DASS-21 Based Psychometric Prediction Using Advanced Machine Learning Techniques

Jayshree Ghorpade-Aher *, Ahbaz Memon, Snehalraj Chugh, Abhishek Chebolu, Prajakta Chaudhari, and Janhavi Chavan

Dr. Vishwanath Karad MIT World Peace University, Pune, India; Email: ahbazmemon0@gmail.com (A.M.), snehalchugh2016@gmail.com (S.C.), abhishek.chebolu@gmail.com (A.C.), prajakta.p.chaudhari@gmail.com (P.C.), janhavi.a.chavan@gmail.com (J.C.)
*Correspondence: jayshree.aher@mitwpu.edu.in (J.G.-A.)

***Abstract*—The Depression, Anxiety, and Stress Scales (DASS) are 3-factor paradigms; the research is about understanding an individual's instant, psychic response to the epidemic in evaluating their mental health. Anxiety, stress, and depression in adults over the age of eighteen are the primary subjects of this study. In computing behaviourism, it's critical to analyse individuals' emotional responses to strain and forecast how they will respond. 39,775 responses from individuals were obtained for the research data who took part in experiments between 2017 and 2019. When connecting attributes to anxiety levels, nearly seven distinct advanced Machine Learning (ML) based classifying algorithms were applied for experimentation. Further Artificial Intelligence (AI) techniques were used to substantially increase the accuracy of the research work. Ada-boost categorising with Support Vector Machine (SVM) base and SVM classification algorithms had a maximum predictive performance of 96% in Machine Learning models. The model acquired the highest results and the quickest time to prepare the entire collection after using Deep Learning (DL) algorithms.**

***Keywords*—DASS-21, Adaboost, support vector machine classifier, psychology, convolution neural network, naïve bayes**

I. Introduction

The Depression, Anxiety, and Stress Scale (DASS-21) is a method for gathering information that may be leveraged to create a best-fitting method for the detrimental consequences of despair, anxiety, and strain on mental health [1]. Depression is an emotional condition characterised by a general lack of happiness as well as other symptoms such as sadness, an absence of enjoyment or joy, low motivation, a sense of shame or inadequacy, disrupted rest or eating habits, and trouble focusing. Furthermore, depressive episodes often occur under stressful circumstances. These problems may become chronic or recurrent, interfering with people's lives and making it difficult for them to function normally [2].

Bodily manifestations of anxiety, such as negative ailments, acts, and behaviours, are shown as its cause. Anxiety-related symptoms such as heart palpitations, a racing heart, and breathing difficulties may have their origins in subconscious factors [3]. These days, as obstacles are developing throughout every profession, worry and pressure have become more severe concerns. That has a devastating effect on everyone's happiness and health. It has a devastating effect on a person's ability to learn, their efficiency at work, and their interactions with others [4].

One's ability to deal with stressful situations and unpleasant emotions is a hallmark of personal well-being. Emotional and mental health assessment is an important but under-researched area in cognitive Machine Learning, which is the subfield of cognitive neuroscience that uses quantitative research models to examine human thinking. Suicide is a devastating result of mental health problems, including anxiety, nervousness, and pressure [5]. Exploring more complete combinations of sentiments, such as everything that an individual could recall from a former emotional journey, as well as the sequencing of such emotional experiences, could improve scientific comprehension of mental health [6].

Research shows that the DASS exhibits strong intrinsic coherence, internal reliability, and effectively distinguishes between the characteristics of stress as well as despair more effectively than that of other assessments currently in use. Therefore, the DASS is a good tool for evaluating the symptoms of anxiety, distress, and stress in clinical as well as non-patient populations. Utilizing a large database and a learned system, the researchers hope to get a better understanding of which techniques perform best for visualising anxiety, panic, and tension in real time [7]. The following relevant questions were addressed as a result of the research:

- What are the best techniques to use with this dataset?
- What kind of human behaviour caused DAS?
- Identifying those who are depressed, anxious, or stressed
- What measures are used by DAS employees?

Using advanced Machine Learning techniques and various ensemble methods, helped to understand the reliability of predictions, where the ensemble approaches are very effective in decreasing the model's variation [8].

Manuscript received August 10, 2022; revised September 22, 2022; accepted October 14, 2022; published June 25, 2023.

It is feasible to eliminate the discrepancy among results by combining various models to generate a single estimate and then selecting that estimate from each of the other assertions made by blended algorithms [9]. To further enhance precision, artificial neural network techniques were implemented. Specifically, people's attitude and the overall level of observed mental health, mostly during the epidemic [10] are primary goals of this study.

II. Literature Survey

For the acquisition and consequent categorization of information from blogs, several academics are focused mostly on the prediction of stress and anxiety using Machine Learning techniques such as Relational Frame Theory (RFT), Logistic Regression (LogR) [11], Support Vector Machine (SVM), and Convolution Neural Network (CNN). The state-of-art research study of these techniques was used to analyse the data [12].

Information on six hundred and thirty elderly patients, of which five hundred and twenty were under special attention, was individually obtained from the said Medical University of Calcutta, West Bengal. It was found that RFT achieved the highest prediction performance of 91% and 89% for the datasets of one hundred and ten and five hundred and twenty people, respectively, upon implementing distinct algorithms including the Multiple Layer Perceptron, Naive Bayes, RFTs, Bayesian Networks, etc. [13].

Patterns of post-traumatic anxiety and depression were examined. The Hidden Markov Model (HMM) identified the changes in the likelihood of Post-Traumatic Stress Disorder (PTSD) [14]. Clinical depression was found to affect 31% and 25% of the whole sample, respectively. A DT classification method was used to analyse the tweets of one hundred and thirty-five individuals. There was a 92% success percentage for predicting the mortality rate in the population.

It is being observed that the effectiveness of several Machine Learning (ML) algorithms for the identification of diseases varies according to the circumstance; no one method has now been proven to be the best across all situations. In this experiment, these techniques were used to detect indicators such as nervousness, worry, and tension.

Denovan studied how strain is experienced by college learners in the UK. Five hundred and twenty-four participants from UK universities participated in the research, which looked at variables such as component architecture, internal consistency, convergence accuracy, and sexual identity affine [15]. In the process, four alternative component frameworks were evaluated. The reliability and repeatability of the DASS-21 were shown as well. A human's psychotic level, which may be gleaned through viewer material, is particularly useful for identifying unfavourable behavioural characteristics. People's psychotic levels may be determined via learning algorithms.

III. Dataset Description

The DASS-21, devised by P. Lovibond with S. Lovibond, is just a condensed variation of the longer DAS Scale. On one 4-point answer measure, the three-typed report accurately measures the existence and severity of depressive, anxious, and stressed feelings in the candidate. There are seven questions mentioned in Table I, on every subscale, and the overall total spans between 0 and 21. Anxiety, stress, and depression are more likely to be present in those with a higher score and in Table II, the symptoms they assess [16].

There were fewer questions, a clearer structural model, and lower interactor relations in the 21-scaled version of the measurement than in the 42-item form. DASS-21 results have yet to be replicated in a different sample, but still, the 21-item edition is way better than the whole 42-item DASS in certain cases.

During 2017 and 2019, an aggregate of 39,775 surveys were completed in various ways. 42 items in DASS42 proper format make up the dataset. Respondents are rated on a scale of one to four and categorised as follows:

1 = "Did not apply to me at all"

2 = "Applied to me to some degree, or some of the time"

3 = "Applied to me to a considerable degree, or a good part of the time"

4 = "Applied to me very much, or most of the time"

TABLE I. Questions Asked Being Divided in Each Category for DASS 21

	Anxiety	Depression	Stress
1	Dryness of Mouth	Couldn't Experience the positive feeling	Found hard to wind down
2	Difficulty in Breathing	Difficult to work up the initiative to do things	Overreact to situations
3	Experience Trembling	Nothing to look forward	A lot of nervous energy
4	Worried about panic and make a fool of themselves	Felt downhearted and Blue	Getting Agitated
5	Close to Panic	Unable to become enthusiastic	Difficult to Relax
6	Aware of the action of the heart in the absence of physical exertion	Felt wasn't worth much as a person	Intolerant to getting what I was doing
7	Felt scared without any good reason	Felt life was meaningless	Touchy

TABLE II. Subscales of Depression, Anxiety and Stress and Symptoms They Assess

	Subscales		
	Depression	**Anxiety**	**Stress**
Symptoms	Inertia	Excitation of ANS	Difficulty to relax
	Anhedonia	Musculoskeletal effects	Nervous excitation
	Dysphoria	Situational anxiety	Easy agitation
	Lack of interest	Subjective anxiety	Irritability
	Self-depreciation		
	Devaluation of life		
	Discouragement		

The results for DAS were determined by summing the numbers linked with the responses to every question throughout the section and then multiplying by two. The final ratings were categorised into Normal, Moderate, Mild, Severe, and Extremely Severe. It is impossible to define a person's personality without using some kind of proof or testing [17]. One may tell a lot about someone's character by looking at their combination of Big Five Personality Qualities [18]. The questions that were asked are divided into categories (as in Table III).

TABLE III. TIPI DETAILS AND WHAT EACH OF THEM REPRESENTS

TIPI Questions	TIPI Details
TIPI - 1	Extraverted, Enthusiastic
TIPI 2	Critical, Quarrelsome.
TIPI 3	Dependable, Self-disciplined
TIPI 4	Anxious, Easily upset
TIPI 5	Open to new experiences, Complex
TIPI 6	Reserved, Quiet
TIPI 7	Sympathetic, Warm
TIPI 8	Disorganized, Careless
TIPI 9	Calm, emotionally stable
TIPI 10	Conventional, Uncreative

The TIPI items were rated "I see myself as:" _____ such that

1 = Disagree strongly
2 = Disagree moderately
3 = Disagree a little
4 = Neither agree nor disagree
5 = Agree a little
6 = Agree moderately
7 = Agree strongly

In Tables II and IV, the Symptoms and Percentile of DAS are shown respectively.

TABLE IV. REPRESENTATION FOR TIPI DETAILS

DASS Rating	Z-Score	Percentile	Depression	Anxiety	Stress
Normal	< 0.5	0–78	0–9	0–7	0–14
Mild	0.5–1.0	79–87	10–13	8–9	15–18
Moderate	1.0–2.0	88–95	14–20	10–14	19–25
Severe	2.0–3.0	96–98	21–27	15–19	26–33
Extremely Severe	> 3.0	99–100	> 28	> 20	> 37

IV. DATA COLLECTION

The information gathering procedure is detailed in the subsection. Following a week of survey responses, nearly 200 comprehensive responses were collected, and no items remained unanswered. All of the individuals were very pleasant and eager to respond to questions in an open and truthful way. They were all gathered for the purposes of validating and predicting the active data in Fig. 1.

Males outnumbered girls 54.4% to 54.6% of those who took part in the research study. A permission agreement completed by all participants guarantees that their personal information is kept private and secure.

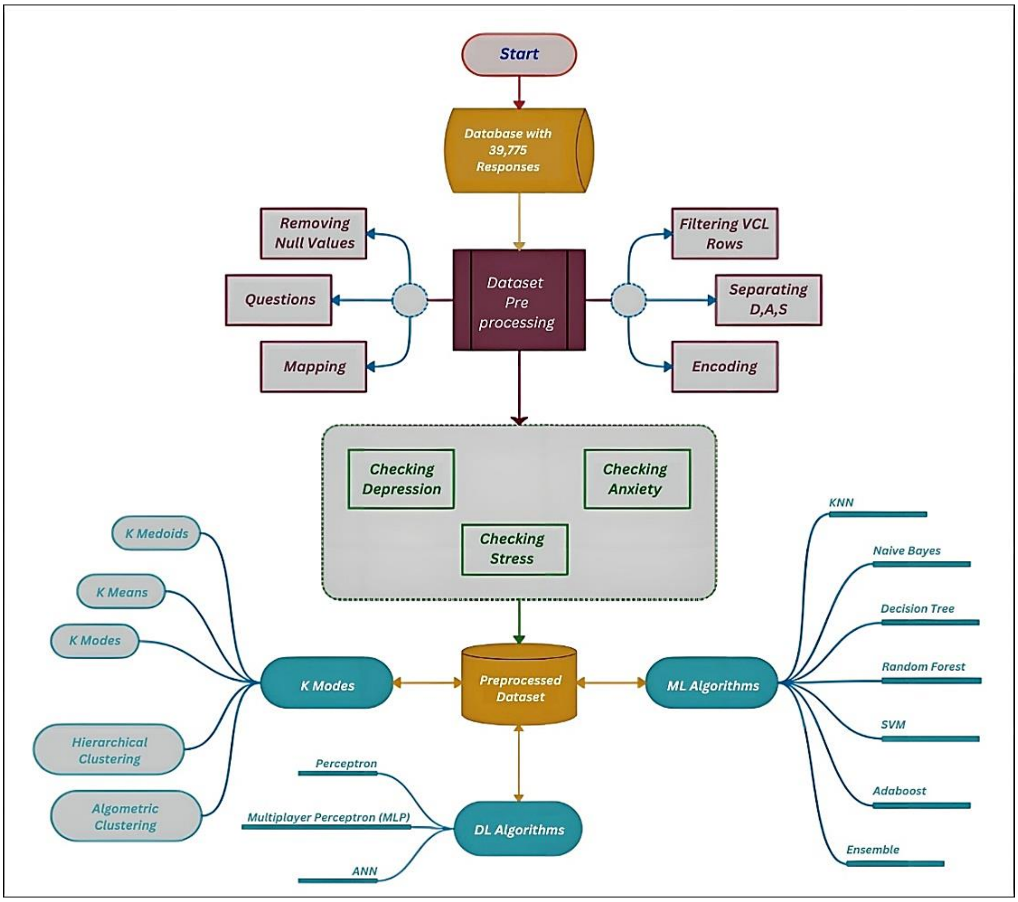

Figure 1. The processes involved in collecting data, pre-processing images, passing them through the CNN model, and resulting in refined images with projections.

V. DATA PREPROCESSING

It is just the characteristic value that is included in the data set. It is necessary to include noisier information in order to improve the dataset's overall quality [19]; nevertheless, such individually adjusted data doesn't provide the whole picture of a character's personality or distress levels.

The dataset was organised in accordance with the experimental questionnaire responses. Items like "faith," "vote," "marriage," "ethnicity," and "orientation" were excluded from the sample, to ensure that the research study wasn't really skewed toward any certain answer. To create several of the most significant actions, the school, urban, sex, and age columns were preserved. Upon that, the age column was encoded too. In the database, the questionnaires and responses were separated into three categories: anxiety, depression, and stress.

VI. TASKS PERFORMED

For Deep Learning as well as Machine Learning techniques, Jupyter Notebooks versions 4.5 and Google Colab have been used. This forecasts the proportion of people who are experiencing feelings of anxiety, nervousness, and despair, depending on the degree of the condition. The DL and ML algorithms have their workings explained in the subsequent sections [20].

VII. ALGORITHM, METHODS, AND TECHNIQUES

The advances in Machine Learning techniques with base estimators such as SVM and an ensemble model [21] are applied as depicted in Fig. 1. Further analysis was done using Deep Learning Models like Perceptrons, Multilayer Perceptrons, and Artificial Neural Networks. Parameters are internal configuration variables of a model. They are essential to the model in order to generate accurate predictions. In this particular instance, a set of input were received from the survey with questions based on DASS-21. There are seven inputs for each depression, stress, and anxiety question, which assess the degree of severity associated with that particular question. The responses were given on an ordinal scale that ranged from 0 to 3, with the higher number referring to a greater degree of severity of the individual's anxiety, stress, or depression. The inputs play a significant part in assisting the Machine Learning model in categorising the levels of the individual's mental health problems as well as the severity of those difficulties [22].

A. Machine Learning Models

k-Nearest Neighbors: The *k*-Nearest Neighbors (*k*-NN) technique is a parameter-independent categorization approach that uses the closest *k*-training dataset as input. In a characteristic set with a classification model, where trained samples have been represented as variables in a multivariate characteristic field, according to the statistics, *k* might range from 0 to 1. Trials with *k* as well as the method's greatest efficiency yielded the best results, so it was set to 5 for all our testing, and it worked well for us.

Naïve Bayes: Naïve Bayes (NB) algorithm is built on a probability technique and has already been utilised in several domains with modifications for a long time. NB relies solely on the Bayes' Theorem for parameterization. This predictor supervises ML processes based on the Bayes rule as well as the assumption that variables are computationally self-reliant. Eq. (1) is the Naïve Bayes formula:

$$p\left(\frac{H}{D}\right) = \frac{p(H)p\left(\frac{D}{H}\right)}{p(D)} \tag{1}$$

Decision Tree: When it comes to predicting issues, the Decision Tree (DT) approach of ML is a good fit since the tree-based patterns are straightforward to understand and reliable. Both classification models and repressors are included in this collection. It is used in the framework of categorization in information retrieval. To verify a learning model, DT has proved useful when dealing with quantitative and qualitative data.

Random Forest: An ensemble classifier called Random Forest (RF) is used for this study. It's natural for RF to function well in multi-class situations, which may also include a combination of quantitative and qualitative characteristics. In order to train RF, it creates a large number of DTs, so each of them has a mean and/or median regression [23].

Support Vector Machine: Many data gathering fields, including medical imaging, bioinformatics, and imaging data management, rely on the Support Vector Machine (SVM) as a classification. Although it has been shown to be more effective in several situations, the learning procedure is more time-consuming in others. A quick approach for learning the SVM has indeed been presented, termed SMO. This classifier's remarkable categorising power and presenting accuracy has recently been used in many situations for its capacity to divide information evenly into two different categories, with such a maximum range among both.

Adaboost Classifier: An AdaBoost classifier is a meta-estimator that starts by matching classification to the initial database, followed by training further replicas of the classification to the very same samples, with the scores of erroneously identified instances changed so that succeeding filters concentrate mostly on challenging situations. Although it's tough to train, the Adaboost plus SVM-based element classification is widely seen as breaking the Enhancing concept due to an unbalanced distribution of variety and efficiency compared to the conventional classification method. When learning with a weighted training set of data, the Adaboost classifiers in the study use an SVM classification algorithm that changes the kernel function variable value, which gradually decreases over time.

Ensemble: An Ensemble is a ML approach in which many ensembles are built using identical methods. Combining and repeating to construct several subsets of the original information, a technique known as bagging, reduces the probability of predicting error by increasing the amount of information available for learning [24]. For the most part, it uses a probabilistic aggregating technique to integrate the results of several base learners, which are trained concurrently and separately from each other.

B. Deep Learning Models

Perceptron: It is a neural net component (an artificial neuron) called a Perceptron, which performs particular calculations in order to identify characteristics or company information inside the inputs. Perceptrons are used as a structural element in artificial neural networks. This approach allows neurons to understand and interpret each piece of information sequentially throughout preparations. To build a perceptron, data input, parameters, and an activation function, along with a weighting factor is needed. When it receives a set of numeric values called connection weights, it processes the data using the weights and biases. The weights of these sources are then multiplied by the weights of such values. The bias is subsequently applied towards the resulting combination.

Multilayer Perceptron: A Multilayer Perceptron has a non-linear input-to-output mapping that identifies it. There are inputs & outputs as well as more than one hidden layer of perceptron layered together in a deep network. In contrast to the perceptron, neurons here may employ whatever activation function they desire, as long as it has a cutoff. It is a feed-forward neural technique since the intakes and starting values are summed and exposed to the function in the same way as the perceptron, which comes under this category. Because of this, every linear combo is transmitted to the following layer. Every layer feeds the succeeding one using its own internal display of information, with the outcome of its calculation. The back propagation training approach is used to learn the MLP. It is possible to use MLPs to estimate any functional form and to resolve issues that are not differentiable. For the most part, MLP is used for classifying and recognising patterns as well as forecasting and approximating outcomes.

Artificial Neural Network: ML, AI, and DL may all benefit from Artificial Neural Networks (ANN), which mimic the natural mind's ability to identify connections and resolve obvious issues. The node level of an ANN consists of a layer of input, more than one hidden layer, as well as a layer of output. A weight as well as a threshold get assigned to every layer, which can be thought of as a perceptron. When a node's output exceeds a certain level, it gets active and begins delivering information to the channel's next tier. Therefore, there is no way for the subsequent component to get any information. ANNs are most typically employed to solve a broad range of issues. In order to increase the depth of the structure, addition of hidden layers is performed.

$$Z = Bias + W_1X_1 + W_2X_2 + \cdots + W_nX_n \quad (2)$$

In Eq. (2),

- Depiction of ANN is denoted by the symbol Z.
- The beta coefficients are often known as the weights (W).
- The inputs or independent variables are denoted by the letters X.
- The bias or intercept is denoted by W_0.

VIII. Clustering Methods

Cluster analysis is useful for making sense of a dataset by revealing its inherent structures. Its objective is simply to segregate the information by splitting it under a collection of meaningful categories. Clustering is quite significant as it provides the inherent classification amongst some of the unstructured variables provided. Even as the name implies, it is a grouping of things according to their similarities and differences. Its performance relies on the methodology and the detection of previously unknown correlations. Datasets are clustered into homogenous groups using clustering. The qualities shared by the members of clusters are very comparable. The cluster centres were regenerated to identify the maximum split by adjusting the number of the K selected cluster centres by an incremental method and recalculating the key points. The same as K-medoids and K-means.

A. K-Means

Unsupervised Machine Learning is the easiest method for solving the cluster issue. The K-means method divides "n" sightings into "K" groups, in which each observable corresponds to a group, with the said cluster's closest average acting as a model. Items will be grouped into k clusters of resemblance by the method.

Elbow Method: The use of k-means clustering as a categorization method is suggested in order to distinguish between distinct intensities of categories. It was established that the best clustering algorithm, the value for k is 3 as indicated in the following Fig. 2.

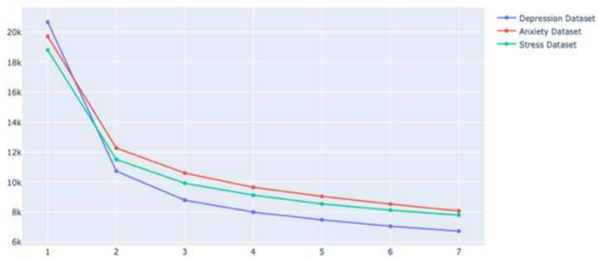

Figure 2. Elbow method for K-means clustering.

B. K-Mediods

Depending on the given dataset, the K-medoids method reduces the K-means computation shortcomings. Using a resemblance measure and attributes specific to every database, K-medoids segmentation divides the information into groups. The values are chosen as the centres using the K-medoids. To enhance clusters, the medoids are chosen at random from N items to create a K grouping, and the remainder of the dataset is distributed to a nearby group based on its proximity to the medoids.

C. Elbow Method

Silhouette Method: Using various k-values, the averaged silhouette of all measurements is calculated using the silhouette technique. Over such a variety of feasible possibilities for *k*, the ideal count of groups *k* would be the number that improves the averaged silhouette. Silhouette measures whether a comparable point is close to the other values within its own group whenever contrasted with the rest of those in the same group. Clusters in Fig. 3 are found to be "3".

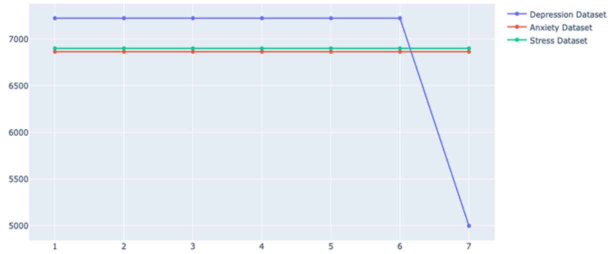

Figure 3. Elbow method for *K*-means clustering.

D. K-Modes

Unsupervised ML methods such as *K*-modes cluster analysis are used to group categorised data. It takes advantage of the differences among the data points. The more different the measured values are, the closer they are to each other.

Elbow Method: Variations of *K* are used to create a scoring formula throughout this technique. The graph was plotted to search for an elbow point where the algorithm provides a number for *K* values as shown in Fig. 4.

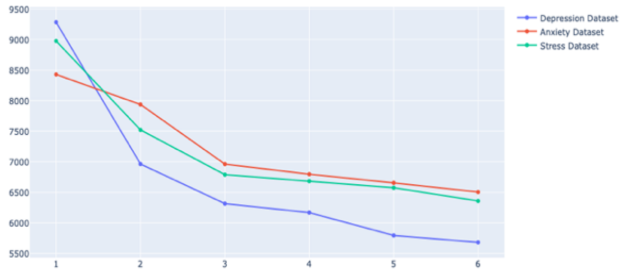

Figure 4. Elbow method for K-modes clustering.

Silhouette Method: This approach calculates the length between locations in one group and those in the remaining groups. Then there are the silhouette plots to choose *K*. *K*-modes have a tendency to become wider as the batch size grows. In our instance, the number of clusters was set at 3.

E. Hierarchical Clustering

The goal of hierarchical clustering analytics is to create a tree-like architecture depending on the hierarchies of groups. A dendrogram's primary purpose is to determine the optimal method for grouping things into groups. The depth of the dendrogram denotes the sequence wherein the groupings are connected in the dendrogram, Figs. 5–7. It is a synopsis of distance measures, and like with other abstracts, some errors occur. As a result of this data loss, the dendrograms just at the bottom are the most reliable, revealing which elements are really related. Even though there's little evidence supporting the result, their analysis generally suggests the proper clustering results.

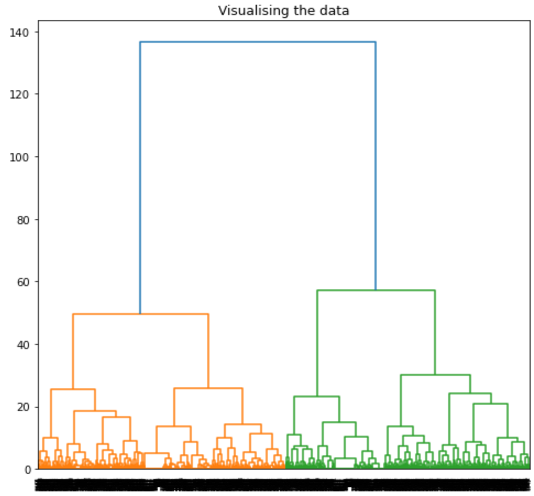

Figure 5. Dendrogram representing depression.

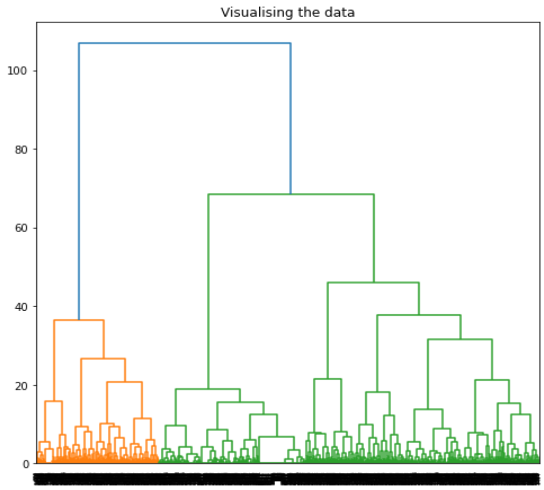

Figure 6. Dendrogram representing anxiety.

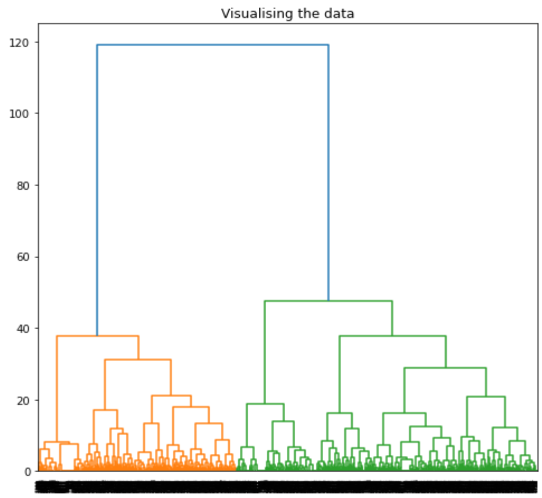

Figure 7. Dendrogram representing stress.

IX. Experimental Results

There are seven alternative categorization methods shown in Fig. 8. All the classifiers have shown a significant accuracy of more than 80% for each set.

However, the research study was able to get a 100% accuracy rate for SVM, shown in Table V, and Adaboost. The classifiers are all based on the same algorithm, although they are tailored to various situations. Hence, the accuracy measures may have produced variable outcomes.

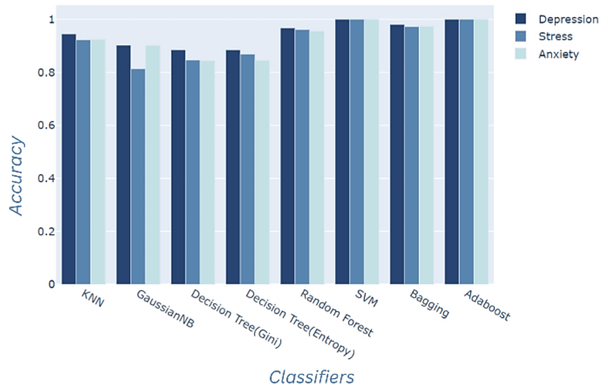

Figure 8. Performance metrics of different ML algorithms.

TABLE V. PERFORMANCE ANALYSIS FOR SVM MODEL BASED ON DAS

Factors Parameters		Performance Metrics			
		Precision	Recall	F1-Score	Support
Depression	0	1.00	1.00	1.00	1911
Depression	1	1.00	1.00	1.00	868
Depression	2	1.00	1.00	1.00	1794
Depression	3	1.00	1.00	1.00	1206
Depression	4	1.00	1.00	1.00	2761
Depression	Accuracy			1.00	8540
Anxiety	0	1.00	1.00	1.00	2084
Anxiety	1	1.00	1.00	1.00	594
Anxiety	2	1.00	1.00	1.00	1652
Anxiety	3	1.00	1.00	1.00	1030
Anxiety	4	1.00	1.00	1.00	3180
Anxiety	Accuracy			1.00	8540
Stress	0	0.95	0.99	0.97	2957
Stress	1	0.83	0.83	0.83	1085
Stress	2	0.87	0.85	0.86	1546
Stress	3	0.93	0.90	0.92	1742
Stress	4	0.98	0.96	0.97	1210
Stress	Anxiety			1.00	8540

The performance criteria utilised for assessment include precision, recall, F1-score, and accuracy [25]. Eqs. (3)–(6) are used to analyse the performance metrics, with True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) observations.

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

$$F1-Score = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{5}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

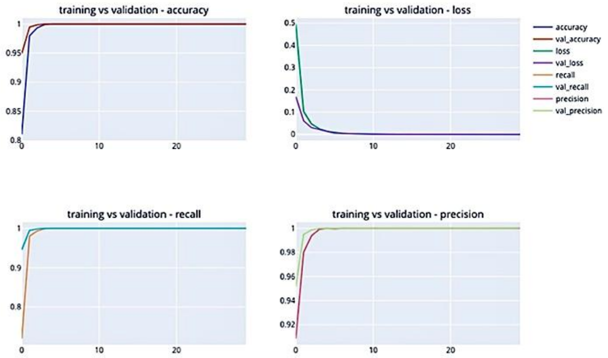

Figure 9. Training vs validation plots for different metrics in deep learning.

Fig. 9 explains the optimization of metrics such as accuracy, recall, precision, and loss over training. It also compares the training *vs* validation metrics over each epoch during training. The figure above explains the optimization of metrics such as accuracy, recall, precision, and loss minimization over training. It also compares the training vs validation metrics over each epoch during training.

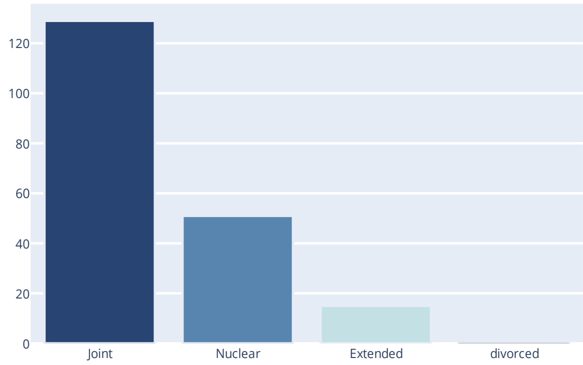

Figure 10. Family type of respondents.

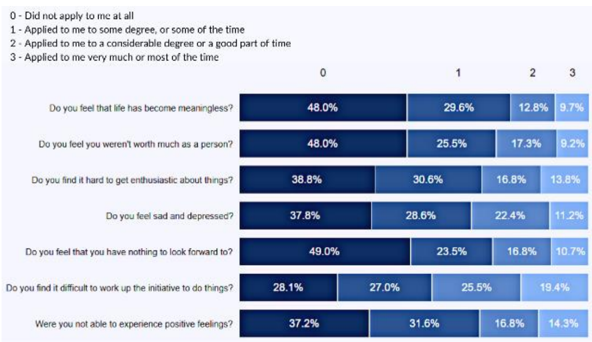

Figure 11. Feelings towards above questions.

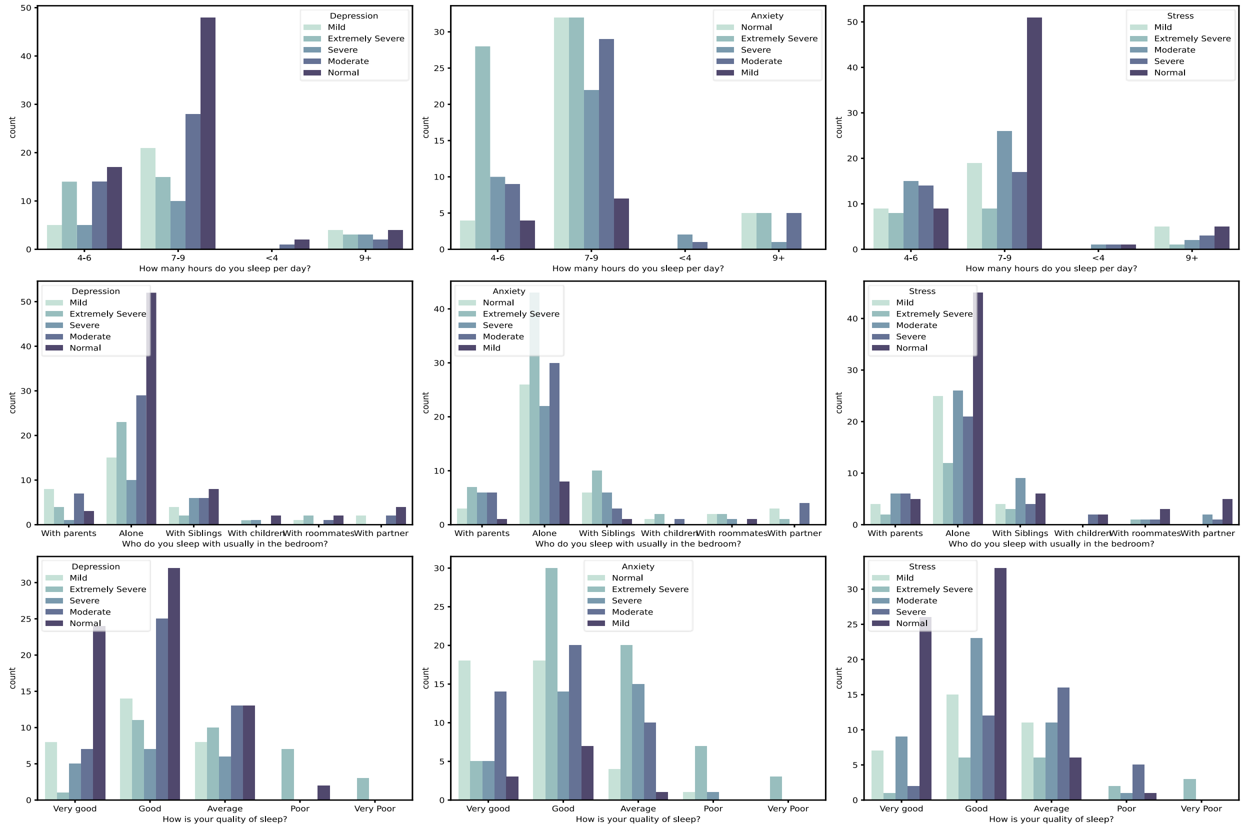

Figure 12. Sleep in comparison to the Mental Health (DAS).

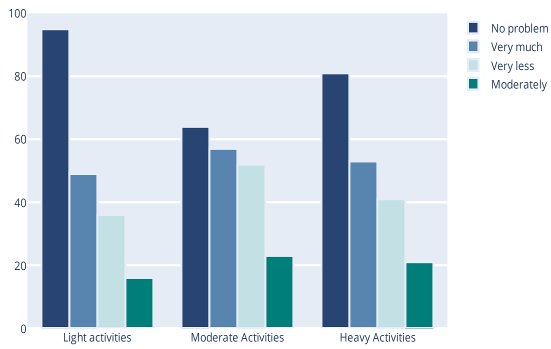

Figure 13. How much of a problem people faced due to their health while performing these different kinds of activities?

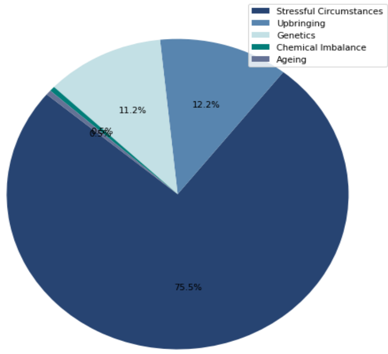

Figure 15. Reasons causing DAS.

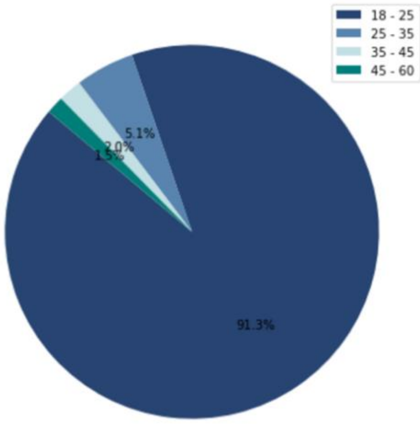

Figure 14. The age group of the respondents of our live dataset.

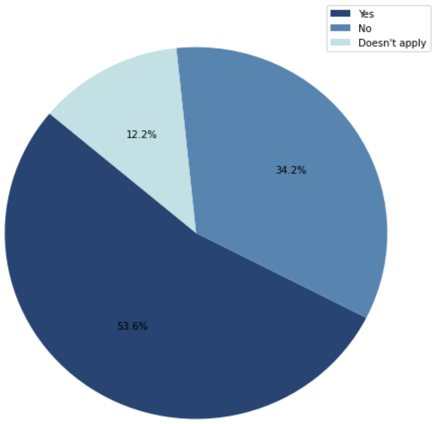

Figure 16. Mentality towards the treatment.

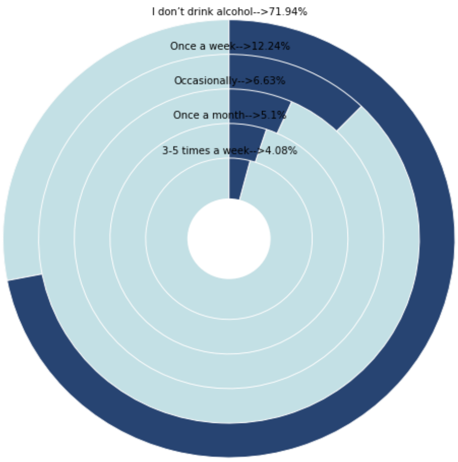

Figure 17. Alcohol consumption.

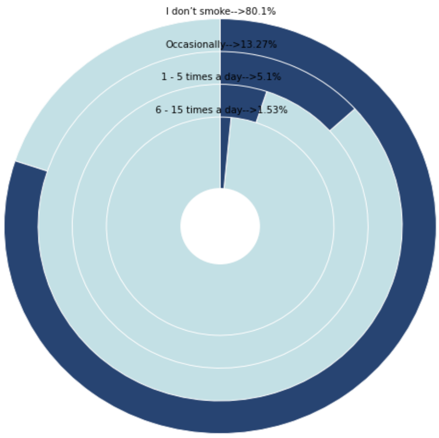

Figure 18. Consumption of cigarettes.

TABLE V. HOW MUCH OF A PROBLEM PEOPLE FACED DUE TO THEIR HEALTH WHILE PERFORMING THESE DIFFERENT KINDS OF ACTIVITIES?

	Activity	No problem	Very much	Very less	Moderately
0	Light activities	95	49	36	16
1	Moderate Activities	64	57	52	23
2	Heavy Activities	81	53	41	21

From the above-described results, as shown in the experimental figures (Fig. 10 to Fig. 18) along with Table VI, it is possible to assert a few outcomes. The experimental results and comparative analysis with various advanced ML and AI techniques demonstrated the following:

- On this dataset, Deep Learning methods function better and more efficiently than Machine Learning techniques.
- The time taken for training was less than ML since the dataset is huge.
- In the comparative study for the best activation function to be used for the output layer, it was observed that the combination of *relu* in SoftMax and the Hidden layer gave the most optimum result since it is a multiclass dataset.
- It provided us with an accuracy, precision and recall of 1.0 for each target class.
- More the depression, stress and anxiety if the sleep range lies between 4–7 hours of sleep.
- DAS strikes more towards a person sleeping alone and least with their children.
- People who are more outgoing are more likely to feel a moderate amount of stress.
- People who score lower on extraversion have less anxiety.
- High levels of stress are more readily perceived by those who score higher on the openness-to-experience scale.
- People mostly think of DAS being caused least through chemical imbalances and ageing and more than 30% of them do not prefer to go to a psychiatrist to help them out.

As stated, through the above study, these were some important contributions to the field of research by using Machine Learning and deep learning to the concept of DASS-21.

X. CONCLUSION

Seven Machine Learning approaches and distinct Deep Learning models were used to predict the three different degrees of anxiety, depression, and stress. The application base for the DASS-21 is to utilise this psychometric test to measure DAS with reliability and validity for a variety of settings and groups of individuals in order to get an understanding of their mental states. The implementation of Machine Learning techniques and artificial intelligence could be used to understand and receive responses from individuals on a trained dataset by utilising a chatbot on a website, which would then provide a live analysis of the individuals' mental health in relation to the DAS. A standardised survey was used to gather information on the most prevalent clinical symptoms of tension, sadness, and anxiety. The DASS-21's depressed and anxious categories were reasonably highly linked with the self-depressive grading system and also the competitive state anxiety questionnaire. The DASS-21 is indeed a useful and trustworthy scientific instrument that may be employed to swiftly screen individuals for depression, anxiety, and stress.

Information from DASS-21, acquired between 2017 and 2019, was used for all of the approaches. SVM with Adaboost Classifying using the base estimator as SVM outperformed than the others, with 96% accuracy across all three parameters, following the deployment of all approaches. People's behavioural qualities and perceived stress levels may be linked using a ML technique. A comparison analysis of ML and DL models was conducted after practising with both. Most of the classes were correctly predicted by the neural network after trial with the ANN model. The ML algorithms performed worse than the ANN network since they were unable to find the optimal solution.

The techniques with improved performance were applied to the actual database, which included information from people of many races, religions, social classes, and ages, among others. The variables provided by the users were used to figure out their DAS level. Tests like these could one day help us discover a person's psychological attributes [21, 22] at the beginning of each school year or whenever they choose to complete them. They'd be able to stay on top of things this way. Afterwards, rather than utilising Google forms to collect data, a webpage or chatbot may be used to get information from individuals across the globe with only a single click. Higher efficiency may also be achieved by using the concept of linking behavioural tendencies and stress levels to global contextual performance.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Abhishek, Janhavi conducted the research; Snehalraj, Prajakta worked on gathering and collection of data; Ahbaz, Jayshree provided insights on the methodology; Abhishek, Ahbaz worked on reviewing the experimental results; Snehalraj, Janhavi wrote the original draft; finalizing, reviewing and editing was handled by Jayshree, Prajakta. All authors had approved the final version.

References

[1] A. Akin and B. Çetin, "The Depression Anxiety and Stress Scale (DASS): The study of validity and reliability," *Educational Sciences in Theory and Practice*, vol. 7, no. 1, p. 260, 2007.

[2] C. Zanon, R. E. Brenner, M. N. Baptista, *et al.*, "Examining the dimensionality, reliability, and invariance of the Depression, Anxiety, and Stress Scale-21 (DASS-21) across eight countries," *Assessment*, vol. 28, no. 6, pp. 1531–1544, 2021.

[3] B. Singh, K. P. Prabhuappa, S. Eqbal, *et al.*, "Depression, anxiety and stress scale: Reliability and validity of Hindi adaptation," *Int. J. Educ. Manage. Stud.*, vol. 3, pp. 446–449, 2013.

[4] B. Chhetri, L. M. Goyal, M., Mittal, *et al.*, "Estimating the prevalence of stress among Indian students during the COVID-19 pandemic: A cross-sectional study from India," *Journal of Taibah University Medical Sciences*, vol. 16, no. 2, pp. 260–267, 2021.

[5] G. Giorgi, L. I. Lecca, F. Alessio, *et al.*, "COVID-19-related mental health effects in the workplace: A narrative review," *International Journal of Environmental Research and Public Health*, vol. 17, no. 21, 7857, 2020.

[6] L. Li, Q. Zhang, X. Wang, *et al.*, "Characterizing the propagation of situational information in social media during covid-19 epidemic: A case study on Weibo," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 556–562, 2020.

[7] V. Chamola, V. Hassija, V. Gupta, *et al.*, "A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact," *IEEE Access*, vol. 8, pp. 90225–90265, 2020.

[8] S. W. Long, R. J. Olsen, P. A. Christensen, *et al.*, "Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area," *MBio*, vol. 11, no. 6, pp. e02707–e02720, 2020.

[9] B. Salzberger, T. Glück, and B. Ehrenstein, "Successful containment of COVID-19: The WHO-report on the COVID-19 outbreak in China," *Infection*, vol. 48, no. 2, pp. 151–153, 2020.

[10] F. S. Vahidy, A. L. Drews, F. N. Masud, *et al.*, "Characteristics and outcomes of COVID-19 patients during initial peak and resurgence in the Houston metropolitan area," *JAMA*, vol. 324, no. 10, pp. 998–1000, 2020.

[11] N. Zhu, D. Zhang, W. Wang, *et al.*, "A novel coronavirus from patients with pneumonia in China, 2019," *New England Journal of Medicine*, vol. 382, no. 8, pp. 727–733, 2020, doi: 10.1056/NEJMoa2001017

[12] C. Li, Y. Chen, and Y. Shang, "A review of industrial big data for decision making in intelligent manufacturing," *Engineering Science and Technology, an International Journal*, vol. 29, 101021, 2021.

[13] A. Priya, S. Garg, and N. P. Tigga, "Predicting anxiety, depression and stress in modern life using machine learning algorithms," *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020.

[14] S. Yang, W. Ni, X. Dong, *et al.*, "Intention mining in medical process: A case study in trauma resuscitation," in *Proc. IEEE International Conference on Healthcare Informatics (ICHI)*, 2018, pp. 36–43.

[15] M. Shevlin, O. McBride, J. Murphy, *et al.*, "Anxiety, depression, traumatic stress and COVID-19-related anxiety in the UK general population during the COVID-19 pandemic," *BJPsych Open*, vol. 6, no. 6, p. e125, 2020, doi: 10.1192/bjo.2020.109

[16] A. O. Coker, O. O. Coker, and D. Sanni, "Psychometric properties of the 21-item Depression Anxiety Stress Scale-21 (DASS-21)," *African Research Review*, vol. 12, no. 2, pp. 135–142, 2018.

[17] Y. Shi, L. Zhu, W. Li, *et al.*, "Survey on classic and latest textual sentiment analysis articles and techniques," *International Journal of Information Technology and Decision Making*, vol. 18, no. 4, pp. 1243–1287, 2019.

[18] T. Zhang, B. Xu, F. Thung, *et al.*, "Sentiment analysis for software engineering: How far can pre-trained transformer models go," in *Proc. IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2020, pp. 70–80.

[19] C. Yuan and H. Yang, "Research on K-value selection method of K-means clustering algorithm," *Multidisciplinary Scientific Journal*, vol. 2, no. 2, pp. 226–235, 2019, https://doi.org/10.3390/j2020016

[20] R. Jacobucci and K. J. Grimm, "Machine learning and psychological research: The unexplored effect of measurement," *Perspectives on Psychological Science*, vol. 15, no. 3, pp. 809–816, 2020, https://doi.org/10.1177/1745691620902467

[21] K. F. Yuen, Y. Cao, X. Bai, *et al.*, "The psychology of cruise service usage post COVID-19: Health management and policy implications," *Marine Policy*, vol. 130, 104586, 2021, https://doi.org/10.1016/j.marpol.2021.104586

[22] R. Tandon, "Covid-19 and mental health: preserving humanity, maintaining sanity, and promoting health," *Asian J. Psychiatr.*, vol. 50, 2021.

[23] J. Ghorpade-Aher and B. Sonkamble, "Effective feature selection using ensemble techniques and genetic algorithm," in *Proc. Sixth International Congress on Information and Communication Technology*, 2022, https://doi.org/10.1007/978-981-16-2380-6_32

[24] World Health Organization. (2022). Covid-who-1460284. [Online]. Available: https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/covidwho-1460284?lang=en

[25] F. N. Al-Wesabi, H. Alsolai, A. M. Hilal, *et al.*, "Machine learning based depression, anxiety, and stress predictive model during covid-19 crisis," *CMC-Comput. Mat. Contin.*, pp. 5803–5820, 2022.