

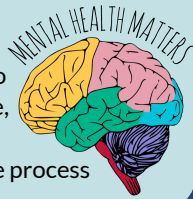
# MENTAL HEALTH ANALYSIS USING MACHINE LEARNING

## Problem Statement

Using the Depression, Anxiety and Stress Scales (DASS-21) questionnaire in order to predict anxiety, depression and stress using various classification algorithms.

## Motivation

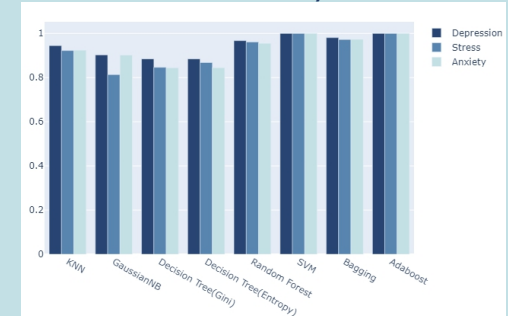
As a result of modern lifestyles, Many people are suffering from various forms of psychological health problems. Anxiety, depression, and stress all have certain characteristics. Generally, psychiatrists measure anxiety, depression, and stress using questionnaires such as the DASS42 and DASS21, since people who suffer from anxiety, depression, or stress are typically unwilling to share their symptoms. Therefore, we thought that with the help of machine learning we can ease the process



## Methodology

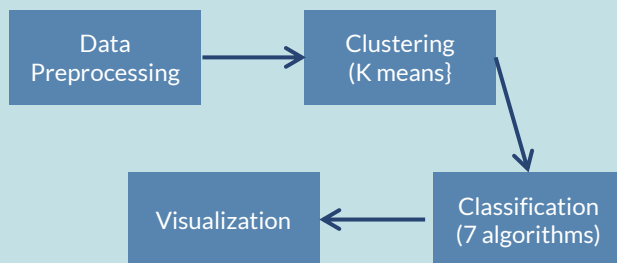
- Dataset: kaggle dataset \_depression-anxiety-stress-scales-responses, which were the DASS 42 form's responses of people around the globe.
- Rows: 35000+ rows
- We mapped the DASS42 to DASS21 to reduce the no. of questions.
- Tools
  - Python
  - Sklearn library for machine learning models
  - Plotly for visualization
- Algorithms Used
  - KNN
  - Naive Bayes
  - Decision Tree
  - Random Forest
  - SVM
  - Voting Classifier
  - AdaBoost with SVM as base

## Accuracy

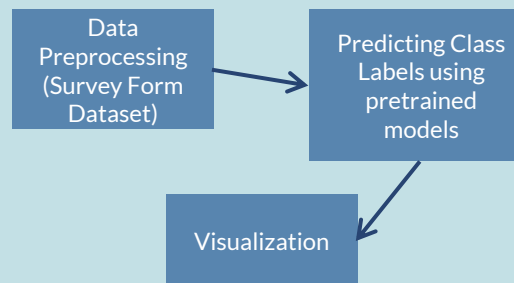


Classification models were trained on the kaggle dataset to generate models for each algorithms so that the best algorithm can be later applied for predicting mental health levels in the survey we conducted using google forms.

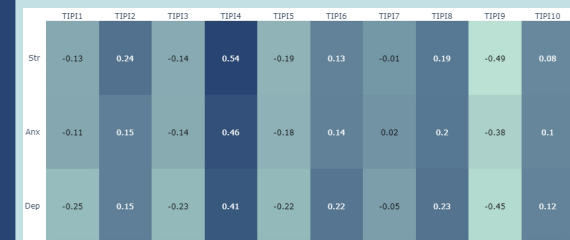
## Steps taken for Kaggle Dataset



## Steps taken for Survey on Google Form



## Correlation Heatmap



The questionnaire also asked about Ten Personality Questions. The above chart shows the correlation of depression, anxiety and stress score. Eg TIP14 (Anxious, easily upset) is positively correlated with the DAS score and whereas TIP11 (Extraverted, enthusiastic) is negatively correlated with the DAS score.

## Conclusion

We used 7 machine learning models to predict five severity levels of anxiety, depression, and stress. The models were KNN, Decision Tree, Naive Bayes, Random forest, SVM and 2 ensemble algorithms. AdaBoost and SVM approach gave 100% accuracy for depression, anxiety and stress.

## 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>

## Group No. 11

PB11 Prajakta Chaudhari  
PB27 Snehalraj Chugh  
PB Janhavi Chavan  
PB Harshit Srivastava