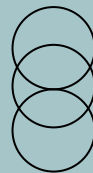
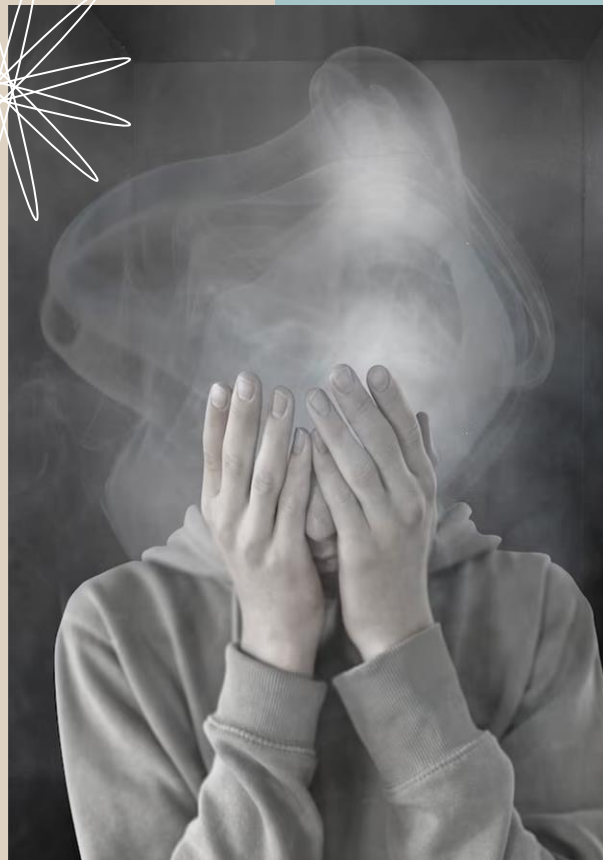
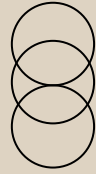


Emotion Recognition Machine Learning Approaches for Depression Detection

Adrián Reviriego

Supervisor: Ralitza Raynova





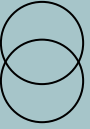
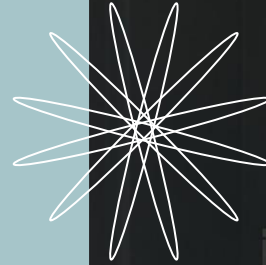
Depression



Data Science

“Data, I think, is one of the most powerful mechanisms for telling stories. I take a huge pile of data and I try to get it to tell stories.”

- *Steven Levitt*



Technologies



Python



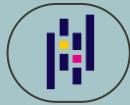
Scickit Learn



Numpy



Matplotlib



Pandas



NLTK

Datasets



CEASE Dataset

Dataset made from suicide cards. Each card is divided in phrases and labeled with an emotion.



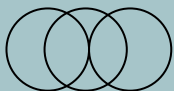
Kaggle Dataset

Dataset made from tweets from the web. Each tweet has been labeled with a emotion.

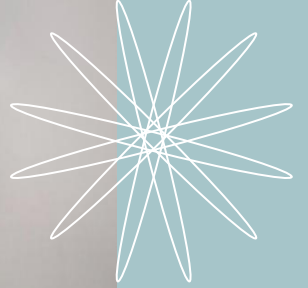
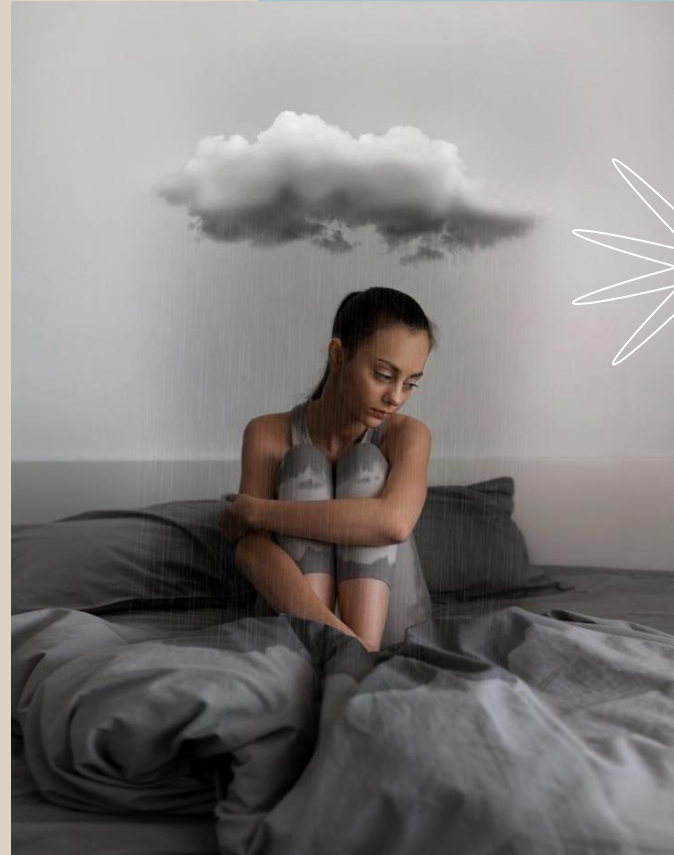


GitHub Dataset

Dataset done gathering tweets from the web. Each tweet has an ID and an emotion associated.



Results



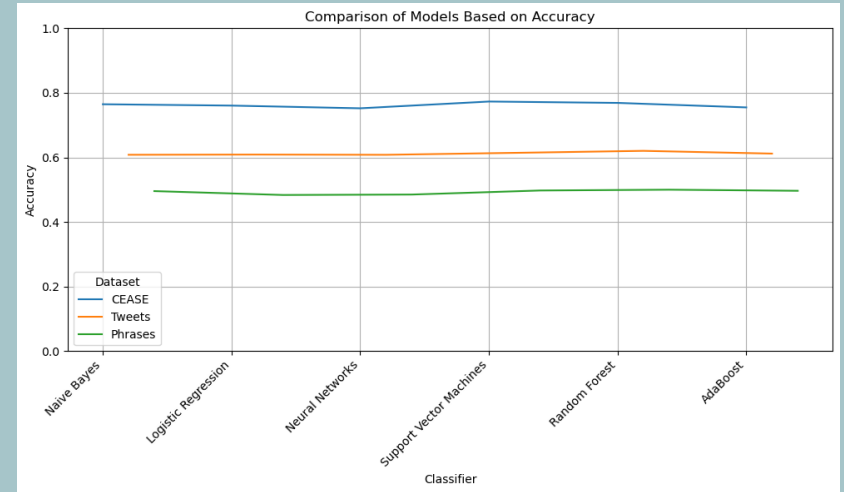
Accuracy



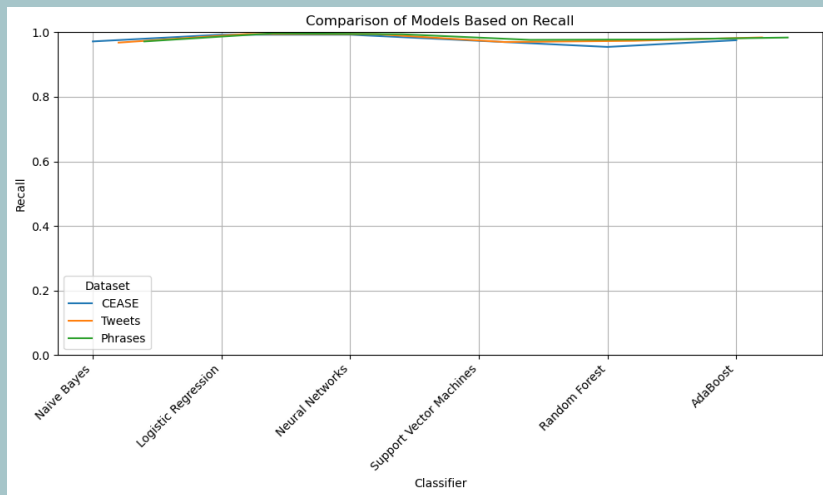
Understanding the numbers

The proportion of correctly predicted instances out of all instances.

Best Accuracy -> Support Vector Machine



Recall



Understanding the numbers

Measures how often the model correctly identifies a comment from a person with depression.

Best Recall -> Logistic Regression

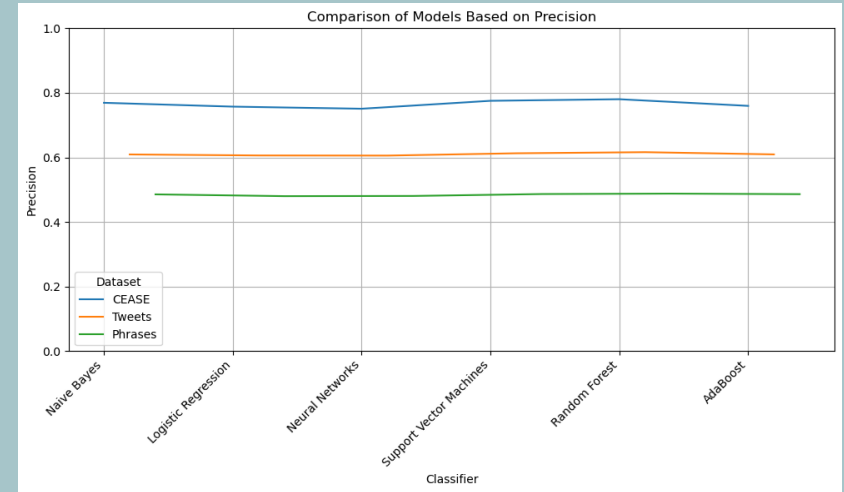
Precision



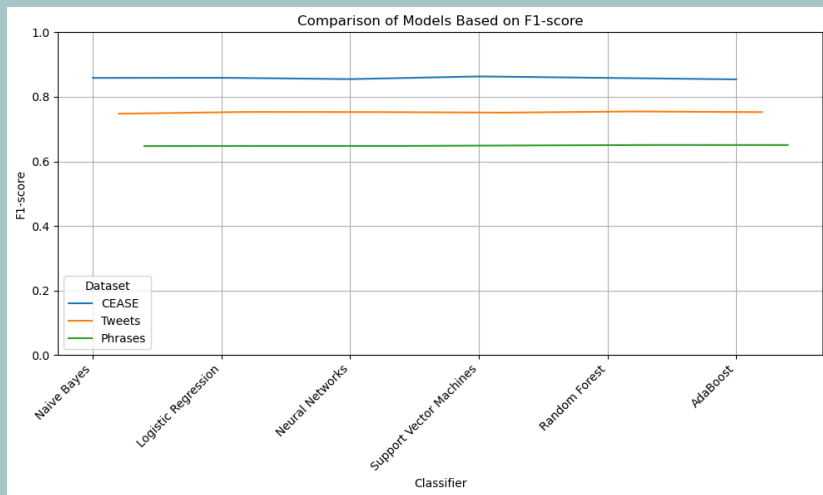
Understanding the numbers

Measures how often, out of all the people predicted to have depression, the model correctly identifies someone with depression.

Best Precision -> Random Forest



F1-Score



Understanding the numbers

The harmonic mean of precision and recall.

Best F1-Score -> Support Vector Machine

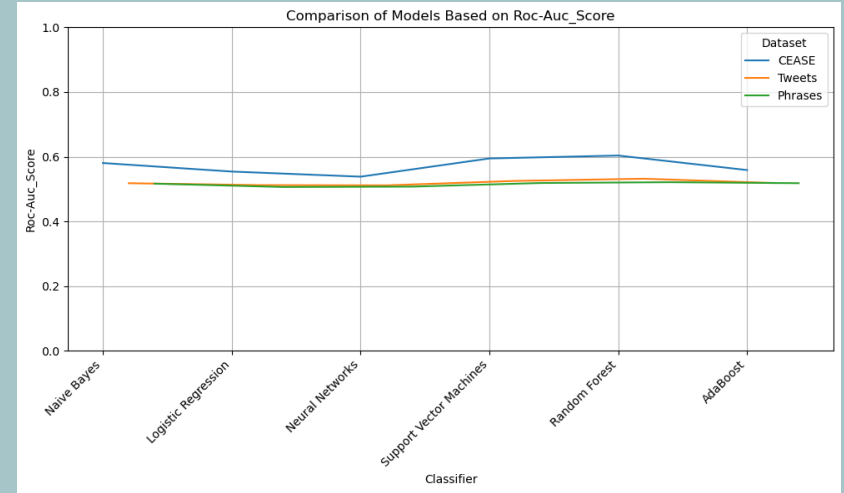
ROC-AUC Score

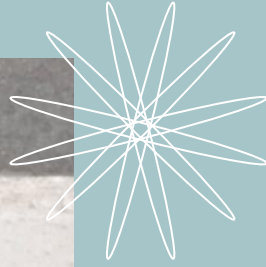
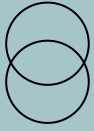


Understanding the numbers

Measures the ability of the model to distinguish between classes.

Best ROC-AUC Score -> Random Forest





Achievements and Future Directions



Thanks

Do you have any questions?