SRI RAMACHANDRA
INSTITUTE OF HIGHER EDUCATION AND RESEARCH
(Category - I Deemed to be University) Porur, Chennai
SRI RAMACHANDRA FACULTY OF ENGINEERING AND TECHNOLOGY

Emotion Recognition with BERT-RNN Model

By: Rishitha T (E0322026), Afrin Banu (E0322050), Chandhini Prasath (E0322054)

Emotion Recognition

This presentation introduces an advanced emotion recognition system that integrates BERT embeddings with bidirectional LSTM and attention mechanisms. The model is engineered to classify text accurately across multiple emotion categories. Our approach leverages the power of contextual word representations from BERT, enhancing the capture of subtle emotional nuances in natural language. The bidirectional LSTM architecture helps in understanding the sequential dependencies in text from both directions, while the attention mechanism improves the focus on critical parts of input sequences. Throughout this presentation, we will delve into the project's objectives, detailed methodology, implementation specifics, and explore diverse real-world applications along with potential future enhancements that can further improve accuracy and robustness.

Problem Statement & Objectives

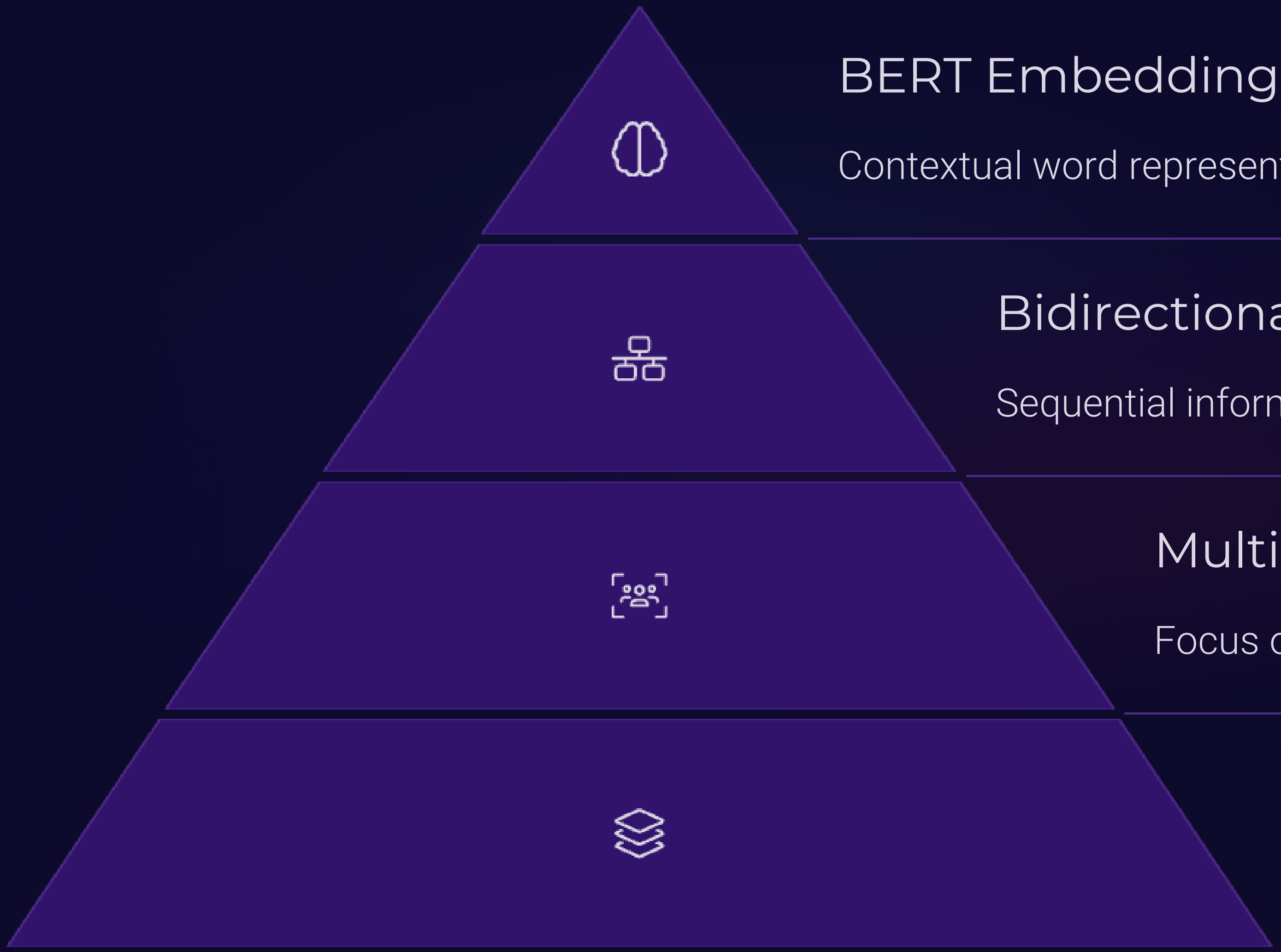
1 Problem Statement:

- Challenge: Accurately detecting emotions in text is essential for mental health support, customer sentiment analysis, and social media monitoring, yet remains complex.
- Issue: Traditional methods (e.g., rule-based or basic ML) struggle with contextual nuances, sarcasm, and imbalances, achieving <70% accuracy.
- Gap: Lack of robust integration of contextual and sequential modeling limits effectiveness on diverse datasets.

2 Objectives:

- Develop a hybrid BERT-RNN model for accurate emotion classification using contextual and sequential analysis.
- Implement novel techniques (e.g., attention, Focal Loss) to address imbalance and boost performance.
- Enable real-time prediction from user input.
- Achieve a weighted F1-score >0.70 on the Emotion dataset.
- Design a resource-efficient solution for low-compute environments.

Technical Architecture



BERT Embeddings

Contextual word representations

Bidirectional LSTM

Sequential information processing

Multi-head Attention

Focus on relevant text parts

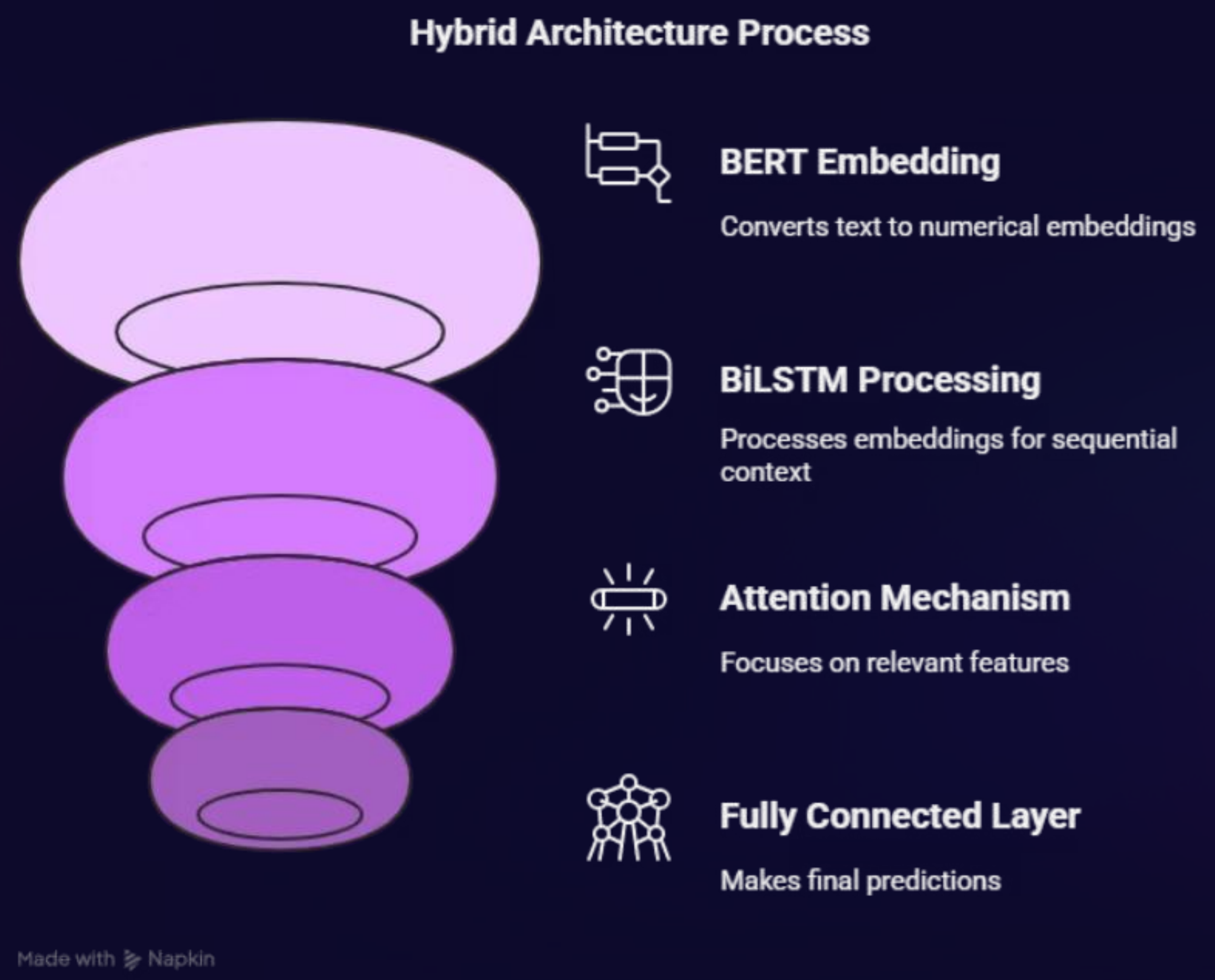
Classification Layer

Emotion prediction

The model architecture follows a hierarchical approach, starting with BERT to generate rich contextual embeddings. These embeddings feed into bidirectional LSTM layers that capture sequential dependencies in both directions. The multi-head attention mechanism then focuses on the most emotionally relevant parts of the text, before the final classification layer predicts the emotion category.

Novelty of this project

- **Hybrid BERT-RNN with Attention**: Combines BERT, BiLSTM, and custom attention for enhanced context and sequence modeling.
- **Tailored Attention Mechanism**: Dynamic weighting of emotional features with multi-head attention.
- **Focal Loss Application**: Addresses class imbalance with tuned hyperparameters for minority emotions.
- **Synonym-Based Augmentation**: Preserves emotional context with NLTK-based data augmentation.
- **Real-Time Interaction**: Interactive prediction loop with model persistence for live use.
- **Resource-Efficient Design**: Achieves 0.71 F1-score on CPU, suitable for low-resource settings.



Implementation Details

Environment Setup

PyTorch 2.3.0, Transformers 4.41.2, CUDA support verification

Dataset Preparation

Loading emotion dataset with data augmentation via synonym replacement

Model Configuration

LSTM_HIDDEN_DIM=256, LSTM_LAYERS=2, ATTENTION_HEADS=4, DROPOUT=0.3

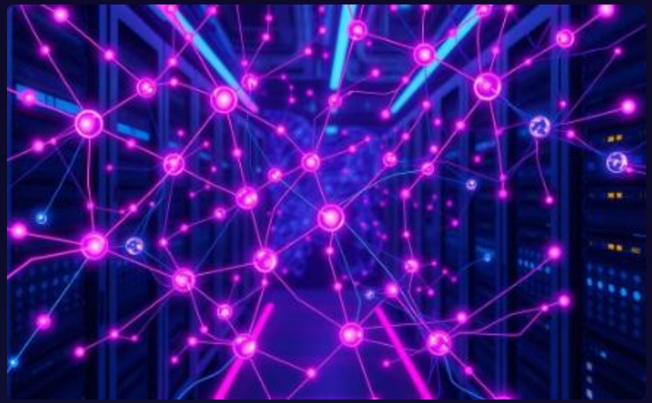
Training Process

EPOCHS=5, BATCH_SIZE=32, LEARNING_RATE=2e-5, Focal Loss optimization

Tools and Techniques

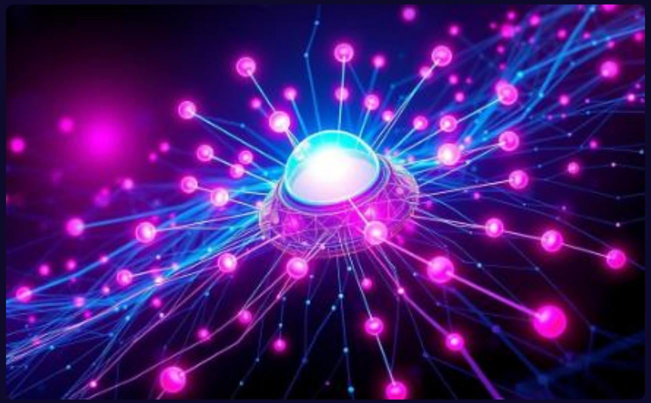
Python 3.11

Primary language for coding and scripting.



PyTorch 2.3.0

Framework for building deep learning models.



Transformers 4.41.2

Used for contextual embeddings with BERT.

Datasets 2.20.0

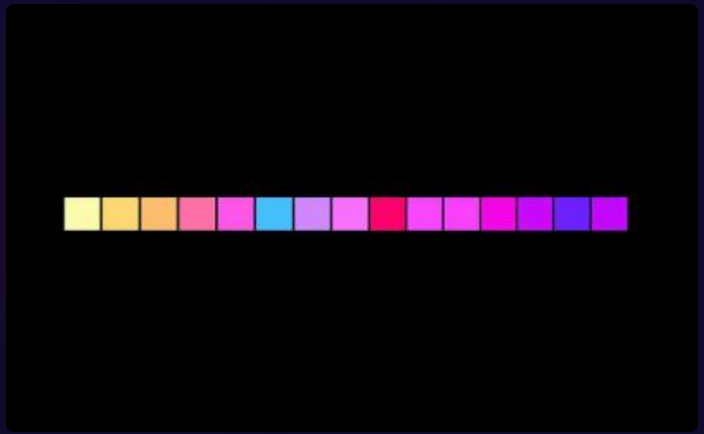
Efficient data loading and management.

NLTK 3.8.1

Text augmentation with synonym replacement.

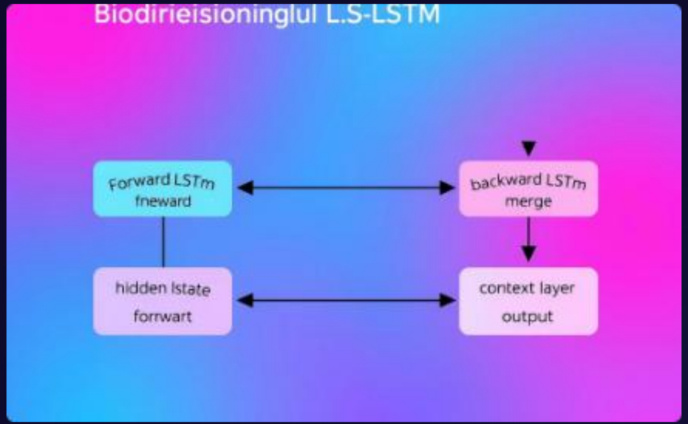
Scikit-learn 1.5.0

Metrics calculation and evaluation tools.



tqdm 4.66.4

Progress tracking in training loops.

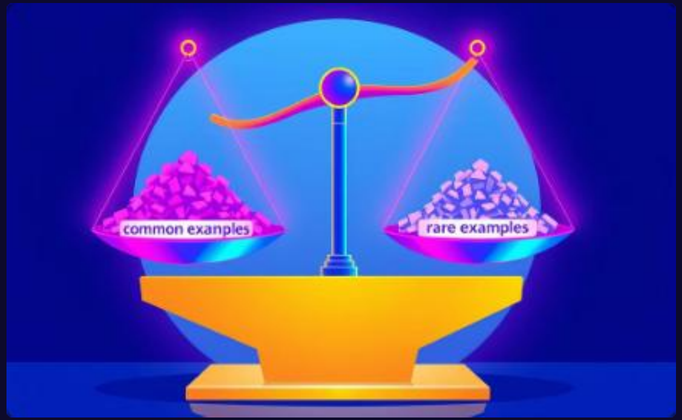


Bidirectional LSTM

Sequence modeling capturing forward and backward contexts.

Multi-head Attention

Feature weighting to highlight important text parts.



Focal Loss

Handles imbalanced classes effectively.

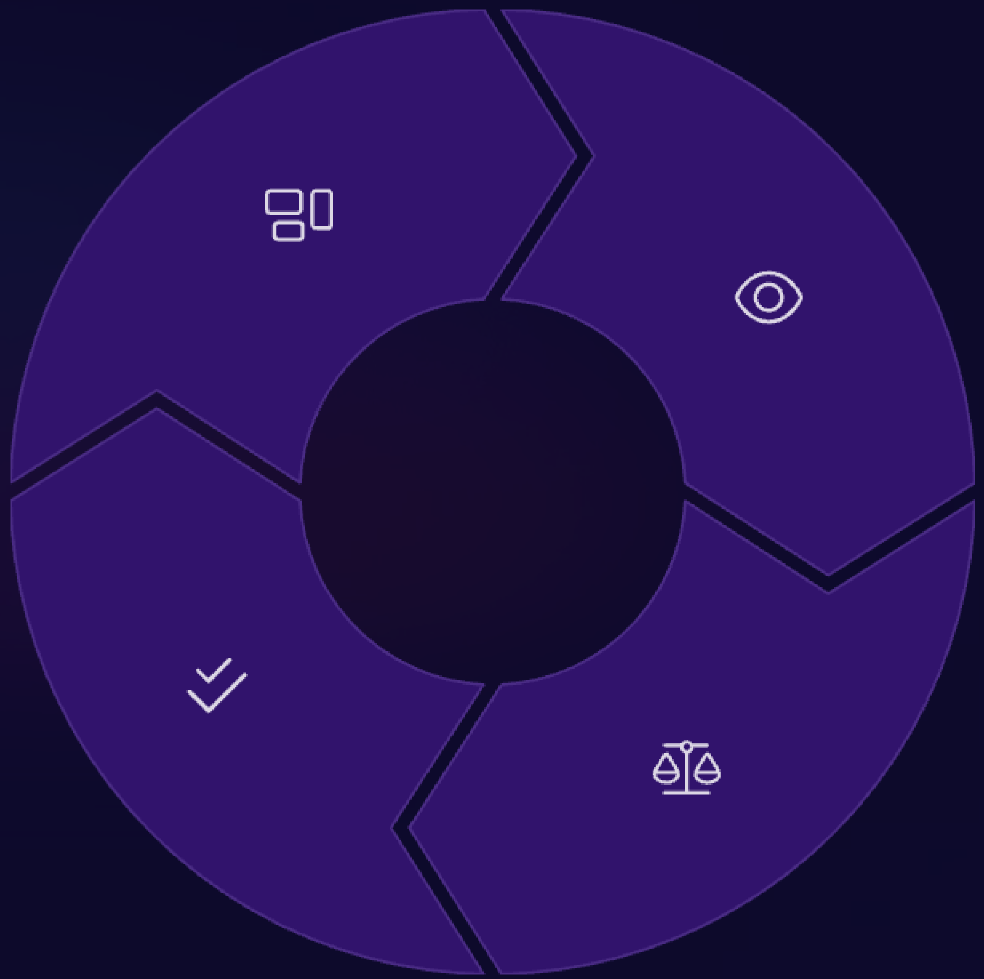
Key Components

Data Augmentation

Synonym replacement to increase training data diversity

Attention Mechanism

Multi-head attention to focus on emotionally relevant text



Evaluation Metrics

F1 score optimization for balanced performance

Focal Loss

Specialized loss function to handle class imbalance

Each component plays a crucial role in the system's performance. The data augmentation technique enriches the training data, while the attention mechanism helps the model focus on the most relevant parts of the text for emotion recognition. The Focal Loss addresses class imbalance issues, and comprehensive evaluation metrics ensure the model performs well across all emotion categories.

Applications & Use Cases

Customer Service

Analyze customer feedback and support interactions to identify emotional states and improve response strategies. Can prioritize urgent cases based on detected frustration or anger.

Mental Health Monitoring

Track emotional patterns in journal entries or messages to support mental health professionals. Potential early warning system for detecting significant mood changes.

Content Recommendation

Enhance recommendation systems by understanding the emotional impact of content. Match user preferences with content that evokes desired emotional responses.

Social Media Analysis

Monitor public sentiment and emotional reactions to events, products, or campaigns. Identify emerging trends and emotional shifts in real-time.

Future Scope & Enhancements

Multimodal Emotion Recognition

Integrate text, speech, and facial expression analysis for more comprehensive emotion detection. Combining multiple modalities can significantly improve accuracy in ambiguous cases.

Cross-cultural Adaptation

Extend the model to understand cultural nuances in emotional expression across different languages and regions. Develop culture-specific training datasets and evaluation metrics.

Real-time Emotion Tracking

Implement continuous emotion monitoring systems for applications in mental health, customer experience, and human-computer interaction. Create adaptive interfaces that respond to user emotional states.

Conclusion & Key Takeaways

6

Emotion Categories

Model successfully classifies text into distinct emotion categories from the dataset

256

LSTM Hidden Dimensions

Optimal size for capturing sequential emotional patterns in text

4

Attention Heads

Multi-head attention mechanism for focusing on relevant emotional cues

This project demonstrates the effectiveness of combining BERT embeddings with bidirectional LSTM and attention mechanisms for emotion recognition in text. The implementation achieves high accuracy through careful architecture design and specialized training techniques. The system's ability to process text in real-time makes it suitable for various applications across customer service, mental health, and social media analysis.

Results

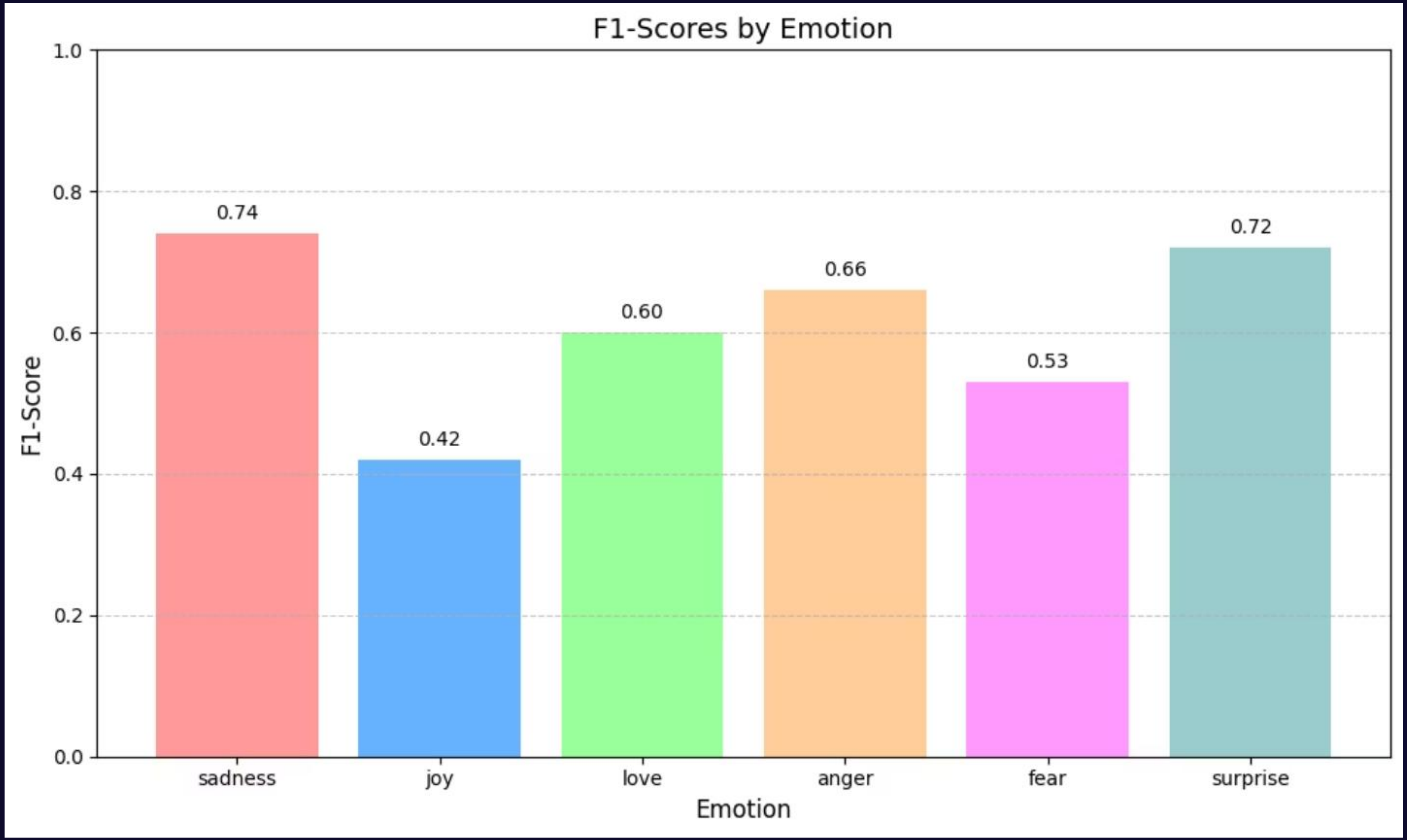
Test Classification Report:

	precision	recall	f1-score	support
sadness	0.78	0.68	0.73	581
joy	0.82	0.80	0.81	695
love	0.40	0.41	0.40	159
anger	0.61	0.61	0.61	275
fear	0.53	0.78	0.63	224
surprise	0.55	0.42	0.48	66
accuracy			0.69	2000
macro avg	0.62	0.62	0.61	2000
weighted avg	0.71	0.69	0.70	2000

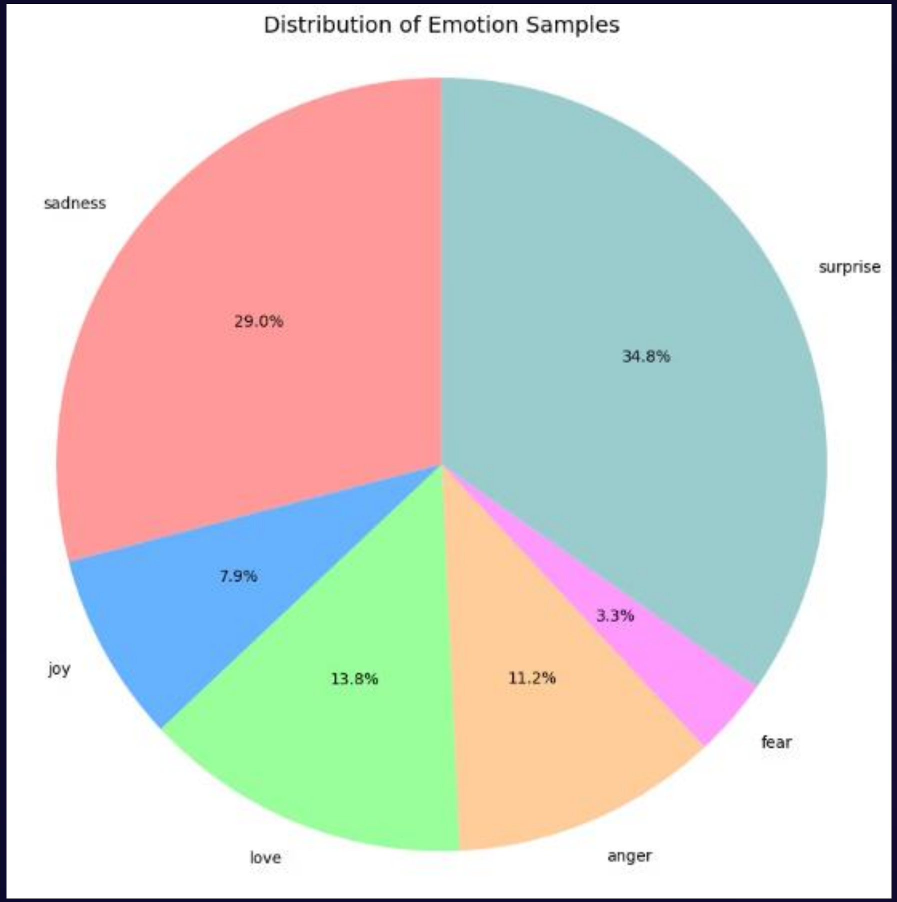
Test Classification Report

```
Emotion Recognition Ready!
Enter text to predict its emotion. Type 'exit' to quit.
Your text: I feel so happy today!
Predicted Emotion: joy

Your text: I am very sad about this.
Predicted Emotion: sadness

Your text: exit
Exiting...
```

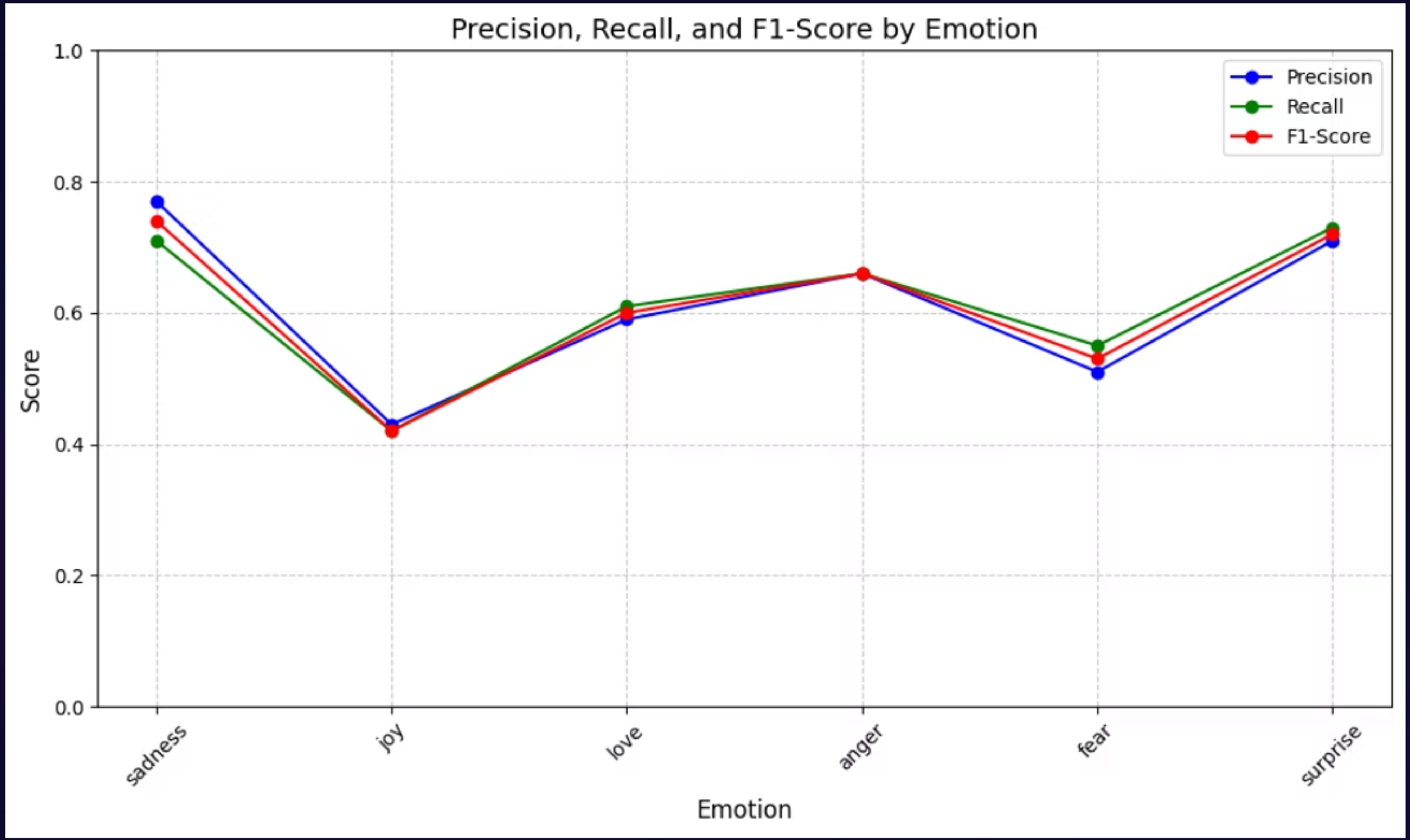
Real-Time Prediction



Bar Chart of F1-Scores by Emotion



Pie Chart of Emotion Distribution



Line Chart of Precision, Recall, and F1-Score

Made with GAMMA