AI-Powered Emotion Detection from Text for Mental Health Monitoring in Digital Communication

P.PRASAD

Abstract

Digital communication is everywhere today, and it gives us a great chance to spot emotional signals that could point to mental health problems early on. This summary suggests creating an AI-based system that can automatically find and understand emotions in chats and emails. It would track how emotions change over time and match them with known psychological measures, giving useful insights to mental health professionals.

The system would use advanced AI models (transformers) that are trained to detect different emotions at once. It would also protect people's privacy by hiding personal information, explain how the AI makes its decisions, and show real-time results on a dashboard to help with early support.

This research hopes to improve emotion-detecting technology, help people who speak less common languages, and create ethical rules for using AI in mental health care.

Keywords: *emotion detection, mental-health AI, transformer models, privacy, affective computing. Digital communication, Early detection, Chats and emails, psychological measures*

1.Introduction

In today's world, people communicate more through **text messages**, **emails**, **social media**, and **chats** than face-to-face. These digital texts often include **emotional expressions** (e.g., sadness, anger, happiness). If we can **understand these emotions**, we can detect **early signs of mental health issues**, such as depression, anxiety, or stress.

This is where **Artificial Intelligence (AI)** comes in. AI can help **analyze text messages** to detect emotions automatically and assist mental health professionals in **monitoring individuals' emotional well-being**.

Emotion detection from text (also called **Sentiment Analysis** or **Affective Computing**) is the process of:

- **Analyzing written text** (like messages, emails, or posts),

- **Identifying emotional content** (like happiness, sadness, anger, fear, etc.),

- And **labelling or scoring** those emotions.

AI uses **Natural Language Processing (NLP)** techniques and **machine learning models** to do this.

Mental-health disorders account for nearly 15 % of the global disease burden, yet timely diagnosis remains a challenge due to social stigma, shortage of professionals, and delayed help-seeking behavior. Simultaneously, individuals now produce vast volumes of text—estimated at 16,000 words per person per week—across e-mail, instant messaging, and social platforms. These digital traces are rich with affective signals that can reveal stress, anxiety, or depressive tendencies long before clinical manifestation.

Recent advances in Natural Language Processing, especially large language models (LLMs), enable fine-grained emotion recognition beyond polarity sentiment. However, existing academic prototypes seldom incorporate longitudinal tracking, domain adaptation, or rigorous ethical safeguards required for deployment in healthcare or corporate wellness settings. This research therefore seeks to bridge the gap between laboratory-grade emotion classifiers and real-world mental-health monitoring tools.

2. Problem Statement

Despite the emotional richness of textual data, most mental health issues remain undetected due to the lack of real-time, scalable tools that can interpret human emotions accurately and ethically. Traditional methods rely heavily on questionnaires or human intervention, which are time-consuming, limited in scale, and reactive rather than proactive.

There is a need for an AI-based system capable of automatically detecting emotional cues from textual communication and correlating them with psychological health indicators for early detection and monitoring.

3. Objectives of the Study

- To design and implement an AI-based emotion detection system from digital text communication.

- To analyze emotional trends over time using NLP techniques and link them to known psychological scales.

- To develop a privacy-preserving framework that ensures user anonymity and ethical usage.

- To provide real-time dashboards and alerts for mental health professionals.

- To create language models adaptable to low-resource and regional languages.

4. Review of Literature

The literature review will cover existing research on:

- Emotion detection in NLP and affective computing,

- Transformer-based language models (e.g., BERT, RoBERTa),

- Mental health applications using AI,

- Ethical frameworks and privacy-preserving techniques in AI,

- Real-time monitoring systems in healthcare.

This review will help identify research gaps and define the novelty of the proposed work.

5. Research Methodology

a) Data Collection

Text data will be sourced from public emotional corpora and (with consent) anonymized user chat/email data.

b) Preprocessing

Includes tokenization, stop-word removal, lemmatization, and language normalization.

c) Emotion Detection Model

- Transformer models (BERT, RoBERTa) fine-tuned on multi-label emotion datasets.

- Multi-label classification to detect co-occurring emotions (e.g., sadness + anger).

d) Temporal Emotional Analysis

- Emotional patterns tracked over time and mapped to psychological scales (like PHQ-9, GAD-7).

e) Dashboard and Alert System

- Real-time visualization for clinicians and caregivers.

- Risk scores and alerts for potential mental distress.

f) Privacy and Ethics

- Text anonymization,

- Explainable AI (e.g., SHAP or LIME),

- Adherence to ethical AI principles and institutional review board (IRB) approval processes.

6. Expected Outcomes

- An AI-based prototype for real-time emotion detection from text.

- An interactive dashboard showing emotion trends and mental health insights.

- Support for multiple Indian languages and dialects.

- A set of ethical and technical guidelines for using AI in mental health monitoring.

7. Scope of the Research

- The research focuses on emotion detection from text-based digital platforms (e.g., chats, emails).

- Only textual (not audio/video) inputs will be considered.

- The system will be designed to handle multilingual text, including low-resource languages.

- Ethical AI practices and privacy regulations (e.g., GDPR-like compliance) will be integrated.

8. Existing System

The current systems used for AI-powered emotion detection from text for mental health monitoring typically focus on basic sentiment analysis or emotion classification. These systems have the following characteristics:

Features:

- Sentiment Analysis: Categorizes text as Positive, Negative, or Neutral.

- Basic Emotion Detection: Detects only a few emotions like joy, anger, sadness, etc.

- Machine Learning Models: Use traditional ML algorithms such as Naive Bayes, SVM, or decision trees.

- Datasets: Use generic datasets like Twitter Sentiment Corpus, EmoLex, or ISEAR which are not specific to mental health.

- Short-Term Analysis: Analyze messages individually, not tracking emotional trends over time.

- Lack of Context Awareness: Cannot understand sarcasm, hidden emotions, or complex mental states.

Limitations:

- Cannot detect early signs of mental health disorders.
- Limited accuracy due to lack of personalization and context.
- Not suitable for private communication like emails and chats.
- Rarely compliant with privacy standards (e.g., GDPR).

- Not designed for real-time monitoring or intervention.

9. Proposed System

The proposed system enhances emotion detection using advanced NLP and deep learning models tailored for mental health prediction in private digital communication such as emails and chats.

Features:

- Fine-Grained Emotion Classification: Detects a wide range of emotions (e.g., fear, guilt, hopelessness, etc.).

- Transformer-Based Models: Uses state-of-the-art architectures like BERT, RoBERTa for deep semantic understanding.

- Multi-label Learning: Identifies multiple emotions in a single message.

- Temporal Tracking: Monitors emotional trends over time to detect patterns linked to depression or anxiety.

- Customized for Private Communication: Trained on anonymized and consented chat/email datasets.

- Ethical & Privacy-Aware: Implements anonymization, user consent, and is compliant with standards like GDPR and IEEE P7000.

- Human-in-the-Loop: Involves mental health professionals in the decision-making loop for validation and support.

Advantages:

- Higher accuracy and better emotional depth.
- Can detect early warning signs of mental health issues.
- Ethical and privacy-conscious for sensitive environments (workplace, clinics).
- Useful for building emotion-aware dashboards for real-time support or alerts.

10. Conclusion

This research aims to bridge the gap between basic emotion detection and proactive mental health monitoring by leveraging advanced AI models. By focusing on emotion-aware text mining in digital communication, the system can provide early alerts, support timely interventions, and contribute significantly to the domain of mental health technology. The integration of privacy-preserving mechanisms and ethical practices ensures that the proposed framework is both effective and socially responsible

11. References / Bibliography

1.   Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*.

2.   Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL*.

3.   Calvo, R. A., Milne, D. N., & Hussain, M. S. (2017). Natural language processing in mental health applications. *JMIR*.