Impact on Human Mental Behavior after Pass through a Long Time Home Quarantine Using Machine Learning

Imrus Salehin[a*]**, Sadia Tamim Dip**[a]**, Iftakhar Mohammad Talha**[a]**, Ibrahim Rayhan**[a]**, Kanij Fatema Nammi**[a]

[a]Department of Computer Science and Engineering, Daffodil International University, Dhaka 1207, Bangladesh
Email: imrus15-8978@diu.edu.bd

Abstract: In the present situation, COVID-19 is a very common and dangerous issue in the whole world. Ensuring our healthy mental state is very essential at the period of COVID-19. But as a result of being in the home quarantine for a long time, people are going to notice a mental change such as stress, depression, mood swing. We proposed an RHMCD model which helps us to reach our required goal. This model contains machine learning algorithms. We examined our work with Naive Bayes classifiers, Support Vector Machine, and logistic regression. For gaining the report of mental conditions we used the sentiment analysis technique. For measuring the level of depression we also used a decision tree approach.

Index Terms: Human mental behaviour, Home quarantine, Machine learning, Decision tree, Time factor.

1. Introduction

In the last 100 years, 2020 has become a special year for the human race. This year, the whole world is under threat from a virus called Covid-19. Home isolation is a special way to protect ourselves from this virus. However, people who have been isolated at home for a long time are experiencing a mental change [1]. From medical data and data collection from people of different quarantines, we analyzed the change in mindset and its impact on our research. More than 10.1 million people are now infected with the virus, according to world meter records [2]. Nowadays most people are living at home for self-isolation and safe living. In this paper, we present the additional effects of home quarantine through data mining and machine learning approaches. We focused on changes in behavior, mental imbalances, and its effects through our research. We used the regression analysis method, Support vector machine (SVM), Naive Bayes classifier algorithm for calculating accuracy value. The main contributions of the research which are briefly as follows:

- We have studied, the major changes in mental imbalances, behavioral changes for the long time period of home isolation. We proposed an RHMCD (Rapid Human Mental Change Detection) model.
- In this research paper, we have worked on the neuro molecular system of the brain as well as the time factor.
- The proposed hybrid method of detecting human depression levels, stress, and mental balance are the most significant for this study. We used the decision model to discover the level of human frustration caused by quarantine.
- In this research, we are describing the collaboration of the medical sector and machine learning. We also reveal the harmful and mental disorder through home isolation which is very indispensable to our upcoming days.

2. Literature Survey

The incidence of coronavirus is increasing day by day. Mental health is one of the major impacts. But we do not have a headache with this important issue. And no one in our society talks about it and the number of psychologists is insignificant. But because of unhealthy mental health, many syndromes like depression are seen in our society all the time. Corona Virus is disrupting our social life and has a huge impact on our mental health. Its effects on children from long-term confinement at home are even greater, with adults making them even more upset by angry voices. Providing psychosocial counseling and supporting bereaved family members in such a situation is Substantial. Children can react

to stress in various ways. More clingy, nervous, retreating, angry, or angry, such as Scrambling, bedwetting, etc. This is the right time when we all need to do more research on mental health and hire more professional doctors. Guaranteeing our good mental wellbeing is critical [3]. An analysis of hospital workers who may have come into touch with SARS showed that having been quarantined was the most predictive predictor for signs of acute stress disorder shortly after the quarantine time (9 days) ended. Even 3 years later, the effects of becoming quarantined were an indicator of signs of posttraumatic stress in hospital staff. Quarantine impact is long term and its effect is long-lasting. Quarantine is essential for the epidemic we are going through. In the meantime, we need to pay close attention to mental health. Everyone should be made of the side effects of quarantine and continue to provide psychologist advice to all. Because we need healthy mental health to get rid of the crisis we are living in. This is a good time to focus on mental health [4]. COVID-19 has now turned into an epidemic. As a result, our normal life is being disrupted. As a result of the epidemic of our school, college, Varsity has been closed. Because of this, we are facing a lot of losses, the future is becoming uncertain, and especially the impact on university students is increasing. Students are facing depression and stress as their future is uncertain. In our paper, we show that its impact is becoming more visible among the age group of 20 to 25 years students. As a consequence, day by day, their mental health is worsening. To keep our future health, our main concern is to ensure healthy mental health [5]. Isolation is a system that separates infected people and quarantine, where the symptoms of the infection appear. Earlier in history, during the epidemic, such a situation had become a quarantine. It reminds us of the history of yellow fever, Cholera, and Pelage. The COVID-19 outbreak began in December 2019. With 118,000 incidents and 4,291 casualties in 114 countries, it has become a big global concern. It was declared an epidemic by the World Health Organization (WHO). Those involved in health care are often neglected. The purpose of the umbrella review is to assess quarantine-related mental health effects and to separate them from current evaluations. Their mental health system also faces various obstacles. In our paper, we show that it is important to bring them under the service first [6]. 264 million persons suffer from depression, according to the World Health Organization. Depression is one of the most used terms of mental health in people currently in quarantine. Depression levels are predicted using SVM and Naïve Bayes classifications. Classifiers are applied to the 9 symptoms of depression as recognized by the American Psychiatric Association Diagnostic and Statistical Manual (DSM-IV). With its help, we can predict our current state of mental health. They built a model that uses two separate classifiers to define the UGC: Support Vector Machine (SVM), and Naive Bayes. In our paper we use Naive Bayes, SVM, Logistic Regression for more accuracy and more efficient result. From this, we can know how much damage is being done to our mental health [7]. Through algorithms, we can differentiate people with mental illness from others. In the context of the psychological patient dataset, the purpose of this analysis is to allow a comparative comparison of the Naïve Bayes and K-Nearest Neighbor classifiers. Useful information can be given by the derived data trends to preventing psychiatric illness and attempting to provide appropriate mental health care. Through this, all the steps to be taken for mental health. If we don't think about mental health, it will destroy us. Quarantine has a huge impact on mental health, In our paper, we can see how terrible its effects are by applying the algorithm. The convergence of the medical sector and machine learning is represented in our study. We rely on the prediction of dangerous and psychiatric illness that is important to our world and culture that is coming [8].

3. Proposed Model

The proposed multi-featured model depicted in Fig 1. there are 3 steps in total. The first step involves the collection and combination of various data.
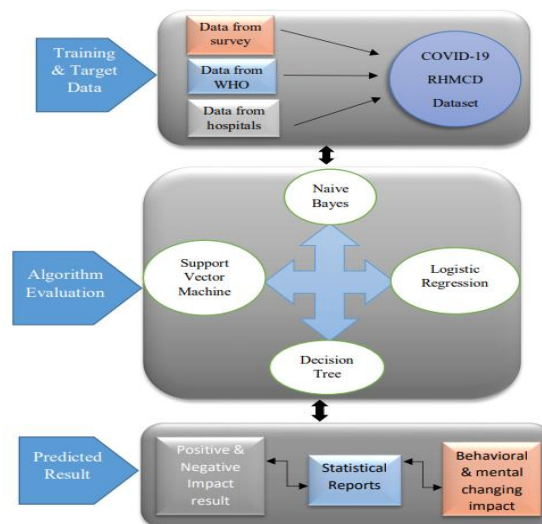

Fig.1. A short RHMD model review about our study

The second step mainly designed with machine learning algorithms. From Naive Bayes and Decision tree, we get the sentiment analysis report and depression level report. SVM and Logistic Regression also help us to get the classification report result. At the end of the study, we represent the predicted accuracy then we ensure the model value and working power. It provides the result of the mental imbalance and changes during the COVID-19 period and long-term home quarantine stage.

4. Methodology

The desired target class can be easily determined from the dataset through machine learning. The value of the target class can be accurately inferred from different dataset classes. From our survey data, we determine their level of mental health from the data set of the main impact of mental health. The value is determined by analyzing these levels with the scale of the main level [9,10,11].

We have given 4 criteria for mental health:

- Very Low (0-40%)
- Low (41-60%)
- Medium (61-80%),
- High (81-100%)

By comparison, we get the value of the Impact of quarantine. Target class → Quarantine Impact

$$\varphi: X \rightarrow Y \tag{1}$$

Determine as $\varphi(X) = Y$
X=Value of Quarantine Impact
Y=Level of mental health

$$\varphi(X) = \begin{cases} \text{Low Impact} \\ \text{Medium Impact} \\ \text{High Impact} \end{cases} \tag{2}$$

The First step is selection, the main attributes have to be selected for the training data. Ineffectual attributes must be removed from the data set. The work of our paper is split into three sections, the first section is training Data [11]. In this section we need to create a data set for the convenience of our work from the data we have collected from the survey-Age & Gender, Main Impact of Quarantine, Relation between Main Impact and Age. Now all the data sets have to be trained by Naïve Bayes and Logistic Regression Algorithms. To get the Ultimate result in the second section, measure the target data. And tests this value until more reliable detection is obtained. In the last and third section, we implement Python code for better accuracy and separate different mental health data by SVM for more accuracy. We applied three algorithms (Naïve byes, SVM, Logistic Regression) for our work and it gave us the curator value of our work. Naive byes are supervised learning and it is the best algorithm to find out the probability. We prepared the naive Bayes algorithm in our work they did sentiment analysis. We have used SVM for classification of Human Mental Health, under the supervision of Supervised Learning. The last one is Logistic Regression. We have applied it to predict the quarantine impact. From the data we have collected we first find out the probability, then we find the accurate class, and its effect is predicted.

4.1. Naïve Bayes Classifier

Naïve Bayes is the most prosperous Algorithm to find out the probability condition and the probability of the condition according to the work technique. The impact of mental health on the attributes of the main impact data set is distributed the value of accurate quarantine impact by the naïve Bayes algorithm. Naive Bayes Algorithm serves based on the Bayes theorem. Most commonly used for Naive Bayes is Probability and Sentimental Analysis [18]. Algorithms have been used to diagnose the effects of mental health on quarantine.

$$P(c|x) = P(x|c)P(c)/P(x) \tag{3}$$

P(c|x), where c is target class and x is Attributes, P(c|x) known as Posterior Probability of class.
P(c) is the antecedent Probability of class.
P(x) is the antecedent Probability of Predictor.
P(x|c) is the probability of the given class of Predictor.

For the analysis of our data at first, we need to remove Punctuations, Numbers, and Special Characters then we remove short words from our data set. Tokenization also done at this period. Using wordCloud we find the frequently used words.

```
combi['covid_data'] = combi['covid_data'].str.replace('[^a-zA-Z#]+',' ')
combi['covid_data'] = combi['covid_data'].apply(lambda x: ' '.join([w for w in
x.split() if len(w) > 3]))

tokenized_covid_data = combi['covid_data'].apply(lambda x: x.split())
```

For the analysis of our data at first, we need to remove Punctuations, Numbers, and Special Characters then we remove short words from our data set. Tokenization is also done at this period. We also use the bag-of-words method which is a representation of text that describes us the occurrence of any words within a document [17].

4.2. Support Vector Machine

SVM is also part of Supervised Learning [13]. SVM does the operation of classification in our work. Through hyperplane and margin, we have separated human mental health for the convenience of our work. SVM has separated the classes of Quentin Mental Health and Normal Mental Health From that we got accurate value.

$$\gamma_i = (w.x_i + b)y_i \tag{4}$$

$y\_i \in \{-1,+1\}$, $⟦y⟧\_i$=Impact or No Impact
$\gamma\_(i)$, is Margin of data Pint
$sign(w.x\_i+b)$, is Probability.

4.3. Logistic Regression

Another algorithm we use for training our data set is Logistic Regression. It is also a well-used algorithm in predictions used in machine learning under Supervised learning [14]. Logistic Regression establishes the relation of variables to the effects obtained from our survey. It then produced our output based on independent variables (mental health of the healthy person) and dependent variables (during quarantine mental health) probability. Reached our goal with the best model by logistic regression

$$P(Probability\ of\ Quarantine\ Impact\ 0\ 0r\ 1) = 1/1 + e^{-(All\ Impact)} \tag{5}$$

If $P() \geq 0.5$ then Impact 1 (Quarantine Impact)
else Impact 0(No Quarantine Impact).

4.4. Decision Tree

To find out the impact of quarantine and for measuring the level we use the decision tree model. The decision tree is a supervised machine learning algorithm that helps us to make a decision about anything. We use the decision tree to find out if segregation really affects our human health. We measure it with a level. We define those levels as Very Low Impact, Low Impact, Medium Impact, and High Impact. If we want to develop any system, and it works properly and we can get a good result then we need to collect data from different persons and different places. Then we can have got totally different values and best accuracy. Many people are disturbing and change their mood, habits, Interest during Home Quarantine. Our Standard dataset from separate areas has been used for proper testing and results. We just take some most important attributes (total question 13) for our testing.

$$SplitInfo_A(D) = -\sum_{j=1}^{V} \frac{|D_j|}{|D|} \times log_2 \left(\frac{|D_j|}{|D|}\right) \tag{6}$$

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo_A(D)} \tag{7}$$

Decision tree algorithm uses information gain and Gini index technique to split the root node. The splitting process continues until it finds the last child node [15].
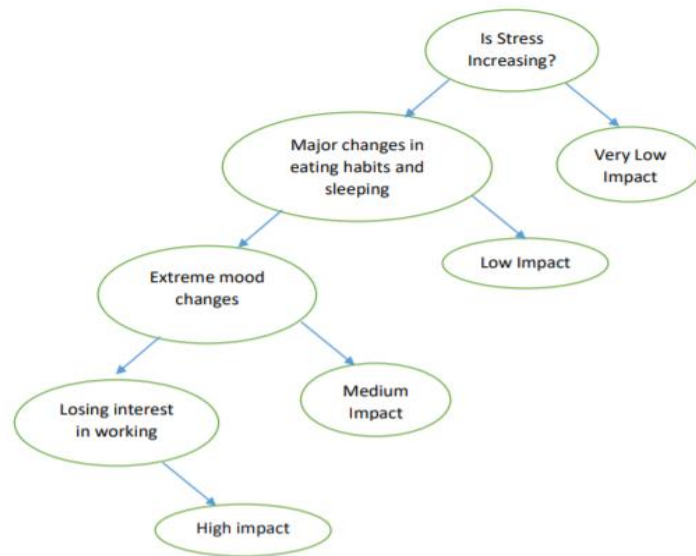
Fig. 2. Decision Tree with impact category

Using this model, we can clearly explain the level of our quarantine impact through this decision tree. There may be several changes in our Human Mental Behaviour in this Quarantine, but these four changes are estimated to be significant challenges. Based on these four changes, the scale of quarantine's impact on mental health behaviour is ranked. The first change is stress increasing, if this change is manifested in someone it is caught on a very low impact scale. The next move is to notice a change in eating habits and sleeping, accompanied by a cycle of low impact. In this process, the scale of impact increases when the frontal signs appear. If 4 major changes are revealed in someone, then it is caught on the scale of high impact see Fig 2.

5. Data Collection

Data Collection it is the process of collecting data from various fields. For data collection, we have created a Google Form and we post it on social media. Due to the lockdown, social media is becoming the only means of communication for people. Facebook, Instagram, WhatsApp, YouTube has become a close friend to all due to the pandemic situation. For this reason, we design our google form using 13 questions set from where we can collect our data. We also collect our data physically. Around 10 thousand data we collect for our research. There we got the most common problem because of Covid-19(Corona Virus). We got a very good response from people.

5.1. Data Processing

If we want to develop any system, and it work properly and we can get good result then we need to collect data from different persons and different places. Then we can have got totally different values and best accuracy. Many people are disturbing and change their mood, habits, Interest during Home Quarantine. Our Standard dataset from separate areas has been used for proper testing and results. We just take some most important attributes (total question 13) for our testing.
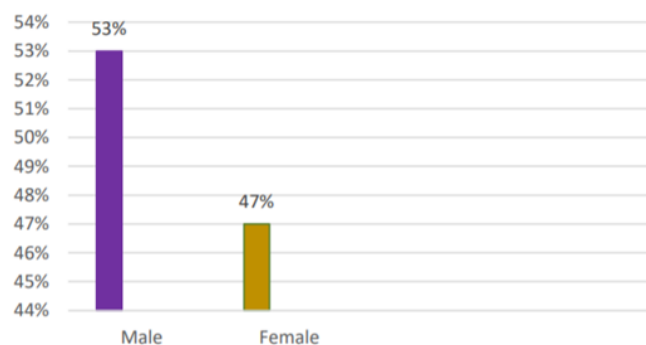

Fig. 3. Showing percentage of gender

There we got many data from student, housewife, businessman and corporate members. And we used different level to collect data. It is very helpful to used different level to identity all of them. In our data collection process we separately calculate our data age, gender and accuracy. In our total data there have 57% make and 43% female data shown in Table1 and Fig 3.

Table 1. Showing gender in different age

Age	Gender	Male	Female
16-20	Children	32%	68%
20-25	Teenager	51%	49%
25-30	Adult	55%	45%
30-above	Most Adult	33%	67%

In our survey, we give different ages and calculated male and female percentage according to our given ages. The data was collected under 30 and above year age. Our survey completed by starting age 16 to 30 above. Using age, we can find out which age are the most affected in this quarantine time. Also, can find out which gender. Table 2 shows the symptoms and criteria levels.

Table 2. Symptom and criteria level

Symptom	Ctieria
Stress	There we check, you can cope with your daily life or not.
Depression	For this home quarantine, how you feel and how you act.
No interest	You don't have any interest to do work or anything
Pervious mental disorder	If your any pervious family member have any mental disorder or problem
Change mood	Your mood change extremely and you feel something serious like suicide

5.2. Data Analysis

We are using Naïve Bayes, support vector machine and linear regression algorithm for this dataset. It gives us a good feedback and improve and evaluation our work. Using this Algorithms, we find good accuracy for our papers. Accuracy hold papers correct prediction of any situation. Actual results came when we can find good accuracy for this dataset [8].

We got 4036 response from peoples through social media. Out of 4036 peoples 2300 peoples were age [20-25]. There were 825 peoples age [16-20], 512 peoples were age [25-30] and others 399 peoples were age 30 above. In this different age peoples, we got different accuracy for male and female both of them.
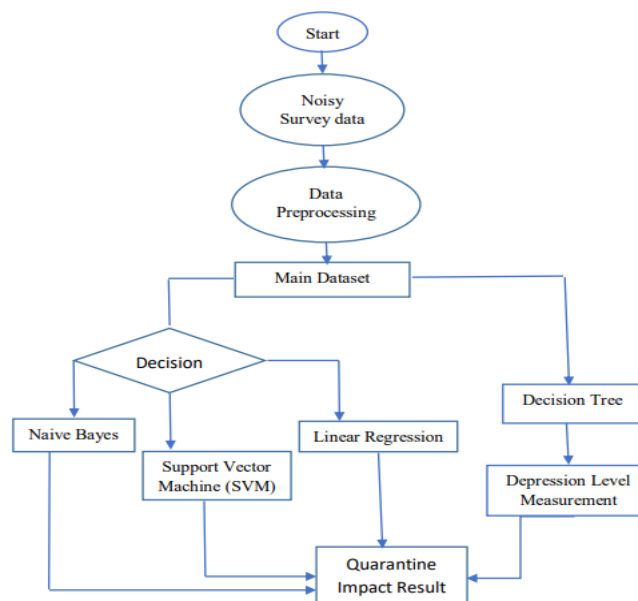

Fig. 4. A diagram Flowchart

Mainly we make some questions for this mental health survey, but we take only the most important data from our survey called the Main dataset. If we need any checking or identifying any data for that he/she has a problem or not then we searching our dataset. All the data in the dataset are correctly calculated using our selected algorithms and applying python code for 100% correct accuracy. Fig 4. Shows that how we can pre-process our data and then we can calculate the accuracy. Then this accuracy tells us to have any mental illness or not.

6. Result

At present home-quarantine is a totally new situation for us. Times have changed now. Now we have to stay home most of the time. Because of this, our daily routine has changed. As a result, we have to keep pace with many new conditions. Being isolated has also changed our mental state. That is why quarantine is becoming a source of frustration for us. We performed a machine learning technique to find out what kind of conditions we have to face at this time and what will be our mental conditions during this pandemic period. We performed 3 individual classifiers to find out the analysis result see Fig 5.
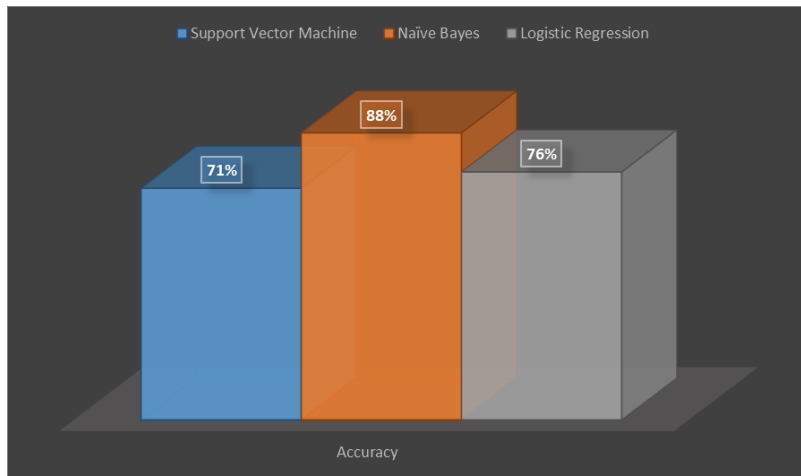

Fig. 5. Algorithm performance result

Homelessness, unemployment, poor living conditions, financial problems all play a huge role in changing human health. We use the emotion analysis technique to understand all the feelings so that we can fully understand one's mental state at this pandemic stage [16].
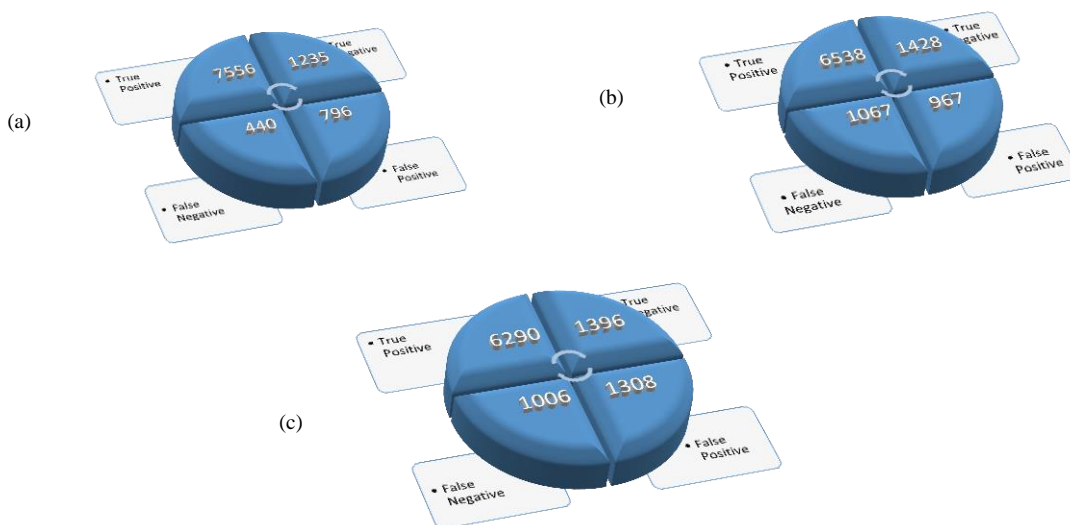

Fig. 6. Confusion matrix of (a) Naïve Bayes, (b) Logistic Regression, (c) Support Vector Machine

Using this technique we achieved 88% accuracy. From the Fig 6., we can see that this algorithm can classify correctly 7556 data. The other two methods got 76% and &71% accuracy and from the confusion matrix, we see that the algorithm correctly classified 6538 and 6290 data. Using sentiment analysis techniques we can understand what is going on in someone's mind. We get the most negative impact of the people we have examined. This indicates that being quarantined for long periods of time changes our thinking, changes our way of talking and behaviour, which can have an adverse effect on our human health, mental health and also our social life. These changes leads us to irrational mood-swings, inattention, changes in habits, and many other problems.
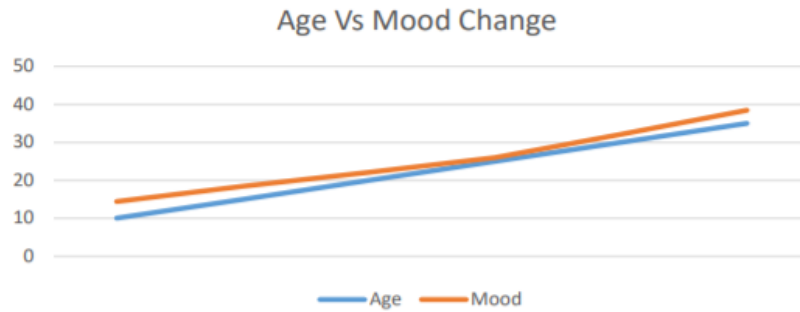

Fig. 7. Age vs Mood Change graph

We also examined a relation between age and mood-swing. We can see that mood swings are most evident in people between the ages of 20 and 30. The mood-swing problem is less common with old-age see Fig 7.
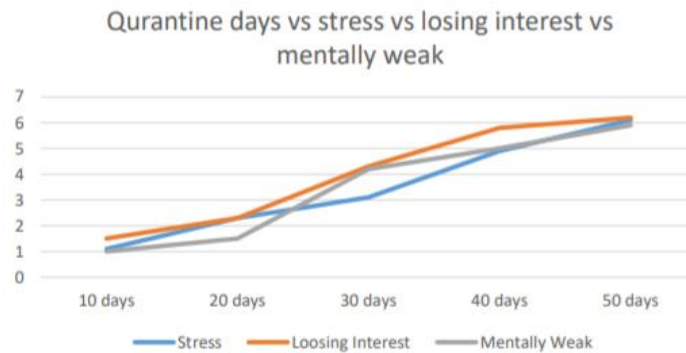

Fig. 8. Days vs Overall Condition graph

Staying in quarantine for long periods of time increases stress, leads to loss of focus on work, and changes in the human condition shown in Fig 8. As the days go by, these 3 problems also become apparent, which adversely affects our personal life or social life.

7. Discussion

Corona is one of the few epidemics in human history. Many are dying because of this carnivorous. We are in a different situation to survive this epidemic. Most of the time we have to stay at home. As a result, our daily routine has changed. These changes affect our lifestyle and also our mental health. Through our research, we have tried to find out what kind of impact the corona condition will have on human health. We used two types of machine learning algorithms in this experiment. We performed sentiment analysis through our data. From which we can get an idea of the human condition of a person and how much impact quarantine has on him. We also use a decision tree algorithm to measure the impact of quarantine.

8. Conclusion

In this paper, by using data mining and machine learning we can easily say that the pandemic we are going through is harming our mental health. As a result of COVID-19, it is very essential to ensure our healthy mental health. Quarantine stress, depression, mood change is most noticeable in people between the age of 20-25. Through the data outline, we see that Quarantine is having an impact on our mental health which is perilous for healthy mental health. The age group of 20 to 25 years plays a significant role in the upgrading of a nation. Assuring their healthy mental

health should be the principal responsibility and the main subject of any community. The goal of this paper is to shed light on the fact that quarantine is having an enormous impact on our mental health.

References

[1] Katsiaryna V. Gris, Jean-Philippe Coutu and Denis Gris*, Supervised and Unsupervised Learning Technology in the Study of Rodent Behavior. Frontiers in Behavioral Neuroscience | www.frontiersin.org. July 2017 | Volume 11 | Article 141

[2] Online :https://www.worldometers.info/coronavirus/.COVID-19 CORONAVIRUS PANDEMIC. Last Access: Last updated: July 07, 2020, 17:53 GMT

[3] Anant Kumar & K. Rajasekharan Nayar (2020): COVID 19 and its mental health consequences, Journal of Mental Health

[4] Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet. doi:10.1016/s0140-6736(20)30460-8

[5] Sahu P (April 04, 2020) Closure of Universities Due to Coronavirus Disease 2019 (COVID-19): Impact on Education and Mental Health of Students and Academic Staff. Cureus 12(4): e7541. doi:10.7759/cureus.7541

[6] Hossain MM, Sultana A, Purohit N. Mental health outcomes of quarantine and isolation for infection prevention: A systematic umbrella review of the global evidence. Epidemiol Health. (2020) DOI:10.4178/epih.e2020038

[7] Aldarwish, M. M., & Ahmad, H. F. (2017). Predicting Depression Levels Using Social Media Posts. 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS). doi:10.1109/isads.2017.41

[8] Dr. E. Chandra Blessie, Bindu George(NCACCT - 2019 Conference Proceedings). A Novel approach for Psychiatric Patient Detection and Prediction using Data Mining Techniques.

[9] Mhambe Priscilla Dooshima, Egejuru Ngozi Chidozie, Balogun Jeremiah Ademola, Olusanya Olayinka Sekoni, Idowu Peter Adebayo. A Predictive Model for the Risk of Mental Illness in Nigeria Using Data Mining. International Journal of Immunology. Vol. 6, No. 1, 2018, pp. 5-16. doi: 10.11648/j.iji.20180601.12

[10] Deziel, M., Olawo, D., Truchon, L., and Golab, L., Analyzing the mental health of engineering students using classification and regression. EDM 2013:228–231, 2013.

[11] Vanlalawmpuia, R., & Lalhmingliana, M. (2020). Prediction of Depression in Social Network Sites Using Data Mining. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). doi:10.1109/iciccs48265.2020.9120899

[12] Cosic, K., Popovic, S., Sarlija, M., … Kesedzic, I. (2020). IMPACT OF HUMAN DISASTERS AND COVID-19 PANDEMIC ON MENTAL HEALTH: POTENTIAL OF DIGITAL PSYCHIATRY. Psychiatria Danubina, 32(1), 25–31. doi:10.24869/psyd.2020.25

[13] Srividya, M., Mohanavalli, S. & Bhalaji, N. Behavioral Modeling for Mental Health using Machine Learning Algorithms. J Med Syst 42, 88 (2018). https://doi.org/10.1007/s10916-018-0934-5

[14] Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and Simple Linear Regression. Radiology, 227(3), 617–628. doi:10.1148/radiol.2273011499

[15] Decision Tree Classification in Python Online Available: https://www.datacamp.com/community/tutorials/decision-tree-classification-python

[16] Wawre, Suchita V., and Sachin N. Deshmukh. "Sentiment classification using machine learning techniques." International Journal of Science and Research (IJSR) 5.4 (2016): 819-821.

[17] Suppala, Kavya, and Narasinga Rao. "Sentiment analysis using naïve Bayes classifier." International Journal of Innovative Technology and Exploring Engineering 8.8 (2019): 265-269.

[18] I. M. Talha, I. Salehin, S. C. Debnath, M. Saifuzzaman, M. N. N. Moon and F. N. Nur, "Human Behaviour Impact to Use of Smartphones with the Python Implementation Using Naive Bayesian," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225620.

[19] I. Salehin , I. M. Talha, M. Saifuzzaman, N. N. Moon and F. N. Nur, (2020) "An Advanced Method of Treating Agricultural Crops Using Image Processing Algorithms and Image Data Processing Systems", on 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA).

Authors' Profiles

Imrus Salehin, studied Computer science and engineering from Daffodil International University. Main Interested research fields are Image Processing, Machine Learning and Data Mining. His also research interests include Data science and Computer vision. Mr. Imrus Salehin is main author of two papers in the international conferences and Journal.

Sadia Tamim Dip, studied Computer science and engineering from Daffodil International University. Main Interested research fields are Image Processing, Machine Learning and Data Mining, Data science and Computer vision.

Iftakhar Mohammad Talha, studied Computer Science and Engineering from Daffodil International University. Main Interested research fields are Image Processing, Machine Learning and Data Mining, Data science and Computer vision. Mr. Iftakhar Mohammad Talha is main author of one paper in the international conferences and one Co-author Journal.

Ibrahim Rayhan, studied Computer science and engineering from Daffodil International University. Main Interested research fields are Image Processing, Machine Learning and Data Mining, Data science and Computer vision.

Kanij Fatema Nammi, studied Computer science and engineering from Daffodil International University. Main Interested research fields are Image Processing, Machine Learning and Data Mining, Data science and Computer vision.