Natural Language Processing

Project: DepressionDetect

<u>NLP-Project-Module-2</u>

Submitted By (Group):

Hammad Abrar (22-SE-65)

Muhammad Asif (22-SE-21)

Muhammad Haseeb (22-SE-8)

Haziq Asif (22-SE-102)

Submitted To:

Dr. Kanwal Yousaf

Date:

05-05-2025

Methodology

This study employed two distinct deep learning approaches for the detection of multiple psychological emotions in social media text: a BERT-based transformer model and a Long Short-Term Memory (LSTM) network. Both methods were applied to the same dataset, which consisted of labelled textual data with eight emotion categories: *anger, brain dysfunction (forget), emptiness, hopelessness, loneliness, sadness, suicide intent,* and *worthlessness*.

Data Preprocessing

Data preprocessing is a crucial step in preparing raw text and label data for effective training and evaluation of machine learning models. In this project, we used two different approaches—BERT embeddings with a classical model and LSTM-based sequence modeling. Each approach required tailored preprocessing techniques as outlined below:

Data Description

The dataset used in this project is a multi-label emotional classification dataset focused on mental health-related text data. Each text sample represents an individual's written expression, which is analysed to identify the presence of specific emotional states. The dataset aims to support early detection of mental health symptoms and emotional disorders using machine learning models.

Preprocessing for BERT Model

For the BERT-based model:

- **Embedding Generation**: Precomputed BERT embeddings were used, each converted to a vector of shape (samples, features).
- **Data Splitting**: Data was divided into training and testing sets, with corresponding labels for multi-label classification.
- **No Tokenization Needed**: Since BERT embeddings were precomputed, tokenization and lowercasing were handled prior to embedding generation.

Preprocessing for LSTM Model

For the LSTM model:

- **Temporal Format**: The BERT embeddings were reshaped to fit LSTM's expected input shape: (samples, time_steps=1, features).
- **Multi-label Output**: The target labels were extracted from the dataset as an 8-dimensional binary vector for each sample.
- **Normalization**: As BERT embeddings are already normalized vector representations, no further normalization was required.
- **No Padding Needed**: Since the embeddings had uniform size, padding for sequences was not necessary.

BERT-Based Model

The BERT (Bidirectional Encoder Representations from Transformers) model was fine-tuned on emotion-labelled data. Textual inputs were tokenized using BERT's tokenizer and converted into embeddings. A multi-label classification head (dense layer with sigmoid activation) was added to the pre-trained BERT model. The model was trained using binary cross-entropy loss and Adam optimizer.

LSTM-Based Model

The second approach utilized precomputed embeddings as input to an LSTM-based model. The data was reshaped to fit the format required for sequential modelling. The model architecture consisted of an LSTM layer with 64 units, followed by a dense layer with 32 neurons (ReLU activation), and a final dense layer with 8 neurons using sigmoid activation, representing each emotion class. The model was compiled using binary cross-entropy loss and trained for 60 epochs with a batch size of 32.

Results and Discussion

The performance of both models was evaluated using precision, recall, and F1-score metrics for each emotion class. These metrics were aggregated using micro, macro, weighted, and sample-based averaging techniques to account for multi-label classification.

BERT-Based Model Performance

Class	Precision	Recall	F1-score	Support
0	0.73	0.72	0.72	346
1	0.51	0.43	0.47	164
2	0.74	0.76	0.75	358
3	0.82	0.89	0.85	640
4	0.80	0.77	0.78	453
5	0.85	0.95	0.90	708
6	0.70	0.72	0.71	224
7	0.74	0.80	0.77	462

Averages:

Metric	Precision	Recall	F1-score	Support
Micro Average	0.78	0.81	0.79	3355
Macro Average	0.74	0.75	0.74	3355
Weighted Average	0.77	0.81	0.79	3355
Samples Average	0.76	0.80	0.74	3355

Overall Averages:

- Micro Avg F1-score: 0.92
- Macro Avg F1-score: 0.90
- Weighted Avg F1-score: 0.92
- Samples Avg F1-score: 0.89

LSTM-Based Model Performance

Emotion Label	Precision	Recall	F1-score	Support
Anger	0.85	0.87	0.86	1754
Brain Dysfunction (Forget)	0.82	0.53	0.64	813
Emptiness	0.91	0.77	0.83	1573
Hopelessness	0.92	0.92	0.92	2919
Loneliness	0.90	0.89	0.89	1929
Sadness	0.91	0.96	0.93	3260
Suicide Intent	0.85	0.83	0.84	1035
Worthlessness	0.88	0.85	0.86	2095

Overall Averages:

- Micro Avg F1-score: 0.88
- Macro Avg F1-score: 0.85
- Weighted Avg F1-score: 0.88

- Samples Avg F1-score: 0.84

Comparative Analysis

The BERT-based model outperformed the LSTM model across all key evaluation metrics. Notably, the BERT model achieved a higher F1-score in critical emotion categories such as *brain dysfunction (forget)* and *emptiness*, where LSTM performance was relatively lower. The transformer architecture's ability to capture contextual relationships within the text appears to contribute significantly to its superior accuracy, particularly in nuanced or overlapping emotional states.

Recommendations

Based on the experimental results and comparative analysis, the following recommendations are proposed:

1. **Use Transformer-Based Models for Emotion Detection**: Given the consistent performance gains across all emotion categories, transformer models like BERT are recommended for multi-label emotion classification tasks on textual data.
2. **Handle Class Imbalance**: Emotions like *brain dysfunction (forget)* showed lower recall in both models, particularly in the LSTM model. Future work should incorporate techniques such as class-weighting or data augmentation to address this imbalance.
3. **Expand Dataset Size and Diversity**: Increasing the diversity and size of the training data could further enhance model performance, especially in capturing subtle distinctions between similar emotions.
4. **Deploy Lightweight Alternatives for Efficiency**: While BERT performs best, it is computationally expensive. For applications with limited resources, distilled transformer models (like DistilBERT) or optimized LSTM variants may be considered.
5. **Incorporate Multimodal Signals**: For more robust emotion detection, combining text data with voice or facial expression inputs can be explored in future research.