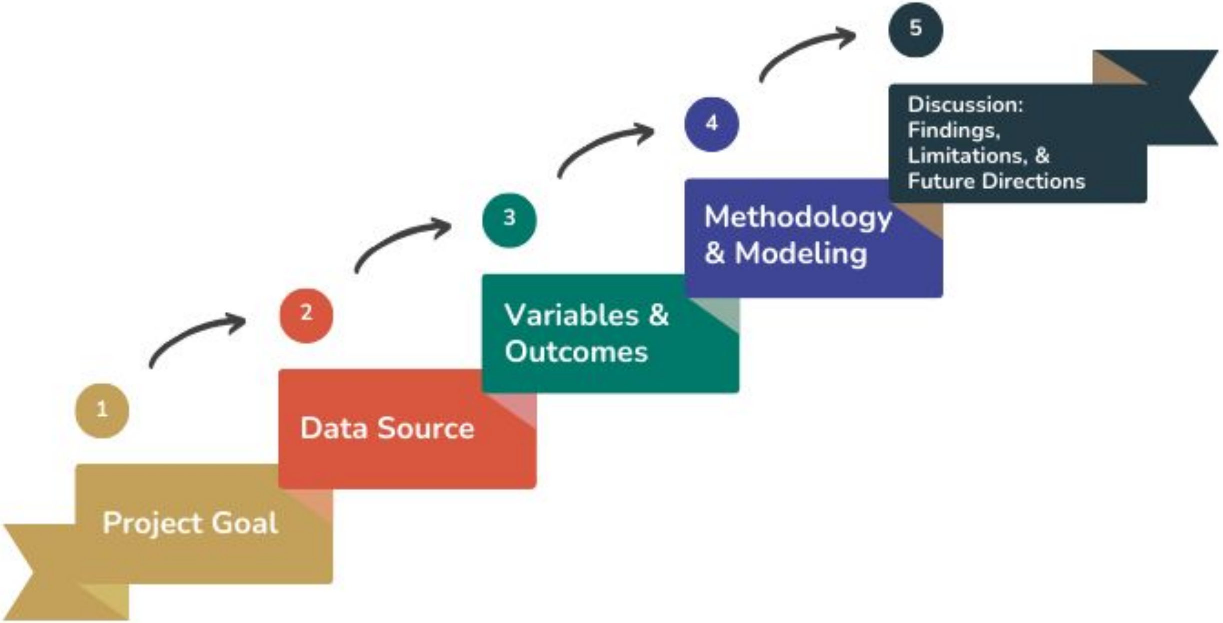
Predicting Mental Health Treatment Decisions From Social Media

Emelie Curl, Alejandra Dashe, Mitch Hamidi, Gabe Khan, & Eunbin Kim

Erdös Institute May-Summer 2024 Boot Camp

Project Overview



1 Project Goal

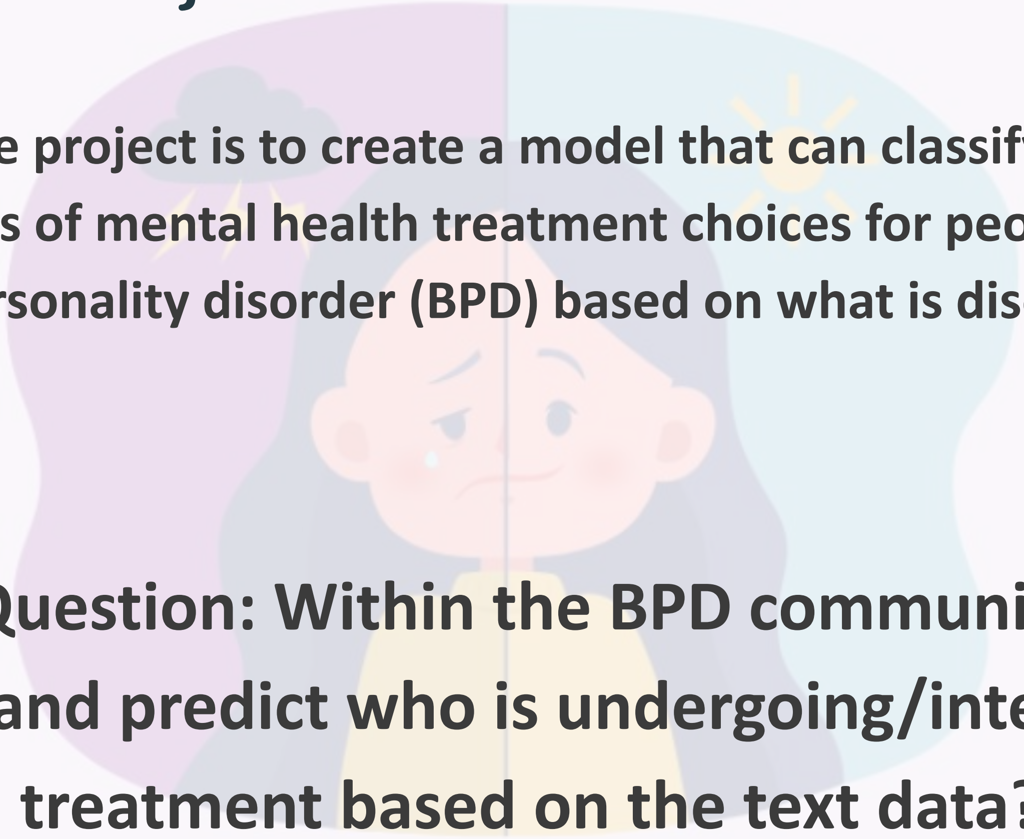
2 Data Source

3 Variables & Outcomes

4 Methodology & Modeling

5 Discussion: Findings, Limitations, & Future Directions

Project Goal and Motivation

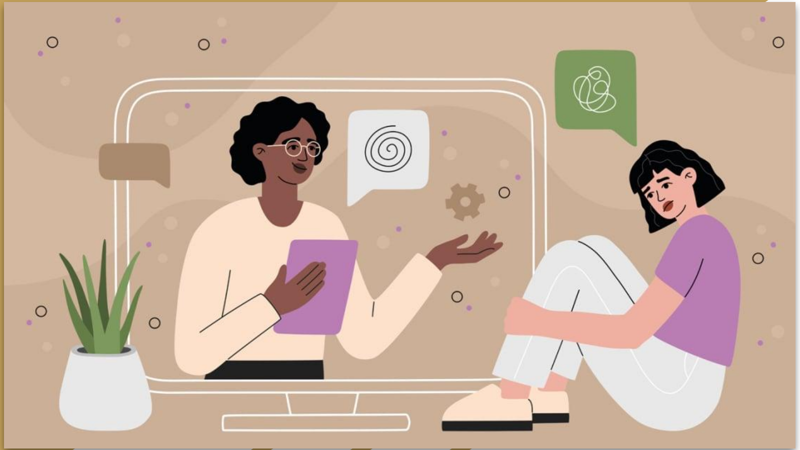


The goal of the project is to create a model that can classify and predict the sentiments of mental health treatment choices for people with borderline personality disorder (BPD) based on what is discussed in their Reddit posts.

Central Question: Within the BPD community, can we classify and predict who is undergoing/interested in treatment based on the text data?

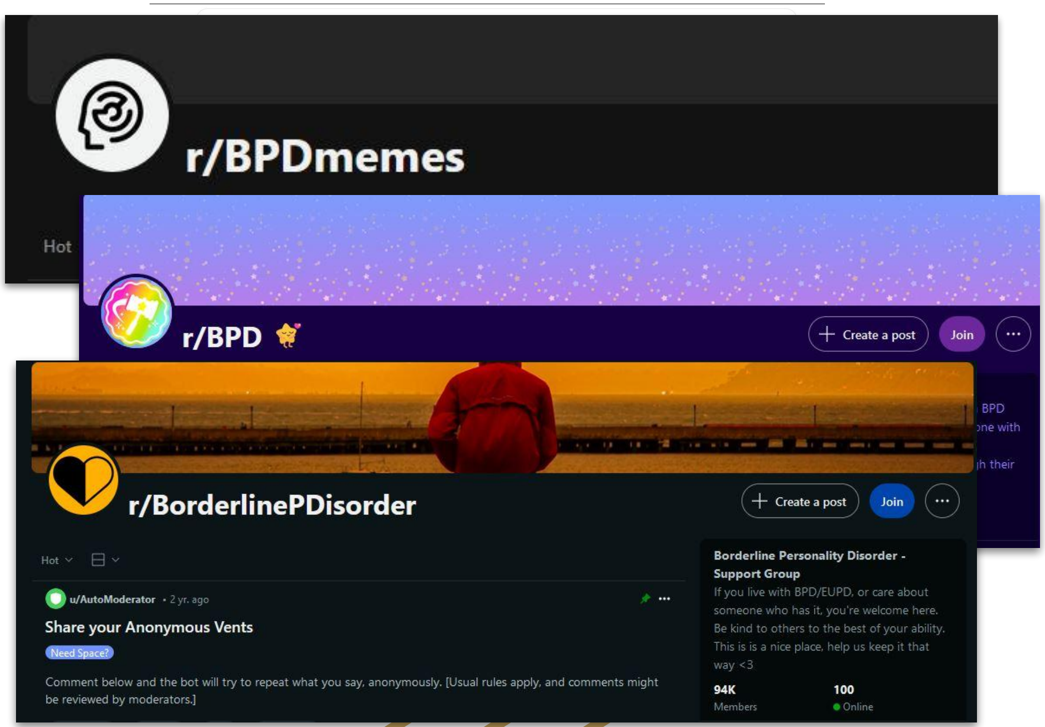
Stakeholders

- Healthcare providers (therapists, psychologists, psychiatrists, social workers, inpatient treatment hospitals)
- BPD patients (end users)
- BPD advocacy groups
- Pharmaceutical companies



Data Source

- Data Sources
    - Kaggle dataset of scraped posts
    - Scraped comments from 8 BPD relevant subreddits using 89 keywords of interest
- Train-Test Split
    - ~500 comments manually coded as relevant/not relevant to BPD treatment scraped from r/BorderlinePDisorder and r/bpdmemes



u/AutoModerator • 2 yr. ago

Share your Anonymous Vents

Need Space?

Comment below and the bot will try to repeat what you say, anonymously. [Usual rules apply, and comments might be reviewed by moderators.]

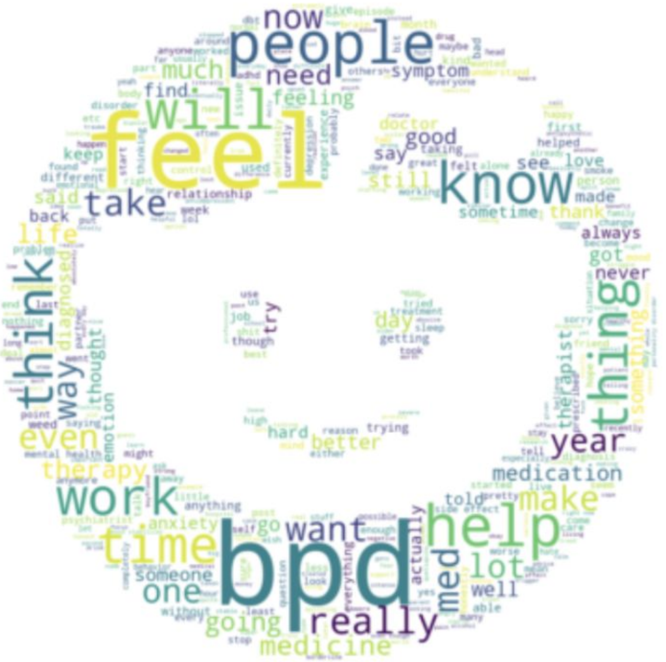
Borderline Personality Disorder - Support Group

If you live with BPD/EUPD, or care about someone who has it, you're welcome here. Be kind to others to the best of your ability. This is is a nice place, help us keep it that way <3

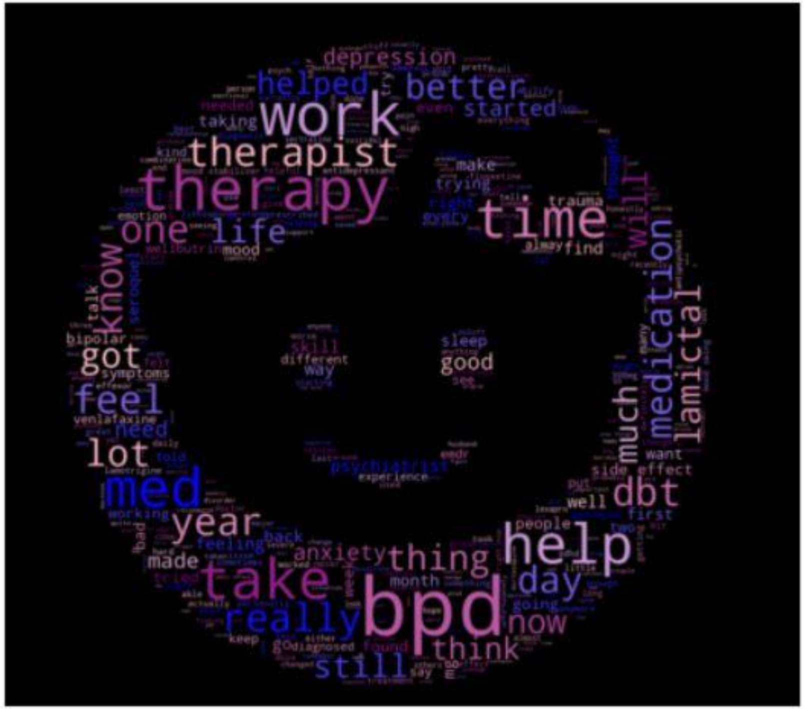
94K	100
Members	Online

Data Distribution

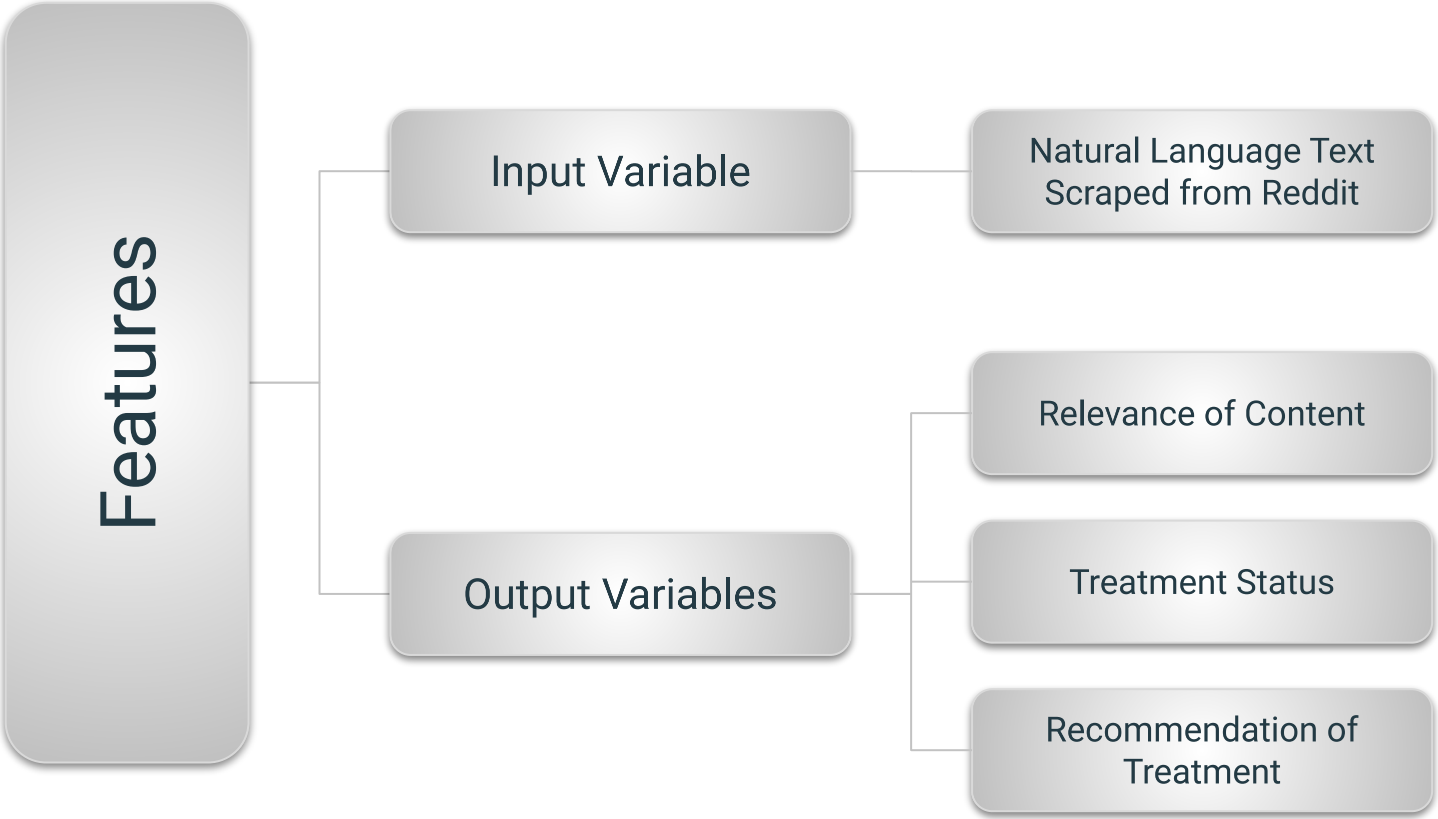
Random Sample of Scraped Reddit Comments



Relevant Scraped Reddit Comments



Features of Interest



Features

Input Variable

Natural Language Text Scraped from Reddit

Output Variables

Relevance of Content

Treatment Status

Recommendation of Treatment

Key Performance Metrics

Accuracy

Accuracy of prediction of relevance of content and treatment status

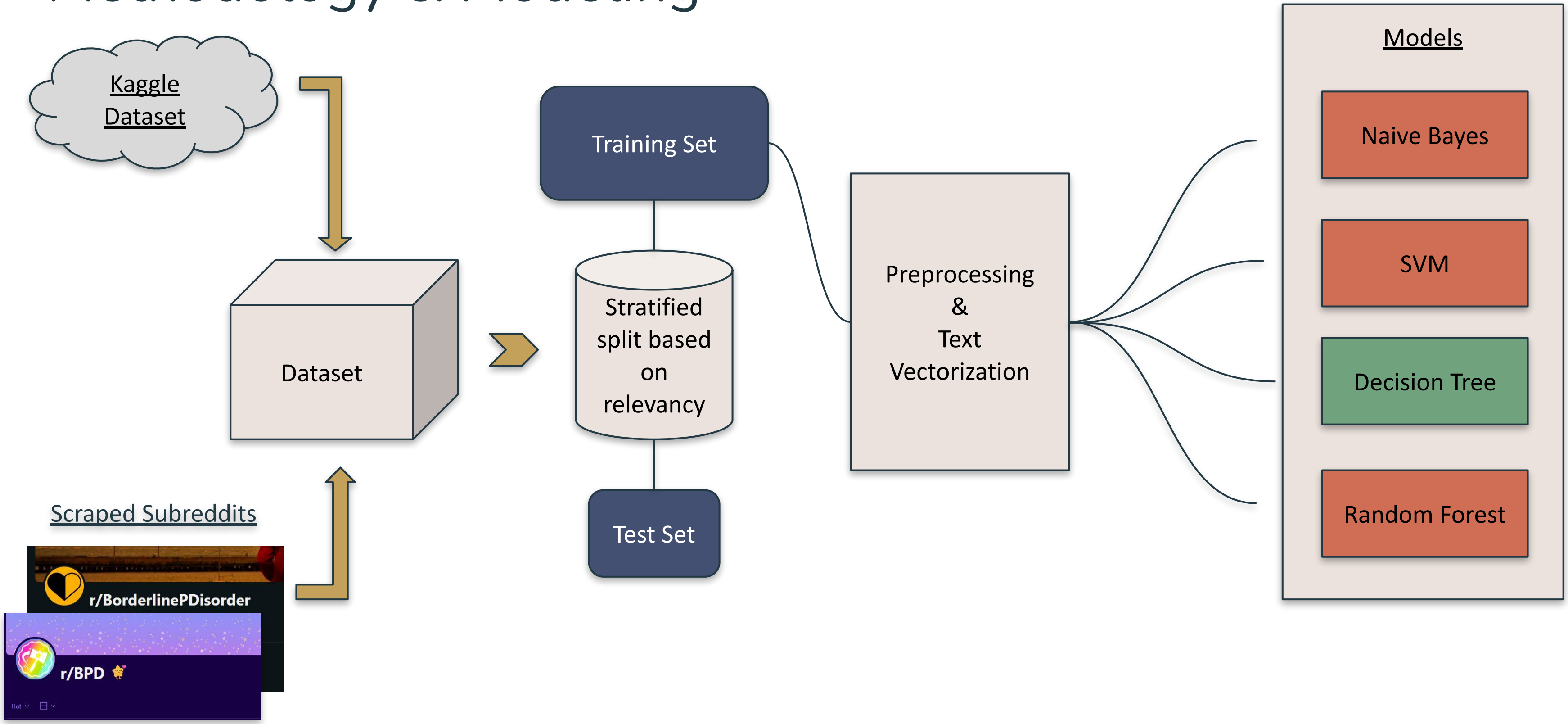
Precision

Proportion of comments marked relevant that were correctly classified in terms of relevance and treatment status

Recall

Proportion of relevant posts actually classified in terms of relevance and treatment status

Methodology & Modeling



Kaggle Dataset

Scraped Subreddits

Dataset

Training Set

Stratified split based on relevancy

Test Set

Preprocessing & Text Vectorization

Models

Naive Bayes

SVM

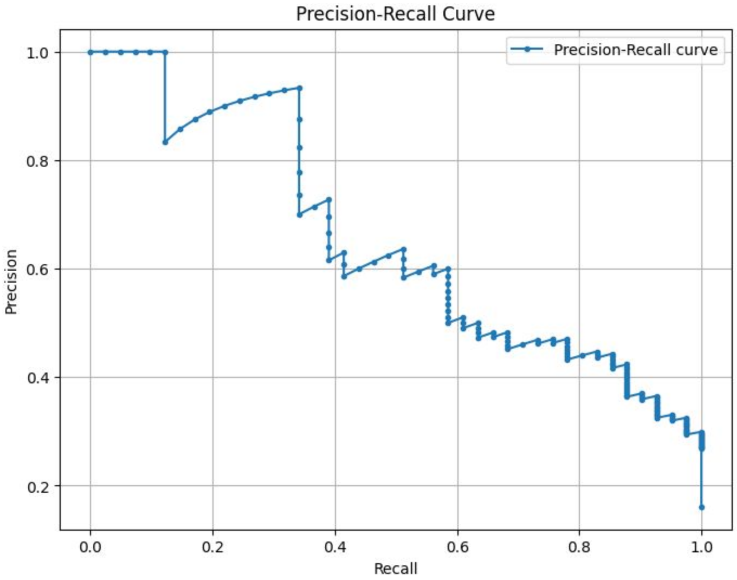
Decision Tree

Random Forest

Results and Outcomes

Relevancy model selection on Kaggle data:

- Baseline: Logistic regression with vectorized text
- Logistic regression with vectorized text and keyword count
- **Decision tree** ⟶
- dilBERT
  - Training data limitations
- ollama
  - Excellent results but 2 mins per post



Results and Outcomes

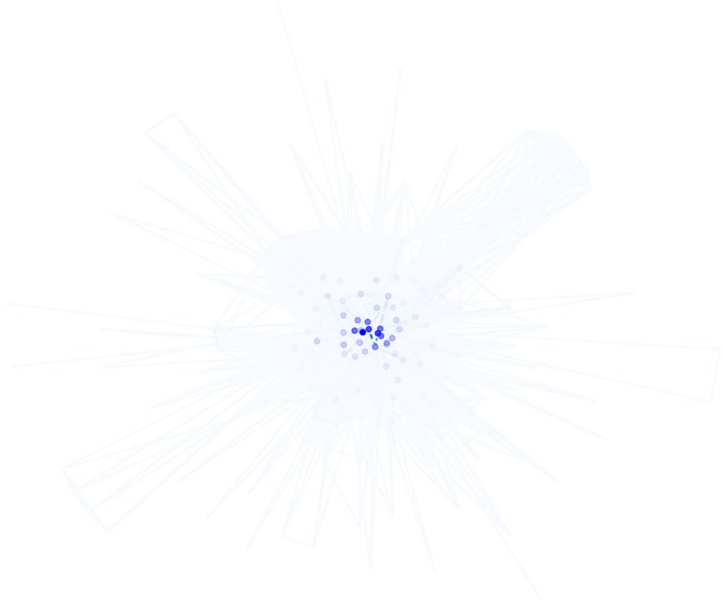
Relevancy model selection on scraped data:

Decision tree model

- Performance: Recall 0.7
- Advantages: interpretable, fast

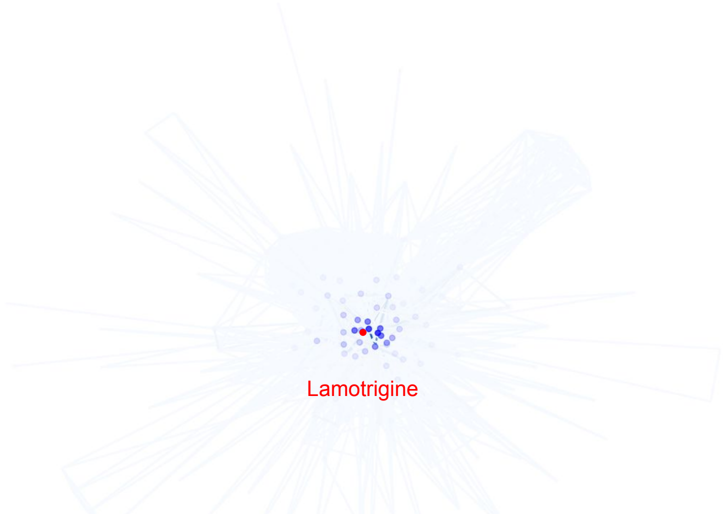
Kaggle Dataset Analysis

- Relevance model in r/bpd over 8 years
- ~240,000 posts ➡ 2557 most relevant posts
- Identified all medications using NER

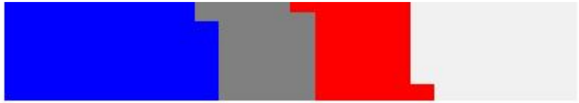


Sentiment Analysis

- Analyze sentiments for medications using GPT-4o + prompt engineering
- Validation was done by manual coding: **95% accuracy** for positive, negative, (neutral+other)
- Lamotrigine: 552 mentions (most frequent)



Sentiment breakdown

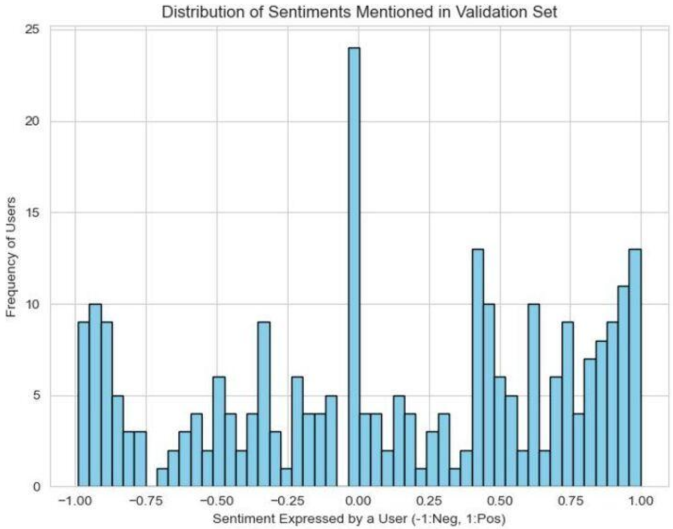


Limitations

- Need more human coded data
- The test set may not be representative actual relevance distribution
- Sentiment analysis lacks nuance

Future Directions

- Adapt to evaluate relevance and sentiment associated to other illnesses
- Build a more nuanced sentiment analysis model
- Test preliminary models on larger data set



Thank You

Special thanks to:

- Steven Gubkin, Erdös Institute Head of Training and Assessment
- Yuchen Luo, Erdös Institute Mentor
- The Erdös Institute