

Mental Health Tracking

Sheryl Deakin

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

The Problem



Limitations of Traditional Mental Health Tracking

- Checklists are rigid, tedious, and impersonal.
- Fails to capture context and implicit mental health markers.
- Inconsistent self-reporting leads to incomplete data.

The Need for a Smarter Approach

- Mental health symptoms are context-dependent & nuanced.
- Rule-based methods fail to detect implicit signals
- Machine learning enables personalized & adaptive tracking

Challenges in Automating Mental Health Monitoring

- Implicit Mental Health Markers
- Lack of labeled data
- Contextual Ambiguity

Real-World Barriers to Effective Therapy

- Therapist shortages and high demand post-COVID
- Therapy intake is slow due to manual symptom discovery
- Patients may drop out if they don't perceive immediate benefits

Classifiers

Import markers for Mental Health



Emotions

Detecting emotions (e.g., anger, sadness, hopelessness).



Symptoms

Identifying mental health symptoms (e.g., fatigue, intrusive thoughts, panic attacks)



Triggers

Recognizing environmental or situational triggers (e.g., arguments, loud noises, financial stress)



Themes

Identifying overarching themes in journal entries (e.g., work stress, relationship issues).



Mental Health Disorder

Extracting common symptoms, triggers and themes for potential mental health disorders based on patterns observed in data.

Datasets

Emotion Classification



Source: Kaggle – GoEmotions Dataset



Size: ~58,000 labeled Reddit comments



Classes: 28 emotions (e.g., anger, sadness, gratitude)



Preprocessing: Tokenization & text cleaning, Conversion into word embeddings



Features: Free-text input with corresponding labeled emotions.

Reddit Mental Health



Source: Mental health subreddits (e.g., r/ADHD, r/depression)



Size: ~100,000+ posts (*estimated*)



Classes: The subreddit it comes from



Preprocessing: Web scraping, Emotion auto-labeling, weak supervision, text normalization & noise removal



Features: Automatically labeled with emotion, and using weak supervision techniques.

Symptoms, Triggers, Themes



Source: Reddit mental health discussions, DSM-5, Mayo Clinic



Size: ~100,000+ posts (from Reddit Mental Health Discussions, 200+ curated keywords)



Classes: Symptoms, Triggers, Themes



Preprocessing: Data Cleaning, Tokenization, Entity Labelling (Weak Supervision), Data Augmentation, BIO Format: Convert weakly labeled text into BIO format for NER training.



Features: Named Entity Labels (BIO format), Contextual Word Embeddings, Part-of-Speech (POS) Tagging, Dependency Parsing, TF-IDF Features

Use Cases



Personal Mental Health Tracking

Users can reflect on emotional patterns over time.



Clinical Support Tool

Helps therapists analyze longitudinal data.



Enhancing Therapy Efficiency

Reduces exploratory sessions, increasing patient retention.



Mental Health Disorder Prediction

Uses Reddit data to detect common symptoms/ triggers per disorder.



Mental Health Research

Enables analysis of emerging trends and supports policy-making.

Thank You!

Methodology (no time)

I had the slide for it and didn't think we needed it based on the piazza post :(



Data Collection

- Scrape Reddit mental health subreddits (e.g., r/ADHD, r/depression, r/anxiety).
- Use GoEmotions dataset for labeled emotion classification.
- Collect medical symptom lists (DSM-5, Mayo Clinic) for distant supervision.



Weak Supervision & Data Labeling

- Expand symptoms, triggers, and themes using BERT embeddings.
- Auto-label Reddit text using distant supervision (matching known patterns).
- Convert text into BIO format for Named Entity Recognition (NER).



Model Training

- Train NER models to detect emotions, symptoms, triggers, and themes.
- Use TF-IDF + Logistic Regression/SVM for baseline models.
- Implement Deep Neural Networks (DNNs), BiLSTM-CRF, and fine-tuned BERT for high accuracy.



Mental Health Disorder Classification

- Extract common features (emotions + symptoms + triggers + themes) per disorder.
- Use clustering & classification models to link patterns to mental health disorders.
- Evaluate performance using precision, recall, F1-score on labeled test data.