Raghav Marwaha

E23CSEU1229
Batch 41

Lab Activity 6: Mental Distress Detector

1. Objective

The aim of this lab is to build a natural language processing model capable of identifying emotional distress in social media text. The goal is to detect negative emotional patterns early and enable mental health support systems to intervene in time.

This aligns with **SDG 3 – Good Health and Well-Being**, emphasizing mental health awareness and early intervention.

---

2. Dataset Information

Dataset link:- https://huggingface.co/datasets/dair-ai/emotion

Detail	Description
Dataset Used	**dair-ai/emotion** (Hugging Face)
Data Type	Social media text samples
Emotion Labels	Joy, Sadness, Anger, Fear, Love, Surprise
Purpose	Emotion classification focused on detecting negative affect linked to distress

For this task, emphasis was placed primarily on identifying distress-related emotions such as **sadness**, **fear**, and **anger**, as these are closely associated with stress, anxiety, and depression.

Data preprocessing included:

- Text normalization
- Tokenization
- Removal of noise (unnecessary symbols, repeated characters)

---

3. Methodology

NLP Processing and Embedding

- Text converted into embedding vectors using pre-trained language models

Classification Model

A **Hybrid Architecture** was used:

Component	Function
BERT Embeddings	Extract deep contextual emotion meaning
1D-CNN	Identify strong emotional phrase patterns
BiLSTM	Learn sequential relationships and emotional flow in sentences

---

4. Training Setup

- Loss function: Cross-entropy
- Optimizer: Adam
- Training monitored using validation performance
- Early stopping and checkpointing applied to prevent overfitting

---

5. Performance Results

(Extracted from the `.ipynb` file)

Metric	FinalSc
Accuracy	94.30%
F1-Score	94.37%
ROC-AUC	98.63%
Validation Loss	0.0789