientific reports*. 2021;11(1):11325. 769
53. Srinivasu PN, Shafi J, Krishna TB, Sujatha CN, Praveen SP, Ijaz MF. Using recurrent neural networks for predicting type-2 diabetes from genomic and tabular data. *Diagnostics*. 2022;12(12):3067. 772
54. Medsker LR. *Hybrid neural network and expert systems*. Springer Science & Business Media; 2012. 775

55. Chicco D. Siamese neural networks: An overview. *Artificial neural networks*. 2021; p. 73–94. 777  
778
56. Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*. 2015;5(2):1. 779  
780  
781
57. Agarwal R. The 5 classification evaluation metrics every data scientist must know. *towardsdatascience.com*. 2019;17. 782  
783
58. Narkhede S. Understanding auc-roc curve. *Towards Data Science*. 2018;26:220–227. 784  
785
59. Atique MMAB, Hoque MN, Uddin MJ. Employee Attrition Analysis Using CatBoost. In: *International Conference on Machine Intelligence and Emerging Technologies*. Springer; 2022. p. 644–658. 786  
787  
788
60. Mandrekar JN. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*. 2010;5(9):1315–1316. 789  
790  
doi:<https://doi.org/10.1097/JTO.0b013e3181ec173d>. 791
61. Dhar BK, Ayittey FK, Sarkar SM. Impact of COVID-19 on Psychology among the University Students. *Global Challenges*. 2020;4(11):2000038. 792  
793

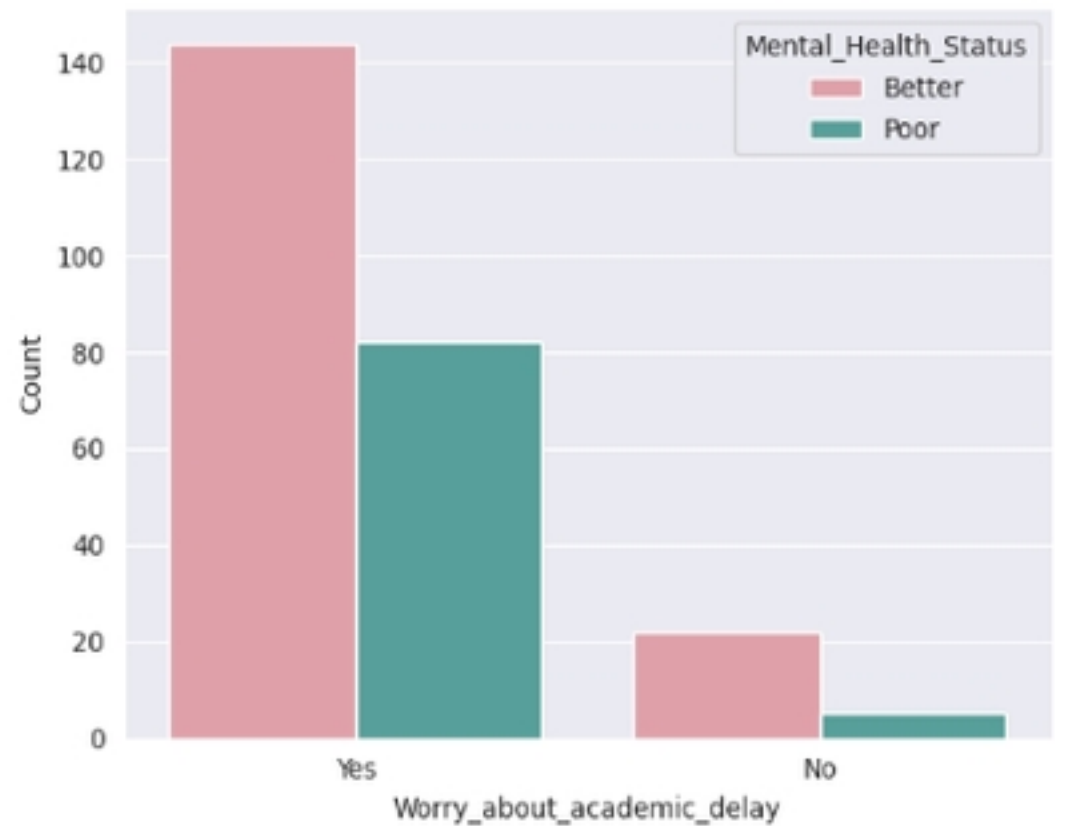
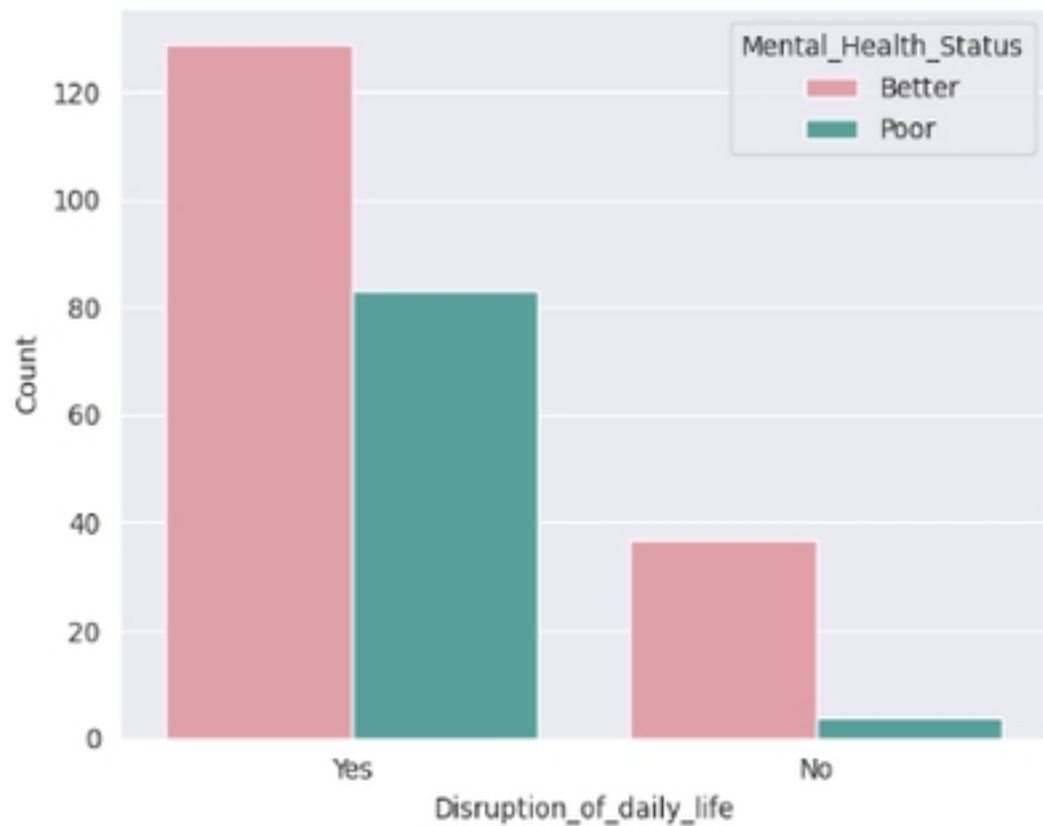
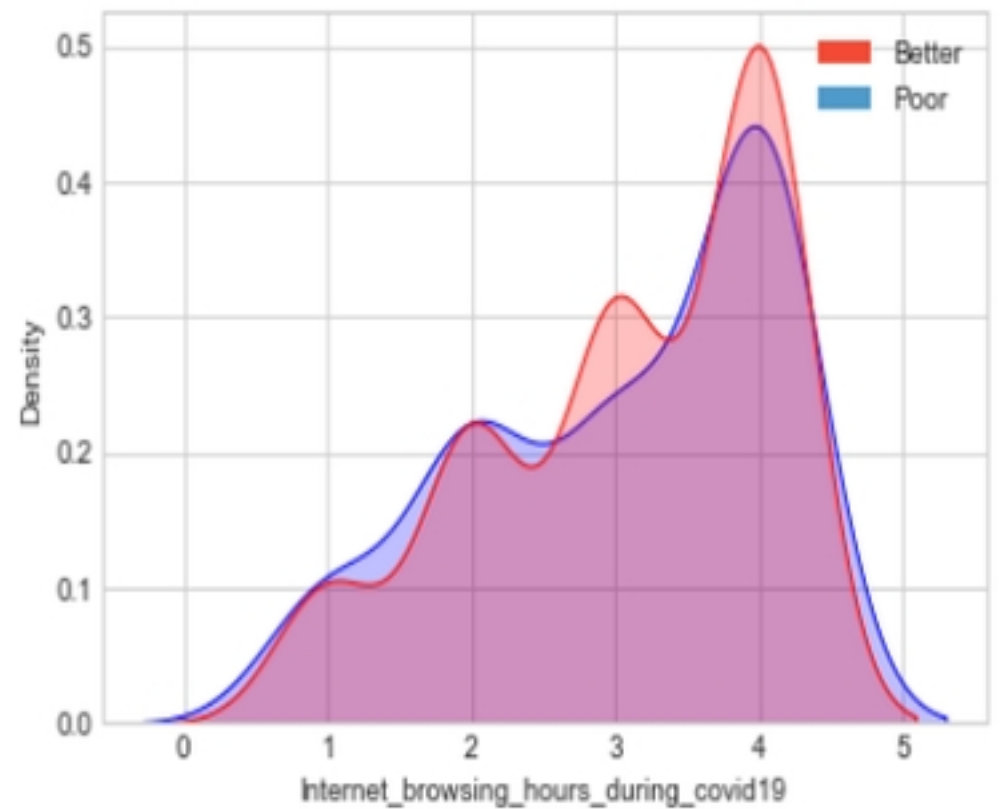
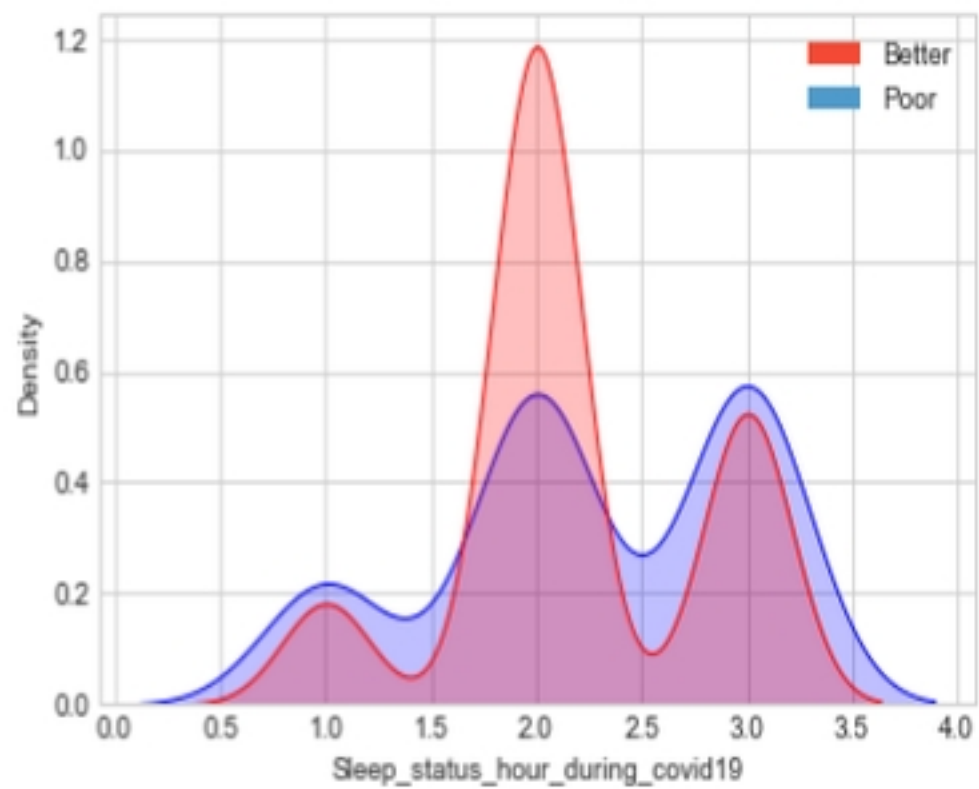


Figure 2A

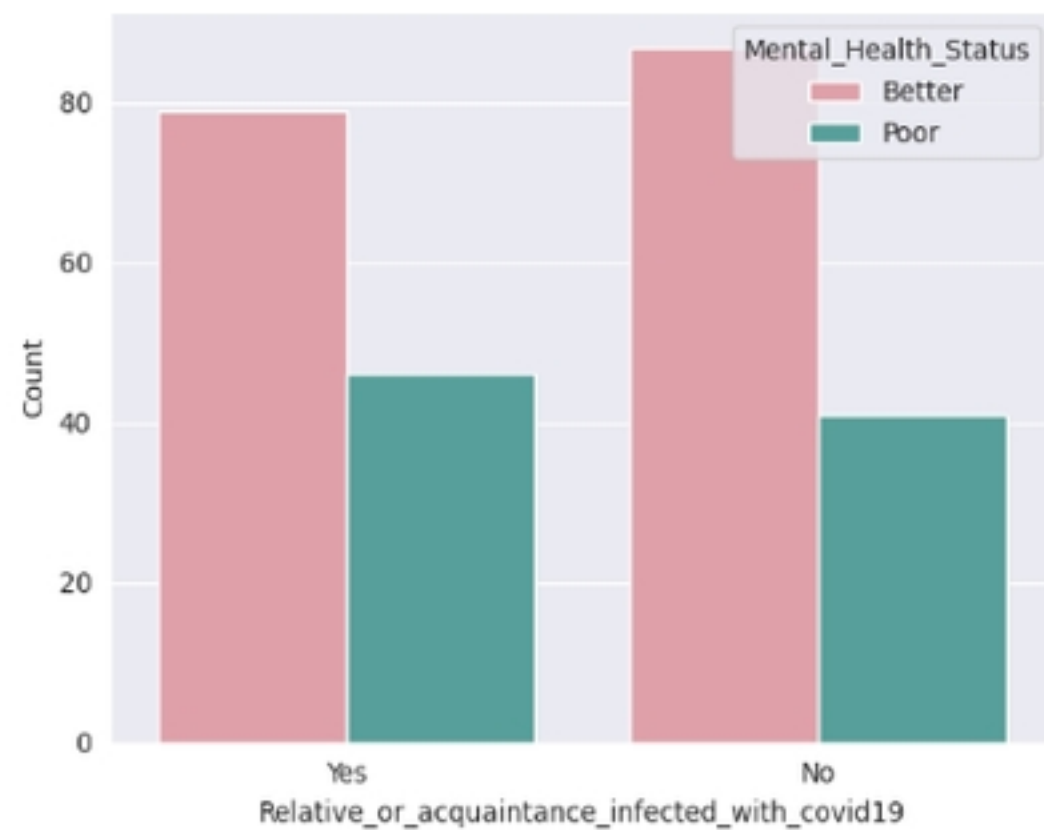
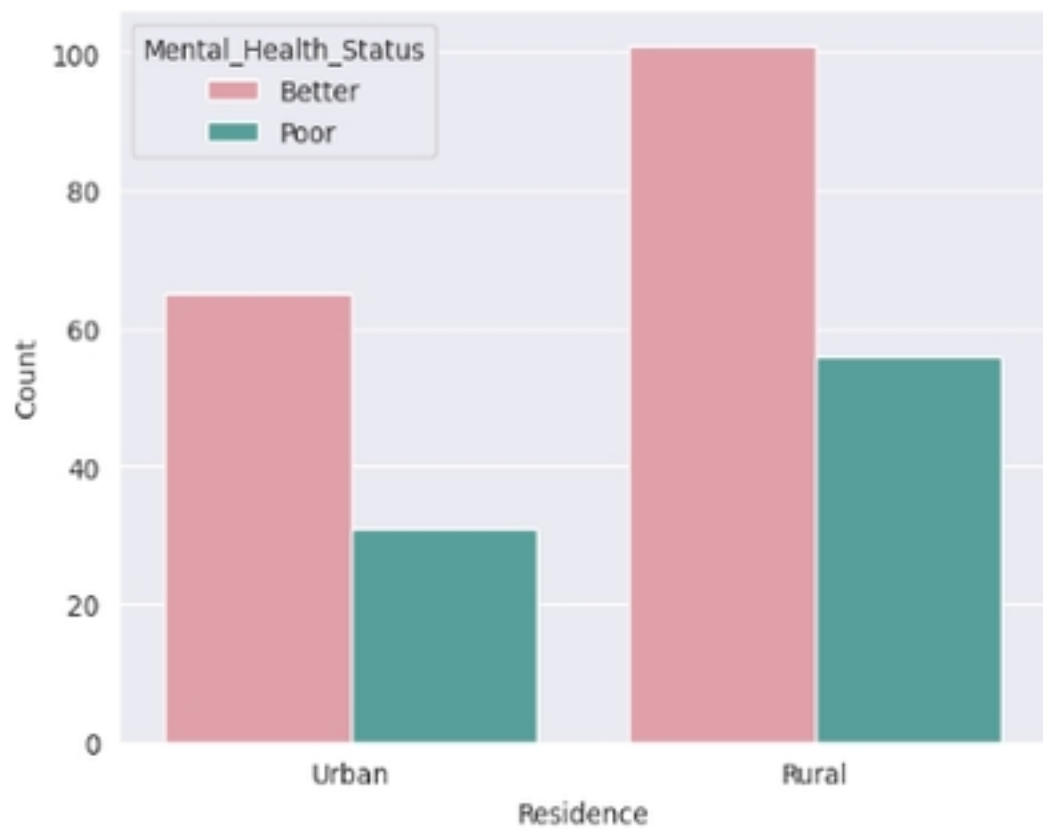
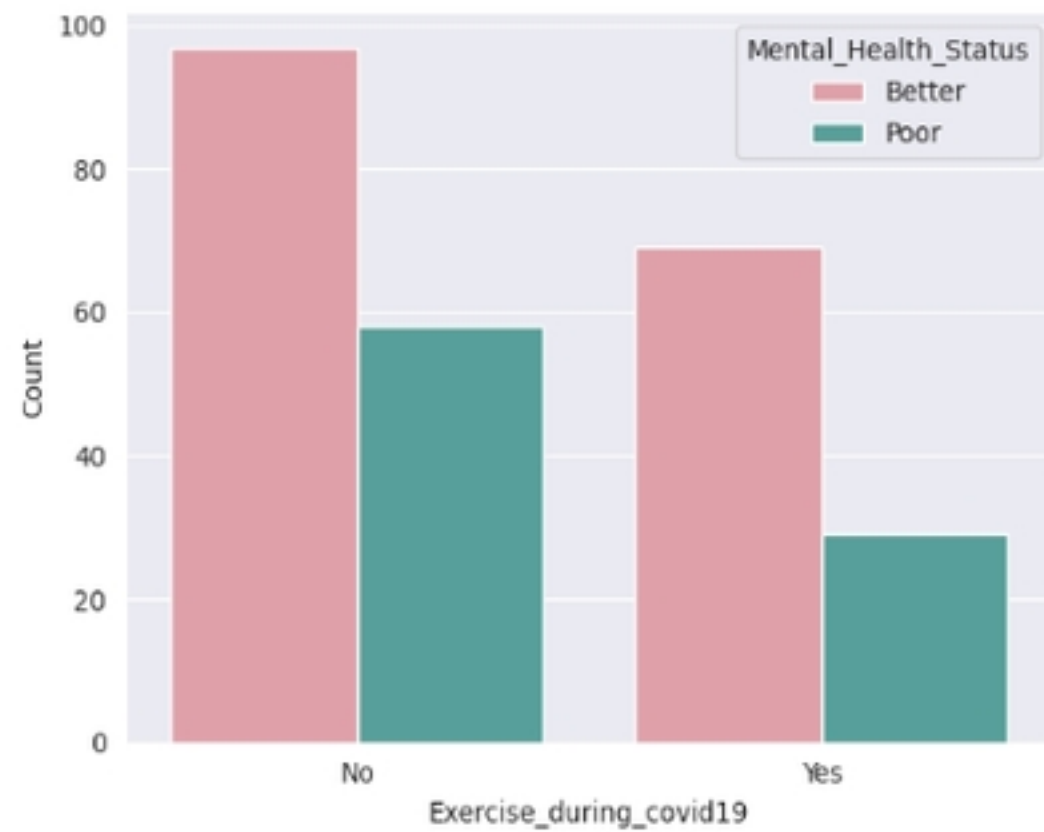
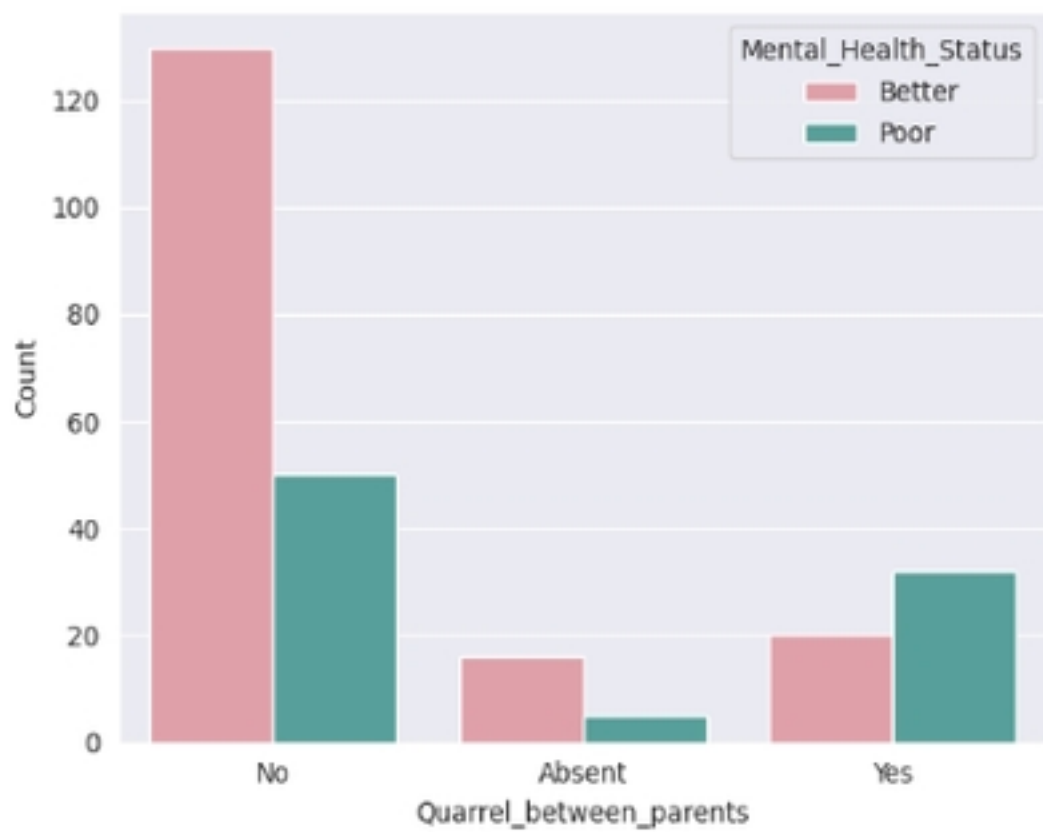


Figure 2B

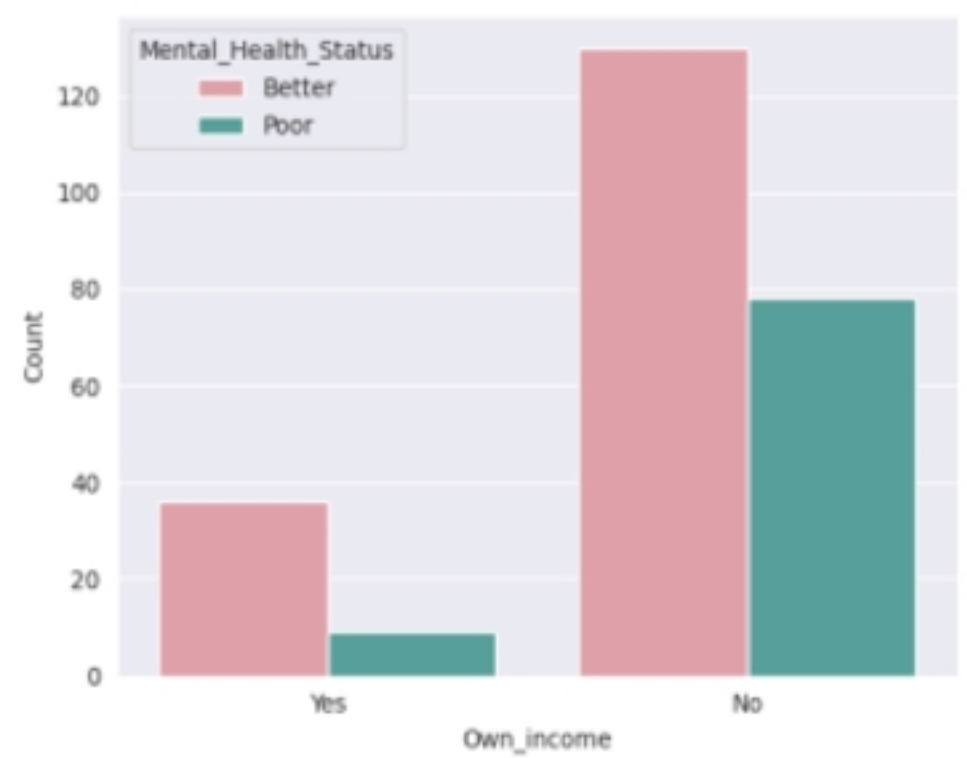
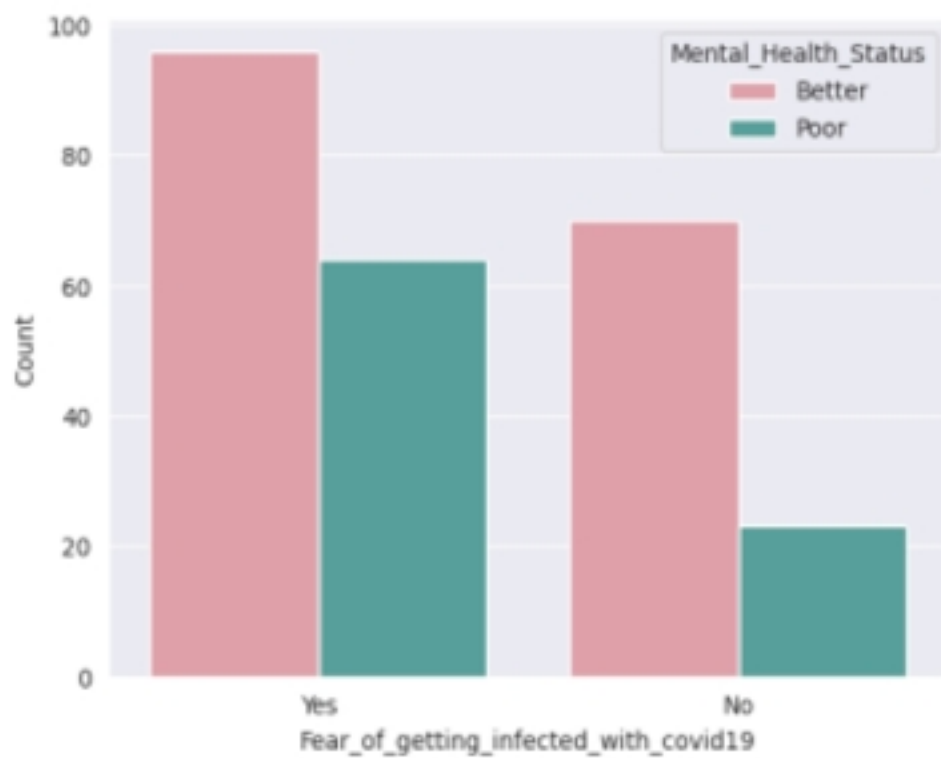
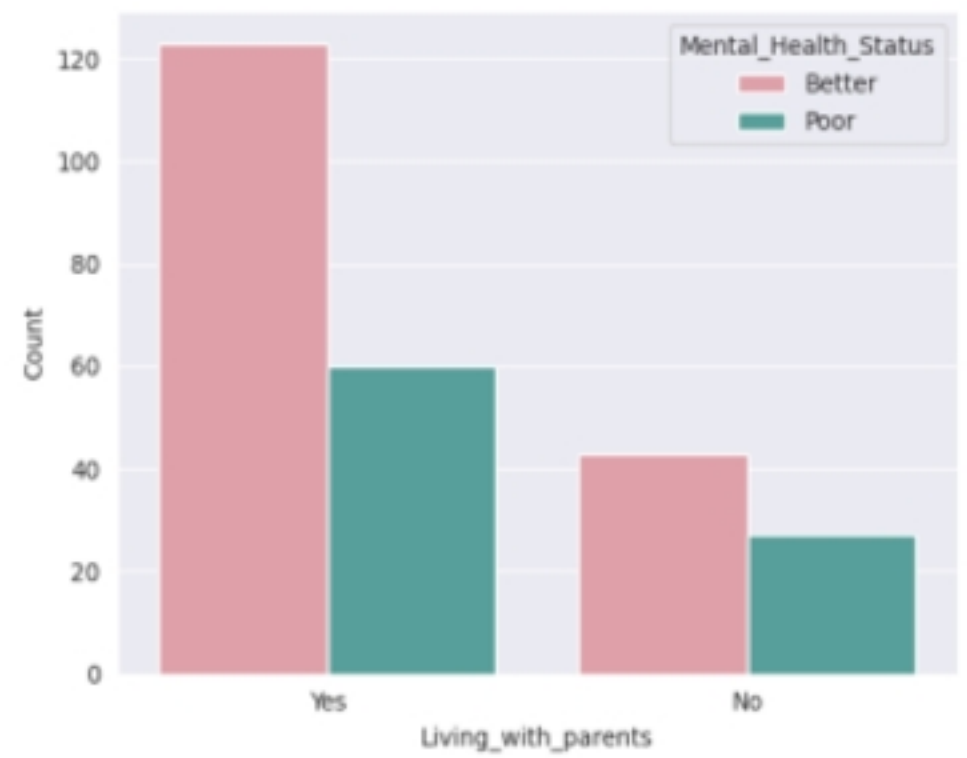
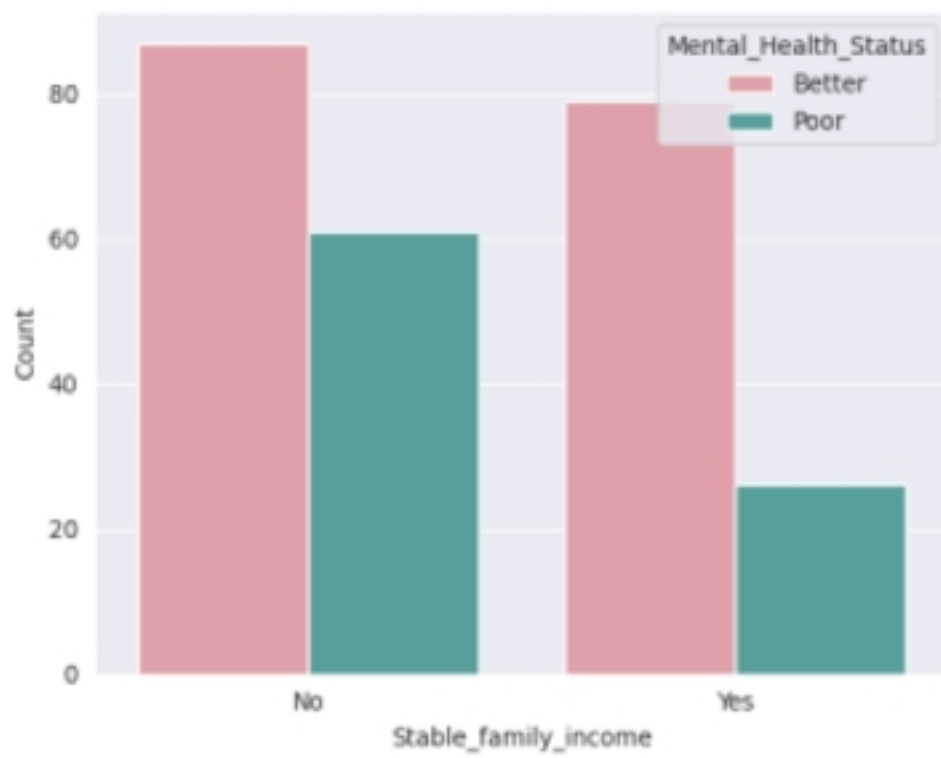
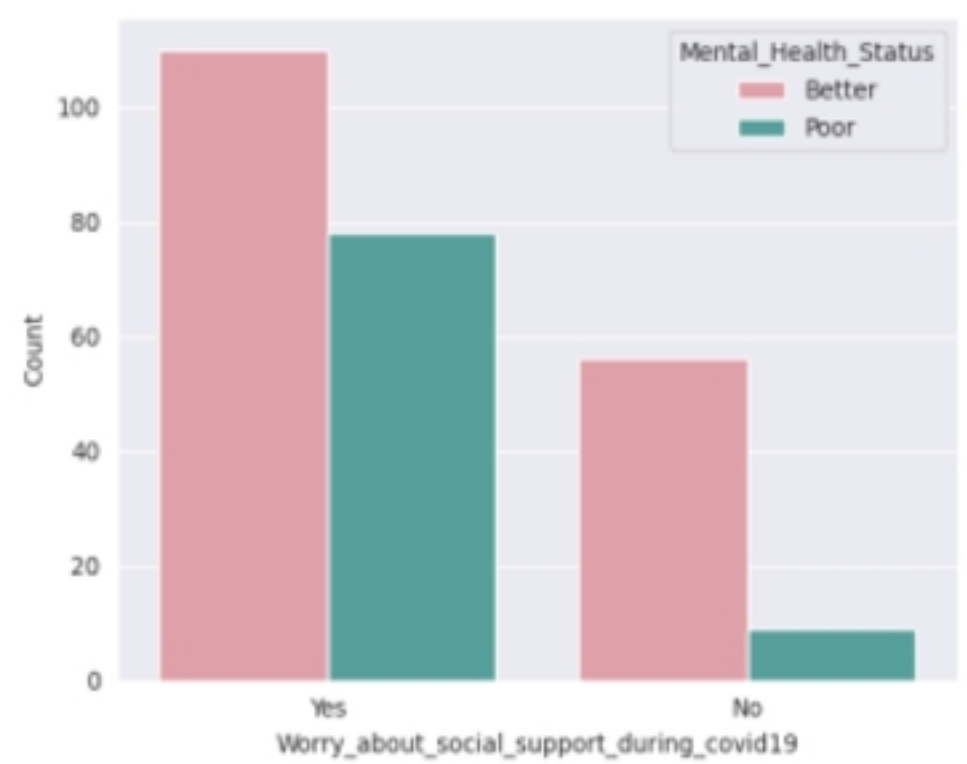
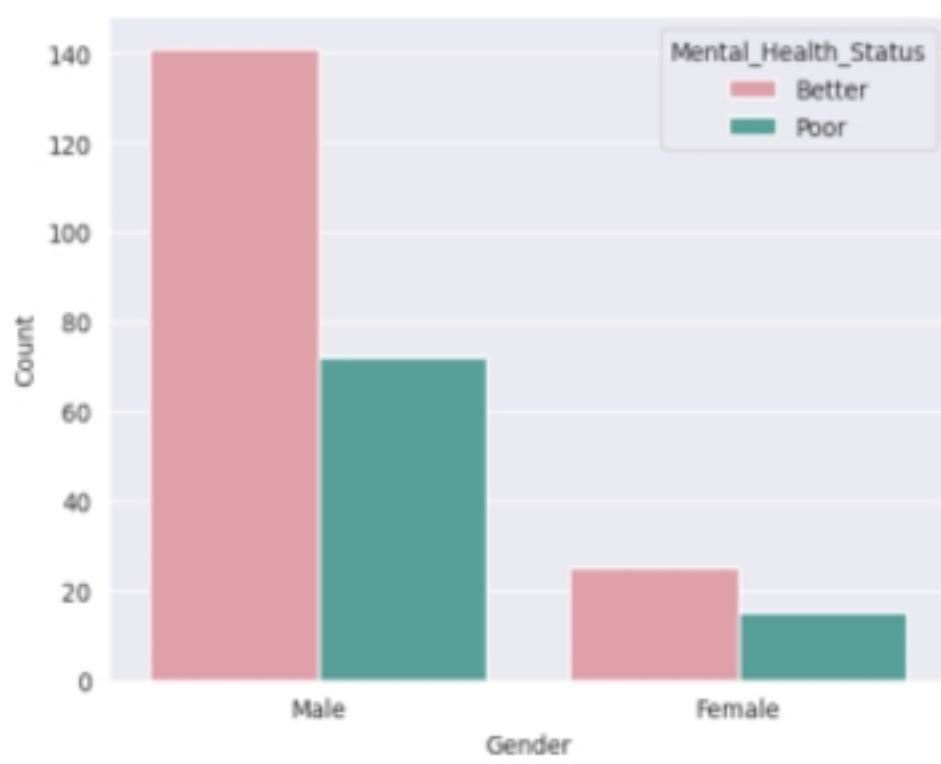


Figure 2C



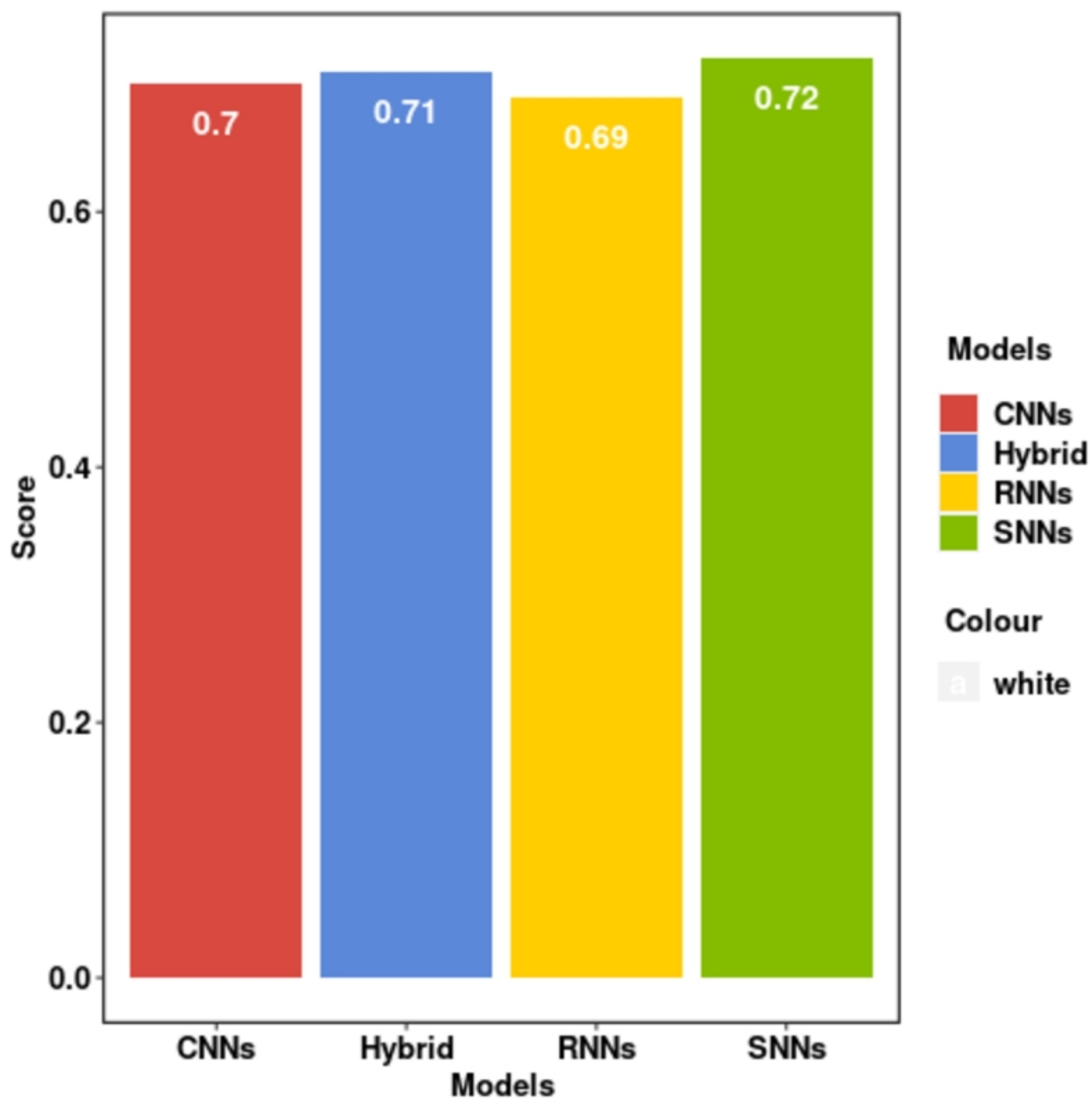


Figure 6

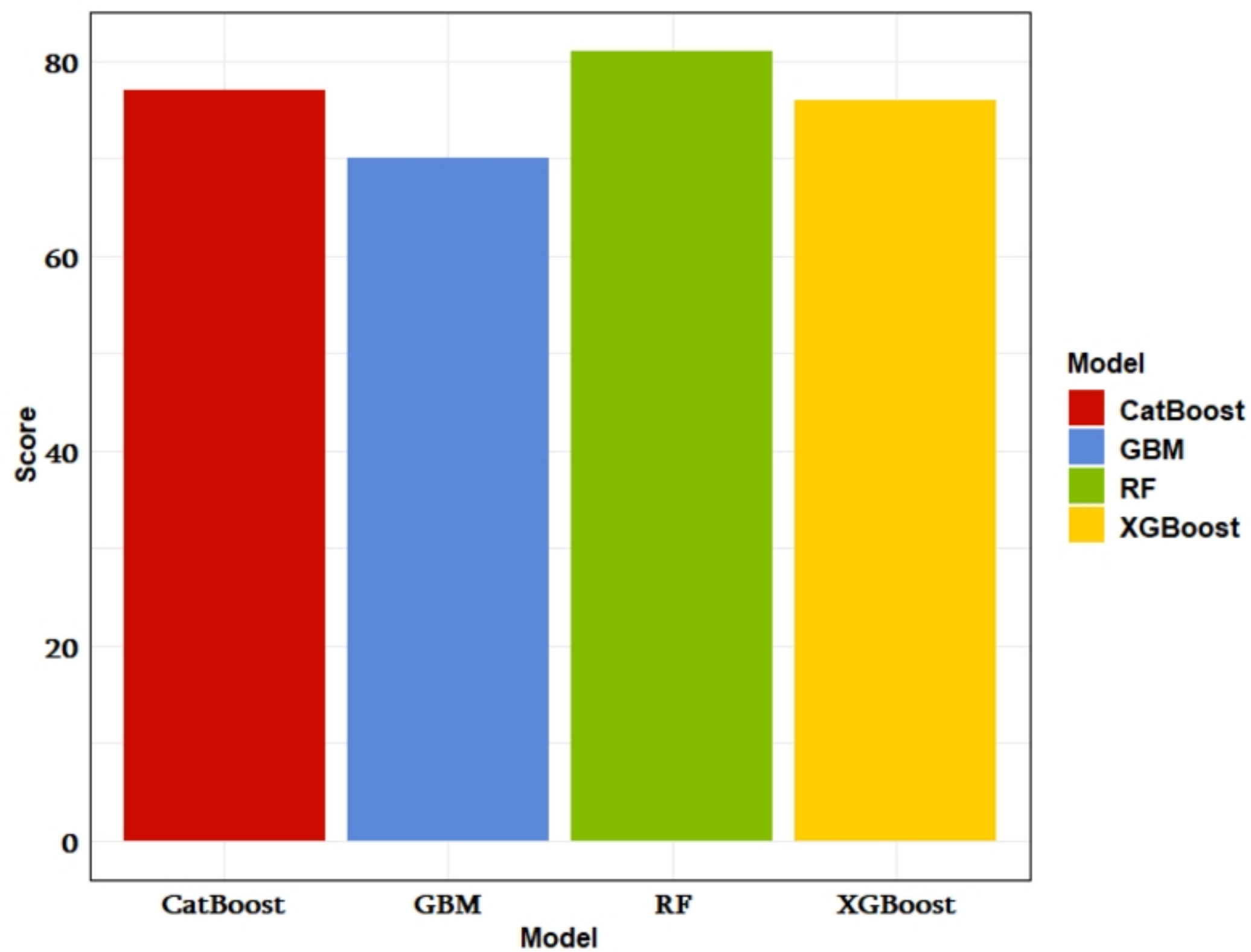


Figure 5

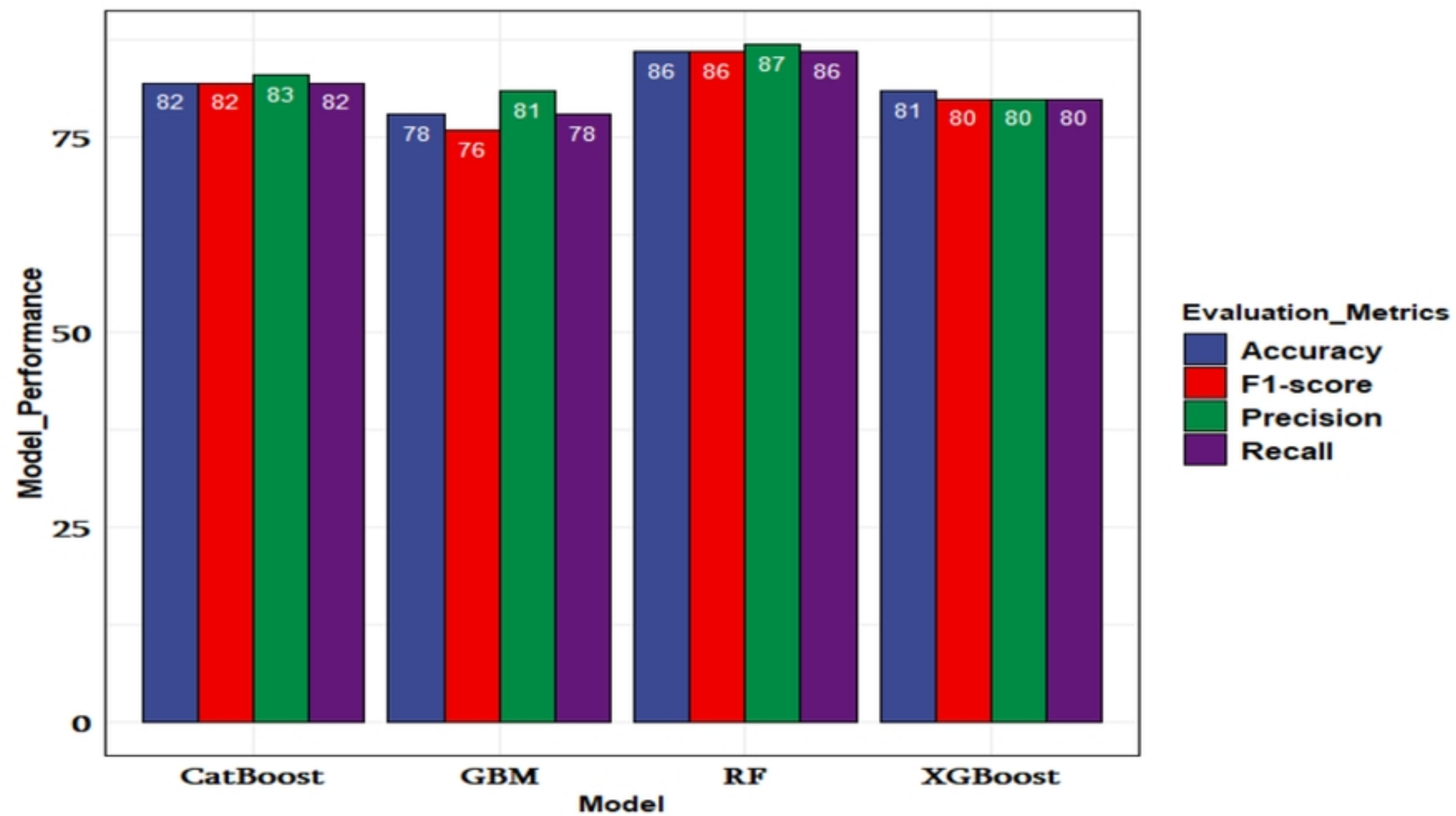


Figure 4

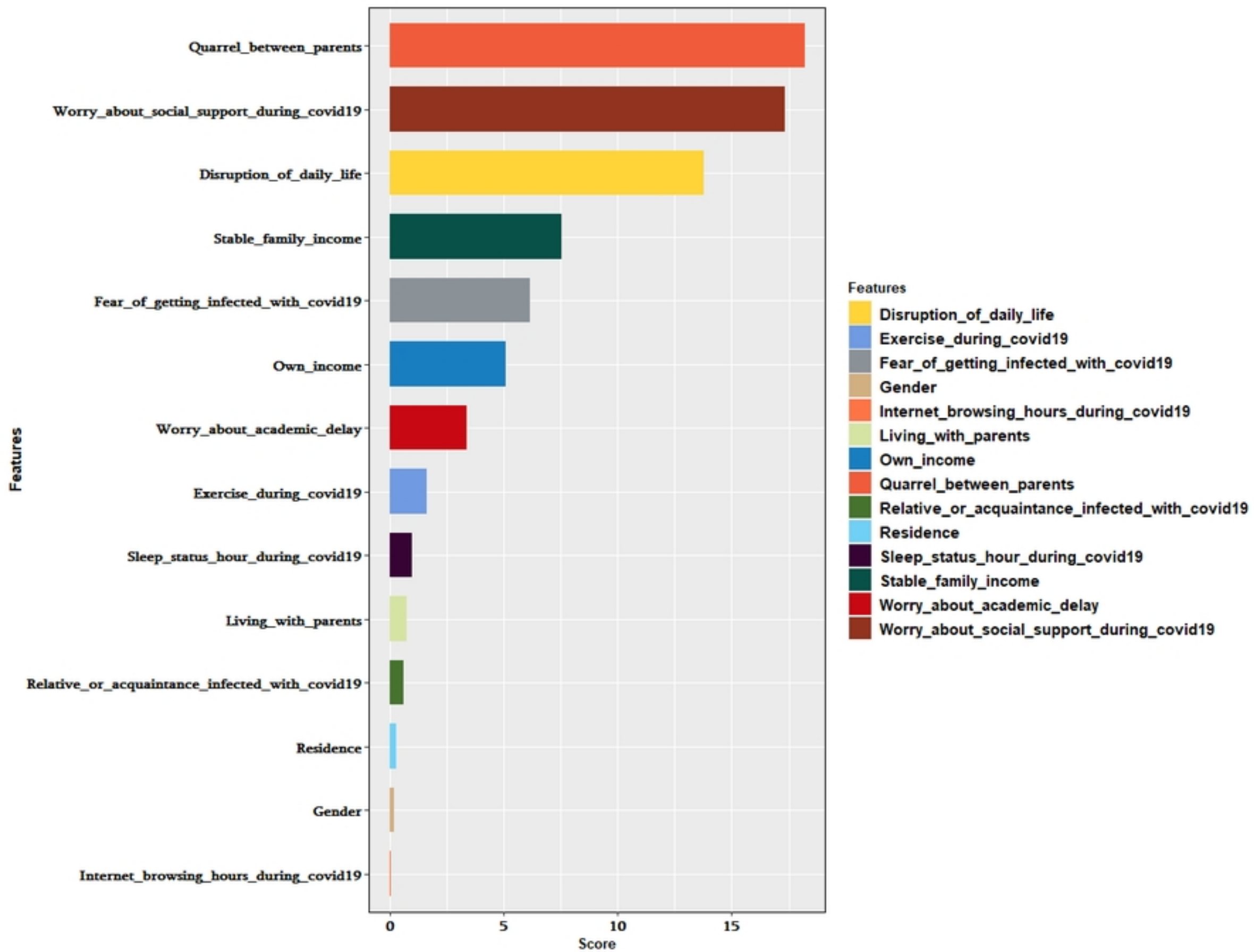


Figure 3

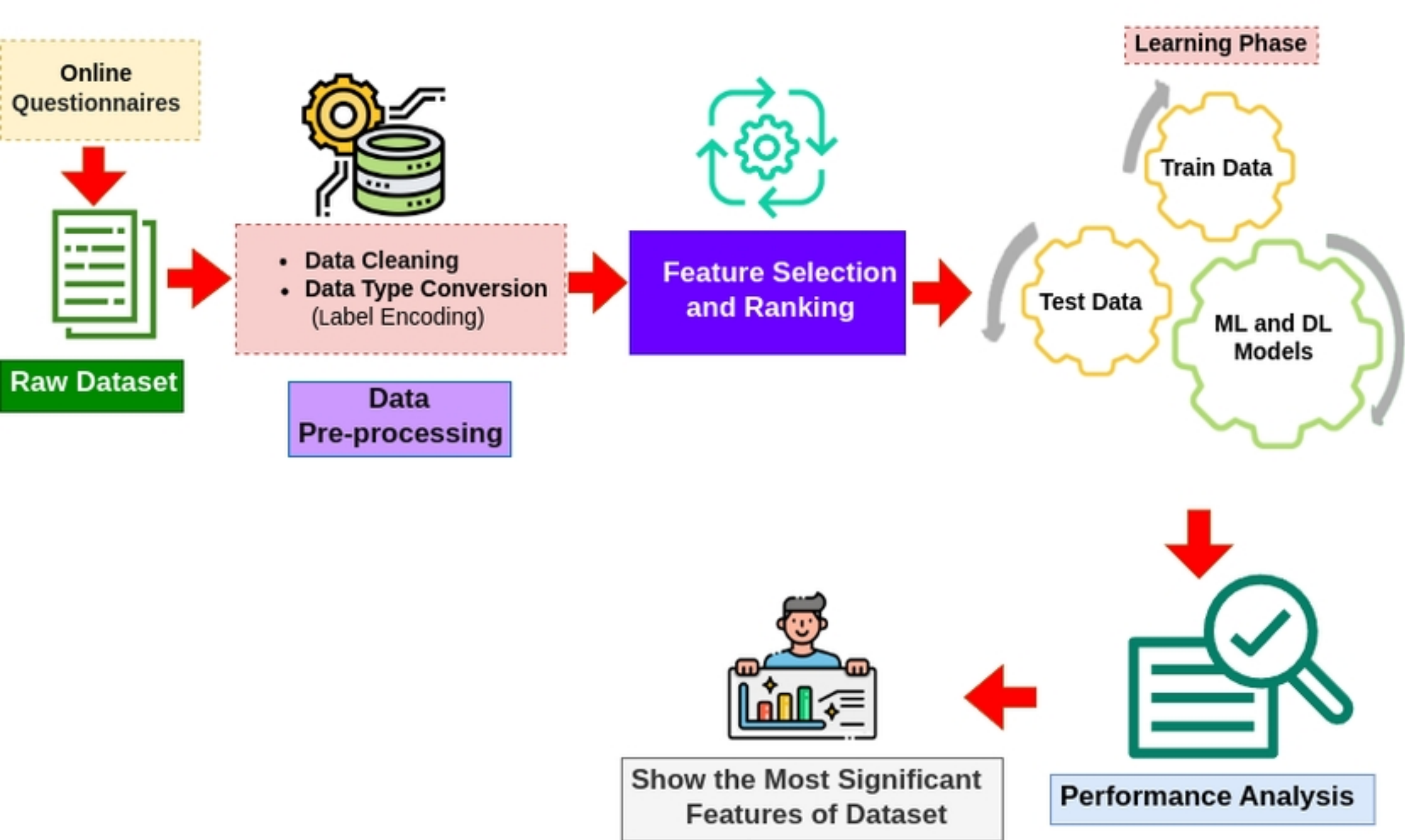


Figure 1