MIT-WPU
॥ विद्यानिर्मितं सुखम् ॥

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

AML - PRESENTATION

DASS-21 based Psychometric prediction using ML Models

Name: Snehalraj Chugh - 1032181182
Ahbaz Memon - 1032180046
Abhishek Chebolu - 1032180316

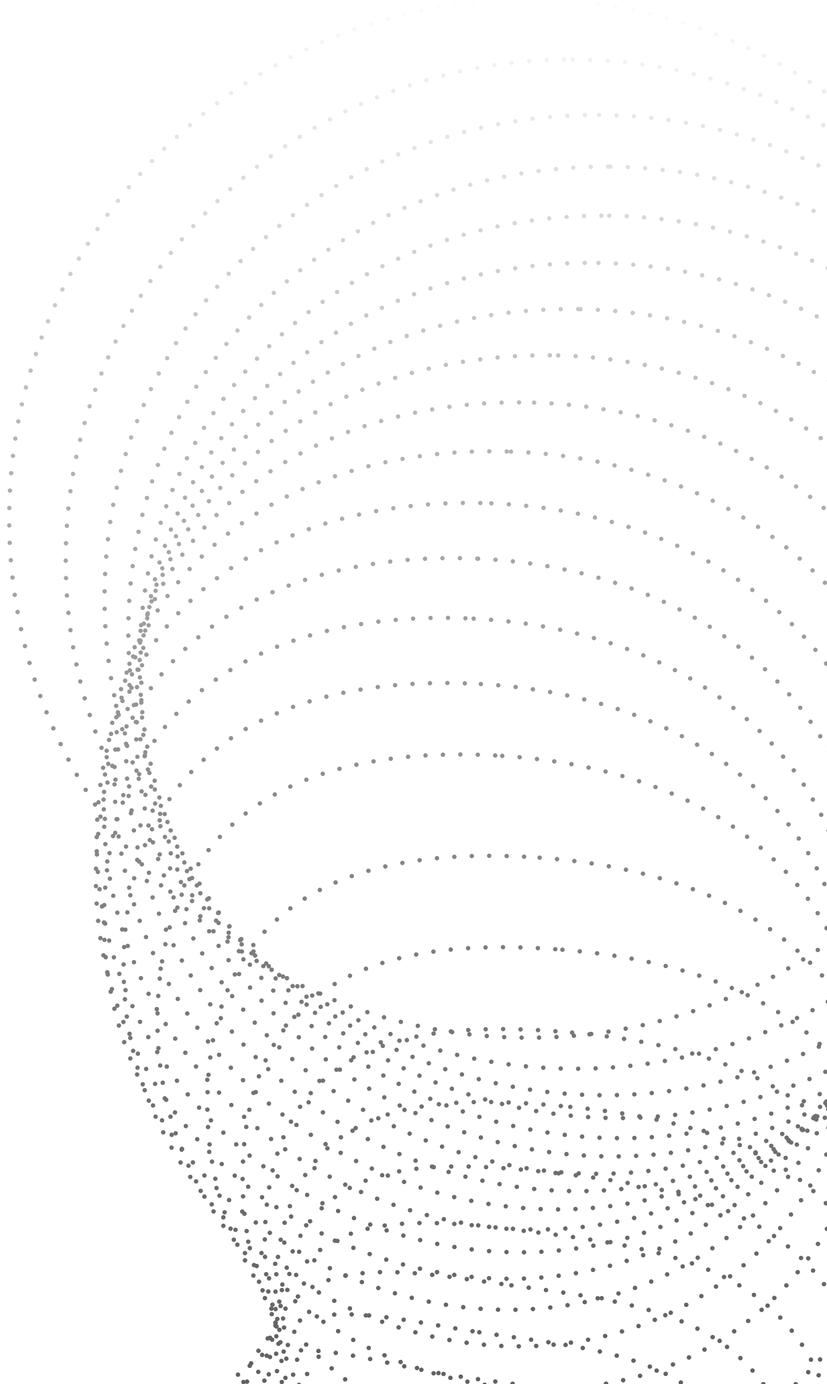
Branch: Final-Year | B.Tech
Year: Computer science Engineering

PROF. JAYSHREE AHER

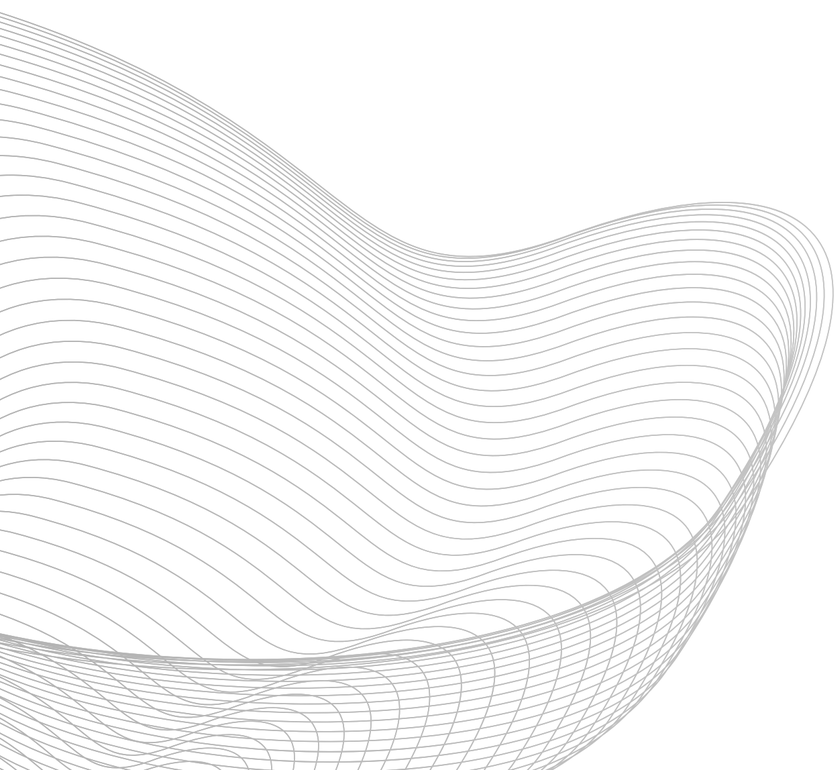
Discussion Points

AN OVERVIEW OF OUR RESEARCH

- Problem Statement
- Objectives
- Research Contribution
- Tools used
- Dataset Description
- Data Pre-processing
- System architecture
- Algorithm
- Output
- Conclusion
- References



Problem Definition

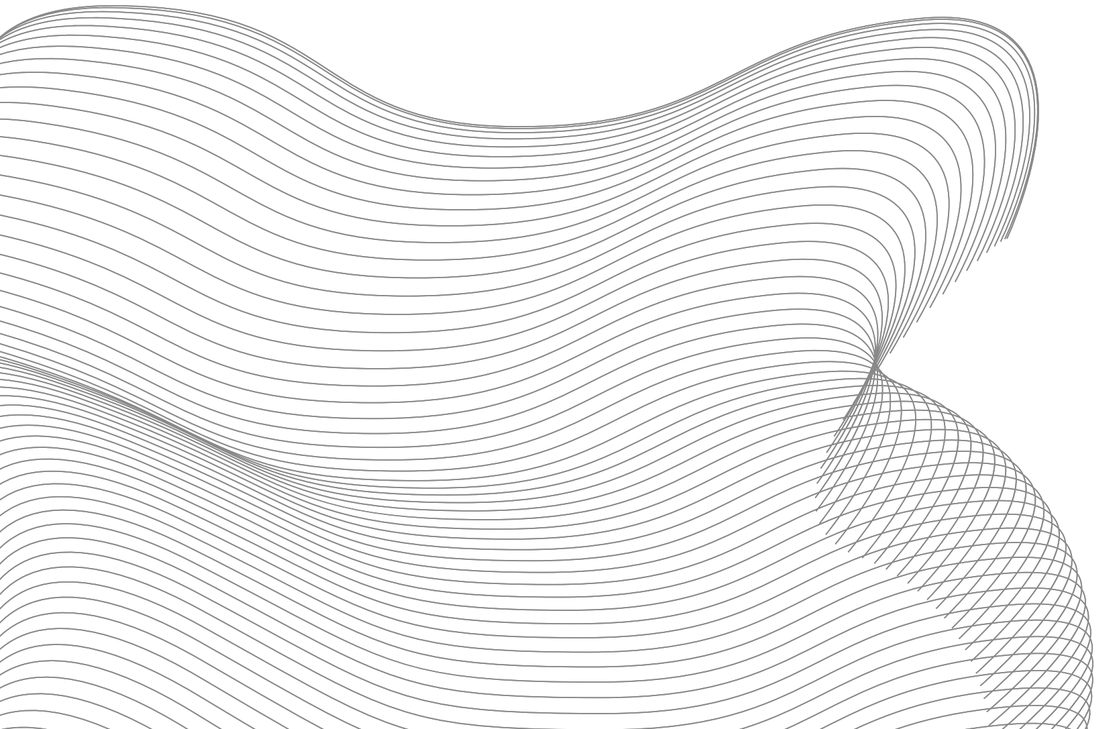


The Depression Anxiety and Stress Scales-21 (DASS-21) involves a simple structure first-order three-factor oblique model, with factors for depression, anxiety, and stress. The purpose of our study is to take use of the DASS-21 scale for use in the mental health screening in people around the globe when 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 have applied seven different machine learning based classification models that are built for mapping the traits with stress scales.

Objectives



- **01** Which are the best Algorithms for the dataset
- **02** What behaviour caused the DAS in the people
- **03** Identification of the people having depression, anxiety or stress
- **04** What scales are the people over DAS
- **05** Achieving high accuracy using just the ML algorithms instead of Neural networks

Tools Used

NUMPY

NumPy offers comprehensive mathematical functions, random number generators, linear algebra routines, Fourier transforms, and more.

PANDAS

Pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

SEABORN

Seaborn is a Python data visualization library based on matplotlib.

SCIKIT LEARN

It features various classification, regression and clustering algorithms including support-vector machines, etc

PLOTLY

Plotly provides online graphing, analytics, and statistics tools for individuals and collaboration

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).

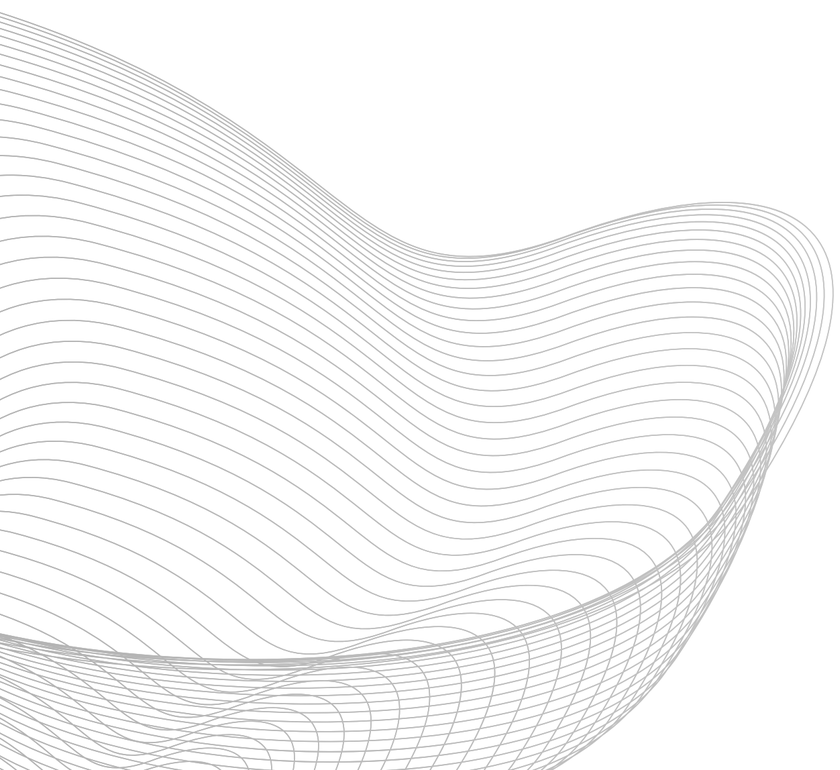
TOTAL VALUES

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

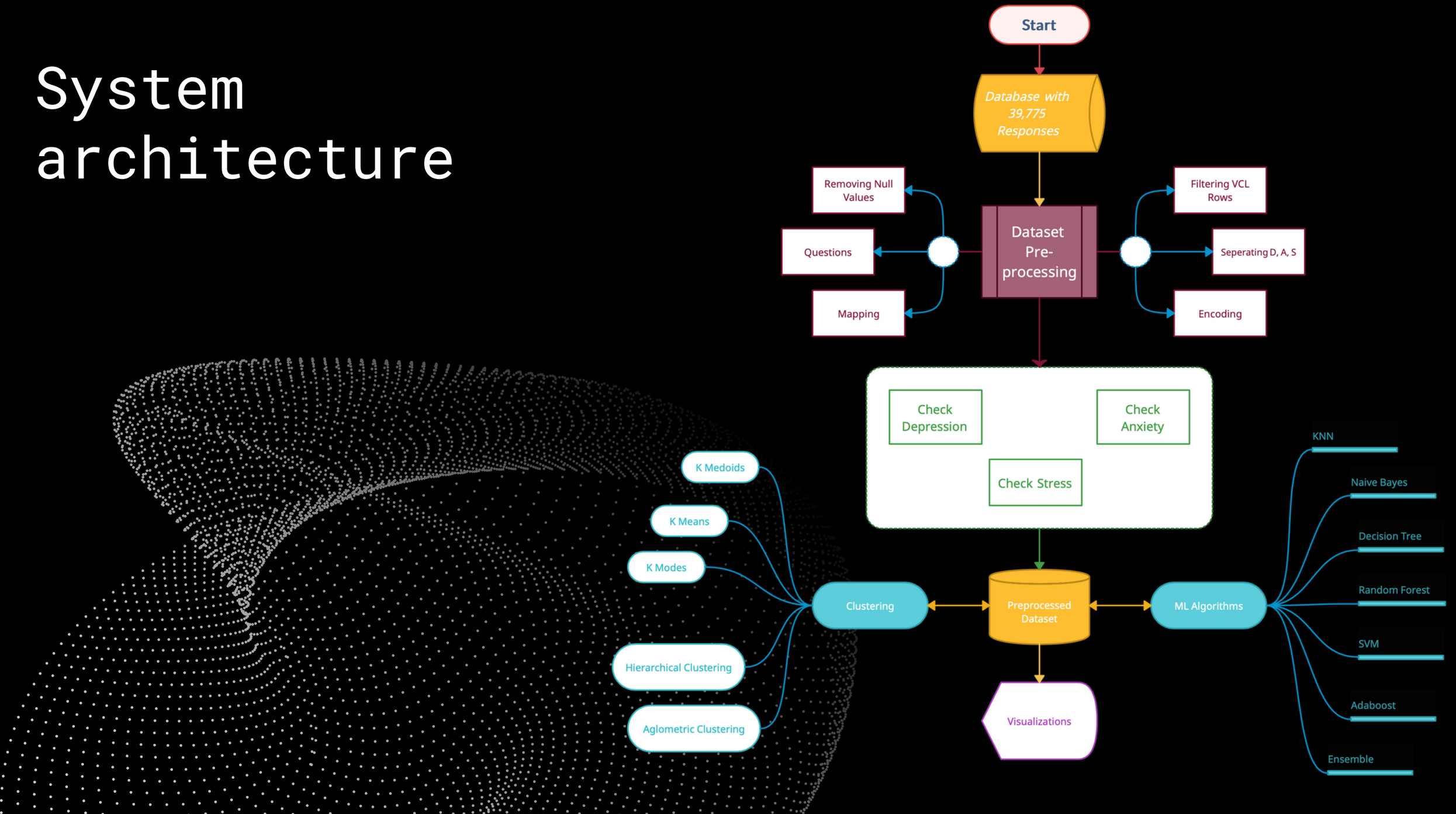


The dataset consists of the personality traits score. Openness-to-experience, Extraversion, Agreeableness, Neuroticism and Conscientiousness scores are determined using pre-set formulas and the no hard rule-based data cleaning or pre-processing is performed over the data. Therefore, the noisy data is 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



ML Algorithms Used

- K NEAREST NEIGHBOURS
- NAÏVE BAYES
- DECISION TREE
- RANDOM FOREST
- SVM
- ADABOOST
- ENSEMBLE

Clustering Methods

- K-MEANS
- K-MEDOIDS
- K-MODES
- AGLOMETRIC
- HIERARCHICAL CLUSTERING

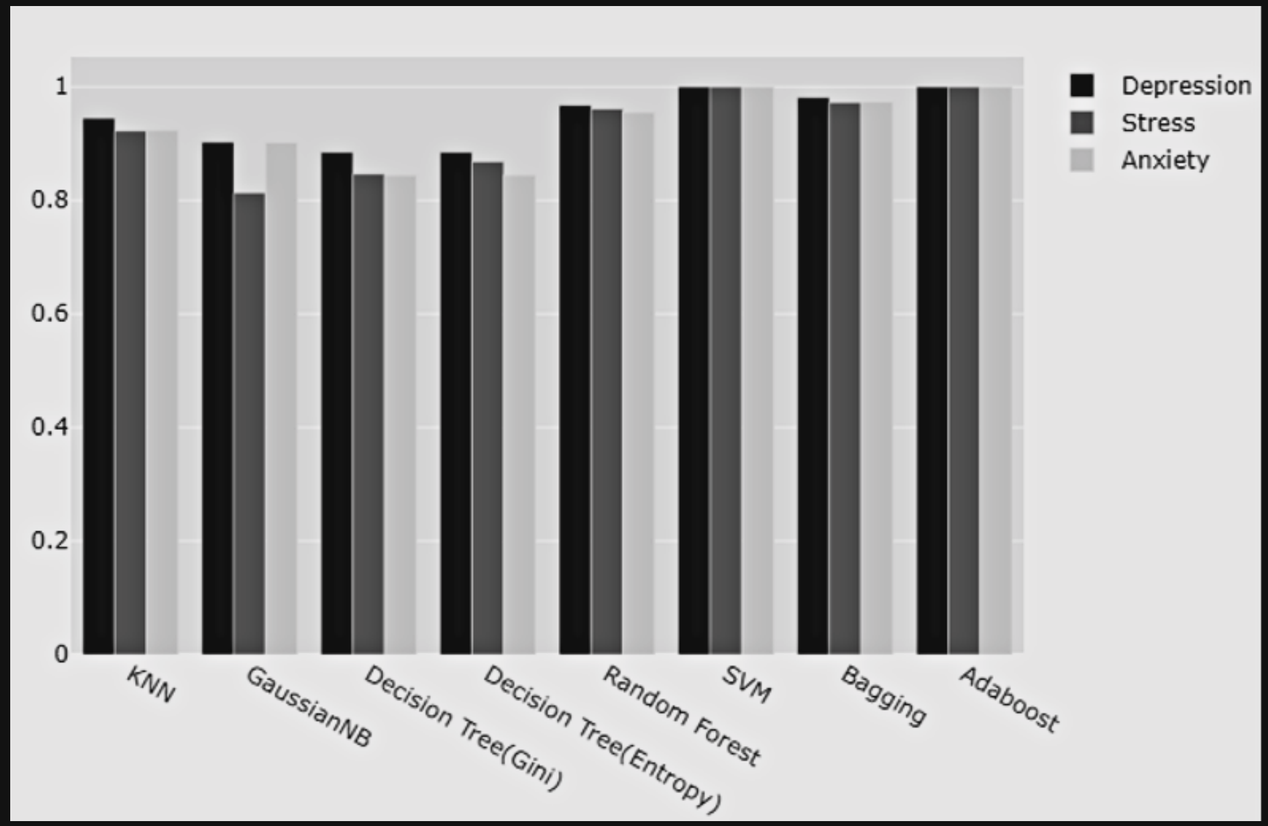
Output

OUR PROGRESS VS. OUR GOALS

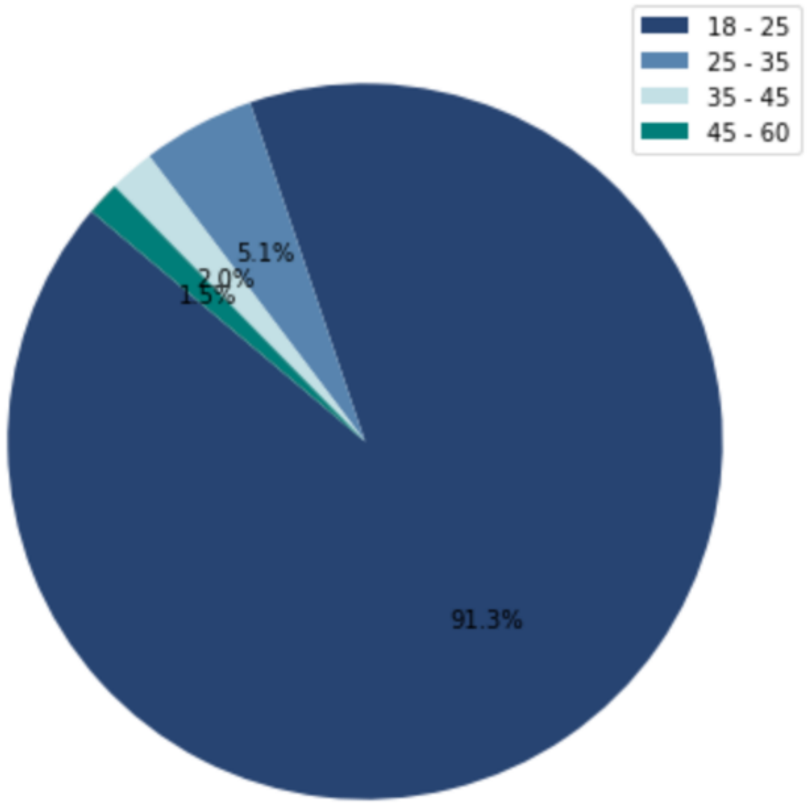
OUTPUT

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.

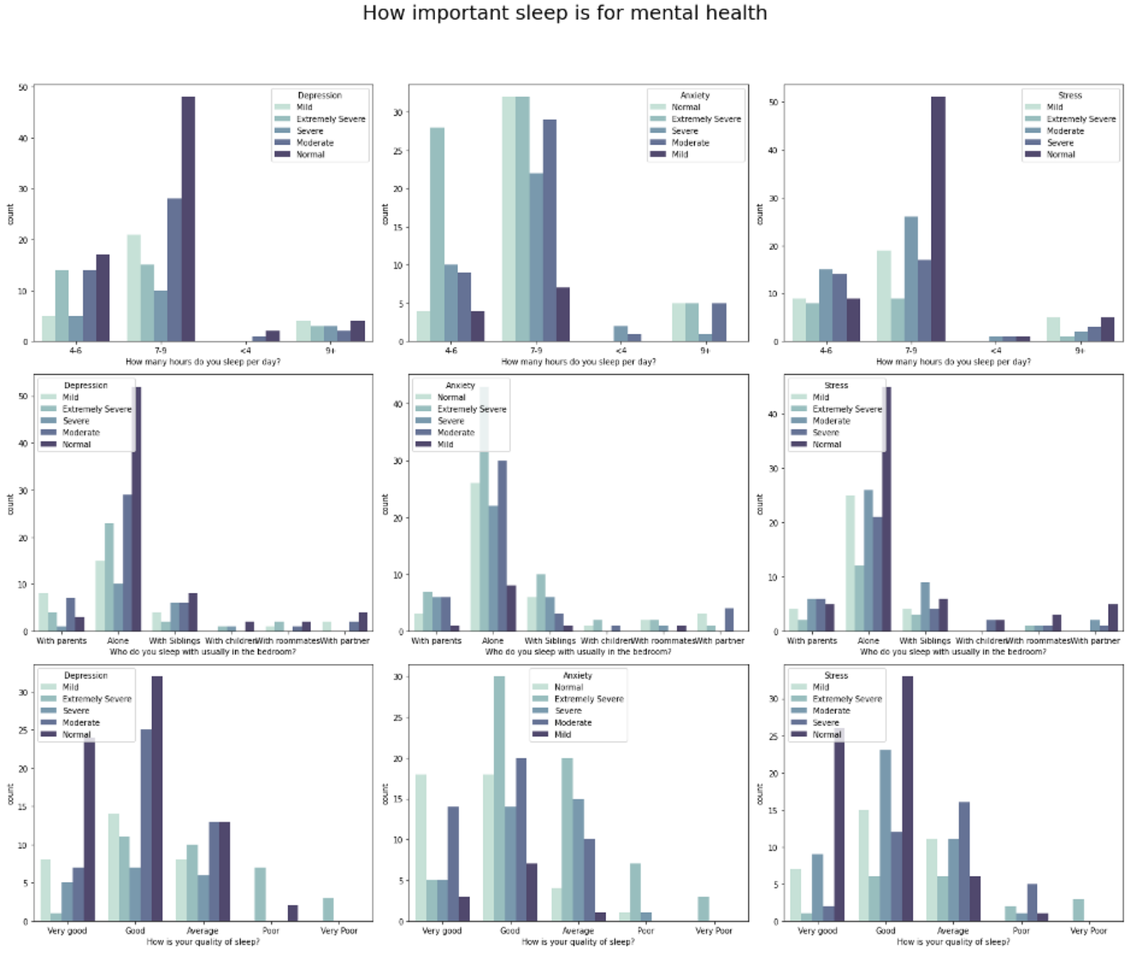
Precision=TP / (TP + FP)
Recall = TP / (TP + FN)
F1 – score = (2 * Precision * Recall)/ (Precision + Recall)
Accuracy = (TP + TN)/ (TP + TN + FP + FN)



Visualizations

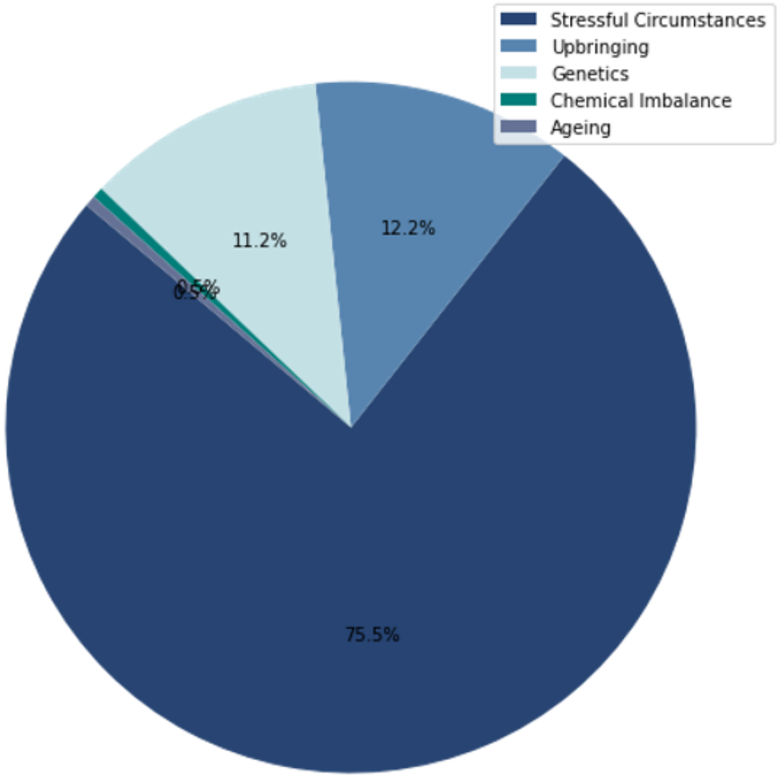


The age Group of the respondents of our live dataset

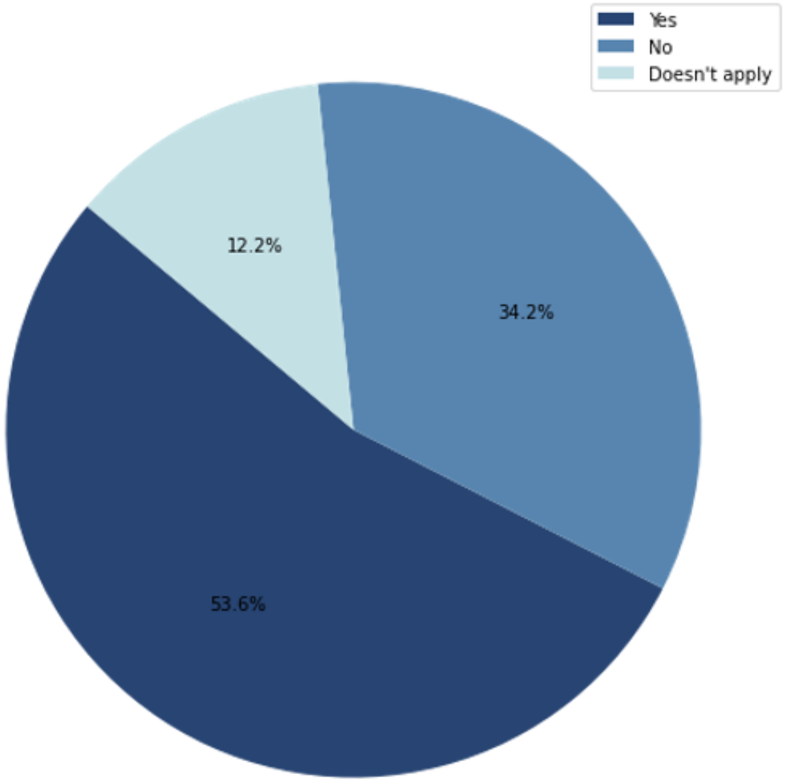


Sleep in comparison to the Mental Health (DAS)

Visualizations

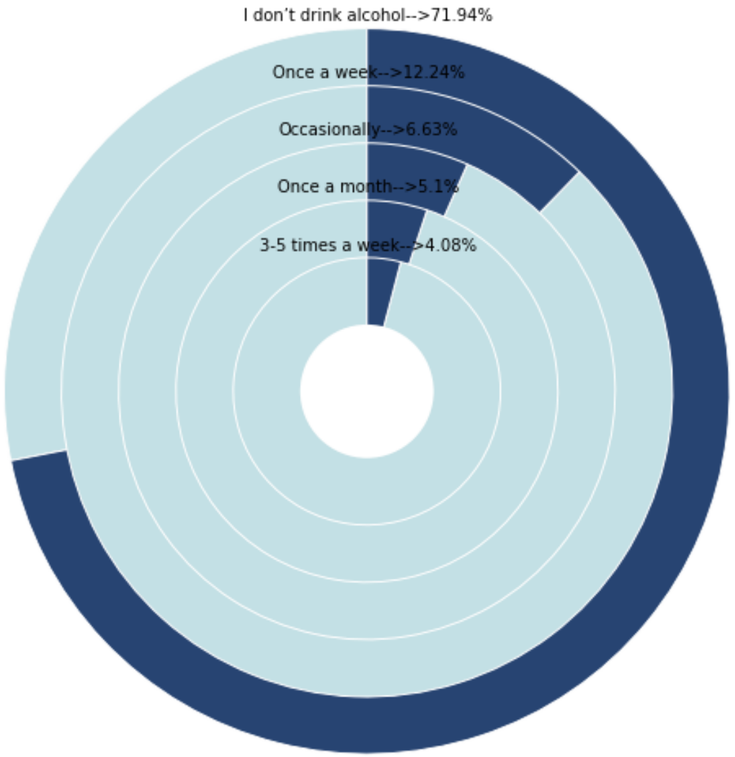


Reasons causing DAS

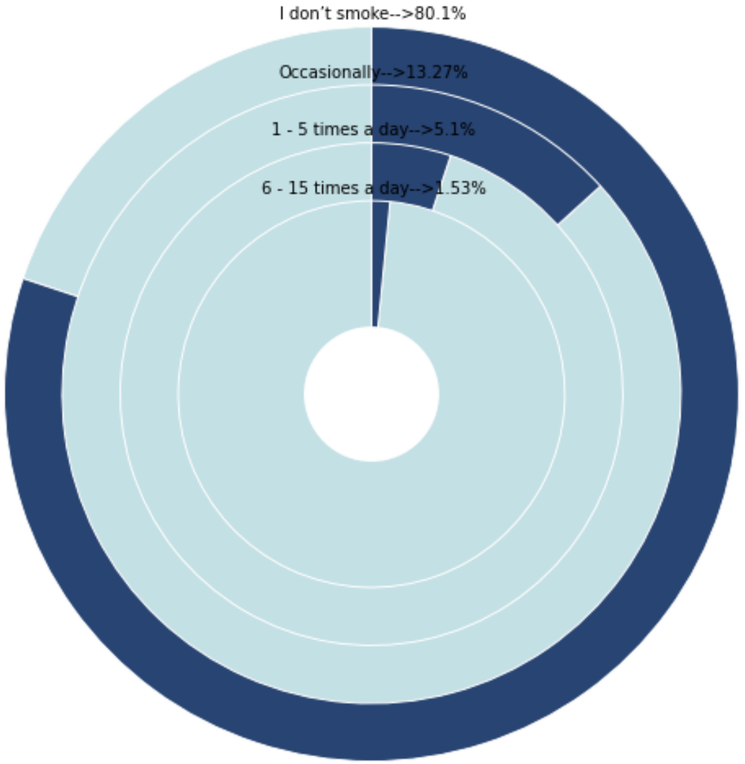


Mentality towards the treatment

Visualizations



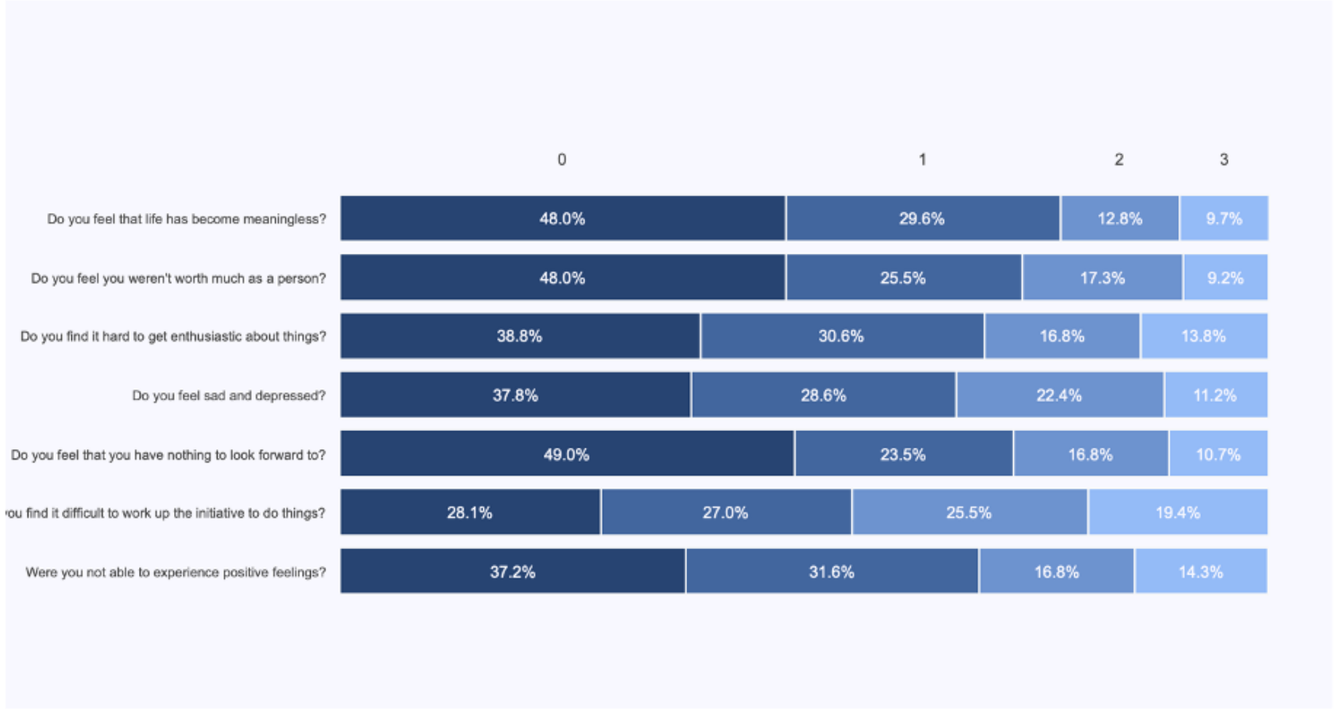
Alcohol consumption



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 100% accuracy for all the three parameters. We have presented a machine learning based approach to map between the personality traits and perceived stress scale of individuals.

Onto which after understanding which ML algorithm performs perfect, we started with working for our Live dataset retrieved from various caste, creed, status, habits, age groups, etc and found out based on the parameters which they answered onto which level of DAS were they on.

References

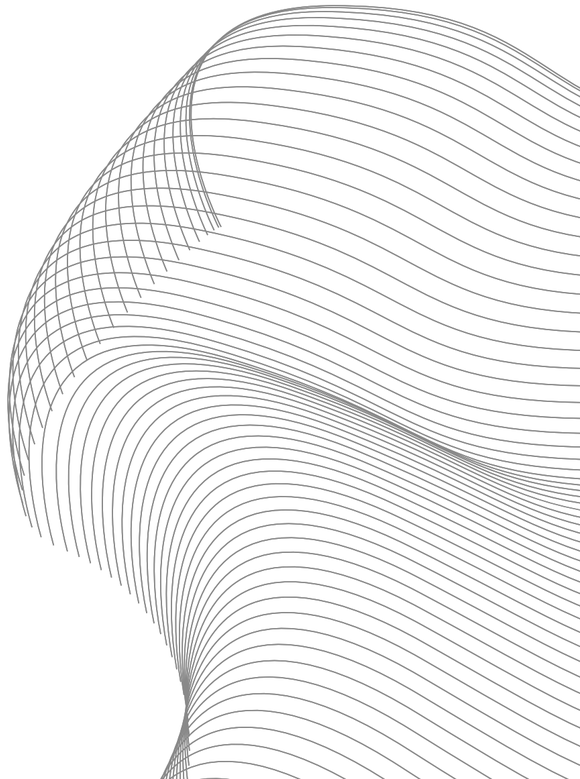
[1] Mary, S. A., & Jabasheela, L., "**Performance Evaluation of Simulated Annealing based Feature Selection Approach on Depression Dataset"**

[2] Coker, A. O., Coker, O. O., & Sanni, D. (2018)., "**Psychometric properties of the 21-item depression anxiety stress scale (DASS-21)**. *African Research Review, 12(2), 135-142.*

[3] Marouf, A. A., Ashrafi, A. F., Ahmed, T., & Emon, T. (2019). **A Machine Learning based Approach for Mapping Personality Traits and Perceived Stress Scale of Undergraduate Students**. *International Journal of Modern Education & Computer Science, 11(8).*

[4] Glöckner, A., Michels, M., & Giersch, D. (2020). **Predicting Personality Test Scores with Machine Learning Methodology: Investigation of a New Approach to Psychological Assessment.**

[5] Gomez, R., Stavropoulos, V., & Griffiths, M. D. (2020). **Confirmatory factor analysis and exploratory structural equation modelling of the factor structure of the Depression Anxiety and Stress Scales-21.** *PloS one, 15(6), e0233998.*



Our gratitude to you for all you have done **Jayshree Ma'am**, we truly appreciate you and the time you have spent helping us in various occasions.

Thank you very much for this **AML** course while we enjoyed every minute of this lecture as well as learnt and practices marvellous ideations throughout.

Thank you!

Thank you to our **Respondents** and **Participants** for taking the time to complete the survey for and submitting it back to us.

Your participation has been important to us and your feedback and answers has helped us improve the research/process for **Psychometric Predictions using ML**

Thank you!