Machine Learning Model to Analyze Mental Health

Aditya Mehta[a], Lakshay Chawla[b], Mehul Rekhi[c], Akanksha Kochhar[d]*

[a]*Maharaja Agrasen Institute of Technology,Plot No 1, Sector 22, PSP Area, Delhi, 110085, India (delhi.adityamehta@gmail.com)*
[b]*Maharaja Agrasen Institute of Technology,Plot No 1, Sector 22, PSP Area, Delhi, 110085, India (lakshaychawla13@gmail.com)*
[c]*Maharaja Agrasen Institute of Technology,Plot No 1, Sector 22, PSP Area, Delhi, 110085, India (rekhi.mehul2000@gmail.com)*
[d]*Maharaja Agrasen Institute of Technology,Plot No 1, Sector 22, PSP Area, Delhi, 110085, India (akankshakochhar@mait.ac.in)*

Abstract: As we all know, mental health problems in India are not recognized well So, we set our aim as to create a platform that will be accessible to all, and one does not have to face any criticism for expressing his/her feelings. They can just go to this platform, fill in a couple of forms and get help from professionals.

1. Introduction

Mental health is serious topic, which is often being ignored, especially in Indian societies. This topic is in conversation and gathering steam in the recent times. The December 2019 coronavirus disease outbreak has caused a huge number of people getting quarantined because of getting affected. Quarantine, in general, means that you're not allowed to come near vicinity of anyone if you're affected. Often, quarantine is not associated with a true to life experience for the person undergoing that phase. Getting separated form loved ones, losing freedom, not knowing when the phase and illness will end and moreover boredom creates enormous effects. Number of reported suicides have increased1, anger levels have increased, and lawsuits brought2 following the growth in previous outbreaks.

During such hard times, coronavirus being the way it is, quarantine must be the only option. But it has been suggested in various studies that quarantine brings with it negative connotations and weird psychological effects. It has been observed that although all of this is bad firsthand experience, the worrying part is the extension of these effects even after quarantine sometimes even a year later. – although the studies highlighting such effects are not very much. [4,5] This suggests that special care must be taken with such cases.

Considering the advent of technology, the best way to make a good progress in the situation would be to bring the helpline in the hands of everyone. It's a well-known fact that the barrier to smartphone acquirement is
Decreasing. So, enabling an individual to know about his/her mental health would be a great boon for the society. In traumatic situations, people don't want to talk to anyone but resort to being lonely and spending time alone which is mostly spent on mobile phones. So, bringing help directly where it is most effective would be a good option.

2. Requirement Analysis

User side – To keep the project accessible to the greatest extent, we decided to make a website, that means, anyone on any device can use this with just a web browser being the requirement.

For development –

- HTML/CSS/JS/BOOTSTRAP
- React
- Django
- Python
- Postgresql

3. System Design

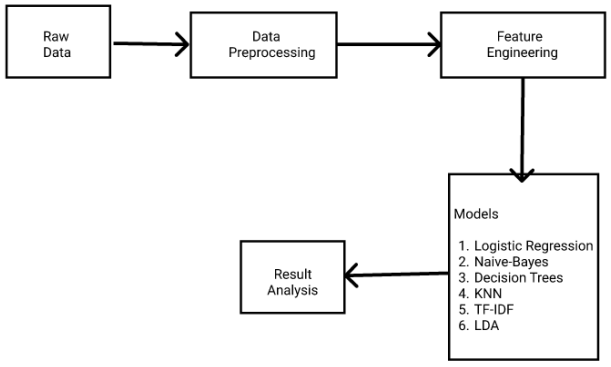

Fig.1- Data flow in machine kearning

3.1 Raw data

Considering the topic is really region specific and we can't use existing datasets, we created a survey form with help of psychologists and collected our own data by surveying actual people.

3.2 Pre-processing

Extracting all the raw data, importing it in the python script, we cleaned the data, removing null values and making it such that we can perform operations on them.

3.3 Feature Engineering

Upon analysing the data, we saw tha splitting data in 3 sections would be useful. Applied different operations to make the data useful and easy to use, then went on to implement models.

3.4 Machine Learning algorithms (Models)

Trying to learn new machine learning models, we stumbled upon many of them, at the end deciding to give a chance to these models, as they seemed apt for our data.

3.4.1 Logistic Regression –

Logistic regression is a type of algorithm where it determines from the given training data a logistic function that fits the data relatively accurately. Then on new data, it applies that function and gives a certain outcome. Generally, the outcome of this algorithm is binary, that is one of two outcomes – 1 or 0.

3.4.2 Naïve Bayes –

As students, we learnt Naïve Bayes theorem, which was use in probability.

The work we are doing here is almost the same, so this was instantly s3elected. Using that theorem, a model is created that is helpful in classification of binary data outcome.

3.4.3 KNN –

K-Nearest neighbors, as the name suggests, this algorithm stores the previous data and upon new data entry it checks nearest k neighbors, the proximity of the new data point to one of the existing outcomes reveals its nature. Specifying the K – that is the number of neighbors, we can customize the results

3.4.4 Decision Trees

Decision trees are like inverted trees. The model creates a tree of certain depth, which we can specify, to try and fit the training data making entry exit gates. This is the only algorithm till now which can be used for regression as well, but we do not need that feature, we'll just try to fir our classification data.

3.4.5 Tf – Idf –

Tf – Idf is a method used to process language, commonly known as "Term frequency – Inverse Document Frequency". This is a method which analyzes the whole data and then try to classify which data is relevant and which is not. Then using clever mathematic functions create numerical version of linguistic data

3.4.6 LDA –

LDA stand for "Latent Dirichlet Allocation". We are using this model to create two-way outcome model. It will take values from Tf – Idf function and upon analyzing divide it in two sections (Good or Bad, in our case).

3.5 Result Analysis

Now, running our data through all the algorithms and checking the results various times, that is changing training and testing data multiple times, we got the following results.

Table 1 – Accuracy scored of machine learning models.

Model Name	Accuracy
Logistic Regression	88.8
KNN	86.8
Multinomial Naïve Bayes	25.3
Decision Trees	84.3

Of all the algorithms, Logistic regression came out on top with a blistering 88.8% accuracy. But this also has accuracy not satisfactory to us, to improve upon that we decided to make our custom model that was a combination of a custom function, logistic regression model and Tf-Idf /LDA model.

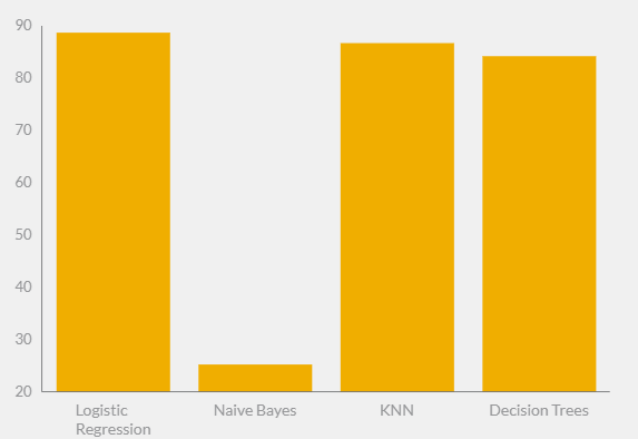

Fig.2- Comparison graph of different machine learning models

Finally, after we were satisfied with the model, we deployed all our work on a website – 'psychepath.xyz'.

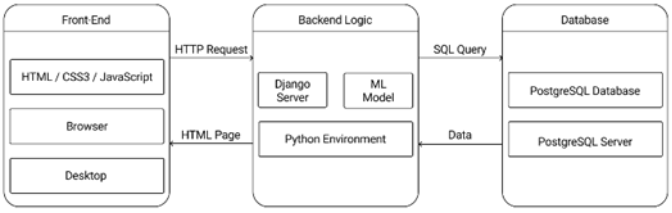

Fig.3- Architecture of Website

In the website as you can see in the architecture diagram, the frontend webpage takes the required inputs from the user, transfers it to the backend Django server using react JS which in turn calls the custom machine learning model that we built and gives the output. If the user is tagged for potential mental risk, he/she is given a certified medical test PHQ-9 to confirm the mental state. On the result, the user is directed to a professional practitioner if required.

4. Literature Survey

Taking upon the difference between the effect of this grave situation on people who were quarantined and not, some studies were made. [3,4,5,6,7]. The most obvious pick was a hospital. Looking upon the staff who was quarantined for 9 days or more showed symptoms of not very good shape. The were diagnosed with acute stress disorder.

Expanding on the research, the same staff reported multiple instances of early exhaustion, not willing to socialize, inability to sleep and what not. It was also noticed that concentration levels were on a decline. In another study7, post traumatic stress symptoms were found in staff even after 3 years. So, the takeaway here is that quarantine indeed takes upon even on the best of us.

5. Conclusion

We noticed from the readings and observations, that indeed something must be done to improve the situation regarding mental health issues. Building upon that we got in touch with some psychologists and developed a form for tagging. The form had been set such that the person filling the form wouldn't feel filling a mental health checkup, giving him/her some friendliness. On the other hand, it had enough information for us to process. We built a custom model to tag the person, following on to a certified medical test which would come up only if the model tags you. After getting the form submitted, you can contact professionals from the site only, without needing to face societies stigma. We hope for a better society and a Earth being a better place to live for people facing mental health issues.

Acknowledgement

We as a team would like to thank our mentor/guide Ms. Akanksha Kochhar to help up along the whole journey. Journey becomes smoother if you have enough motivation. Our main source of motivation was Dr. Namita Gupta, head of Department, Computer Science and Engineering in our college, who constantly gave us useful feedback and strength to complete the project.

References

Barbisch D, Koenig KL, Shih FY. Is there a case for quarantine? Perspectives from SARS to Ebola. Disaster Med Public Health Prep 2015; 9: 547–53.

Miles SH. Kaci Hickox, Public health and the politics of fear. Am J Bioeth. 2015; 15(4):17-9

Bai Y, Lin C-C, Lin C-Y, Chen J-Y, Chue C-M, Chou P. Survey of stress reactions among health care workers involved with the SARS outbreak. Psychiatr Serv 2004; 55: 1055–57.

Jeong H, Yim HW, Song Y-J, et al.

Mental health status of people isolated due to Middle East respiratory syndrome. Epidemiol Health 2016; 38: e2016048.

Liu X, Kakade M, Fuller CJ, et al. Depression after exposure to stressful events: lessons learned from the severe acute respiratory syndrome epidemic. Compr Psychiatry 2012; 53: 15–23.

Taylor MR, Agho KE, Stevens GJ, Raphael B. Factors influencing psychological distress during a disease epidemic: data from Australia's first outbreak of equine influenza. BMC Public Health 2008; 8: 347.

Wu P, Fang Y, Guan Z, et al. The psychological impact of the SARS epidemic on hospital employees in China: exposure, risk perception, and altruistic acceptance of risk. Can J Psychiatry 2009; 54: 302–11.

Twende J, Martin G.Gender differences in associations between digital media use and psychological well-being: evidence from three large datasets.J Adolesc. 2020; 79:91-102

Iannotti RJ, Janssen I, Haug E, et al.Interrelationships of adolescent physical activity, screen based sedentary behavior, and social and psychological health. Int J Public Health. 2009;54:191–198.