Data in Brief 54 (2024) 110376

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

RHMCD-20 dataset: Identify rapid human mental health depression during quarantine life using machine learning

Nazrul Amin[a], Imrus Salehin[b,c,*], Md. Abu Baten[a], Rabbi Al Noman[d]

[a] *Department of Computer Science and Engineering, Northern University Bangladesh, Bangladesh*
[b] *Department of Information Convergence Engineering, Pusan National University, Busan, Republic of Korea*
[c] *Department of Computer Science and Engineering, Daffodil International University, Dhaka 1207, Bangladesh*
[d] *Department of Electronics and Telecommunication Engineering, Rajshahi University of Engineering and Technology, Rajshahi, Bangladesh*

A R T I C L E   I N F O

Article history:
Received 7 January 2024
Revised 19 March 2024
Accepted 28 March 2024
Available online 5 April 2024

Dataset link: The RHMCD-20 datasets for Depression and Mental Health Data Analysis with Machine Learning (Original data)

Keywords:
Mental health
Machine learning
COVID-19 quarantine
Isolation effects

A B S T R A C T

The RHMCD-20 dataset offers a thorough investigation of the dynamics of mental health in Bangladesh while under quarantine. The structured survey that was distributed to different demographic groups yielded a dataset that included a wide range of variables, such as age, gender, occupation, and stress levels. Predictive modelling, understanding the effects of quarantine on the workplace and society, and intergenerational insights are all greatly enhanced by this dataset. The dataset allows intelligent algorithms to be developed by bridging the gap between machine learning and healthcare. Although sampling bias is one of the limitations of correlation analysis, it does improve understanding. This presents opportunities for improving precision in mental health management, fostering interdisciplinary collaborations, and creating dynamic forecasting models. Researchers and policymakers can benefit greatly from the RHMCD-20 dataset, which offers nuanced insights into mental health experiences during quarantine and informs evidence-based interventions and policies. groundwork for innovative methodologies, steering

* Corresponding author at: Department of Information Convergence Engineering, Pusan National University, Busan, Republic of Korea.
E-mail address: imrus15-8978@diu.edu.bd (I. Salehin).

https://doi.org/10.1016/j.dib.2024.110376

the trajectory of informed decision-making in dynamic energy landscapes.

Specifications Table

Subject	Data Mining, Statistical Analysis, mental healthcare, AI
Specific subject area	Categorical Data Analysis for Identify Rapid Human Mental Health Depression During Quarantine Life Using Machine Learning.
Data format	Raw, Filtered.
Type of data	Filtered measurement data is stored as .csv files.
Data collection	The RHMCD-20 dataset was gathered in Bangladesh using a structured survey approach that covered a wide range of age groups and professional backgrounds. We included teenagers, college students, housewives, corporate professionals, and people in a variety of professions using established methods. A number of thoughtfully constructed questions on mental health during quarantine were included in the survey, which was distributed via online and in-person channels. These inquiries were designed to gather a thorough grasp of the experiences of the participants, encompassing elements linked to behavioral changes, stress, frustration, family history of mental illness, physical changes, mood swings, coping strategies, interests at work, and social interactions. In order to further our understanding of the topic, the survey was designed to measure the mental health difficulties that individuals in Bangladesh experienced when they were isolated for an extended period of time.
Data source location	Department of Computer Science and Engineering Northern University Bangladesh & Daffodil International University, Dhaka, Bangladesh. Khilgaon, Uttara and Bashundhara residential area, Dhaka, Bangladesh.
Data accessibility	Repository name: Mendeley data Data identification number: 10.17632/pxjmjyfdh2.1 Direct URL to data: https://data.mendeley.com/datasets/pxjmjyfdh2/1

1. Value of the Data

- The RHMCD-20 datasets are highly valuable due to the crucial information they offer in understanding and dealing with various aspects of mental health and overall well-being while under quarantine. Every characteristic adds to a comprehensive picture of participants' experiences and provides insightful information.
- Researchers can use these datasets to create predictive models and learn more about factors of isolation, such as age, gender, employment, and number of days spent indoors, that affect mental health. The details on mounting stress, frustrations during the quarantine, habit changes, and coping mechanisms offer important data points for researching the psychological effects of prolonged isolation.
- Including past mental health issues, weight changes, and mood swings is beneficial It makes it possible to investigate intergenerational mental health and identify risk factors for specific treatments. Using important markers of participants' mental health during quarantine, researchers use this data to identify trends and create specialized interventions for those feeling emotional disorders during long-term isolation.
- The dataset provides insights into participants' work-related experiences, including changes in work interest and struggles with social interactions. This information is essential for understanding the broader impact of quarantine measures on individuals' professional and social lives.
- This dataset is a vital resource for researchers investigating the relationship between machine learning and healthcare. It combines historical benchmarks and real-time metrics in the field of clinical data, especially mental health. It enables the development of intelligent algorithms and complex predictive models to improve knowledge and assistance for mental health issues.

2. Background

The RHMCD-20 dataset originated from a thorough investigation of the dynamics of mental health among people in Bangladesh during quarantine. As opposed to datasets that concentrate on particular sectors or industries, this dataset aims to explore the complex emotional and psychological terrain that is shaped by prolonged periods of isolation. Providing a thorough grasp of the complex difficulties that people from a variety of backgrounds and occupations encounter while under quarantine is the goal. Its potential to advance mental health care and research makes this dataset significant. The dataset captures the diversity and depth of experiences by compiling data from a broad range of participants, such as professionals, housewives, college students, and teenagers. From demographic information to more in-depth questions about stress, frustrations, coping strategies, and habit changes, the dataset's questions are carefully crafted to reveal important details. The dataset has a greater impact when AI technologies are integrated. The objective is to not only record and capture experiences, but also to open the door for creative approaches to mental health treatment. The dataset attempts to provide customized interventions and support strategies that are in line with the changing field of mental health research and the particular difficulties brought about by Bangladeshi quarantine experiences by utilizing artificial intelligence. The RHMCD-20 dataset is essentially the result of a deliberate attempt to connect the fields of AI and mental health, promoting a better comprehension of and ability to address the intricacies of psychological well-being during isolation.

3. Data Description

While compiling the RHMCD-20 dataset, we took care to include information from a wide range of sources, including teenagers from Bangladesh, college students, housewives, professionals from businesses and corporations, and other people. We guarantee representation from a range of life stages and professional backgrounds with our inclusive approach. The dataset attempts to provide a complete picture of the mental health of the Bangladeshi populace during quarantine by combining viewpoints from various ages, genders, and professional domains. This wide variety of participants adds to the dataset's richness and depth and enables a detailed investigation of the experiences of mental health among a range of people dealing with the difficulties of prolonged isolation in Bangladesh.

A total of 1000 participants took part in the survey; of these, 825 responded, yielding an approximate 82.5% response rate overall. The percentage of missing responses (175) divided by the total number of participants yields the missing response rate, which comes off to be 17.5%. This suggests that 17.5% of the respondents did not fill out the survey.

Our dataset's questions are sincerely formed to reveal important facets of people's experiences during quarantine, providing crucial information for mental health treatment and research. Starting with personal information such as age, gender, and profession, these questions were foundations for the participants' background. One of the key factors affecting mental health during quarantine is the degree of social isolation, which is explored in the question about the number of days spent indoors. To capture emotional and behavioral shifts and provide important insights into the psychological impact of isolation, questions about increasing stress, frustrations in the initial weeks of quarantine, and major changes in eating and sleeping habits aim to capture emotional and behavioral shifts, offering valuable insights into the psychological impact of isolation. The exploration of family mental health history and changes in body weight introduces a generational and physical health perspective. More specific mental health dimensions are assessed through questions about extreme mood swings, difficulties coping, and loss of interest in work. Lastly, the inquiry into feeling mentally weak during interactions sheds light on social and emotional challenges. This dataset serves as a comprehensive tool for mental health care professionals and researchers, facilitating tailored interventions, support strategies and understanding of the multifaceted nature of mental health experiences during quarantine. The questions address a variety of behavioral, psychological, and demographic aspects, emphasizing how these components are interrelated under the particular conditions of a worldwide pandemic.

Demographic Information: To gather basic demographic information, the first questions ask about age, gender, and occupation. Understanding the various circumstances and backgrounds of the survey respondents requires knowledge of these specifics.

Quarantine and Behavior Modifications: The next queries focus on the effects of quarantine, including the amount of time spent indoors and any notable changes to sleeping and feeding schedules. These investigations add to a better understanding of behavioral adaptations by illuminating the lifestyle changes people have made in response to the pandemic.

Stress and Emotional Well-Being: In order to better understand participants' emotional well-being, the survey asks questions on how much stress they felt they had been under during the first two weeks of quarantine in particular. The emotional and psychological aspects of the respondents' experiences are further explored with questions concerning abrupt mood swings, difficulties managing, and feelings of mental weakness in social situations.

Physical Health Changes: By asking the same question again about variations in body weight during a quarantine, the need of keeping an eye on physical health outcomes which may be strongly linked to mental health is highlighted. This simultaneous emphasis on psychological and physical factors offers a comprehensive viewpoint on people's well-being during the pandemic.

Work and Social Engagement: The survey's final section looks at shifts in interest in work-related topics, possible drops in interest, and emotions of mental fragility in social situations. These inquiries delve into how the pandemic has affected the participants' social and professional lives, providing a window into the wider ramifications for their everyday routines and mental well-being.

4. Experimental Design, Materials and Methods

A number of factors are included in the data analysis: age, gender, occupation, days spent indoors, growing stress, frustrations during quarantine, changes to habits, History of mental health, fluctuations in weight, mood swings, difficulties coping, interest in work, and social weakness. The diagram illustrates the relationship between each variable by visualizing it. By using correlation analysis to identify possible dependencies and patterns in the dataset, variables that may have an impact on mental healthcare can be better understood. The correlation analysis results for the primary data variables are shown in Fig. 1.

Fig. 1. Total responses and missing responses of the partitioners.

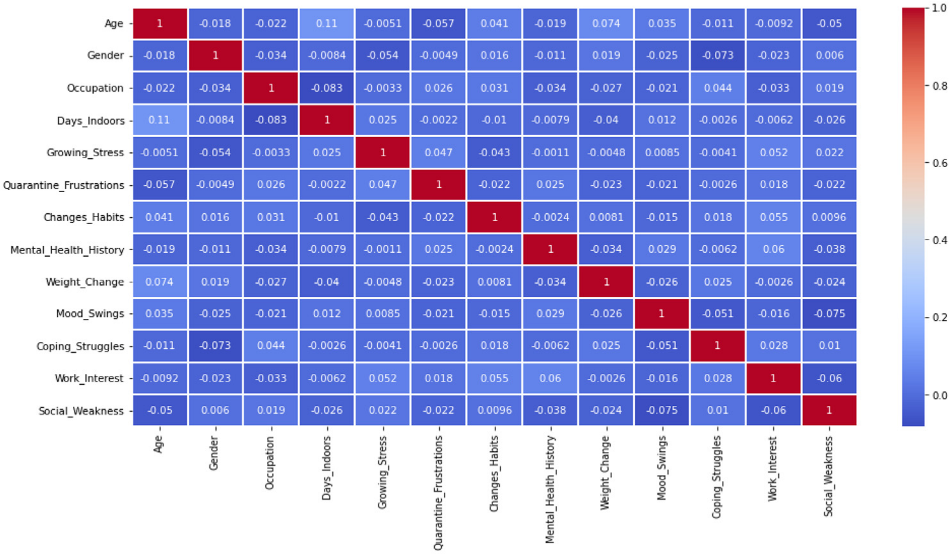

Fig. 2. Heat map displaying the correlations between the dataset variables.

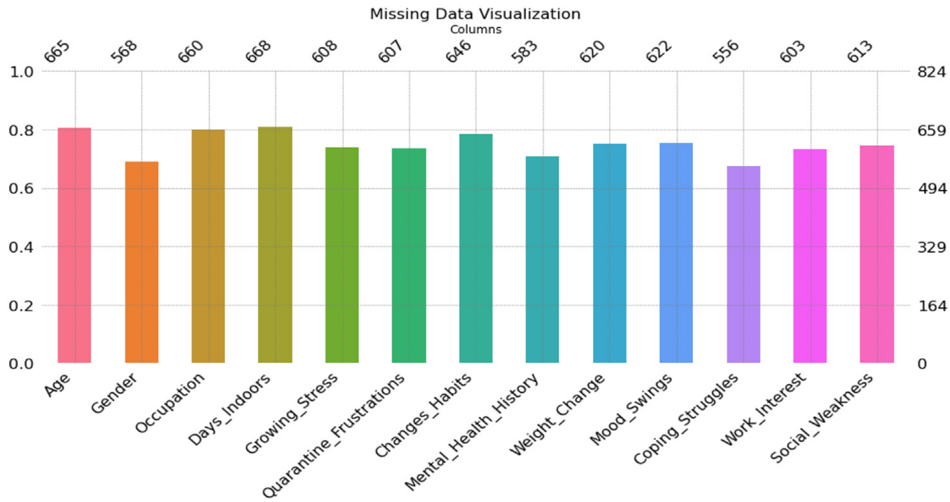

Fig. 3. Showcases the absence of values of the dataset.

A thorough depiction of the absent data is given in Fig. 3. This analysis focuses on the specific cases in which these critical variables show gaps in the available data. Finding any discrepancies in the dataset requires the use of this image. A more thorough dataset evaluation is ensured and issues with data completeness are greatly addressed by this kind of visual representation. Missing values in crucial variables are filled in during the data cleaning and missing value filling procedure. By imputed with the average number, it is addressed. Using the mean imputation approach, the computed averages were used to fill in the gaps left by missing values for each variable. The choice to utilize the average was made with the intention of preserving the dataset's general central tendency (Fig. 2).

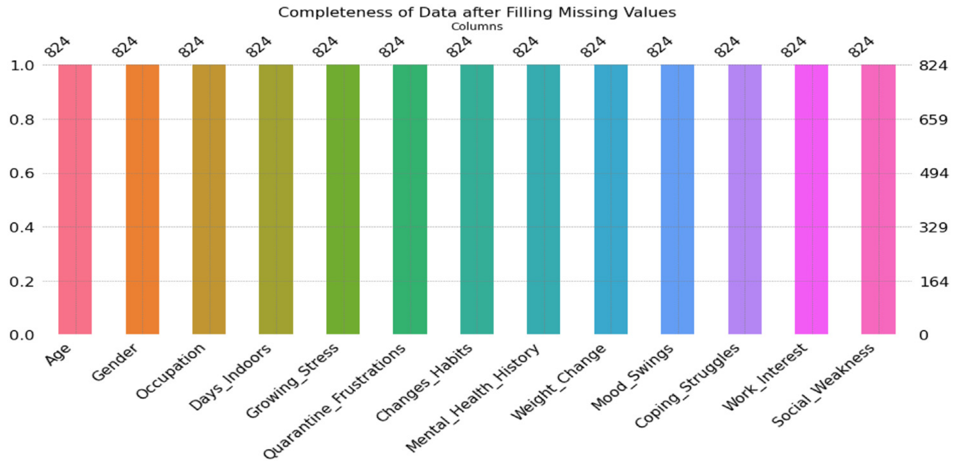

Fig. 4. Showcases the absence of values after filling the dataset.

Fig. 4 presents the cleaned dataset graphically after imputation. It is crucial to remember that mean imputation introduces bias if the data is consistently missing and assumes randomness in the missing data pattern. To guarantee the validity of ensuing statistical analyses, careful evaluation of the suitability of mean imputation for various variable types and awareness of potential effects on statistical analyses are crucial.

By splitting the data into intervals, or bins, and showing the frequency of occurrences within each bin as bars, the pie chart illustrates the distribution of the data. A pie chart could show different aspects of this dataset, including:

- **Age**: The age distribution of respondents in the quarantine satisfaction survey is shown in a pie chart. Notably, 26.5% of respondents are between the ages of 20 and 25, 26.3% are between the ages of 25 and 30, and 24.0% are older than 30. The youngest group is made up of 23.2% of people who are between the ages of 16 and 20.
- **Occupation**: The distribution of mental health occupations during the quarantine survey is shown in the pie chart. The largest group is made up of business professionals, who make up 22.6%, and housewives, who make up 20.4%. 19.3% were students, and 18.0% were others.
- **Days-Indoors**: Different lengths of time spent indoors during quarantine are depicted in the pie chart: 22.8% spent 32–60 days, 20.4% spent 15–30 days, 17.6% spent more than two months, and 20.3% spent 1–14 days. Above 18% of them also went outside each day, indicating a range of reactions to extended periods spent indoors.
- **Other Variables**: The distribution of other parameters, including Growing Stress, Quarantine Frustrations, Changes in Habits, Mental Health History, Weight Change, Mood Swings, Coping Difficulties, Work Interest, and Social Weakness, can also be shown using a pie chart.

This dataset's pie chart is used to graphically depict the frequency of occurrence of particular values or ranges within various variables, providing insights into the distribution patterns and frequencies of those values within the dataset. A pie chart that shows the frequency of different levels within the dataset is used in Fig. 4 to illustrate the distribution of mental health data variables. In the RHMCD-20 Dataset about experiences with quarantine, we put into practice a mapping procedure for categorical variables. The objective of this mapping is to provide numerical representations for qualitative data, including age, gender, occupation, mood swings, and days spent indoors. Given that numerous algorithms require numerical input, this transformation is essential for quantitative analysis, machine learning, and statistical

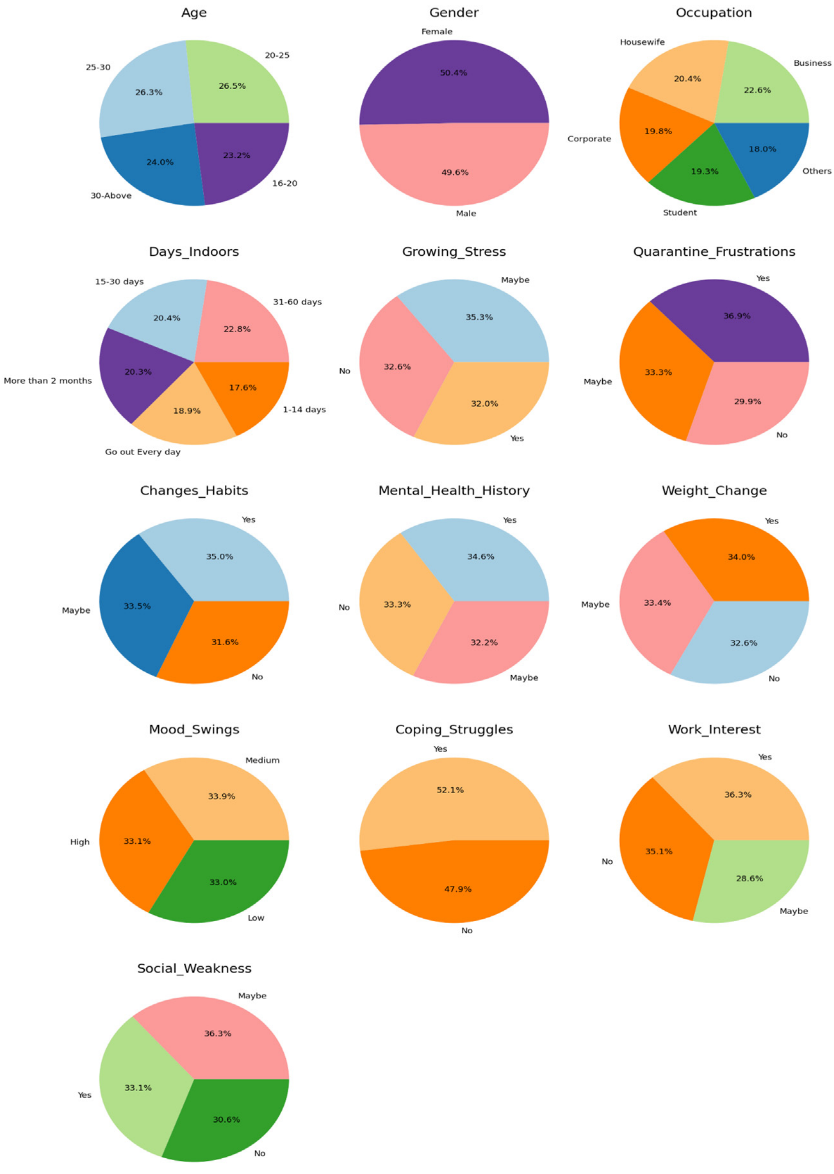

Fig. 5. Distribution of variables age, gender, occupation, growing stress, quarantine frustrations, changes habits, mental health history, weight change, mood swings, coping struggles, work interest, and social weakness.

modeling [1]. We make it possible to examine patterns, correlations, and trends within the dataset by giving numerical values to categorical labels [2]. Researchers can gain valuable insights into the factors influencing mental health during quarantine thanks to this numeric representation, which also improves the dataset's compatibility with a variety of analytical techniques (Fig. 5).

Researchers can visit the Mendeley Data repository and use the search function to find the particular dataset of interest if they are interested in accessing the dataset hosted on Mendeley Data. Researchers can examine specifics, variables, and related documentation after locating the

Table 1
An outline of the variables measured in the RHMCD-20 dataset.

Name of data column	Description
Age	Represents the age of the participants.
Gender	Indicates the gender of the participants.
Occupation	Represents the participant's occupations.
Days_Indoors	Indicates the number of days the participant has not been out of the house.
Growing_Stress	Indicates the participant's stress is increasing day by day.
Quarantine_Frustrations	Frustrations in the first two weeks of quarantine.
Changes_Habits	Represents major changes in eating habits and sleeping.
Mental_Health_History	A precedent of mental disorders in the previous generation.
Weight_Change	Highlights changes in body weight during quarantine.
Mood_Swings	Represents extreme mood changes.
Coping_Struggles	The inability to cope with daily problems or stress.
Work_Interest	Represents whether the participant is losing interest in working.
Social_Weakness	Conveys feeling mentally weak when interacting with others.

dataset. Usually, Mendeley Data offers alternatives to get the dataset straight from the repository. The dataset's thorough details on age, gender, occupation, and mental health indicators can provide researchers and policymakers with profound insights into the complex landscape of mental well-being during the COVID-19 epidemic. Through the use of sophisticated analytical techniques, researchers are able to identify risk factors for stress and coping and forecast the course of mental health. By enabling tailored mental health care initiatives, comparative investigations across demographics offer a comprehensive knowledge of intergenerational changes. Decisions made by policymakers can be supported by evidence, and they can carefully plan the distribution of resources and create policies. Workplace-related factors guided the creation of policies that promote resilience and mental health by illuminating the pandemic's effects on professional spheres. Additionally, the information improves public awareness campaigns by facilitating customized messaging to address particular issues and bolstering community-wide initiatives to support mental health (Table 1).

This table provides a comprehensive overview of the mapping process, converting categorical survey responses into numerical values for each variable, enhancing the dataset's analytical utility [3].

Several popular machine learning algorithms can be used to analyze and predict mental health in this dataset. This includes binary classification algorithms like SVM and Naive-Bayes, as well as Random Forest and Logistic Regression. With replies to questions about age, gender, occupation, and mental health during the COVID-19 pandemic, the RHMCD-20 Dataset is a useful tool for a range of analysis. Machine learning algorithms can be used in predictive modeling to predict stress levels based on behavioral and demographic data, which can help with individualized support plans and early intervention. By examining factors including job interest, coping mechanisms, and behavioral changes, workplace and societal impact evaluations can be carried out, offering employers and legislators valuable information for developing well-being programs. When factors like stress, coping mechanisms, and the impact of disorders from earlier generations are taken into account, intergenerational comparisons can also highlight trends in mental health experiences among various age groups. The multidimensional nature of mental health during the pandemic is revealed by these analyses, which gives researchers and policymakers the information they need to develop support programs and treatments that are specifically targeted at particular behavioral and demographic traits (Table 2).

Analysis of Survey Response Rate and one-sample t-test:

In this section, we analyze the response rate to the survey conducted as part of our research study. A total of 1000 participants were invited to participate, and of these, 825 respondents completed the survey, yielding an approximate overall response rate of 82.5%.

The response rate was calculated by dividing the number of respondents (825) by the total number of participants invited (1000), resulting in the observed response rate of 82.5%. However,

Table 2
Categorical to numerical mapping of survey responses in the RHMCD-20 dataset.

Variable	Original value	Mapped value	Question
Age	16–20, 20–25, 25–30, 30-Above	0, 1, 2, 3	How old are you?
Gender	Female, Male	0,1	Gender?
Occupation	Corporate, Others, Student, Housewife, Business	0, 1, 2, 3, 4	What is your occupation?
Days_Indoors	Go out Every day, 1–14 days, 15–30 days, 31–60 days, more than 2 months	0, 1, 2, 3, 4	How many days have you not been out of the house?
Growing_Stress	No, Maybe, Yes	0, 1, 2	Is your stress increasing day by day?
Quarantine_Frustrations	No, Maybe, Yes	0, 1, 2	During the first two weeks of quarantine, did any kind of frustration affect your life?
Changes_Habits	No, Maybe, Yes	0, 1, 2	Major changes in eating habits and sleeping?
Mental_Health_History	No, Yes	0, 1	Has there been any change in body weight during quarantine?
Weight_Change	No, Maybe, Yes	0, 1, 2	Has there been any change in body weight during quarantine?
Mood_Swings	Low, Medium, High	0, 1, 2	Extreme mood changes of?
Coping_Struggles	No, Maybe, Yes	0, 1, 2	Inability to cope with daily problems or stress?
Work_Interest	No, Yes	0, 1	Are you losing interest in working?
Social_Weakness	No, Yes	0, 1	Are you feeling mentally weak when talking with others?

to assess the significance of this response rate and to compare it against a hypothesized value, we conducted a one-sample t-test.

- Formulate Hypotheses:

Null Hypothesis (H_0): The true response rate is equal to the expected response rate (e.g., 80%).

Alternative Hypothesis (H_1): The true response rate is not equal to the expected response rate.

- Calculate the Sample Proportion:

Calculate the sample proportion of respondents who filled out the survey. In this case, the sample proportion is 82.5% or 0.825.

- Specify Expected Response Rate:

Define the expected response rate against which you want to test. Let's say it's 80%, or 0.80.

- Compute the Standard Error:

Calculate the standard error of the sample proportion using the formula for proportions in the below equation:

$$S_E = sqet\left[\frac{[p * (1 - p)}{n}\right] \tag{1}$$

Where:

S_E = Standard Error p = Sample proportion (0.825 in this case) n = Sample size (825 in this case)

- Calculate the T-Statistic:

Calculate the t-statistic using the formula:

$$t = \frac{p - p_0}{S_E} \tag{2}$$

Where: p0 = Expected response rate (0.80 in this case)

- Determine Degrees of Freedom (d_f):

Degrees of freedom = Sample size - 1 = 825 - 1 = 824

The sample proportion of respondents who completed the survey was calculated as 82.5 %. Using this proportion and the expected response rate of 80 %, we calculated the standard error and subsequently the t-statistic. The computed t-statistic was compared against the critical value from the t-distribution table, with 824 degrees of freedom and a significance level of 0.05. From our *t*-test statistical table, it indicates a positive hypothesis rate (0.10 (Two-tailed) and 0.05 (Two-tailed)).

Our cutting-edge mental health prediction modeling makes use of a variety of potent machine learning methods, such as Support Vector Machines (SVM), Random Forest, Logistic Regression, and Naive Bayes. These sophisticated methods provide a strong basis for improving precision and comprehending complex linkages in mental health dynamics. The many perspectives that this algorithmic method offers can be essential in advancing the creation of focused therapies and tactics that promote mental health [4]. We present a systematic four-step modeling methodology that includes data pretreatment, data sampling, data gathering, and the smooth integration of data into machine learning methods. This thorough methodological approach guarantees a solid basis for accurate and dependable analysis and prediction of mental health outcomes. The main objective is to improve mental health [6] management techniques by facilitating more individualized and successful treatments. Using Random Forest, Logistic Regression, Naive Bayes, and SVM, we want to influence the direction of future mental health research. Recognizing the various obstacles presented by the ongoing epidemic, this inclusive approach aims to support the creation of customized solutions that meet the specific requirements of people navigating the intricacies of mental health [5].

Subsequent investigations employing the RHMCD-20 mental health dataset exhibit considerable potential to propel our comprehension and remediation tactics concerning mental health throughout the current COVID-19 epidemic. Researchers may make use of the different properties of the dataset to create more accurate models for predicting mental health outcomes by investigating novel methodologies, such as customized machine learning models and sophisticated statistical analysis. Furthermore, including extraneous variables, such as socioeconomic indicators or contextual variables, may improve the precision and resilience of prediction models. Seasonal impacts, geographic differences in mental health issues, and subtle trends may become apparent with more research into the dataset's temporal patterns and regional variances. The development of novel approaches for more precise and flexible mental health forecasts can result from interdisciplinary collaborations between domain professionals, data scientists, and mental health specialists. Additionally, by using the dataset to investigate new developments in mental health, such as changes in coping techniques, resilience variables, or the effects of intervention tactics, forward-looking models in line with the changing field of mental health will be developed.

In conclusion, the RHMCD-20 curated mental health dataset offers a comprehensive range of characteristics that capture the varied experiences of people throughout the COVID-19 pandemic. This dataset includes important characteristics including age, gender, employment, number of days spent indoors, stress levels, coping strategies, and elements connected to work. The RHMCD-20 dataset analysis provides important insights into the complex dynamics of mental health, which serve as a basis for evidence-based treatments and policy recommendations that are customized to meet the specific requirements of various populations. Researchers and policymakers may use this tool to further our knowledge of mental health issues and build resilience in these unheard-of times.

Limitations

The RHMCD-20 dataset sheds light on the experiences of mental health during Bangladeshi quarantine. Nevertheless, its drawbacks include the possibility of sampling bias brought about by the emphasis on particular groups, the use of self-reported responses, which introduces subjective biases, and a narrow temporal scope that might miss changing conditions. The size of the dataset is indicative of potential for improving representativeness, even though it is informative. The depth of understanding may be impacted by the dataset's lack of certain cultural quirks and contextual information. Although a useful tool for targeted study within predetermined parameters, researchers should take these limitations into account and proceed with caution when interpreting results.

Ethics Statement

The Ethical approval for this study involving survey methods was obtained from the Daffodil International Student Research Board before conducting the research. All procedures performed in this study involving human subjects were by the ethical standards and guidelines of BMRC. Informed consent was obtained from all individual participants included in the study. In cases where explicit ethical approval for survey-based studies is not mandatory per institutional or national regulations, this study adhered to the ethical principles outlined in the BMRC, ensuring confidentiality, voluntary participation, and respect for the participants' rights and privacy. All identifiable data were handled with strict confidentiality and anonymized during analysis to protect the participants.

Data Availability

The RHMCD-20 datasets for Depression and Mental Health Data Analysis with Machine Learning (Original data) (Mendeley Data).

CRediT Author Statement

Nazrul Amin: Investigation, Software, Visualization, Writing – original draft, Writing – review & editing; **Imrus Salehin:** Investigation, Software, Visualization, Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing; **Md. Abu Baten:** Conceptualization, Data curation, Writing – review & editing; **Rabbi Al Noman:** Conceptualization, Data curation, Writing – review & editing.

Acknowledgments

The authors wish to thank members of the Information Convergence Engineering Research Lab and Daffodil International University Research Lab and anonymous referees for their helpful comments on earlier drafts of this paper.

Declaration of Competing Interest

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

References

[1] V. Gudivada, A. Apon, J. Ding, Data quality considerations for big data and machine learning: going beyond data cleaning and transformations, Int. J. Adv. Softw. 10 (1) (2017) 1–20.
[2] S.B. Kotsiantis, I. Zaharakis, P. Pintelas, Supervised machine learning: a review of classification techniques, Emerg. Artif. Intell. Appl. Comput. Eng. 160 (1) (2007) 3–24.

[3] I. Salehin, N. Amin, The RHMCD-20 datasets for depression and mental health data analysis with machine learning ", Mendeley Data V1 (2023) Salehin, Imrus; Amin, Nazrul (2023), "The RHMCD-20 datasets for Depression and Mental Health Data Analysis with Machine Learning ", Mendeley Data, V1, 10.17632/pxjmjyfdh2.1, doi:10.17632/pxjmjyfdh2.1.
[4] D.B. Thieme, G. Doherty, Machine learning in mental health: A systematic review of the hci literature to support the development of effective and implementable ml systems, ACM Trans. Comput.-Hum. Interact. (TOCHI) 27 (5) (2020) 1–53.
[5] C. Decouttere, K. De Boeck, N. Vandaele, Ad-vancing sustainable development goals through immu-nization: a literature review, Glob. Health 17 (1) (2021) 95.
[6] I.M. Talha, I. Salehin, S.C. Debnath, M. Saifuzzaman, N.N. Moon, F.N. Nur, Human behavior impact to use of smartphones with the python implementation using naive Bayesian, in: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, 2020, pp. 1–6.