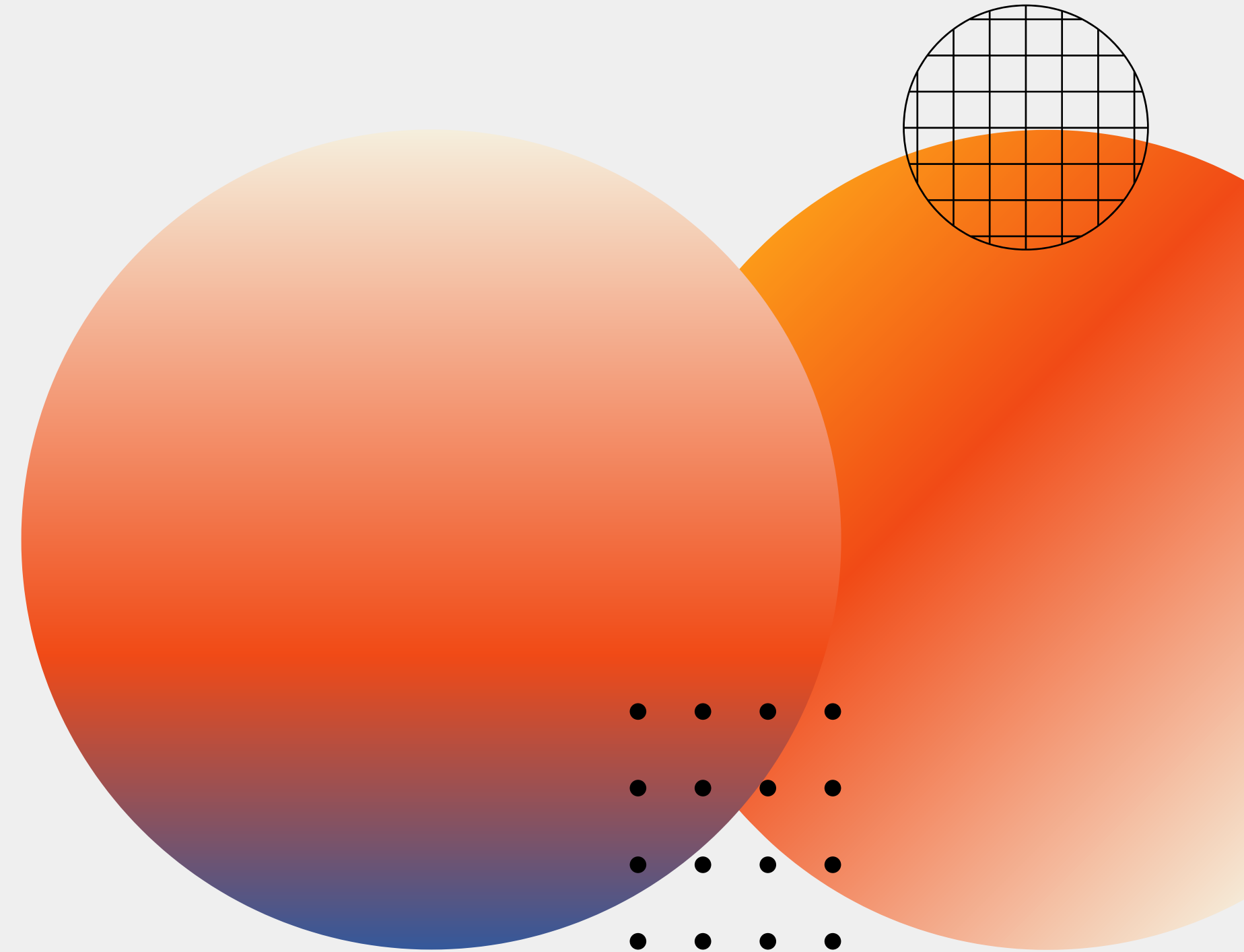


Let's Start

DASS-21 based Psychometric prediction using Deep Learning Models

Name: Snehalraj Chugh – 1032181182
Ahbaz Memon – 1032180046
Abhishek Chebolu – 1032180316
Prajakta Chaudhari – 1032180317

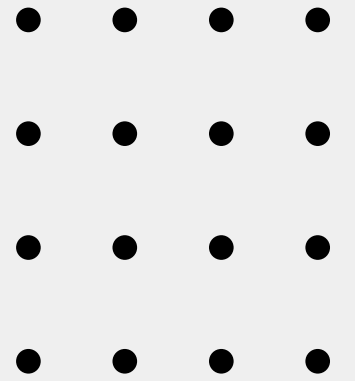
Branch: Final-Year | B.Tech
Year: Computer science Engineering



PROF. JAYSHREE AHER

Contents

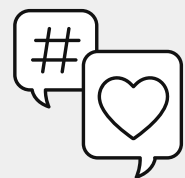
- Problem Statement
 - Objectives
 - Tools to be used
 - Dataset Description
 - Data Pre-processing
 - System architecture
 - Algorithms
 - Methodology
 - Output
 - References
-



Problem Statement



The Depression Anxiety and Stress Scales-21 (DASS-21) involves a simple structure first-order three-factor oblique model, with factors for depression, anxiety, and stress. The purpose of our study is to take use of the DASS-21 scale for use in the mental health screening in people around the globe when they suffer an immediate psychological reaction in the pandemic environment.

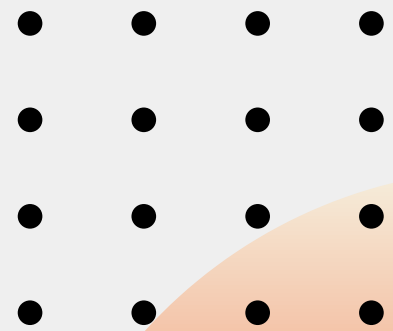


In the area of computational psychology, it is important to understand participants' psychological behaviour using personality traits. This research focuses on the personality traits of people of age group from 18 and above and stress scale they had to face in a pandemic-like situation.



The reliability and validity of DASS and its short-form—the DASS-21 have been widely recognized to assess depression, anxiety, and stress among adults. Because DASS and DASS-21 are developed in Australia and applied to different socio-cultural contexts and types of populations, they have been carefully considered. For the experiment, we have collected data of 200 participants which becomes our live dataset while have trained and tested on a dataset of 39,775 responses collected from 2017 – 2019. We will be applying different deep learning models built for mapping the traits with stress scales, which we will be talking ahead.

Objectives



- 1** Which are the best Algorithms for the dataset
- 2** What behaviour caused the DAS in the people
- 3** Identification of the people having depression, anxiety or stress
- 4** Comparison of Machine Learning and Deep Learning Methods and finding the optimum for the task.

Tools Used

Numpy

Pandas

Keras

Scikit Learn

Plotly

Tensorflow

Dataset Description

DASS-21

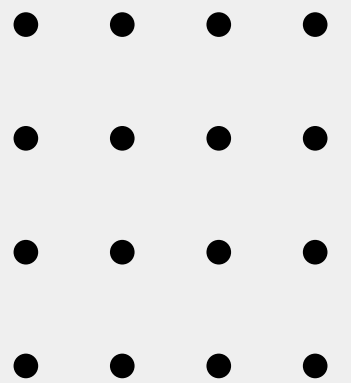
The DASS-21 is the shortened version of the DASS developed by P. Lovibond and S. Lovibond (1995) which was used. Each subscale has seven items and its total score ranges from 0 to 21 points. A higher score indicates higher symptomatology of depression, anxiety, and stress. Three subscales of DASS-21 are named DASS-21-Depression (DASS-21-D), DASS-21-Anxiety (DASS-21-A), and Stress (DASS-21-S).

TOTAL VALUES

In total, total 39,775 instances were collected through online questionnaires between 2017 and 2019 by different methods. The dataset consists of 42 questions taken from the standard form of DASS42. The responses are scaled between 1 and 4. The scores for anxiety, depression and stress were calculated by adding the values associated with the answers to each question of the particular class and then multiplying by 2.

PERSONALITY BASIS

The personality of a human could not be declared without any proper justification or experimentation. According to the Big Five Personality Traits, there have been five traits to identify the overall personality of any individuals named Openness-to-experience, Extraversion, Agreeableness, Neuroticism and Conscientiousness.



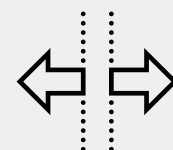
Data Pre processing



The dataset consists of the personality traits score. Openness-to-experience, Extraversion, Agreeableness, Neuroticism and Conscientiousness scores are determined using pre-set formulas and the no hard rule-based data cleaning or pre-processing is performed over the data. Therefore, the noisy data is incorporated within the dataset, which may affect the performance, but the manually altered data will not give the real insight of the personality and stress measurements.



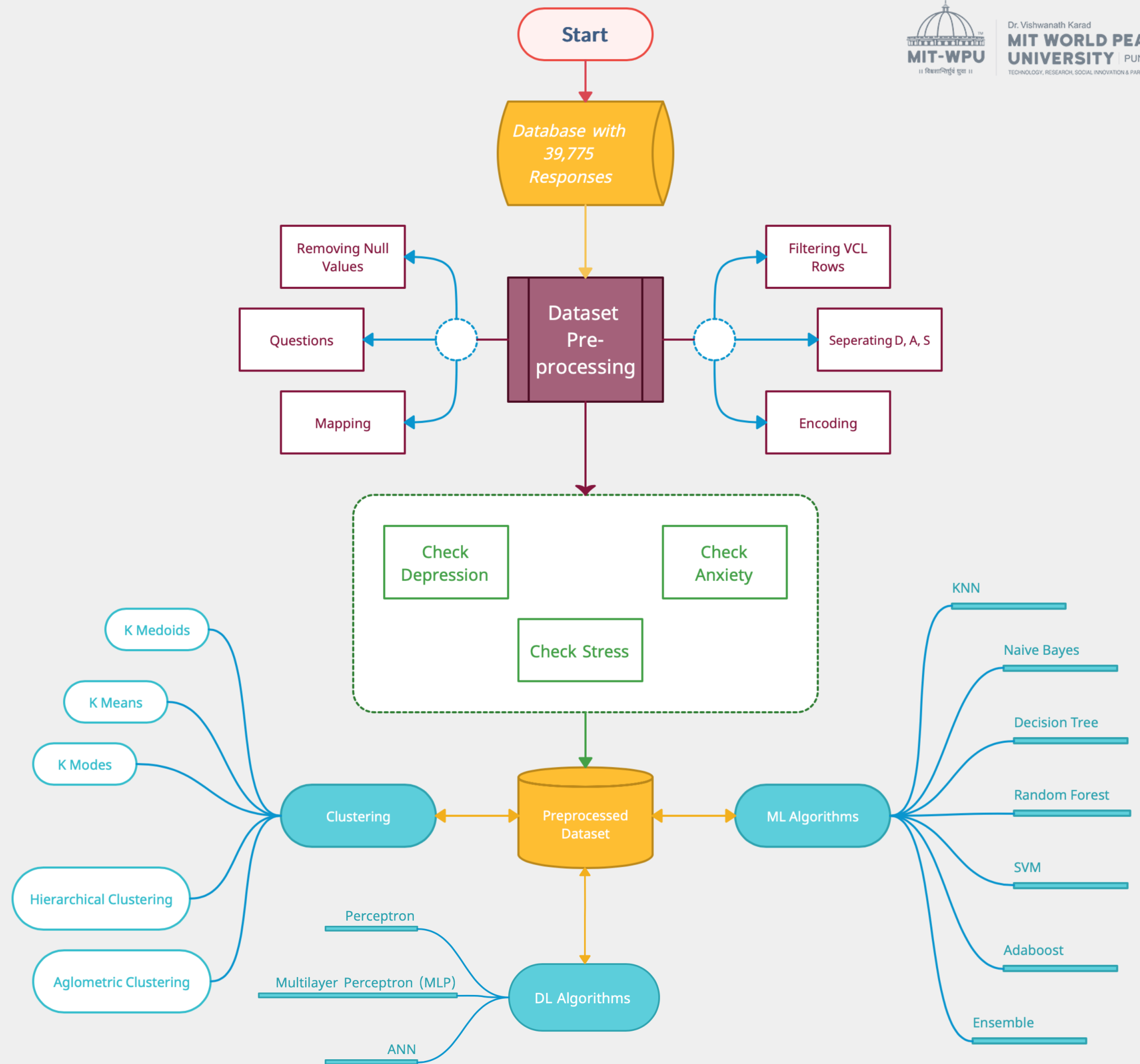
The dataset present with us had various excess answers, which were timed and also had the rank. Thus, we removed them and arranged the values according to the answers we made with the form. On the dataset, we performed cleaning as well as dropped the values which contained excess information about the user such as religion, voted, married, race, orientation, etc, just to make sure our research was not bias over any response. We kept the education, urban, Gender, and age of the person filling the form, to make some of the important activities, over which also applied encoding to the age.



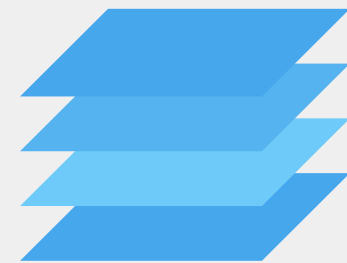
We kept the questions as well as answers for each set separate in the dataset which were Stress, Depression and Anxiety.



System Architecture



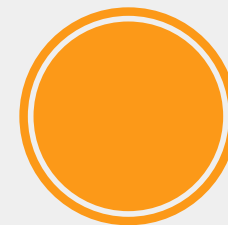
Deep Learning Algorithms



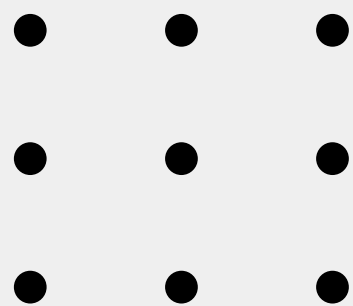
Multilayer Perceptron
(MLP)



ANN



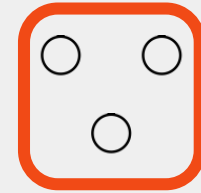
Single layer Perceptron



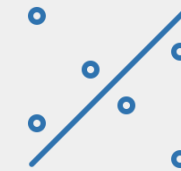
Machine Learning Algorithms



Naïve Bayes



AdaBoost



SVM



Decision Tree



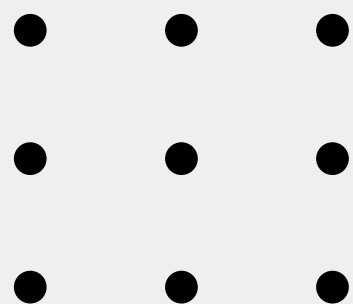
Random Forest



K Nearest Neighbours



Ensemble



Clustering Methods



K-MEANS



K-MEDOIDS



K-MODES



AGLOMETRIC



HIERARCHICAL
CLUSTERING

Methodology

- The first step would be to preprocess the kaggle dataset which has responds of people all over the world. The responses are based on the DASS 42 scale.
- Preprocessing will involve many steps such as dropping unnecessary columns, encoding the dataset, And reducing the set of questions to DASS42.
- We used it as a backend to train and test various deep learning models on this dataset ,with more than 40000 rows.
- The training will be done to compare deep learning models with existing machine learning models such as KNN, SVM,naive Bayes, etc.
- The deep learning models employed will involve basic neural network such as Multilayer Perceptron, and ANN, RNN with the experimentation of their hyper-parameters
- The candidate's levels of depression, anxiety, and stress can then be assessed after classification.
- With the aid of visualisation we can also analyse the factors affecting the candidate who has filled the form according to the DASS21 scale.

Output

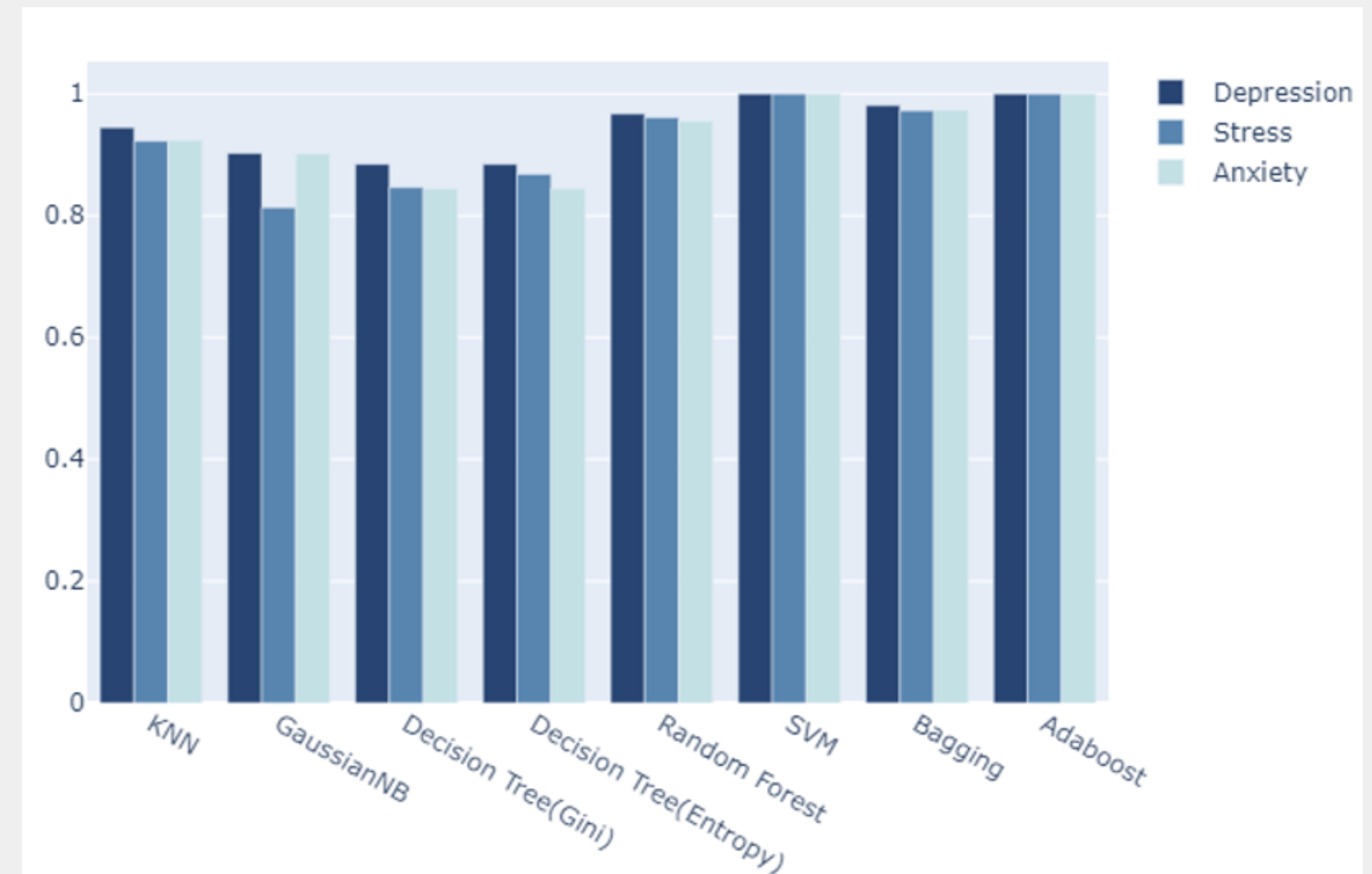
The performance metrics used for evaluation are precision, recall, f1-score and accuracy. The metrics could be calculated using the following equations (1), (2), (3) and (4), where TP is True Positive, FP is False Positive, TN is True Negative and FN is False Negative.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1 - score} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

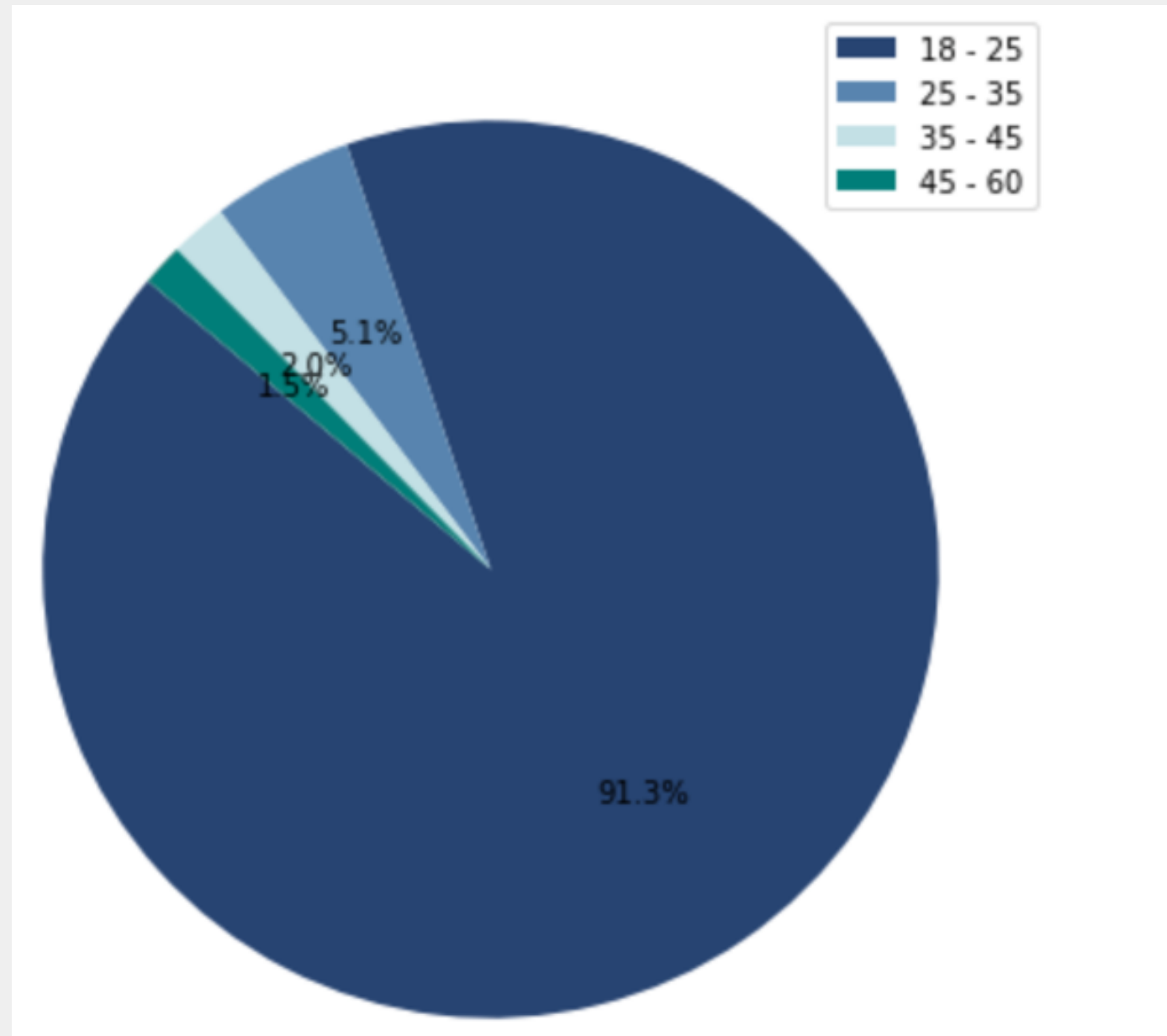


Output

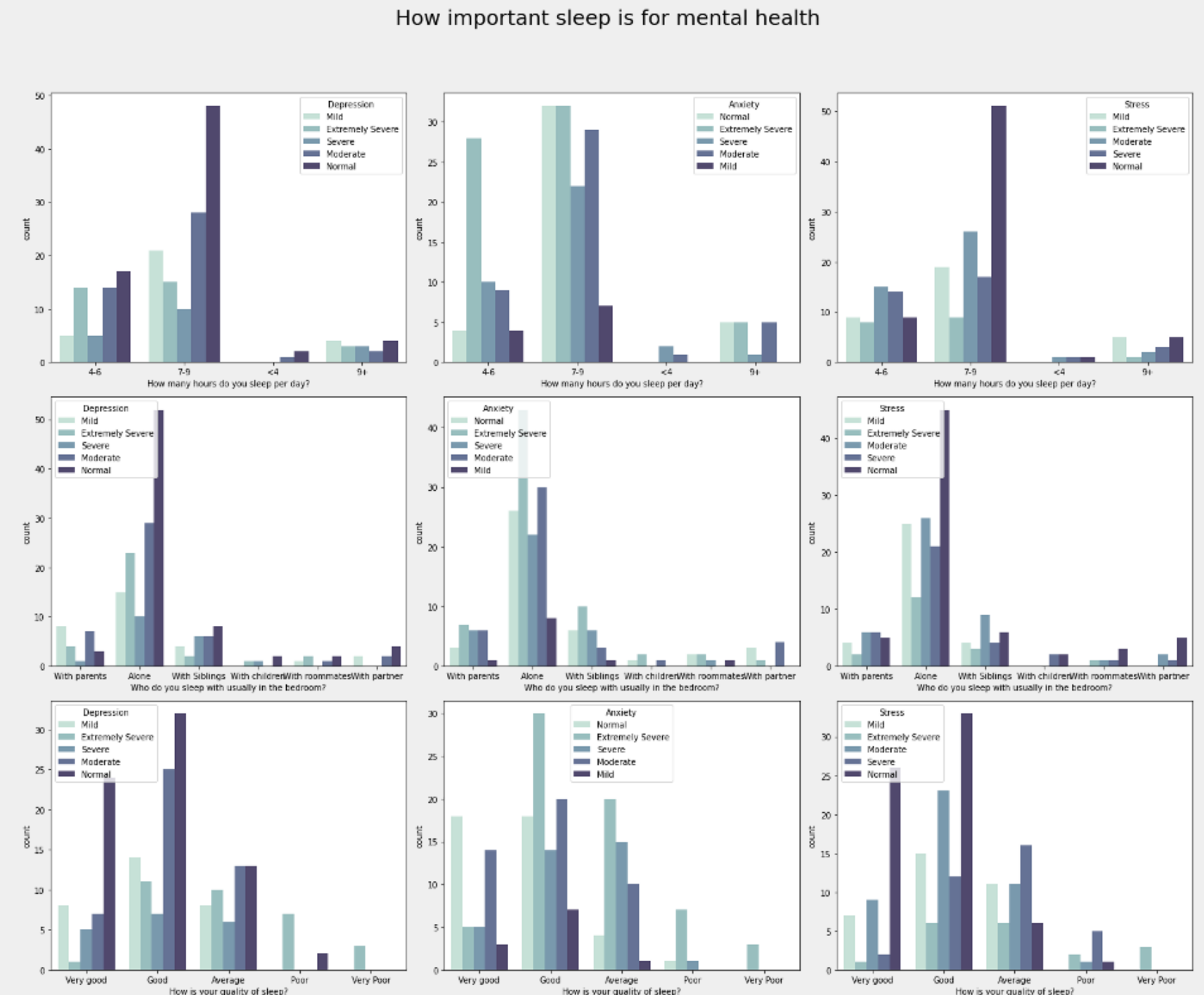
The figure above explains the optimization of metrics such as accuracy, recall, precision, and minimisation of loss over training. It also compares the training vs validation metrics over each epoch during training. The figure above explains the optimization of metrics such as accuracy, recall, precision,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



Visualizations

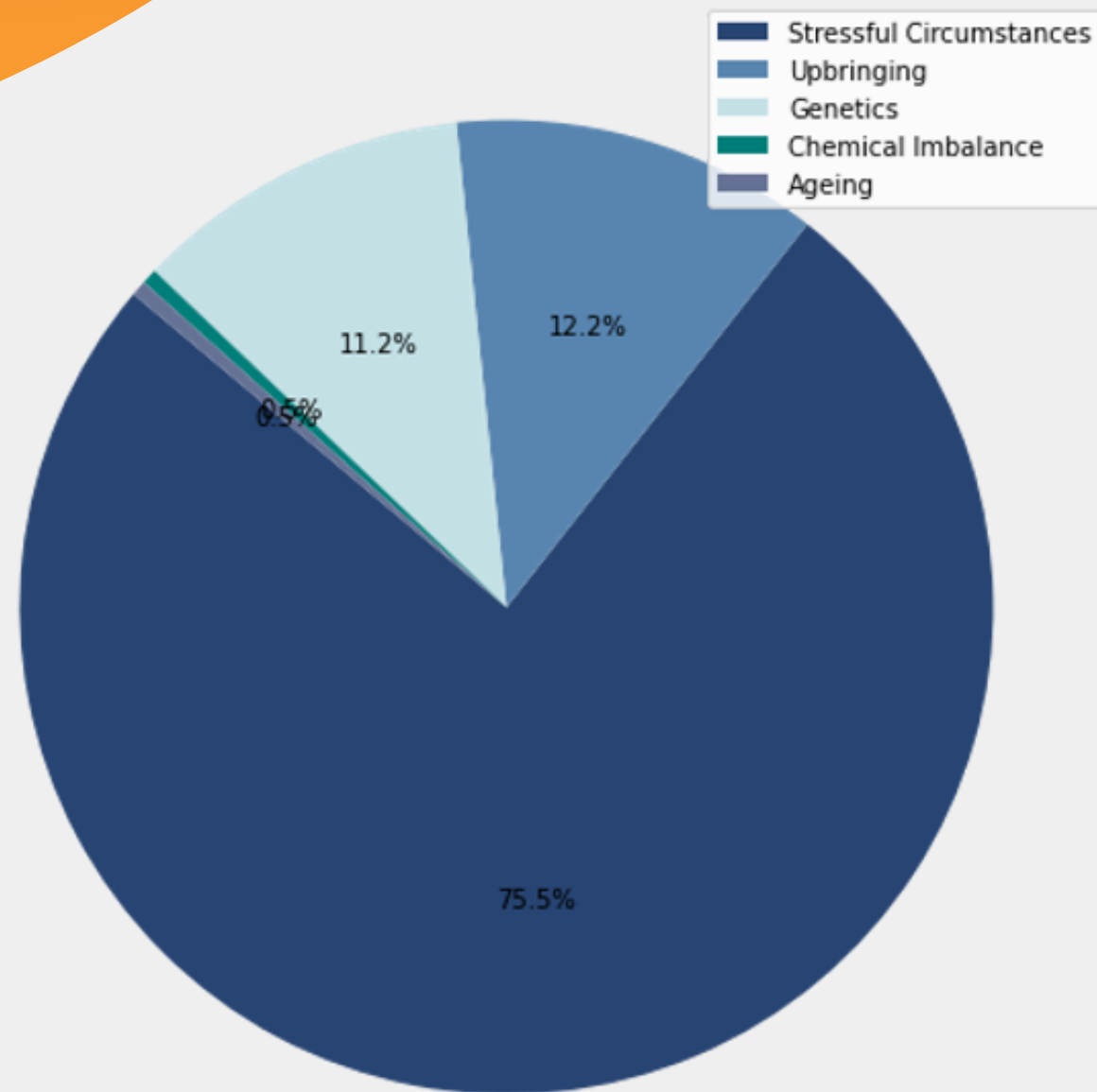


The age Group of the respondents of our live dataset

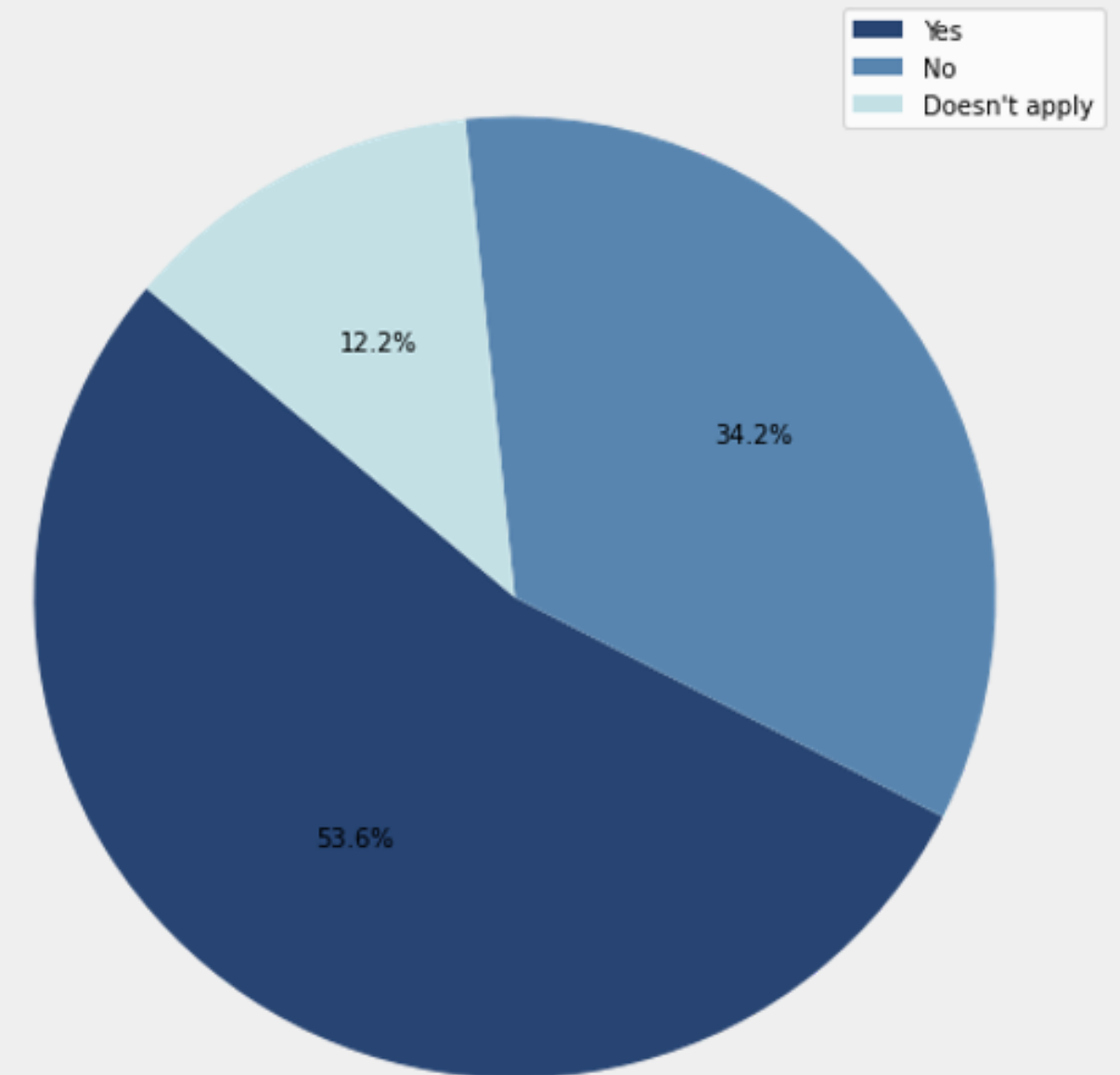


Sleep in comparison to the Mental Health (DAS)

Visualizations

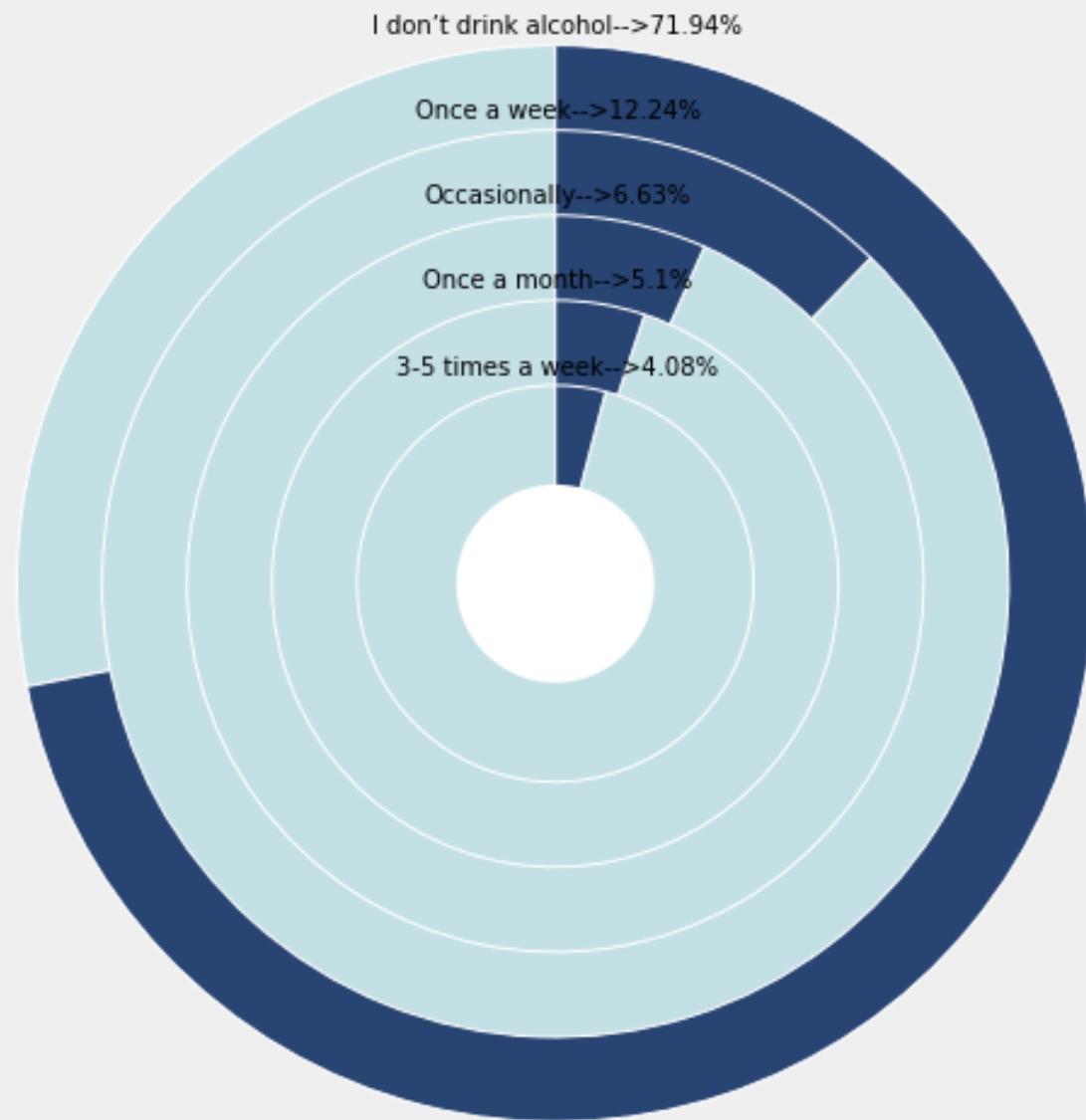


Reasons causing DAS

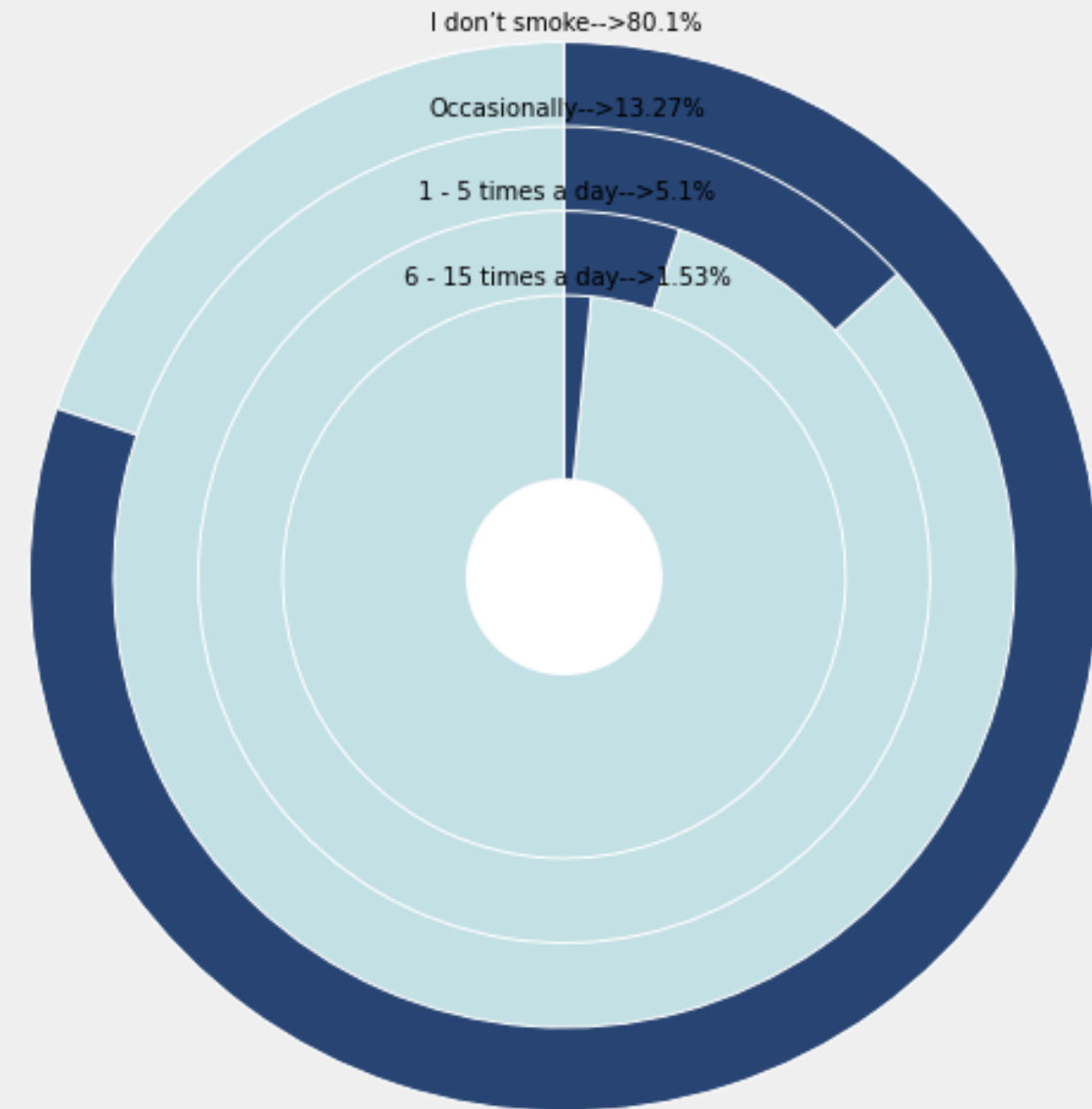


Mentality towards the treatment

Visualizations



Alcohol consumption

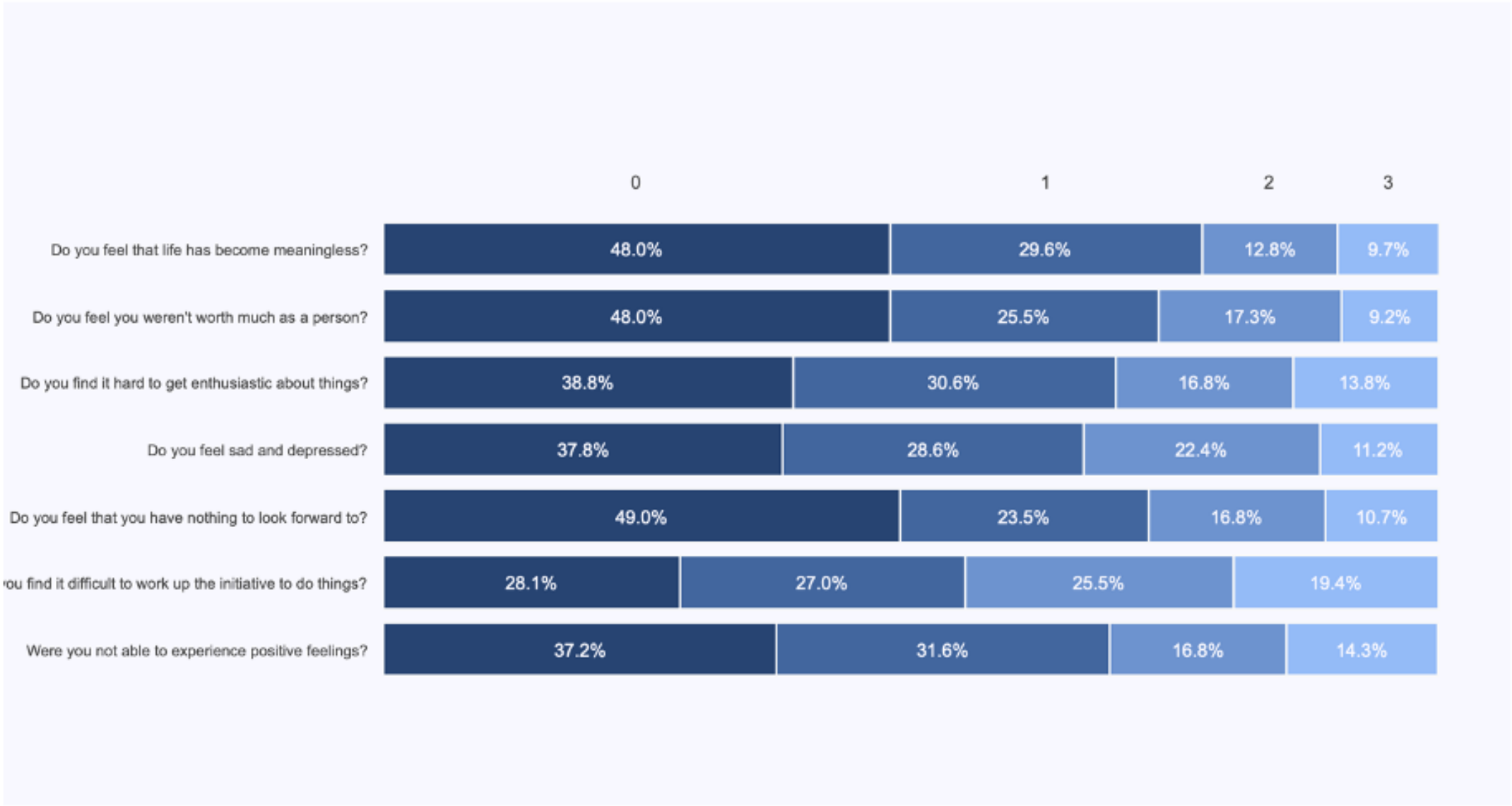


Consumption of cigarettes

Visualizations



0 = Did not apply to me at all
1 = Applied to me to some degree, or some of the time
2 = Applied to me to a considerable degree or a good part of time
3 = Applied to me very much or most of the time



Feelings towards above questions

Results

- 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
- The more extrovert persons are tending to get moderately stressed.
- The persons having fewer score in extraversion are less stressed.
- The persons having higher score in openness-to- experience are tending to perceive high stress.
- 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.

Conclusion



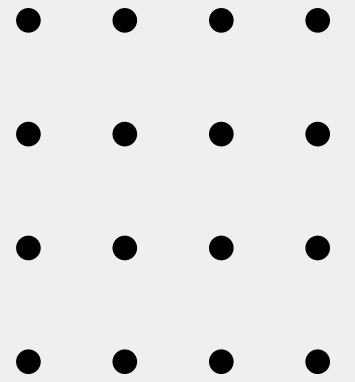
The depressive and anxiety domains of the DASS-21 correlated moderately strong with the self- depression rating scale and the state trait anxiety inventory. Our findings will support other published evidence that DASS-21 is a reliable and suitable research tool useful for quick screening of depression, anxiety, and stress among people.



All the methods were applied to the different databases of DASS21, collected from different years 2017-2019. After application of all the techniques, the results showed that SVM as well as Adaboost Classifying with base estimator as SVM performed better than all the others giving the 96% accuracy for all the three parameters.



Compared to machine learning algorithms deep learning algorithms perform better and efficiently on this dataset. The time taken for training was lesser than ML since the dataset is huge. In the comparative study for the best activation function to be used, we found out that the combination of relu in Hidden layer and SoftMax for output layer, gave the most optimum result since it is a multiclass dataset. We have presented a deep learning based approach to map between the personality traits and perceived stress scale of individuals.



References

- Dataset : <https://www.kaggle.com/lucasgreenwell/depression-anxiety-stress-scales-responses>
 - <https://doi.org/10.1016/j.procs.2020.03.442>
 - <https://doi.org/10.1016/j.procs.2020.04.213>
-

End

Thank you

We would love to answer questions.

