Mental health Prediction

Reem Abanmi - Norah Tamsan

ABSTRACT

Mental health is an integral and essential component, it's the foundation for individual well-being and the effective functioning of the community. Globally the mental health problems are rising the prevalence of mental disorders as per World Health Report is around 13%. with the increase in cases around the world, the need for psychological treatment has increased So we found it useful to provide a service to predict whether a patient should be treated of his/her mental illness or not by building a classification module.

DESIGN

Using the data collected we will predict whether an individual should be treated of his/her mental illness or not by building a classification module and evaluate different methods of classification to find highest accuracy and apply this method to the Testing set. This classification module will help a lot of people and notify them to start psychotherapy sessions before there mental health gets worse.

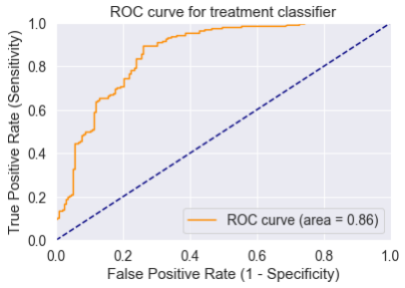
DATA

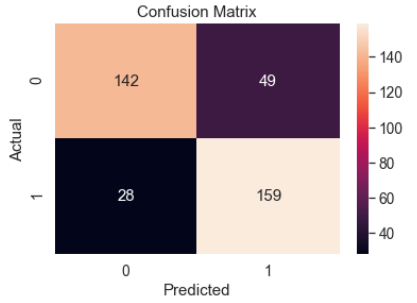
The dataset we are going to use to train our model is a survey done for 1259 people data set. This dataset contains 27 features most of it are categorical e.g. (yes, no, not sure), A few feature highlights include age, gender, and family history of mental illness.

ALGORITHM

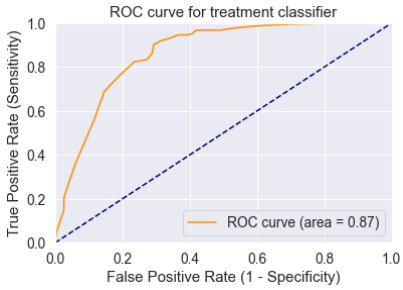
1. We started by loading the data from the csv file
2. Preparing the data by cleaning it which include dealing with missing data and getting rid of unwanted features
3. Feature engineering:
    1. We start by Encoding data by creating a dictionary to map survey answers to numerical values
    2. Combining particular ranges of numeric features like Ranges of Age.
    3. Features Scaling where we scale age because it's extremely different from the other features

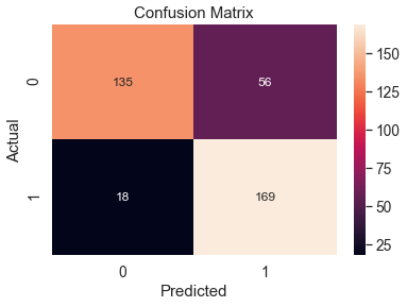
4. Models, we apply more than one classification model to find the best model and evaluate each of them
    1. Logistic Regression classifier



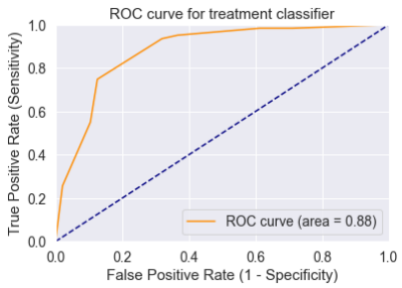


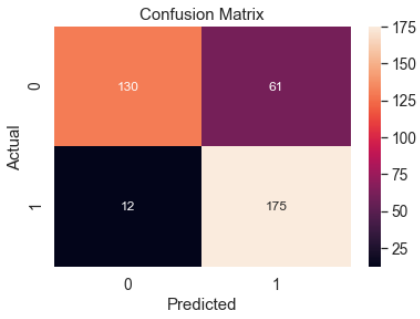
    2. KNN classifier



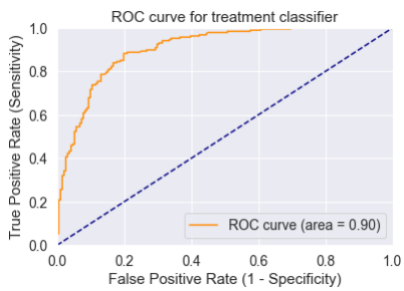


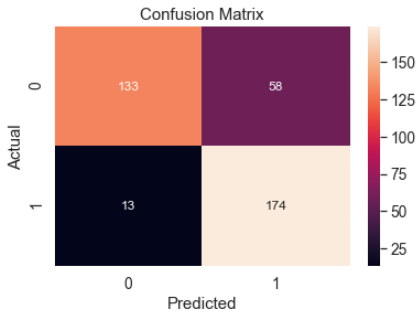
    3. Decision tree classifier



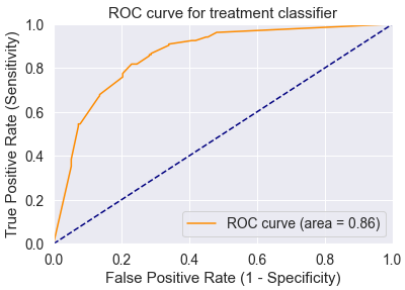


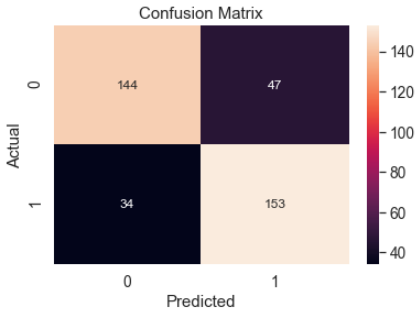
    4. RandomForest classifier



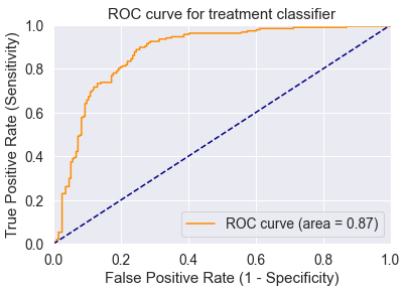


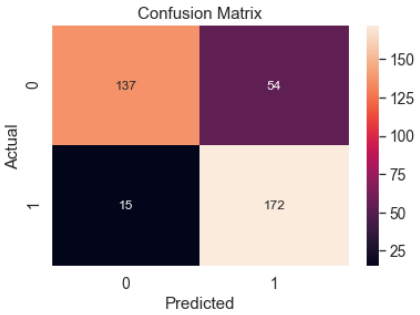
5. Bagging classifier



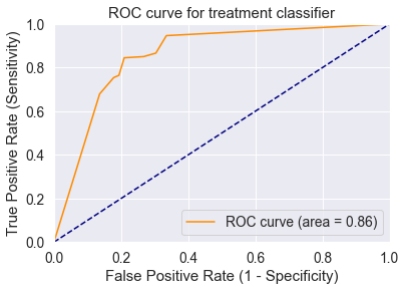


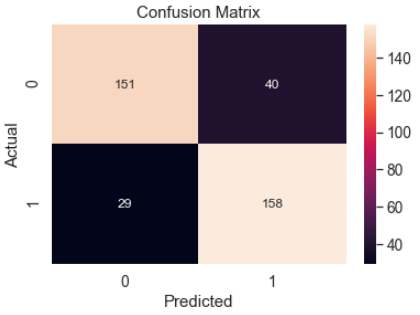
6. Boosting classifier





7. Stacking classifier

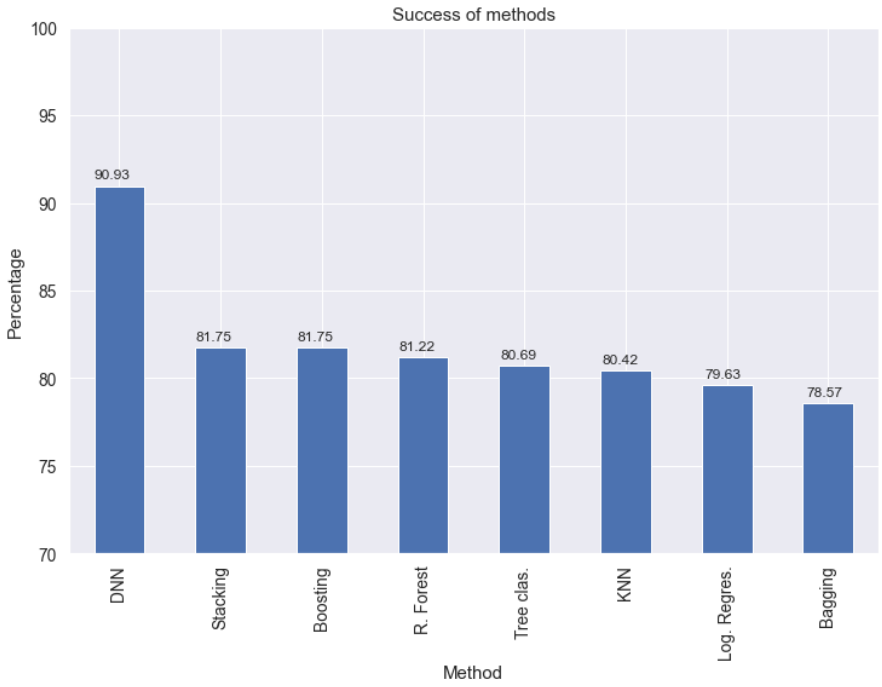




8. Predicting with Neural Network (DNN)

deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers, it gives us an accuracy = 90.93

5. Model evaluation and selection: the data was split into 70/30 and after applying all models we choose the DNN since it has highest accuracy as shown below.



TOOLS

1. Seaborn
2. Pandas library
3. NumPy
4. TensorFlow
5. Matplotlib
6. Scikit-learn
7. SciPy