ADEOLA AYANDEYI
STUDENT, MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

SEMI SUPERVISED LEARNING FOR UNDERSTANDING THE EFFECTS OF COVID-19 ON MENTAL HEALTH FROM TWITTER DATA

More contagious virus variants plunge parts of Canada into the third wave of the pandemic and it likely contributes to rising mental health difficulties. It is speculated that the fourth wave of the pandemic is going to be a mental health wave which will be sustained for a very long time. Now it's crucial importance to understand the pandemic induced distresses.

In this research, we construct a corpus consisting of twitter data as it relates to mental health from Jan 2020 till October 2020 from about 4 million users to analyze their concerns on the effects of virus on mental health. We carried out the following activities to develop a sentiment analysis model which could be exploited to automatically dig the big data and excavate meaningful information to support public policy decisions. To develop the model, first we preprocess and clean the data. Then we compute 1-gram feature which computes the frequency of the words in the corpus related to mental distress. This feature is fed into an extensive list of machine learning classifiers such as K- Nearest Neighbors, Gaussian Process, Decision Tree, Random Forest, MLPClassifier, AdaBoost, GaussianNB, Gradient Boosting and Logistic Regression to predict the sentiments: positive, negative, or neutral. The success of Machine learning algorithms depends on the amounts of the labelled data. However, labelling hug amount of data is time consuming, laborious and burdensome. Hence, we develop a semi-supervised model which outperforms supervised-learning in the context of insufficient labelled data.

This research will help policymakers to select the best model which facilitates analyzing social media data and understand the mental distresses. This tool could be useful for post-pandemic period as well. This is a generic tool which could be exploited to analyze all other mental health issues, and public sentiments in other social issues as well. In addition, this tool will make us better prepared for the future pandemic.

Research Advisor: Dr. Baidya Nath Saha