Project Deliverables

Multi-class classification of Mental Health Disorders:

Mentor TA - Kawshik Manikantan

Problem Statement

This project deals with the classification of mental health disorders from the Reddit mental health dataset. This project is designed to help you learn to differentiate and appreciate models belonging to a varied spectrum as well as solve a critical and rising problem in society.

Requirements

- Implement the following models from scratch and improve their performance to the maximum extent possible:
  – Random Forests
  – SVM
  – CNN
  – RNN
  – distilBERT (only finetune)
- Provide an intuitive and explanatory analysis of the results of the model. It is also required to make an effort to create explainable models.
- You are required to provide an analytical pipeline to differentiate and explain the weakness of each model.

Checkpoint 1

7 October 2023

This checkpoint will focus on understanding the required deliverables for the project and understanding the resources and expected approaches.

Deliverables -

- Setting up the Project Log document.
- Understanding given requirements.
- Pushing the project timeline and deliverables document.

Checkpoint 2

30 October 2023

- Cleaning and preprocessing the dataset.
- Performing exploratory analysis on the data.
- Implementing SVM and training it on the dataset.
- Training Random Forest Classifier on the dataset.
- Improving and analysing the outcomes for the trained models.
- Making the models explainable and analysing weaknesses.
- Setting up basic frameworks for CNN and RNN to work upon in checkpoint 3.

Checkpoint 3

27 November 2023

- Completing RNN and CNN models and analysing them.

- Improving their performance and making them explainable.
- Finetuning distilBERT on the dataset and analysing it.
- Comparative analysis of the trained models on varied parameters with an analytical pipeline.
- Exploring other possible options for improving performance and comparing it with the obtained results.
- Preparing a conclusive report.

Final Deliverables List

- Setting up the Project Log document.
- Understanding given requirements.
- Pushing the project timeline and deliverables document.
- Cleaning and preprocessing the dataset.
- Performing exploratory analysis on the data.
- Implementing SVM and training it on the dataset.
- Training Random Forest Classifier on the dataset.
- Improving and analysing the outcomes for the trained models.
- Making the models explainable and analysing weaknesses.
- Setting up basic frameworks for CNN and RNN to work upon in checkpoint 3.
- Completing RNN and CNN models and analysing them.
- Improving their performance and making them explainable.
- Finetuning distilBERT on the dataset and analysing it.
- Comparative analysis of the trained models on varied parameters with an analytical pipeline.
- Exploring other possible options for improving performance and comparing it with the obtained results.

- Preparing a conclusive report.