

# Blog Post (to be shared with the public)

## Subtitle: Predicting depressive symptoms among Twitter users

Mental health diseases have become one of the major issues the modern world is facing nowadays. People have been diagnosed with depression and only a few of them decide to get treatment or diagnosis.



Twitter users express their thoughts and opinions through tweets. Almost 200 billion tweets are tweeted per year. Predicting the depressive symptoms among Twitter users will contribute to decreasing the consequences of the issue through early diagnosis and enabling early intervention to treat it.

Social media has played a crucial part in the lives of almost everyone in the past decade. It can also provide an unrivaled opportunity to predict depressive symptoms and contribute to early interventions to treat them.

## **Ethical Perspective (#ethicaljudgement)**

Research has been made on different platforms of social media to investigate this phenomena. Although the methods discussed later in this post use human subjects, language communicated in tweets is still considered a proxy for human behavior and a representative of human traits. The aim of such use of data is primarily driven with ethical intentions that aspires to help people and make a change in the community.

Depression diagnosis heavily relies on self-reported symptoms. In other words, no physical or tangible examination is required to make an official diagnosis of depression.

We find ourselves debating between two difficult options: whether to enable the user the freedom of speech and their privacy or to help prevent severe consequences of depression. Taking into consideration the severe impact of depression on the quality of life of people and even the possibility of suicide. In fact, in 2015, the World Health Association reported that 788,000 have died by suicide for individuals between 15 and 29 years old.

Therefore, we can safely rely on social media posts such as tweets of users with reliable data to investigate through machine learning models.

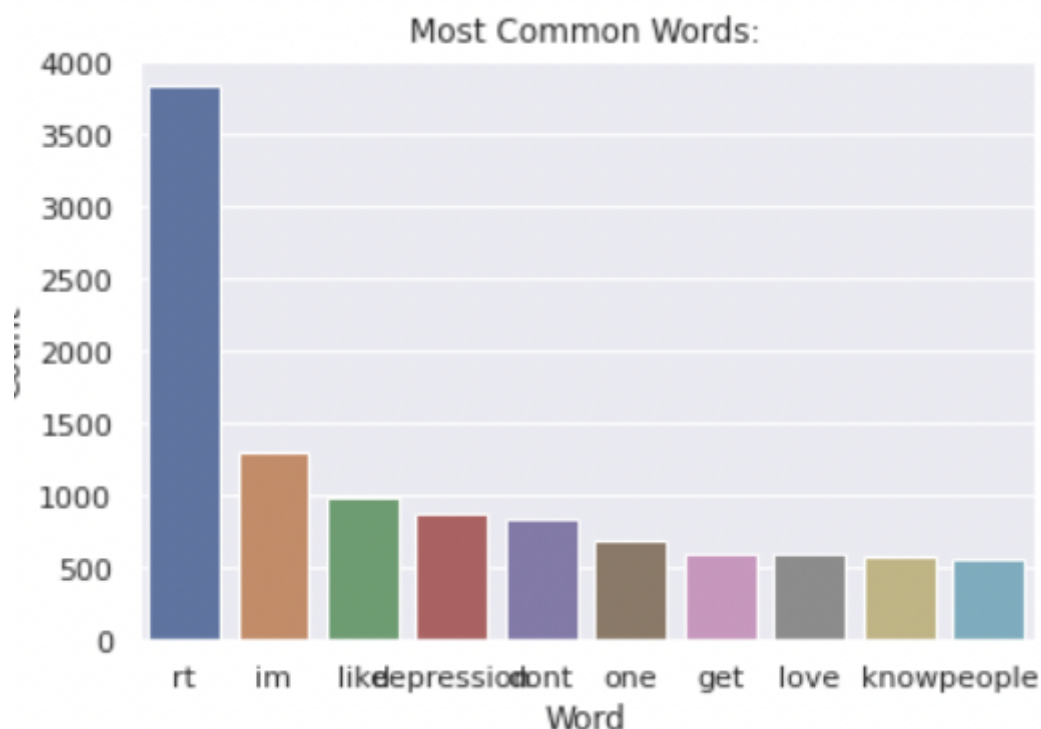
## **Tweets with Depressive Symptoms**

We wanted to create a computational model that captures linguistic clues associated with depression. This tool will help predict the emergence of depression in Twitter users.

The model required acquiring the data from Twitter. The dataset includes metrics such as tweets, retweets, number of followers, etc. They mainly informed us about the user and their past tweets and engagement for each tweet.

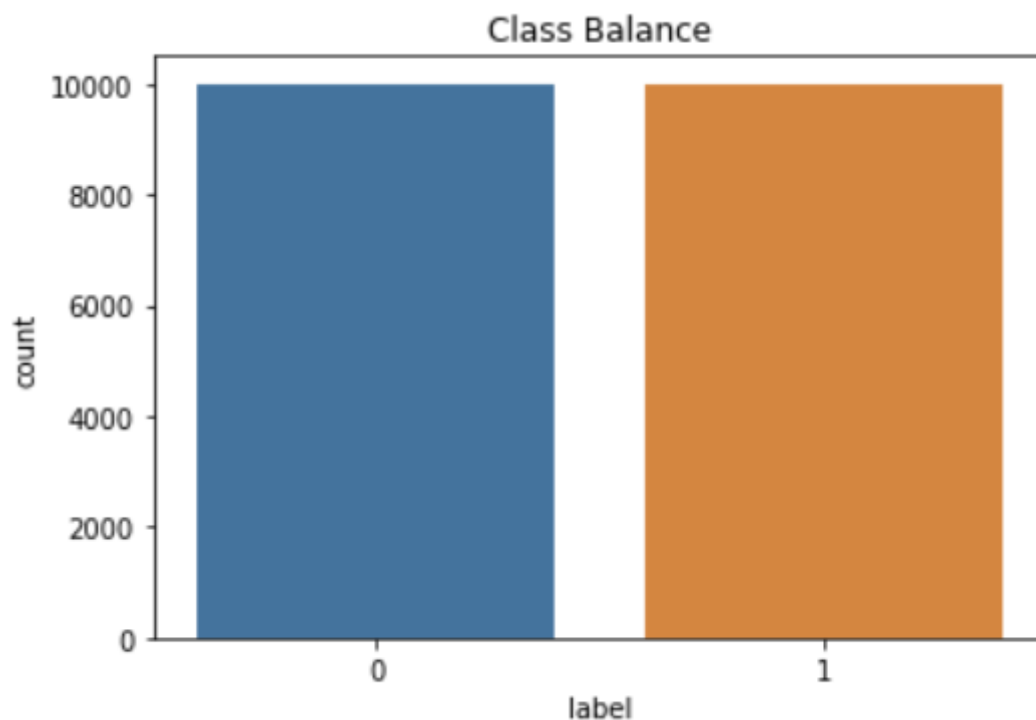
We extract a word-cloud to spotlight the frequent words in the dataset.

We zoom into this word-cloud. The initial analysis shows that the most common words used in these tweets are : [rt, I'm, like, depression, don't, one, get, love, know, people]



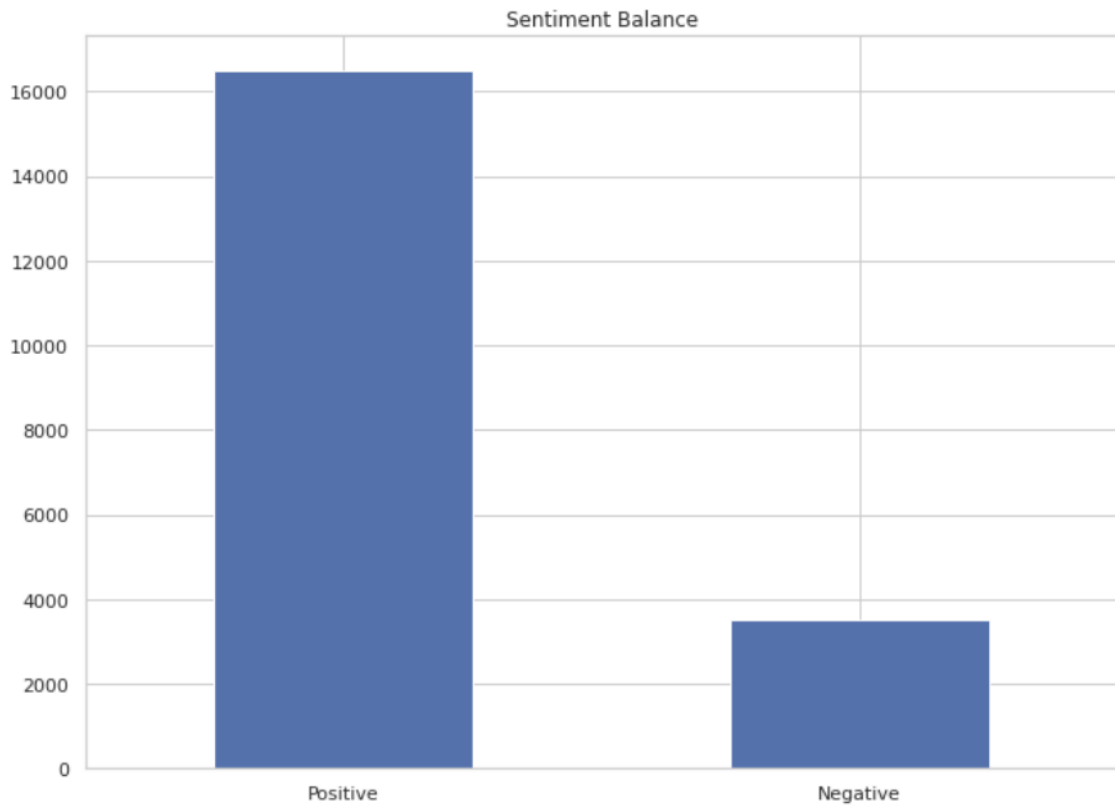
The word “rt” has a very significant frequency in the dataset which highlights that users agree with the what a user tweeted.

We extracted these predictive features to train the model. We start by running a sentiment analysis on the data. It is balanced since there is an equal number of depressed and non-depressed tweets. There will be no issue of class incidence affecting the accuracy of the classification.



We create the term frequency inverse document frequency (TFIDF) matrix to assign each word a value describing its importance in the sentence given its prevalence in the document.

The sentiment balance in the graph below shows that 16000 words within the dataset have a positive connotation while 3500 words have a negative one.



We also used the HashingVecotizer from sklearn to convert each word into a vector, since it is very scalable to large datasets and uses a small amount of memory.

We then trained our model using an ensemble of Random Forest classifiers, since it would be easy and relatively quick to train and test with, and would provide some initial insights into how complex our model need.

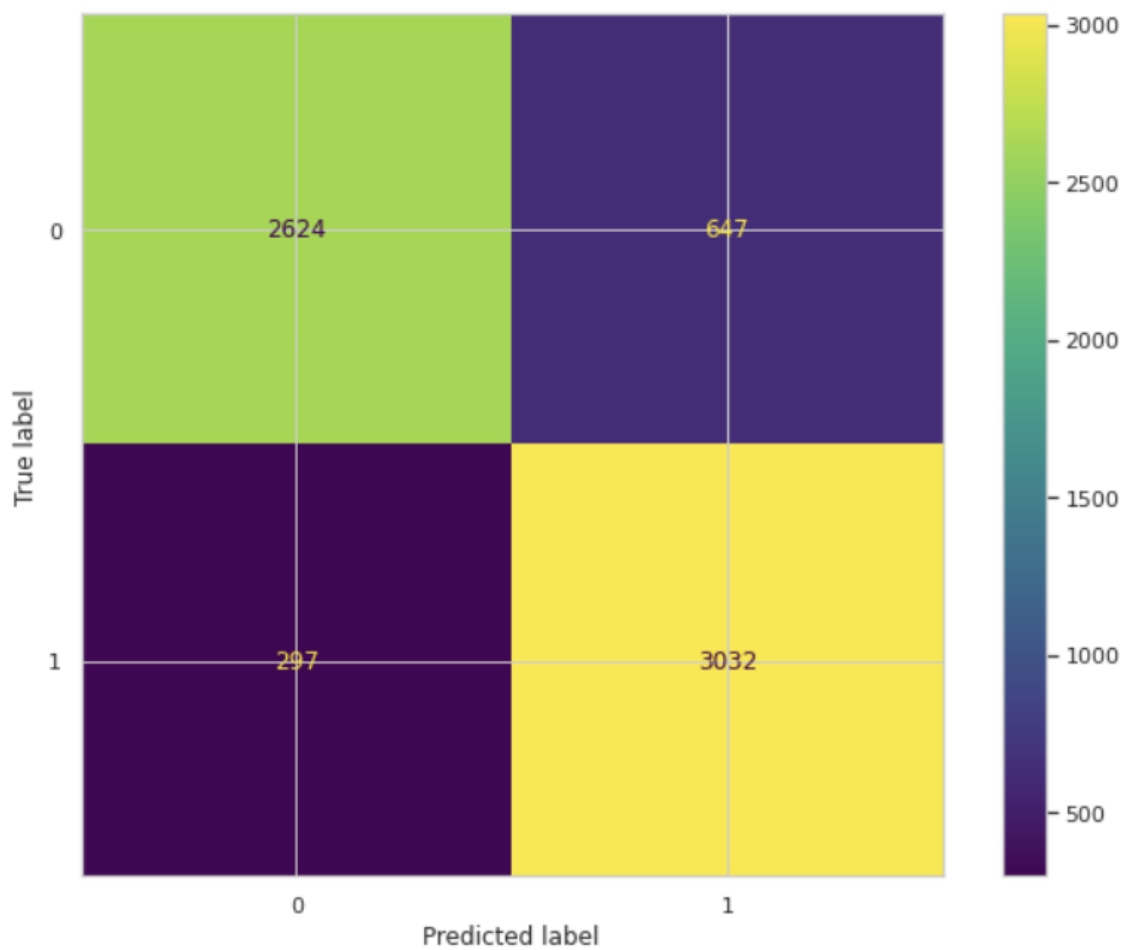
We use Random Forest, Logistic Regression, Naive Bayes and XGBoost.

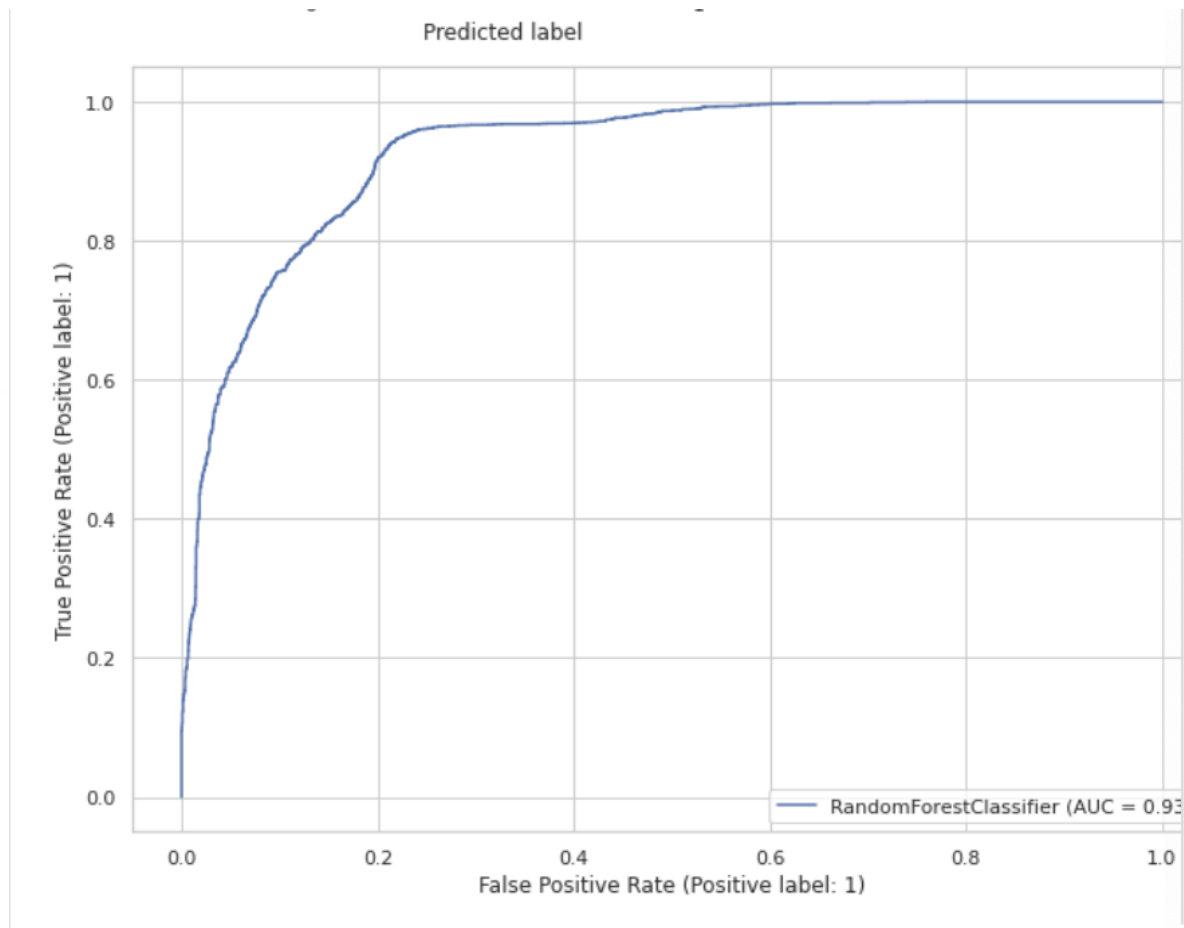
Random Forest:

	precision	recall	f1-score	support
non-depressed	0.90	0.80	0.85	3271
depressed	0.82	0.91	0.87	3329
accuracy			0.86	6600
macro avg	0.86	0.86	0.86	6600
weighted avg	0.86	0.86	0.86	6600

Cross Validation Score: 0.859776119402985

ROC Curve:





Our training cross-validated accuracy score was 85.97% and our testing cross-validated

accuracy was 85.91%. This shows that our model avoids much over-fitting since there is a very small loss in accuracy between our training and test sets.

The confusion matrix for our model and the ROC-curve (below) show that our model has about twice as many false positives as false negatives, which might be a good thing because it would be better to have people go for more mental health check-ins to create an overall more positive mental state, than for us to classify the people with depression as not having depression

Logistic Regression:

Model: Logistic Regression

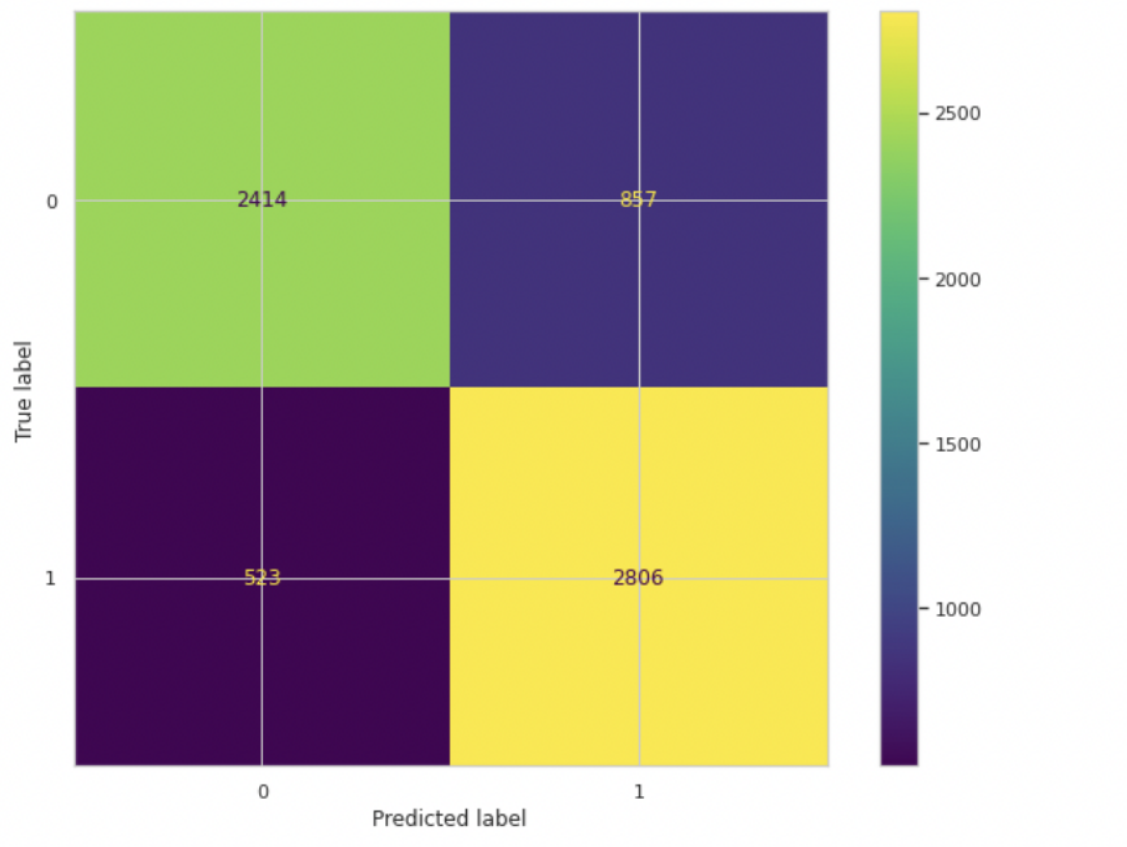
Classification report:

	precision	recall	f1-score	support
non-depressed	0.82	0.74	0.78	3271
depressed	0.77	0.84	0.80	3329
accuracy			0.79	6600
macro avg	0.79	0.79	0.79	6600
weighted avg	0.79	0.79	0.79	6600

Cross Validation Score: 0.7877611940298508

Confusion matrix:

ROC Curve:



The logistic regression had a slightly lower cross validation score of 78%

Naive Bayes:



Model: Naive Bayes

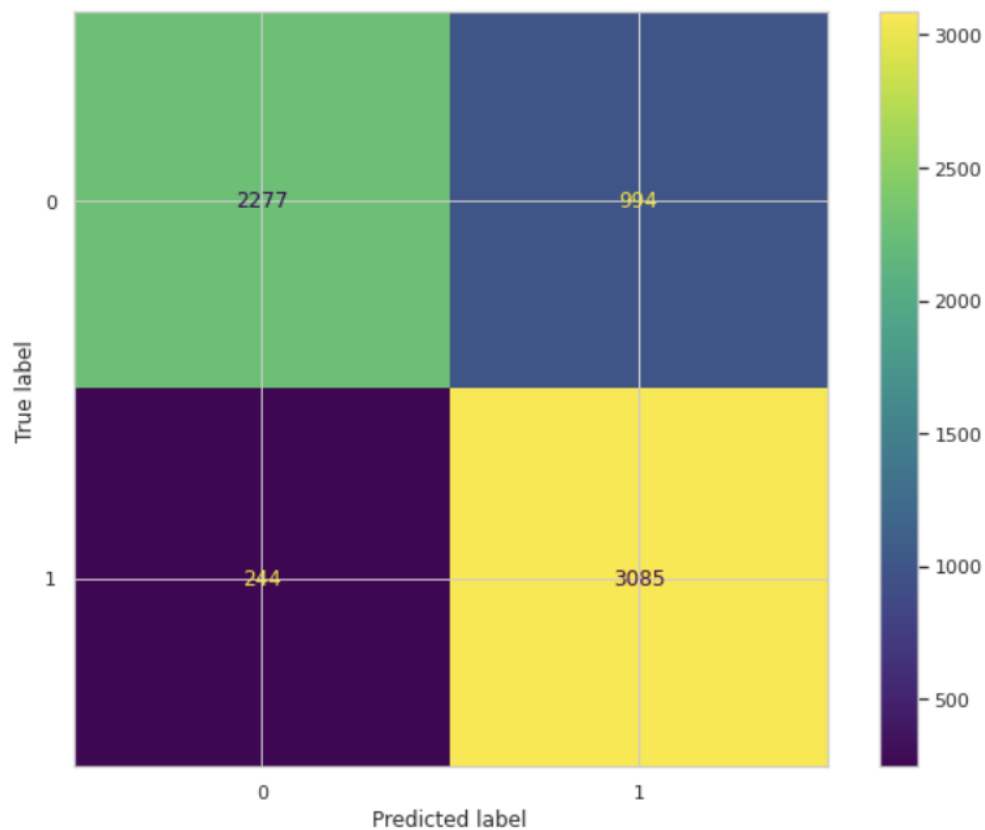
Classification report:

	precision	recall	f1-score	support
non-depressed	0.90	0.70	0.79	3271
depressed	0.76	0.93	0.83	3329
accuracy			0.81	6600
macro avg	0.83	0.81	0.81	6600
weighted avg	0.83	0.81	0.81	6600

Cross Validation Score: 0.8144029850746268

Confusion matrix:

ROC Curve:



We acquired 81% for the cross validation score.

## Further Consideration and Suggestion



It is necessary to highlight that the classification of the words is still computational. Any findings may be limited by the non-specific use of the term “depression”.

The findings are limited to Twitter Users only. Notwithstanding, the framework used is applicable on datasets with similar features from other social media platforms.

Additionally, it would be interesting to see if supporting this dataset with more metrics such as clinical variables. Put differently, depression is also diagnosed across clinical spectrums. It would be revolutionary if participants with a specific type of depression provided information that can broaden the dataset. This will enable a further examination of specific depression classes, and lead to a predictive screening per diagnosis type.

## Conclusion

With a high accuracy across all different models run on the dataset, the findings seem to be accurate and robust that could be interpreted for the necessary aims of the study.