DETECTING SUICIDE IDEATION IN ONLINE FORUMS USING NATURAL LANGUAGE PROCESSING.

PROJECT REPORT:

Group Name: The Mind Benders

Group Members:

- Esther Nyokabi
- Jane Kinuthia
- Joel Lenkinyei Tipape
- Karura Muthoni
- Keith Eugene Ochieng

INTRODUCTION:

This report discusses the development of an NLP model for accurately detecting suicide ideation in online conversations, particularly on Reddit. The aim is to aid governments and mental health organizations in identifying individuals expressing suicidal ideation online and providing timely support to prevent suicide attempts.

OBJECTIVES:

The main objective was to develop a machine learning model capable of accurately detecting posts related to depression and suicide in text, using posts from the "Suicide Watch" and "depression" subreddits. The steps taken include collecting and processing the data, developing baseline models using traditional machine learning algorithms, developing deep learning models, and evaluating and comparing the models.

METRIC OF SUCCESS:

The metric of success for this project is the accuracy score of the developed machine learning models in detecting posts related to depression and suicide in text, with a goal of achieving an accuracy score of 80%. Other metrics used include F1-score, precision, and recall.

DATA DESCRIPTION:

The dataset contains a collection of posts from Reddit, labeled as either "suicide" or "non-suicide". The purpose of this dataset is to develop a machine learning model that can accurately classify posts as either indicating suicidal ideation or not.

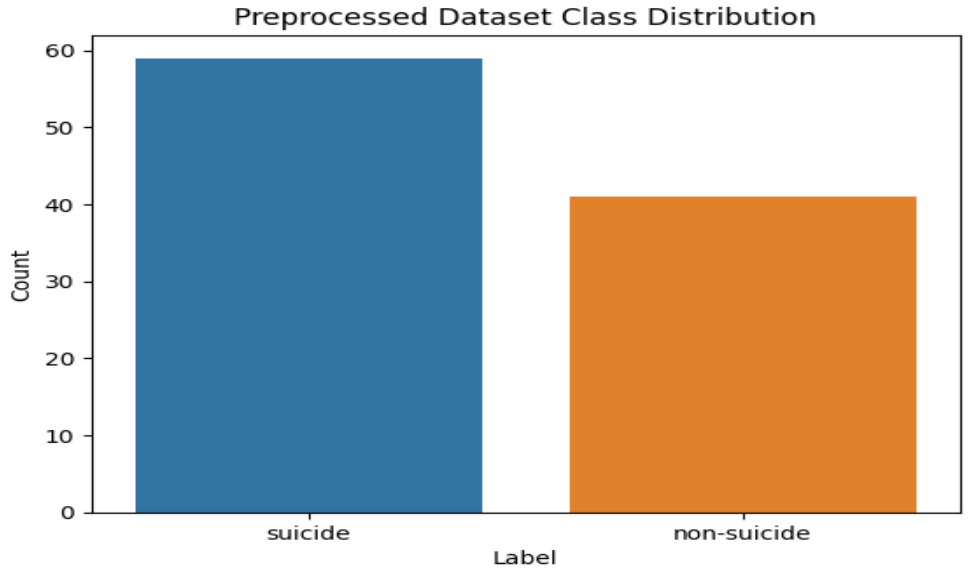
DATA PROCESSING:

Before building the model, the dataset was pre-processed through text cleaning, which involved removing bad symbols, unwanted lines, certain regex patterns, punctuations, and stop words, converting all text to lowercase, and finally tokenizing the text.

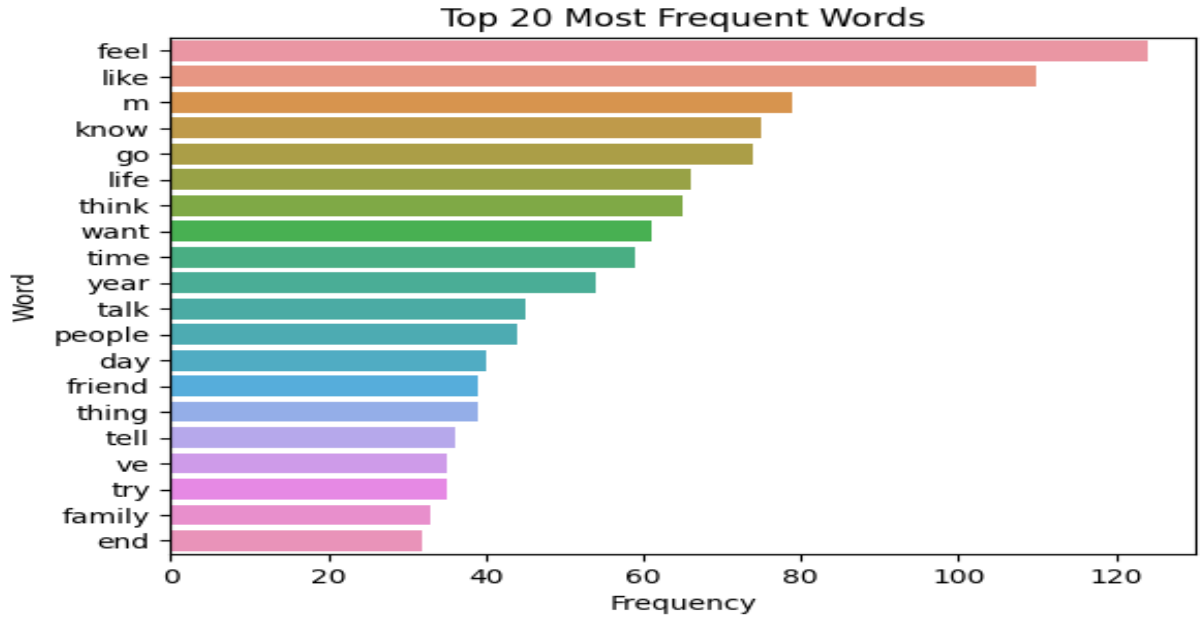
EXPLORATORY DATA ANALYSIS.

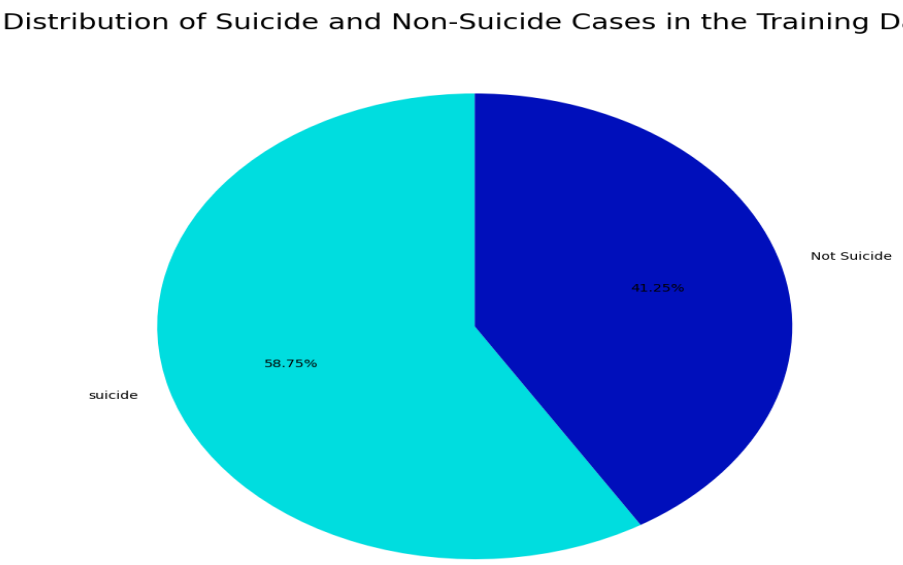
- The dataset contains 232,074 posts, evenly split between the two classes: 116,037 labeled as "suicide" and 116,037 labeled as "non-suicide".
- The average length of the posts in the dataset was 374.0 characters, with a standard deviation of 601.8 characters.
- The minimum length of a post was 20 characters, and the maximum length was 3,762 characters.

Distribution in classes of the preprocessed data.



Bar plot showing the 20 most frequently used words:





MODEL DEVELOPMENT:

The dataset was split into training and testing sets, with 80% of the data used for training and 20% for testing.

A supervised machine learning model was developed using Support Vector Machine (SVM). The model was trained and evaluated using the training and testing datasets, achieving an accuracy score of 75%. The model had a higher recall score for “non-suicide” at 80% and “suicide” at 73%. The F1-score for “non-suicide” was 62% while the F1-score for “suicide” was 81%, with a weighted average F1-score of 76%.

For the Deep learning model, Recurrent Neural Network (RNN) was used to improve the accuracy of the text classifier achieving an accuracy score of 80%. This suggests that it can be an effective tool for identifying individuals who may be at risk for suicide.

CONCLUSION:

- Our study aimed to detect suicide ideation in online forums using natural language processing techniques. The developed deep learning model, RNN, achieved an accuracy score of 80% and an F1-score of 0.81 for the "suicide" class, indicating its potential for accurately identifying posts related to suicidal ideation.
- Further exploration is necessary to identify the critical language feature or patterns that predict suicidal ideation in online forums.

- It is also essential to consider ethical concerns such as privacy, consent, and responsible use and interpretation of the results when using machine learning to analyze language related to suicide ideation.
- It should be emphasized that these models are not a replacement for mental health professionals, but rather a tool to aid in the identification of individuals who may be at risk for suicide.