



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 7, July 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms

Prof. Ramesh, Prof. Sruthi M K, Apoorva B S, Anusha A M, Divya M N, Kavya

Project Coordinator, Asso. Prof, Department of Computer Science Engineering, S J M Institute of Technology,
Chitradurga, India

Project Guide, Asso.Prof, Department of Computer Science Engineering, S J M Institute of Technology,
Chitradurga, India

B.E Student, Department of Computer Science Engineering, S J M Institute of Technology, Chitradurga, India

B.E Student, Department of Computer Science Engineering, S J M Institute of Technology, Chitradurga, India

B.E Student, Department of Computer Science Engineering, S J M Institute of Technology, Chitradurga, India

B.E Student, Department of Computer Science Engineering, S J M Institute of Technology, Chitradurga, India

ABSTRACT: Anxiety, sadness, and stress are typical mental health issues in today's fast-paced society. Machine learning algorithms were utilised to develop predictions regarding anxiety, depression, and stress in this work. Data from the Depression, Anxiety, and Stress Scale questionnaire was gathered from individuals with and without employment from a variety of cultures and groups in order to apply these algorithms (DASS 21). Anxiety, sadness, and stress were predicted to occur at five distinct degrees of intensity using a variety of algorithms. Because they are so precise, these algorithms are ideal for spotting psychological issues before they arise. We discovered that our confusion matrix did not include enough of each class after trying several ways. So, the f1 score was introduced. Among the five models tested, the Random Forest classifier was the most accurate. In addition, the specificity parameter indicated that the algorithms were quite sensitive to negative outcomes.

KEYWORDS: Anxiety and Stress Scale questionnaire Random Forest classifier, Machine Learning Algorithms.

I. INTRODUCTION

People are more driven than ever before, and they seize any opportunity to advance in their careers. Many individuals today believe that anxiety and despair are simply a regular part of the workday since they have become so commonplace. Depression is the most frequent mental illness, according to the WHO. Over 300 million individuals worldwide have it, and many health professionals are investigating the issue due of its severity. A good learning algorithm is needed to distinguish between anxiety, depression, and stress in a machine's mind. As stated by WHO, a healthy individual is one who has a well-functioning brain and body. [1]

In order to diagnose depression, the PHQ is the conventional method. Anxiety, depression, and stress may all be assessed using the Depression, Anxiety, and Stress Scale (DASS 21) [5–6]. When it comes to clinical depression, the most common symptoms are memory loss, difficulty concentrating, inability to make judgments, a lack of interest in once-enjoyable activities like sex and hobbies, as well as eating excessively and gaining weight or not eating at all. You may also have feelings of shame, worthlessness, helplessness, restlessness, and irritability, and you may contemplate taking your own life. People in case [2] said that these symptoms adversely affected their ability to

function in daily activities including school, job, and socialising. A clinical diagnosis relies heavily on these data. Irritability, anxiousness, exhaustion, sleeplessness, gastrointestinal issues, panic, a feeling of impending doom and difficulty concentrating are all symptoms of Generalized Anxiety Disorder (GAD) [3]. Overreacting, experiencing headaches, and getting colds or infections that don't go away are indicators of stress. Other signs include feeling disturbed or irritated and not being able to relax. It is clear that the symptoms of stress and anxiety are similar to those of depression in that they include sleeplessness and chest discomfort as well as exhaustion and an elevated heart rate. As a result, it is difficult for machines to categorise them.

II. LITERATURE REVIEW

1) Mental Disorder Detection : Bipolar Disorder Scrutinization Using Machine Learning

AUTHORS: Ranjana Jadhav; Vinay Chellwani; Sharyu Deshmukh; Hitesh Sachdev

As our lives get more technologically advanced, we find ourselves under greater stress and strain. As a consequence, people are suffering from a variety of mental health issues. Untold numbers of patients with Bipolar Disorder have not been examined to determine whether these anomalies may be used to differentiate them from those with Mood Disorder or healthy controls, nor have they been classified based on how unwell they are overall. The Mood Disorder Questionnaire and Machine Learning are used in this research to determine whether a patient has Bipolar Disorder (MDQ). Using the Decision Tree Classifier, we were able to determine which parts of the data set were most essential and which ones were not. Each test case is assigned a class based on its comparison to the decision factor at each stage (Screened Positive or Screened Negative). Bipolar Disorder may be diagnosed using the Mood Disorder Questionnaire.

2) Intelligent data mining and machine learning for mental health diagnosis using genetic algorithm

AUTHORS: Azar, Ghassan & Gloster, Clay & El-Bathy, Naser & Yu, Su & Neela, Rajasree & Alothman, Israa

The inappropriate therapy is offered when mental health issues are misdiagnosed. Immediate or long-term hospitalisation or even death might result when a client's mental health is irrevocably harmed in this way. Approximately 12 million individuals in the United States are incorrectly diagnosed each year. New research demonstrates that semi-automated systems may aid in the early identification of a person suffering from a mental illness. The DSM-IV-TR, Fourth Edition Text Revision is used to compare a patient's mental health to the mental diseases listed in the DSM-IV-TR. Combining genetic algorithms, categorization data mining, and machine learning technologies, the research creates a semi-automated system. Not all mental diseases can be detected automatically, but a classifier should be aware of all the potential diagnoses that a patient may have based on their symptoms, not just the most likely ones. In this way, the classifier or analyst will be in a position to provide an accurate and pertinent analysis. As a result, a precise prognosis will be provided. The final diagnosis and treatment plan will be made by the analyst.

3) A Framework for Classifying Online Mental Health-Related Communities With an Interest in Depression

AUTHORS: B. Saha, T. Nguyen, D. Phung and S. Venkatesh

People, their families, and society as a whole are all affected by mental illness in a variety of ways. Online communities on social networks allow persons with mental problems to connect with others who have the same condition. If you're seeking for literary clues about mental illness, this database is a gold mine for you. Symptoms of a variety of mental illnesses may co-occur. People who suffer from anxiety disorders, for example, may also experience depression. We are focusing on this mental health issue that occurs simultaneously as we categorise internet groups that are interested in depression. We crawled 620k posts generated by 80k members in 247 online

groups to accomplish this. In our model, we've taken the postings' subjects and psycholinguistic properties and included them. In order to characterise co-occurring online communities related to mental health, we used a machine learning approach to create a joint modelling framework. As a final step, we evaluated our model on the crawling dataset and discovered that it performed better than current state of the art baselines.

4) Detecting Cognitive Distortions Through Machine Learning Text Analytics

AUTHORS: T. Simms, C. Ramstedt, M. Rich, M. Richards, T. Martinez and C. Giraud-Carrier

A growing variety of health-related fields are benefiting from machine learning and text analytics, particularly in the search for disease outbreaks and warning indicators of mental health issues. Cognitive distortion, a symptom and a cause of mental diseases including anxiety, anorexia, and depression, is the focus of our research here instead of behavioural changes. Using the Tumblr API, we retrieved a number of personal blogs and categorised them based on their warped thinking. Machine learning was then used on the vectors that were generated by LIWC. Research reveals that cognitive distortions may be automatically detected from personal blogs with a high degree of accuracy (73%) and a low percentage of false negatives (30.4 percent).

5) Machine Learning Framework for the Detection of Mental Stress at Multiple Levels

AUTHORS: Subhani, Ahmad & Mumtaz, Wajid & MOHAMAD SAAD, MOHAMAD NAUFAL & Kamel, Nidal & Malik, Aamir.

There has been an increase in the prevalence of mental health issues in society, which may make it difficult to carry out daily responsibilities at work. Long-term stress has also been related to a variety of physical and mental health problems. In addition, stress may lead to a variety of health issues such as a stroke, heart attack, and even death. According to the most recent findings in neuroscience, mental stress is primarily directed at the human brain. This is due to the fact that the brain's perception of a situation affects whether or not it is hazardous or stressful. In this case, an objective method of measuring stress levels that takes into consideration how the brain operates might make a significant difference in the severity of the consequences. A machine learning (ML) framework that analyses the EEG signals of persons who are stressed is proposed in this research. The Montreal Imaging Stress Task (MIST) is a well-known approach for causing stress in experiments. Because of how successfully the assignment was completed and how others reacted, there was a high level of stress. Features extraction, feature selection, classification and tenfold cross-validation were all part of the ML framework developed by Dr. Xia. The suggested framework was shown to be 94.6 percent accurate when recognising stress at two levels, and 83.4 percent accurate when identifying stress at more than two levels, according to the findings. This EEG-based machine learning system may, therefore, allow us to objectively detect stress on several levels. The technique might be used to develop a computer-aided stress diagnosis tool.

Existing System

People may now converse with each other more easily because to the rise of the internet and other communication technologies, in particular online social networking sites. Facebook, Twitter, Instagram, and other social media platforms allow its users to share their thoughts and opinions on a wide range of topics and issues. It's fantastic for those who use social networking sites since they may openly discuss and reply to any issue online. The flip side of this is that it allows those in the medical industry to learn more about why a patient has reacted the way they have. The distinctive patterns buried in online communication may be examined using machine learning methods in order to provide this kind of information, which can be used to determine if individuals are feeling "happy," "sad," "angry," "anxious," or "depressed." Studies on the impact of social networks on such things as breakups, mental illness (such depression or anxiety or bipolar disorder), relapses in drug or alcohol abuse or smoking and drinking, sexual harassment, or even suicide ideation are becoming more common.

Limitations of Existing System

- Accuracy of the algorithm is less
- Performance of the model is not good

Proposed Solution

Users' comments are used to determine whether they are sad in this project. In addition to text, images, and audio, users may upload in any of these formats. Depression sufferers may benefit from this project's encouragement messages. Online post services have replaced face-to-face communication as the preferred method of communication. As a result, this programme can recognise when a user is sad and send them uplifting messages. The person in charge of this app will send out inspirational messages to anybody who signs up for it. A user's postings are analysed using the SVM (support vector machine) method, which returns a positive or negative result. Using terms like "depression" in postings is seen as a negative by SVM, which considers it a bad post.

Advantages of Proposed System

- Messages of encouragement may be sent to sad individuals through this initiative. This is accomplished by scanning the comments on other people's blogs for indications of despair. To interact with one other, people increasingly utilise internet post services rather than face-to-face conversation.
- If there are any indications that relevant Facebook users may be depressed, we want to check their Facebook data for that information. Several types of machine learning methods are used in this process. This article discusses the following research issues in light of the study's primary goal.
- Facebook users often use terms like "joy," "sadness," "fear," "rage," or "surprise" to express their emotions in posts or comments.

Methodology

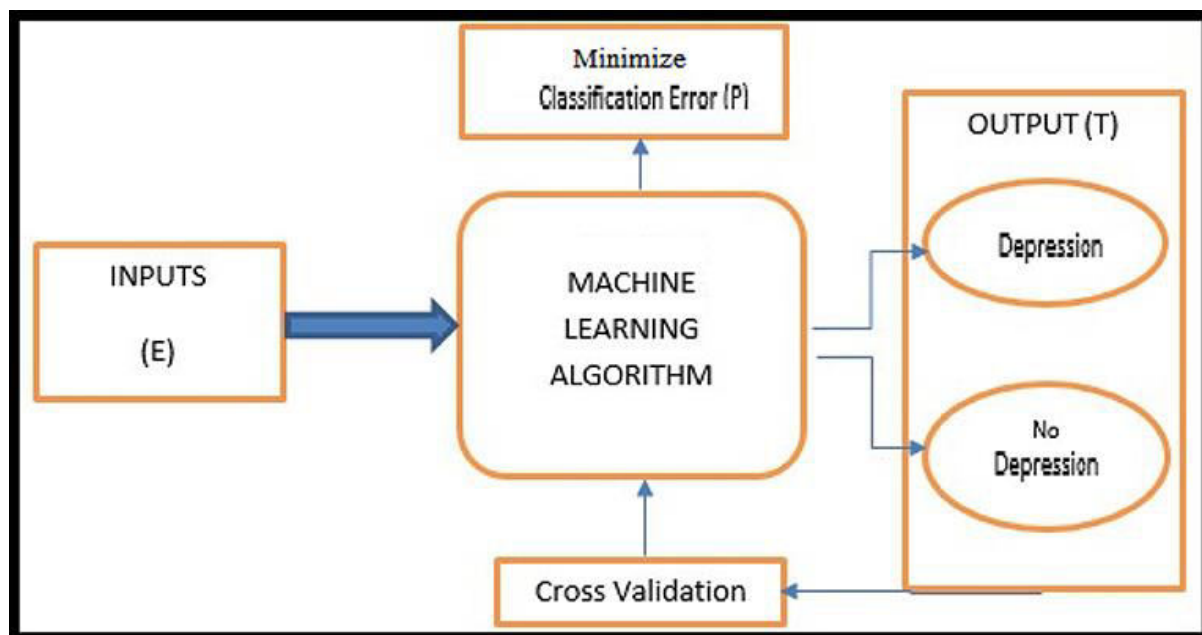


Fig: Architecture Diagram

Anxiety, depression and stress were examined using the Depression, Anxiety and Stress Scale questionnaire (DASS 21). 348 participants shared their data using Google forms, which were then sorted using five different machine learning algorithms: Support Vector Machine (SVM), Decision Tree (RT), Nave Bayes Decision Tree (NBDT), and KNN.

Participants

A total of 348 men and women between the ages of 20 and 60 participated in this research by completing a questionnaire. This group was responsible for a broad variety of tasks, ranging from domestic errands to work obligations.

Questionnaires

In order to gather data, the Depression Anxiety and Stress Scale (DASS-21) was used. The DASS 21 is comprised of a total of 21 questions. Each of the three stressors—stress, anxiety, and depression—has its own set of seven questions. The following are the potential responses for each question, which may be words or numbers: In my case, number 0 didn't apply. Some or most of the time, 1 was true for me. Most of the time, if not always, number two applied to me. True or near-true for me in at least three of them.

Classification

The machine learning techniques were implemented in Python 3.8. The severity of stress, anxiety, and depression symptoms is used to calculate the likelihood that a certain population may have any of these disorders. 70 percent of the data were used for training, while 30 percent were used for testing.

III. RESULTS & DISCUSSION

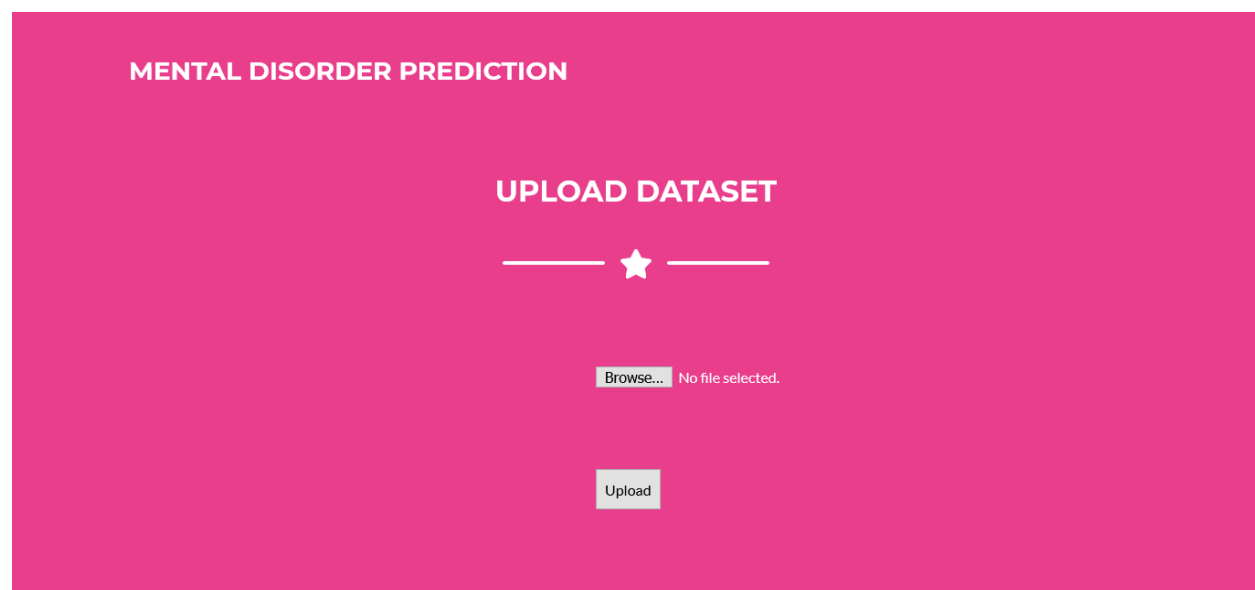


Fig: Upload Dataset

MENTAL DISORDER PREDICTION

PREVIEW



	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	remote_work
Id											
1	27-08-2014 11:29	37	Female	United States	IL	-	No	Yes	Often	Jun-25	No
2	27-08-2014 11:29	44	M	United States	IN	-	No	No	Rarely	More than 4000	No

Fig:Preview Dataset



Fig: Predict mental Disorder

IV. CONCLUSION

In this experiment, machine learning algorithms were employed to determine the severity of anxiety, sadness, and stress at five distinct levels. An anxiety, sadness, and stress questionnaire was utilised to quantify the most frequent symptoms (DASS-21). Decision Tree (DT), Random Forest Tree (RFT), Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) classification algorithms were then utilised (KNN). Although naive Bayes was determined to be the most accurate model, Random Forest was found to be the best model. The f1 score, which is utilised for unbalanced partitioning, was used to choose the best model since this issue caused imbalanced classes. It was discovered that "fear for no clear reason," "life was meaningless," and "hard to relax" were the most significant factors for anxiety, depression, and stress scores. As a result, these criteria were considered to be the most crucial in determining the presence of mental illness.

REFERENCES

- [1] Sau, A., Bhakta, I. (2017) "Predicting anxiety and depression in elderly patients using machine learning technology." *Healthcare Technology Letters* 4 (6): 238-43.
- [2] Tyshchenko, Y. (2018) "Depression and anxiety detection from blog posts data." *Nature Precis. Sci., Inst. Comput. Sci., Univ. Tartu, Tartu, Estonia*.
- [3] <https://adaa.org/understanding-anxiety/depression/symptoms>
- [4] https://www.webmd.com/balance/stress-management/stress-symptoms-effects_of-stress-on-the-body#1
- [5] Oei, T. P., Sawang, S., Goh, Y. W., Mukhtar, F. (2013) "Using the depression anxiety stress scale 21 (DASS-21) across cultures." *International Journal of Psychology* 48 (6): 1018-1029.
- [6] Kroenke, K., Spitzer, R. L., Williams, J. B. (2001) "The PHQ-9: validity of a brief depression severity measure." *Journal of general internal medicine* 16 (9): 606-613.
- [7] Sau, A., Bhakta, I. (2018) "Screening of anxiety and depression among the seafarers using machine learning technology." *Informatics in Medicine Unlocked* :100149.
- [8] Saha, B., Nguyen, T., Phung, D., Venkatesh, S. (2016) "A framework for classifying online mental health-related communities with an interest in depression." *IEEE journal of biomedical and health informatics* 20 (4): 1008-1015.
- [9] Reece, A. G., Reagan, A. J., Lix, K. L. M., Dodds, P. S., Danforth, C. M., Langer, E. J. (2016) "Forecasting the Onset and Course of Mental Illness with Twitter Data." *Scientific reports* 7 (1): 13006.
- [10] Braithwaite, S. R., Giraud-Carrier, C., West, J., Barnes, M. D., Hanson, C. L. (2016) "Validating machine learning algorithms for Twitter data against established measures of suicidality." *JMIR mental health* 3 (2): e21.
- [11] Du, J., Zhang, Y., Luo, J., Jia, Y., Wei, Q., Tao, C., Xu, H. (2018) "Extracting psychiatric stressors for suicide from social media using deep learning." *BMC medical informatics and decision making* 18 (2): 43.
- [12] AlHanai, T., Ghassemi, M. M., Glass, J. R. (2018) "Detecting Depression with Audio/Text Sequence Modeling of Interviews." In *Interspeech* : 1716-1720.
- [13] Ramiandrisoa, F., Mothe, J., Benamara, F., Moriceau, V. (2018) "IRIT at e-Risk 2018." *E-Risk workshop*: 367-377.
- [14] Hou, Y., Xu, J., Huang, Y., Ma, X. (2016) "A big data application to predict depression in the university based on the reading habits." *3rd IEEE International Conference on Systems and Informatics (ICSAI)*: 1085-1089.
- [15] Leightley, D., Williamson, V., Darby, J., Fear, N. T. (2019) "Identifying probable post-traumatic stress disorder: applying supervised machine learning to data from a UK military cohort." *Journal of Mental Health*. 28 (1): 34-41.
- [16] Young, C., Harati, S., Ball, T., Williams, L. (2019) "Using Machine Learning to Characterize Circuit-Based Subtypes in Mood and Anxiety Disorders." *Biological Psychiatry* 85 (10): S310.
- [17] Li, L., Zhang, X. (2010) "Study of data mining algorithm based on decision tree." In *2010 International Conference On Computer Design and Applications IEEE* 1: V1-155.



- [18] Paul, A., Mukherjee, D. P., Das, P., Gangopadhyay, A., Chintla, A. R., Kundu, S. (2018) "Improved random forest for classification." *IEEE Transactions on Image Processing* 27 (8): 4012-4024.
- [19] Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J. P. (2012) "An assessment of the effectiveness of a random forest classifier for land-cover classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93-104.
- [20] Liu, X. Y., Wu, J., Zhou, Z. H. (2009) "Exploratory undersampling for class-imbalance learning." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39 (2): 539-550. [21] Dietterich, T. G. (2000) "Ensemble methods in machine learning." In *International workshop on multiple classifier systems* Springer, Berlin, Heidelberg: 1-15.
- [22] Saitta, L., (2000) "Support-Vector Networks." Kluwer Acad. Publ. Bost. : 273-297.
- [23] Hamed, T., Dara, R., Kremer, S. C. (2014) "An accurate, fast embedded feature selection for SVMs." In *2014 13th International Conference on Machine Learning and Applications* IEEE :135-140.
- [24] Martinez-Arroyo, M., Sucar, L. E. (2006) "Learning an optimal naive bayes classifier." In *18th International Conference on Pattern Recognition (ICPR'06)* IEEE 3: 1236-1239.
- [25] Cheng, J., Greiner, R. (1999) "Comparing Bayesian network classifiers." In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence* Morgan Kaufmann Publishers Inc : 101-108. [26] Taneja, S., Gupta, C., Goyal, K., Gureja, D. (2014) "An enhanced k-nearest neighbor algorithm using information gain and clustering." *Fourth International Conference on Advanced Computing & Communication Technologies* IEEE : 325-329.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details