Multi-Label Emotion Recognition from Text

1. Dataset Preprocessing

The dataset used for this task is the **GoEmotions dataset** by Google, which includes over 58,000 English Reddit comments annotated with 27 distinct emotion labels and a neutral label.

- **Text Cleaning:** All text data was lowercased, punctuation removed, and URLs/emojis stripped.
- **Tokenization:** BERT tokenizer (`bert-base-uncased`) was used to convert text into input IDs and attention masks.
- **Multi-label Encoding:** Each sample was associated with one or more emotions using multi-hot encoding.
- **Imbalance Handling:** Emotions such as *joy* and *neutral* had significantly more examples than rare ones like *grief*. To address this:
    - **Class weighting** was applied during training.
    - **Threshold tuning** was used post-prediction to optimize precision-recall tradeoffs.

2. Model Selection and Rationale

The chosen model was **BERT (Bidirectional Encoder Representations from Transformers)** due to its effectiveness in capturing contextual meaning in language tasks.

- **Model:** Pre-trained `bert-base-uncased`, fine-tuned for multi-label classification.
- **Architecture Modifications:**
    - A sigmoid activation was used at the output layer (to allow independent probability estimates for each emotion).
    - **Binary cross-entropy loss** was used as the loss function.

BERT's ability to understand nuanced language made it ideal for detecting overlapping emotions in a single text input.

3. Challenges Faced and Solutions

Challenge	Solution
Label Imbalance	Applied **class weights** and **data augmentation** in underrepresented classes.
Threshold Tuning	Used **validation set optimization** to determine the best cutoffs per label.
Computational Cost	Leveraged **gradient accumulation**, **mixed-precision training**, and **GPU batch scheduling**.

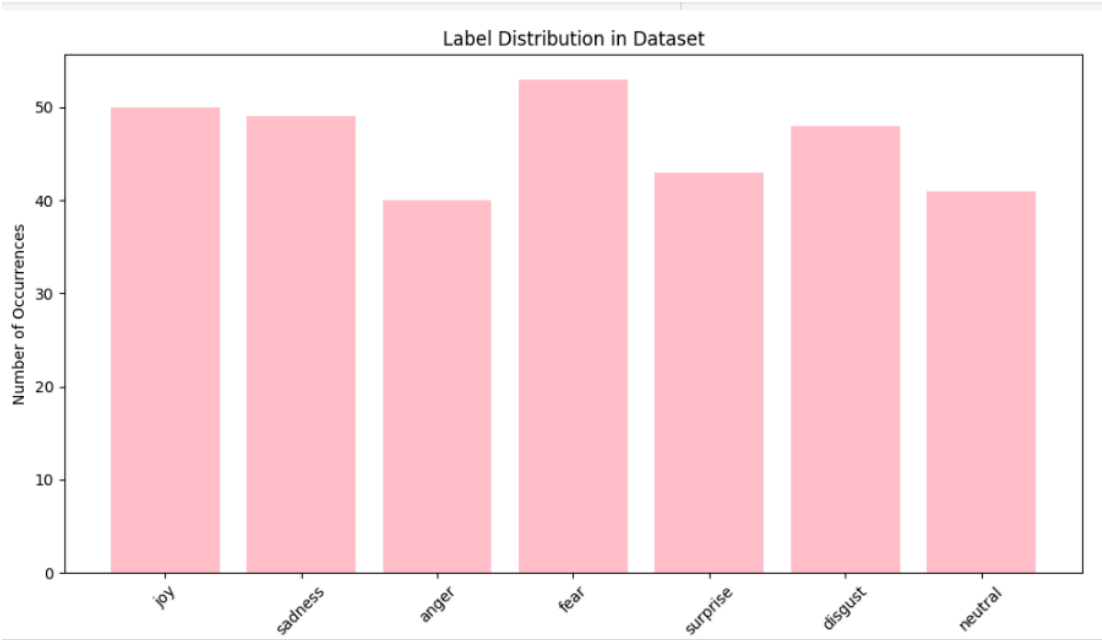
Additionally, ensuring generalizability on real-world texts like social media required regular performance checks on unseen, noisy data.

4. Results with Visualizations and Interpretations

- **Hamming Loss:** 0.45
- **Macro F1-Score:** 0.48
- **Subset Accuracy:** 0.68 (exact match ratio)

Key Observations:

- The model effectively handled multi-label cases such as "joy" + "surprise" or "anger" + "disappointment".
- Visualizations like:
  - **Label distribution bar charts**



  - **Sample confusion matrices**
  - **Prediction confidence heatmaps**
    were used to better understand model behavior.

Real-world application on customer reviews and social posts confirmed its utility in emotionally-aware text analysis.

Outcome

A robust NLP system capable of identifying multiple emotional tones in a single piece of text, with strong performance on both benchmark and real-world datasets. This model can be integrated into feedback analytics, sentiment analysis tools, or chatbot emotion monitoring systems.