Mental Health Tracking: Contextual Emotions, Symptoms, Triggers & Themes Analysis and Mental Health Disorder Prediction

Sheryl Deakin

Abstract

Traditional mental health tracking relies on checklist-based assessments, which are rigid, tedious and fail to capture nuanced experiences. This project proposes an ML-driven system to extract mental health indicators—emotions, symptoms, triggers, and recurring themes—from free-text journal entries. Using Named Entity Recognition (NER) and classification models, the system will identify patterns over time. Data will be sourced from mental health subreddits and labeled using weak supervision. Models include TF-IDF + Logistic Regression/SVM, deep neural networks (DNNs), BiLSTM-CRF, and fine-tuned BERT. The extracted features will be mapped to mental health disorders. This method provides a more personalized and dynamic alternative to checklist-based tracking, improving self-awareness and enhancing mental health interventions.

Introduction

Problem Statement

Mental health disorders significantly impact daily life, productivity, and overall well-being. Accurate symptom tracking is essential for early intervention, treatment evaluation, and long-term mental health monitoring. Traditional mental health tracking methods primarily rely on structured checklists, requiring individuals to manually select symptoms, emotions, and triggers from predefined categories. They are often rigid, impersonal, and fail to capture the complexity and contextual nuances of real-world experiences. Many individuals struggle to track their mental health consistently due to the tedious nature of such methods. Additionally, mental health patterns are highly contextual and are often vague or implicit. Rule-based approaches struggle with such nuances, requiring professional intervention to extract meaningful insights.

Importance of the Problem

The COVID-19 pandemic has exacerbated the mental health crisis, increasing the demand for accessible mental health solutions. However, there remains a shortage of qualified therapists, and traditional therapy sessions often suffer from inefficiencies. Clinicians must gather extensive patient history, identify patterns, and build a diagnostic foundation, a process that can be time-consuming and may lead to patient disengagement. By enabling individuals to self-monitor their mental health patterns, this system aims to streamline clinical assessments, facilitating more efficient therapy sessions to address core issues, improving both treatment outcomes and patient retention in mental health care.

Challenges to Address

1. Implicit Mental Health Markers and Contextual Understanding: Mental health symptoms and triggers are often described indirectly or vaguely. Traditional ML models often fail to distinguish emotional nuances based on context. Detecting such implicit signals requires context-aware NLP models rather than direct keyword matching. **2. Data Scarcity & Labeling:** Unlike structured medical records, mental health discussions are informal and lack standardized labels. The DSM-5 diagnostic criteria, while widely used, require subjective clinical interpretation, making automated labeling challenging. Additionally, no comprehensive labeled datasets exist for mental health symptoms, triggers, and themes. **3. Accuracy/Labeling Uncertainty in Weak Supervision:** The absence of manually labeled training data requires the use of distant supervision. However, inaccurate labeling may introduce noise, affecting model performance.

Applications and Use Cases

Personal Mental Health Tracking: Users can record daily experiences, allowing the system to automatically detect emotions, symptoms, and triggers. Enables self-reflection and progress tracking over time. **Clinical Support Tool, Enhancing Therapy Efficiency & Patient Retention**: Helps clinicians identify recurring emotional patterns and assess treatment effectiveness. Reduces exploratory sessions, allowing therapy to focus on actionable interventions. Many individuals drop out of therapy due to a lack of perceived progress. By visualizing mental health trends and highlighting key triggers and symptoms, the system can increase engagement and adherence to therapy. **Mental Health Disorder Prediction Using Reddit Data**: Subreddit discussions often contain self-reported experiences from individuals with mental health conditions. By analyzing

text from r/ADHD, r/depression, r/anxiety, etc., the system can extract common symptoms and triggers to train classification models for disorder prediction. **Mental Health Research & Policy Making**: Researchers can leverage the extracted data to study emerging mental health trends and develop targeted interventions. Insights can inform mental health policies and resource allocation.

Proposed Project

This project will utilize Natural Language Processing (NLP) and Named Entity Recognition (NER) to extract and classify relevant features from short journal-like entries. Since the goal is to categorize aspects of mental health, classification models will be used across five key tasks using TF-IDF + Logistic Regression/SVM, Deep Neural Networks (DNNs), BiLSTM-CRF and fine-tuned BERT models. The accuracy of each will be compared. The classifiers are **1. Emotion Classification:** Detecting emotions (e.g., *anger, sadness, hopelessness*). **2. Symptom Classification:** Identifying mental health symptoms (e.g., *fatigue, intrusive thoughts, panic attacks*). **3. Trigger Classification:** Recognizing environmental or situational triggers (e.g., *arguments, loud noises, financial stress*). **4. Theme Classification:** Identifying overarching themes in journal entries (e.g., *work stress, relationship issues*). **5.Mental Health Disorder Classification:** Using the previous classifiers and extracting common symptoms, triggers and themes for potential mental health disorders based on patterns observed in data, this classifier will hopefully be able to identify mental health disorders.

Dataset	**Source**	**Size**	**Classes**
Emotion Dataset	Kaggle - GoEmotions Dataset	~58,000 labeled Reddit comments	28 fine-grained emotions
Reddit Mental Health Discussions	Subreddits dedicated to mental health discussions.	TBD based on the scraping process (At least 100,000+ posts).	The subreddit it comes from
Symptoms, Triggers, Themes (Weak Supervision)	Reddit mental health discussions, DSM-5, Mayo Clinic, WHO	~100,000+ posts (from Reddit Mental Health Discussions, 200+ curated keywords	Symptoms, Triggers, Themes

Dataset	**Preprocessing**	**Features**
Emotion Dataset	Tokenization, text cleaning (removing special characters, stopwords), and conversion into word embeddings.	Free-text input with corresponding labeled emotions.
Reddit Mental Health Discussions	Data Collection: Use Reddit API (PRAW/Pushshift), Auto-label emotions using emotion classifier, Distant supervision for symptoms, triggers, and themes, Text normalization. Data Cleaning: Remove low-information posts, handle missing data, and normalize text.	Free-text mental health discussions, automatically labeled with emotion, using the emotion classifier, and using weak supervision techniques for symptoms, triggers and themes.
Symptoms, Triggers, Themes (Weak Supervision)	Data Cleaning: Remove special characters & non-textual data. Lowercasing (unless acronyms or proper nouns like “PTSD”). Remove stopwords. Tokenization: Convert text into subword tokens using BERT tokenization. Entity Labelling (Weak Supervision): Auto-label data based on known symptoms, triggers, themes. Data Augmentation: Expand symptoms/triggers/themes using BERT embeddings. BIO Format: Convert weakly labeled text into BIO format for NER training.	Named Entity Labels (BIO format), Contextual Word Embeddings, Part-of-Speech (POS) Tagging, Dependency Parsing, TF-IDF Features.