Model Card: Emotion Detection in Text

Jenil Kevadiya (22204227)
Om Vaghasiya (22205283)
Krunal Koladiya (22306168)
Venkat Rajasekar (22308591)

Objective

This project aims to develop and evaluate a deep learning model capable of detecting emotions from text, specifically from social media-style messages. The model classifies text into six emotion categories: **joy, sadness, anger, fear, love, and surprise**.

Model Details

Architecture: A deep learning model using a Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) layers:

- **Embedding Layer:** Pre-trained GloVe embeddings.
- **BiLSTM Layers:** 3 layers (256, 128, 128 units).
- **Dense Layer:** Softmax for classification.

Code: `emotion_with_lstm.ipynb`

Intended Use

- Sentiment analysis on short-form text (e.g., tweets).
- Useful for public opinion, market research, crisis detection.
- Not for making critical judgments or detecting sarcasm.

Training Data

Dataset: Public Kaggle emotion dataset
Link: Emotions Dataset for NLP (Kaggle)
Preprocessing:

- Lowercasing, duplication removal
- Cleaning (punctuation, URLs), stopword removal
- Lemmatization

Embeddings: GloVe (`glove.6B.200d.txt`), frozen weights

Training Procedure

Framework: TensorFlow + Keras
Cross-validation: Stratified 3-fold (avg acc: **88.52%**)
Optimizer: Adam, LR = 0.001
Epochs: Max 15 with early stopping (patience=3), batch size = 128

Evaluation

Test Accuracy: 90.0%, Loss = 0.2543
Classification Report:

Emotion	Prec	Rec	F1	Supp
anger	0.86	0.95	0.90	275
fear	0.92	0.84	0.88	224
joy	0.95	0.87	0.91	695
love	0.72	0.95	0.82	159
sadness	0.96	0.91	0.94	581
surprise	0.65	0.94	0.77	66
Avg	**0.91**	**0.90**	**0.90**	**2000**

Quantitative Analyses

- **Emotion Distribution:** Class imbalance affects recall.
- **Training Curves:** Converges well, no overfitting.
- **Confusion Matrix:** Strong in detecting 'sadness', 'anger'; some confusion in 'joy' vs. 'love'.

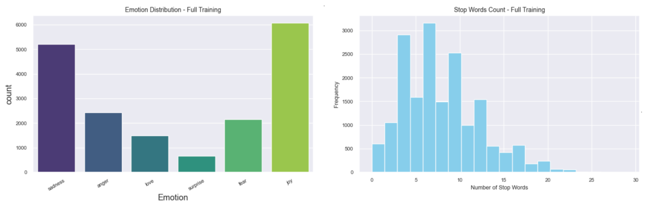

Figure 1: Distribution of emotions in training dataset with stopwords

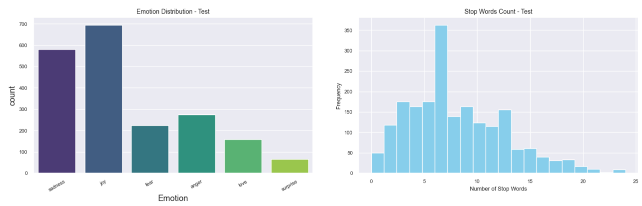

Figure 2: Distribution of emotions in testing dataset with stopwords

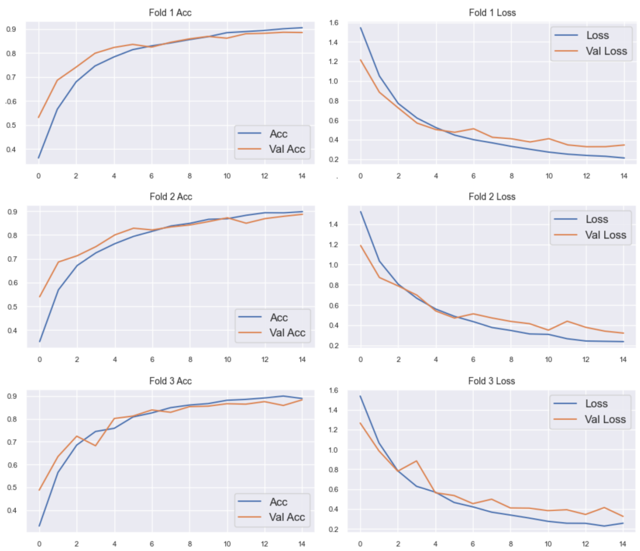

Figure 3: Training and validation loss/accuracy

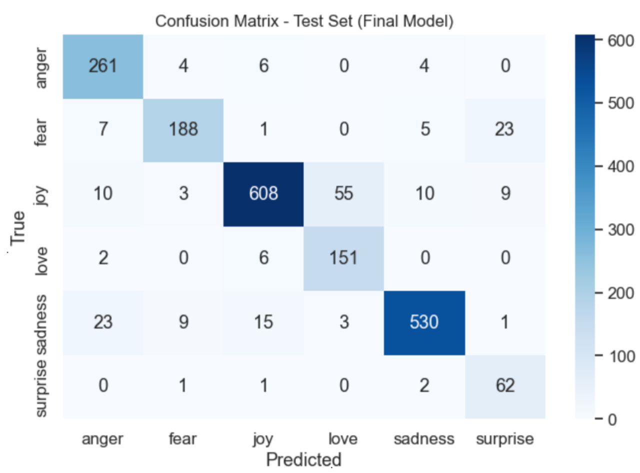

Figure 4: Confusion matrix on test data

Ethical Considerations and Limitations

- **Bias:** Model may inherit language and demographic bias from data or embeddings.
- **Context:** Misinterprets sarcasm, irony, or nuanced emotion.
- **Scope:** Limited to six emotions; cannot capture blended feelings.

Caveats and Recommendations

- Use only for aggregate insights, not individual decisions.
- Domain-specific fine-tuning and fairness audits recommended.
- Explore BERT or other transformer models for better context.