CSE465
Section: 03

DEPRESSIVE TWEETS DETECTION USING DEEP LEARNING MODELS

Submitted by:

Rifat Azim #1912353642

MOTIVATION

- Social media is essential for instant communication and global connectivity, serving as a central hub for social interaction and information sharing in our daily lives.
- Along with sharing our life updates, pictures, memes, writing blogs, some people also share their mental struggles in a subtle way on their preferred social media platform.
- Early detection of people who are struggling with mental health through their social media posts can help them in many ways.
- Which brings us to this project where we are going to build a model that can extract depressive, anxious, or passive-suicidal posts from tons of social media posts and help the sufferers.

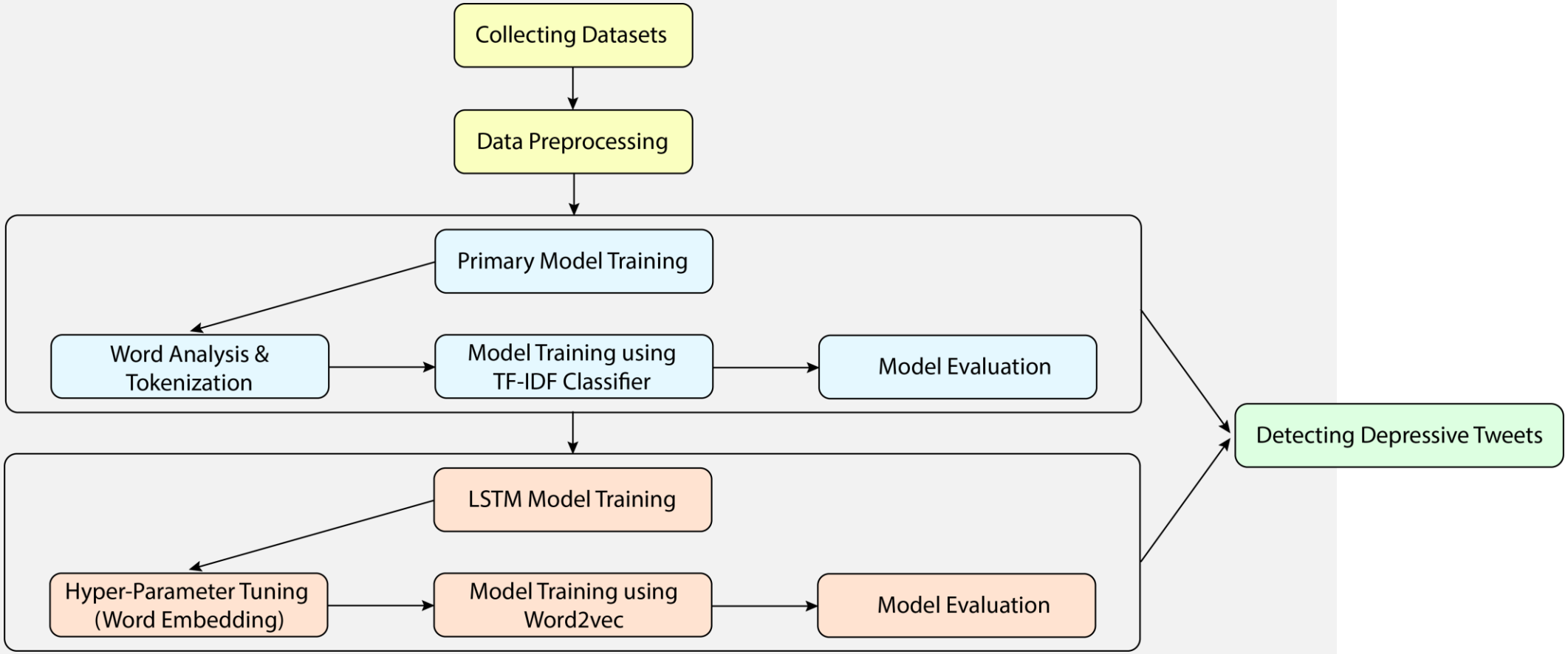
DATASET

For this project, we may need to use multiple datasets to classify depressive posts from normal posts.

1. For sentiment analysis, we're using Sentiment140 dataset which is commonly used amongst the machine learning engineers. (https://www.kaggle.com/datasets/kazanova/sentiment140)
2. Since there is no public dataset available regarding depressive posts, so primarily, we're using a twitter dataset available on github, which is scrapped by a company named TWINT. (https://github.com/eddieir/Depression_detection_using_Twitter_post)
3. To improve generalization, address ambiguity and enhance model performance, we're also using English-contraction dataset. (https://www.kaggle.com/datasets/yetman/english-contractions)

SYSTEM DESIGN

The core my project is based on these ideas:



ACCURACY SCORE

- F1 Score of TF-IDF Classifier

	precision	recall	f1-score
0	0.99553	1.00000	0.99776
1	1.00000	0.98323	0.99154
accuracy			0.99646
macro avg	0.99776	0.99162	0.99465
weighted avg	0.99647	0.99646	0.99645

- F1 Score of LSTM

	precision	recall	f1-score
0	0.99414	0.99456	0.99435
1	0.98179	0.98042	0.98111
accuracy			0.99130
macro avg	0.98797	0.98749	0.98773
weighted avg	0.99130	0.99130	0.99130

ACCURACY SCORE

- Using Naïve Baye's:

	precision	recall	f1-score
0	0.88731	0.99832	0.93955
1	0.99057	0.58172	0.73298
accuracy			0.90142
macro avg	0.93894	0.79002	0.83627
weighted avg	0.91133	0.90142	0.89150

- Using Logistic Regression:

	precision	recall	f1-score
0	0.98512	0.97271	0.97888
1	0.91356	0.95152	0.93216
accuracy			0.96778
macro avg	0.94934	0.96212	0.95552
weighted avg	0.96848	0.96778	0.96801

FUTURE IMPROVEMENT

- We need to use a larger and more diverse dataset that includes multiple languages to improve model generalizability globally.
- We are looking forward to develop the system to perform real-time sentiment analysis, making it suitable for monitoring on-going social trends and help the sufferers.
- Although we already have a very good result, we still need to use more powerful models for bigger datasets in future.
- Collaborate with social media platforms to integrate this model as a tool for identifying depressive patterns among users.
- Facilitate timely interventions by providing users with mental health resources or connecting them with professionals.
- Lastly, we need to use this technology to raise awareness and support mental health initiatives, ultimately benefiting users on a global scale.

THANK YOU