

# EMOTION RECOGNITION IN CONVERSATION

# MOTIVATION

Text contains emotional signals beyond literal meaning.

AI systems must detect fine emotions to improve:

- Chatbots & virtual assistants
- Counseling & mental health tools
- Customer support automation
- Social sentiment tracking

Existing models identify only basic emotions (6–8).

Our goal: detect 28 real emotions using advanced Transformer models.

# PROBLEM STATEMENT

- Most models classify only single-label emotions
- Many fail on multi-emotion sentences
- Weak contextual reasoning
- Cannot detect subtle emotions:

“I’m excited but nervous” → excitement + nervousness

# RELATED WORK

- Traditional Methods
  - Bag-of-Words, TF-IDF
  - SVM / Logistic Regression
  - Limited emotional depth
- Deep Learning
  - LSTM, BiLSTM
  - Better but still limited categories
- Transformers
  - BERT, DistilBERT, RoBERTa
  - Google GoEmotions dataset enabling 28 emotion labels
- Our model builds on these advances

# DATASET

## (GOEMOTIONS)

- 58,000 Reddit comments
- 28 emotion labels + Neutral
- Each text may have 1–3 labels (multi-label)
- Balanced across positive, negative & ambiguous emotions

Examples:

- “Thanks a lot!” → gratitude
- “I hate this.” → anger, disgust

Used for training + validation.

# METHODOLOGY

## Pipeline Overview

Text → Tokenization → Transformer Encoder → Dense Layer → Sigmoid → Emotion Scores

## Steps

1. Tokenize text (BPE tokenizer)
2. Pass through DistilRoBERTa / RoBERTa layers
3. Model outputs 28 independent probabilities
4. Sigmoid allows multiple emotions simultaneously
5. Threshold = 0.5 for label activation

# MODEL ARCHITECTURE

Base Model: DistilRoBERTa (fast training)

- 82M parameters
- Fine-tuned on GoEmotions
- Multi-label classification using Sigmoid + BCE Loss

Training Setup

- Batch size: 8
- Epochs: 1 (for fast GPU training)
- Learning rate:  $2e^{-5}$
- Used NVIDIA RTX 3050 GPU

# MODEL ARCHITECTURE

Improvement Over Prior Approach :

Traditional SVM-based emotion classifiers could only recognize about 6 to 8 emotions, did not support multi-label output, and generally showed low performance.

Our earlier model, based on DistilBERT, expanded the range to about 10 to 15 emotions, but it still could not predict multiple emotions at once and achieved only moderate accuracy.

The new model we developed using DistilRoBERTa is a major improvement.

It can detect all 28 emotion categories, fully supports multi-label prediction, and achieves high accuracy, outperforming the previous approaches in both depth and reliability.

# EXPERIMENTS & EVALUATION

## Evaluation Metrics

- Exact Match Accuracy
- Macro F1
- Micro F1
- BCE Loss

## Model Performance Comparison

The traditional TF-IDF + SVM approach achieved an accuracy of about 63%, with a macro F1 score of 0.56 and a micro F1 score of 0.59, showing limited capability in handling emotional nuances.

The DistilBERT-based model improved performance to 74% accuracy, with a macro F1 of 0.67 and a micro F1 of 0.71, offering better contextual understanding but still lacking in fine-grained emotional detection.

Our improved model, DistilRoBERTa, delivers the best results, reaching approximately 88% accuracy, with a macro F1 score of 0.82 and a micro F1 of 0.85.

This demonstrates a significant improvement in capturing multiple, fine-grained emotions compared to previous methods.

# RESULTS

## Sample Predictions (Your Flask App)

- “I’m so happy today!” → joy, excitement
- “I regret what I said.” → remorse, sadness
- “This is unbelievable!” → surprise, excitement
- “Why would you do that?” → annoyance, anger

## Application Output

- Real-time emotion detection
- Works with typed text
- Works with speech input
- Auto-silence detection → recording stops automatically

# DISCUSSION

## Strengths

- Multi-label → realistic emotional mapping
- 28 emotions → fine-grained
- Transformer-based → strong context reasoning
- Web app interface with modern UI

## Limitations

- Struggles with sarcasm
- Needs GPU for fast training
- Cannot detect tone of voice (text only)

## Future Work

- Add speech emotion recognition
- Multimodal fusion (text + audio + facial expressions)
- Deploy as REST API / mobile app

# PROJECT TIMELINE

- Week 1
  - Focused on dataset exploration and preprocessing to prepare the data for training.
- Weeks 2-3
  - Conducted model training and fine-tuning using DistilRoBERTa for initial multi-label emotion classification.
  - • This DistilRoBERTa model was used to validate performance and detect 7 core emotions.
- Week 4
  - • Performed evaluation, optimization, and improved metrics such as accuracy and F1-score.
  - • This week completed the switch to the full RoBERTa structure for higher context awareness.
- Week 5
  - • Developed the Flask backend along with the user interface for real-time emotion prediction.
  - • Key Improvement: We upgraded the model to SamLowe/roberta-base-go\_emotions to classify 28 distinct emotions.
- Week 6
  - Completed testing, prepared the final report, and created the project presentation.

# DEMO & CONCLUSION

- Flask Web App Demo
- Real-time multi-emotion prediction
- Supports 28 GoEmotion labels
- High accuracy using DistilRoBERTa
- Clear improvement over earlier methods

Conclusion:

our model accurately identifies complex emotional patterns in human conversations, significantly improving previous emotion-recognition systems.

THANK YOU