

Clinical EmotiSupport

Interim Presentation

Team Members:

Yuval Reznik, Lior Zvieli, Yoni Libman

December 19, 2025

Project Review



Motivation

- Patient messages convey emotional cues intertwined with clinical information
- Current triage systems overlook emotional signals
- Results in suboptimal prioritization of patient needs



Goals

- Develop NLP-based system for emotional intensity detection
- Enable emotion-aware clinical workflows
- Support administrative processes with emotional context



Novelty

- Emotion intensity prediction (not binary)
- Application to telemedicine messages
- Privacy-safe synthetic data generation

Previous Work

Title / Year	Task	Methods	Data	Results	Relation to Our Project
GoEmotions (2021) [Cited by 1229]	Multi-label emotion detection	BERT-based classifiers	58K Reddit comments	High multi-label accuracy	Establishes best practices for emotion modeling
ClinicalBERT (2019) [Cited by 1525]	Clinical NLP	Domain-adapted BERT	MIMIC-III notes	Outperforms general BERT	Motivates domain-aware emotion modeling
Emotion Detection in Patient Messages (2020) [Cited by 507]	Healthcare emotion detection	Transformer classifiers	Patient portal messages	Limited by small datasets	Our synthetic data addresses data scarcity

Dataset Overview



Dataset Overview

- Samples: **N = 1132**
- Language: **English**
- Domain: **Clinical & Administrative(564/568,49.8%/50.2%)**
- Format: **JSONL(one JSON object per line)**

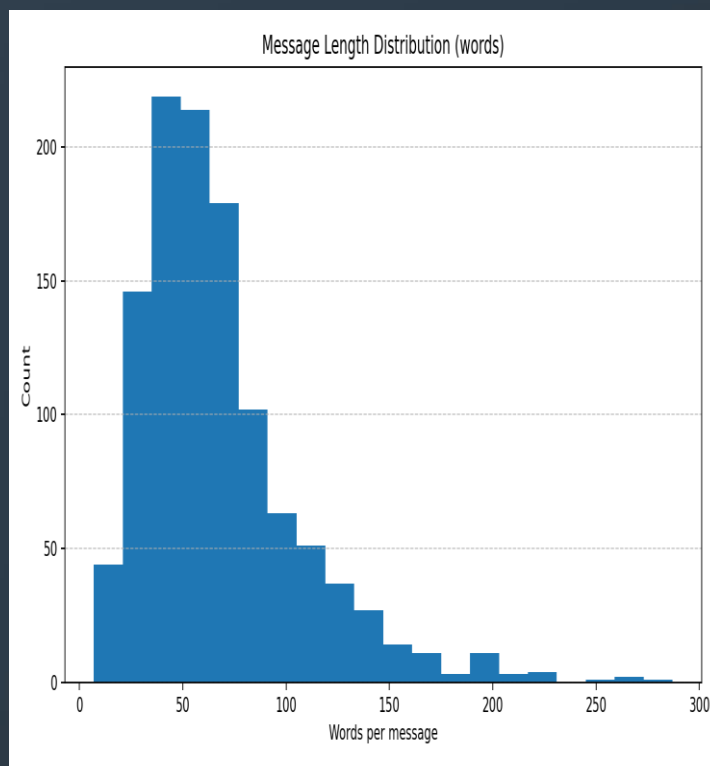


Data Generation

- Model: **DeepSeek-R1:8B**
- Inference: **Ollama**
- Prompts: **Telemedicine communication (patient/caregiver messages; domains clinical/administrative)**



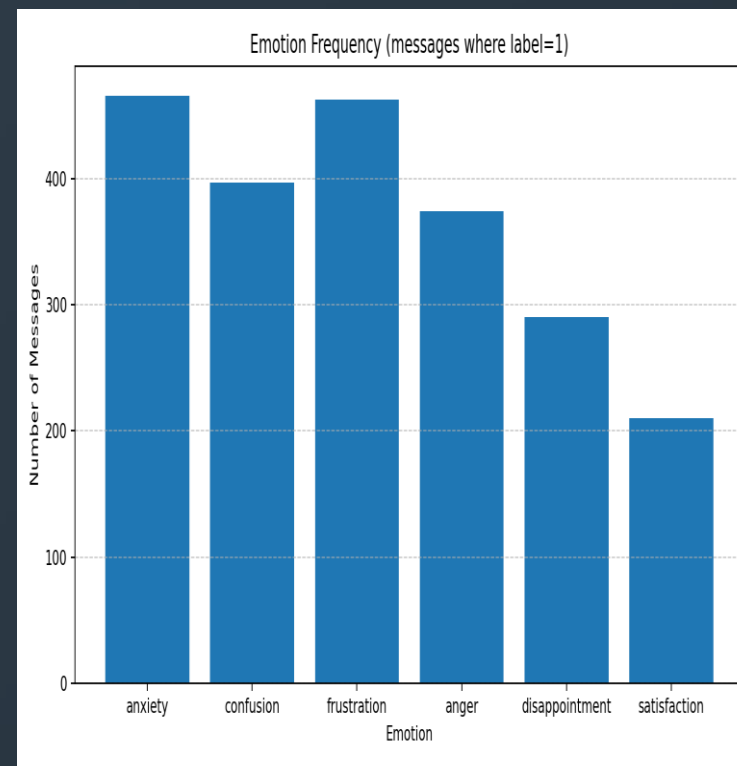
Message Length (words)



• **Min / Median / Mean / Max = 7 / 59 / 67.1 / 287**



Distribution of Emotions



Baseline Solution & Results

Model Architecture

- Architecture: DistilBERT
- Task: Multi-label emotion classification
- Input: Patient/Caregiver message text
- Output: Emotion classification per label ($\in [0,1]$)
- Training: Fine-tuned on synthetic dataset

Overall Metrics

Metric	Precision	Recall	F1-Score
Micro Average	0.8467	0.7341	0.7864
Macro Average	0.8609	0.7507	0.7971

Per-Emotion Analysis

Emotion	Precision	Recall	F1-Score	Support
Anxiety	0.81	0.57	0.67	74
Confusion	0.73	0.81	0.77	58
Frustration	0.94	0.82	0.87	72
Anger	0.76	0.7	0.73	50
Disappointment	1.00	0.71	0.83	49
Satisfaction	0.93	0.89	0.91	28

Evaluation Metrics

Subset Accuracy

0.4765

Neutral Detection(gate)

P 1.00 R 1.00 (24/170
Neutral)

Threshold/Neutral threshold

0.65/0.38

Project Plan & Next Steps



Dataset Expansion

Generate 500–2000 additional samples

IMPROVED ROBUSTNESS



Label Balancing

Target rare emotions explicitly

REDUCED BIAS



Model Upgrade

Evaluate ClinicalBERT / RoBERTa

HIGHER ACCURACY



Multi-task Learning

Joint domain + emotion prediction

BETTER CONTEXTUAL MODELING



Error Analysis

Manual inspection of failure cases

IMPROVED INTERPRETABILITY



Final Presentation

Summarize results & insights

COURSE SUBMISSION