Predicting Onset of Depressive Disorder Using Machine Learning

Srivastava, Varun (School: Redlands High School)

The COVID-19 pandemic had a major impact on mental health due to the increased stress caused by the mental toll of events of the pandemic and isolation from necessary preventative measures such as quarantines and shutdowns. During the pandemic, about 4 in 10 adults in the US have reported symptoms of anxiety or depressive disorder, up from one in ten adults who reported these symptoms from January to June 2019 according to the Kaiser Family Foundation. This research aims to develop a machine learning model that can accurately predict the onset of depressive disorders, presenting a binary classification problem. Using the Behavioral Risk Factor Surveillance System (BRFSS) dataset, multiple machine learning algorithms, from Gaussian Naive Bayes to Multi-layer Perceptron Classifier were tested and analyzed using model metric analysis. The Random Forest machine learning model using Gini criterion was determined to be the most accurate. An easy-to-use iOS and Android application (called Let's Heal) was developed where users answer 25 questions and receive a prediction from the machine learning model based on their response. The results show the model to be about 84.29% accurate when tested on 83,796 samples, with 70,632 correct predictions, demonstrating that the model can be used by schools and doctors to identify the risk of developing depressive disorders and refer them for treatment.

Awards Won: