Establishing digital phenotypes for Mental Health using Artificial Intelligence

Nandita Bhaskhar[1], Scott Fleming[2], Shaimaa Bakr[1], Imon Banerjee[3], Daniel Rubin[4,5,6]

1 Electrical Engineering, Stanford University; 2 Biomedical Informatics, Stanford University; 3 School of Medicine, Emory University; 4 Biomedical Data Science, Stanford University; 5 Radiology, Stanford University; 6 Ophthalmology, Stanford University
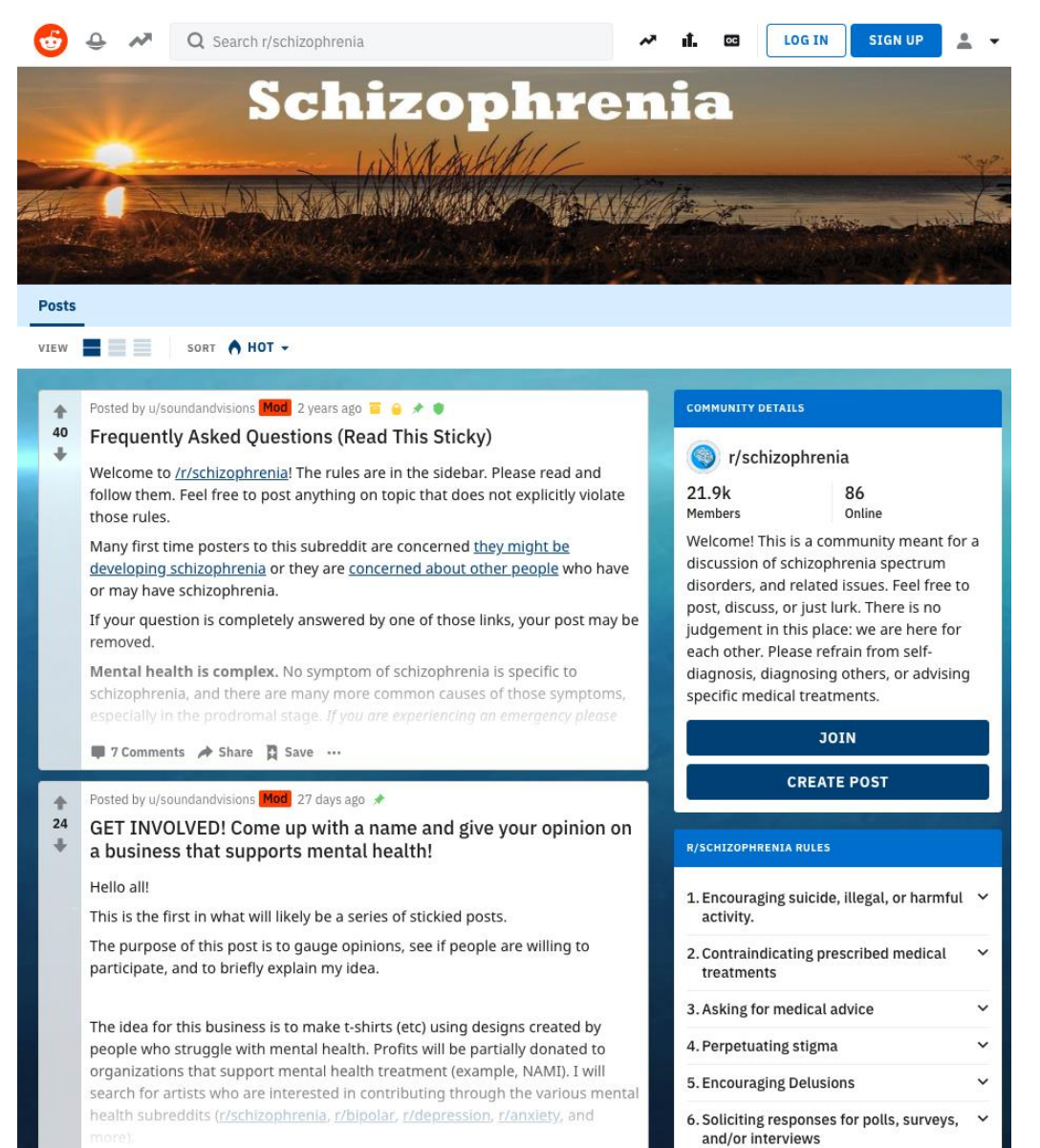
Web: www.stanford.edu/nanbhas
Email: nanbhas@stanford.edu
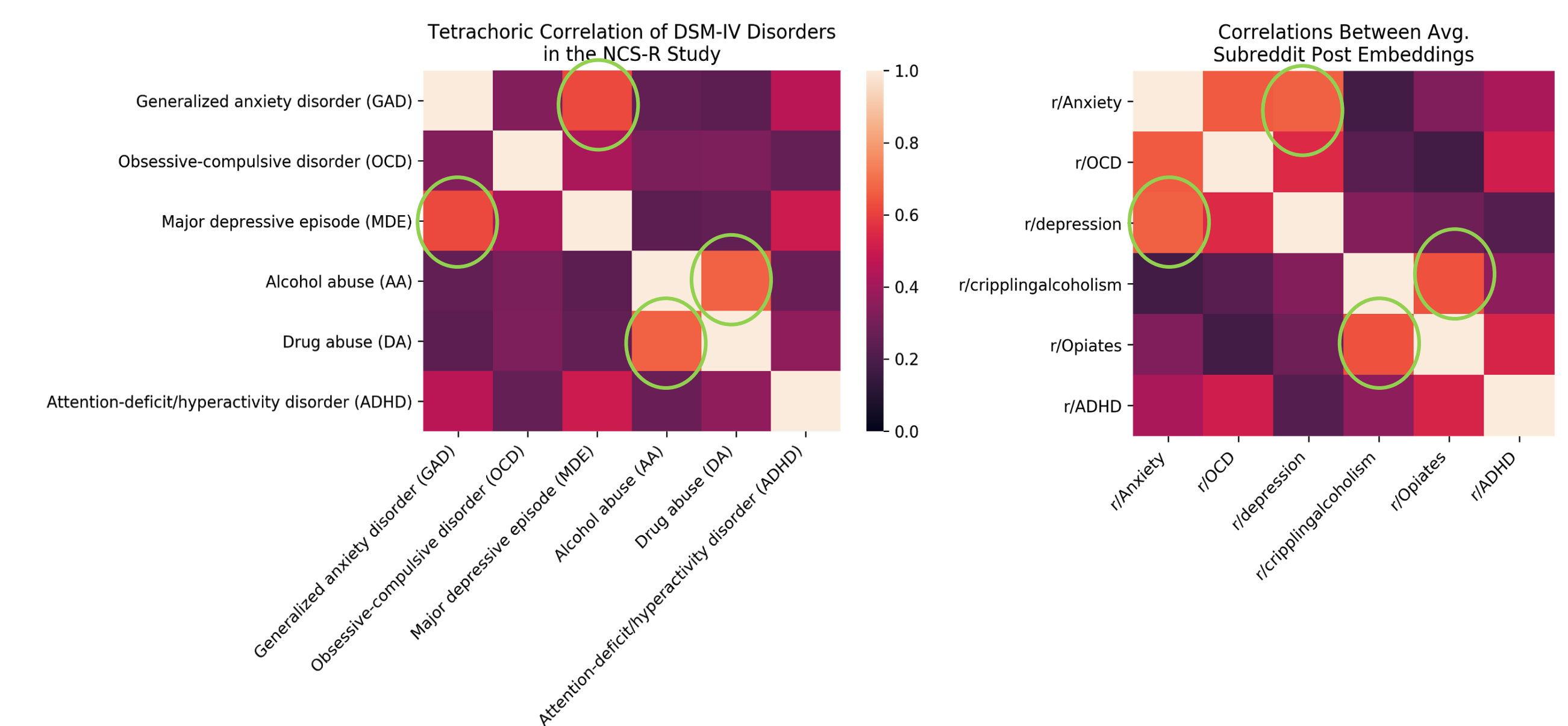**LinkedIn:** linkedin.com/in/nanditabhaskhar/

SCAN ME

Reddit: Social Media Platform CASE STUDY

- **Unique:** users post anonymously
- **Widespread:** over 200 million users
- **Active:** discussions on mental health with plenty of moderators
- **~Organized:** A 'subreddit' is a forum dedicated to a specific topic on Reddit. Names start with the token 'r/', e.g. 'r/ADHD'

- **Caveats? MESSY**
- User-created subreddits, no restrictions
- Several subreddits discussing similar issues
- Posts in subreddits might not conform to the subreddit
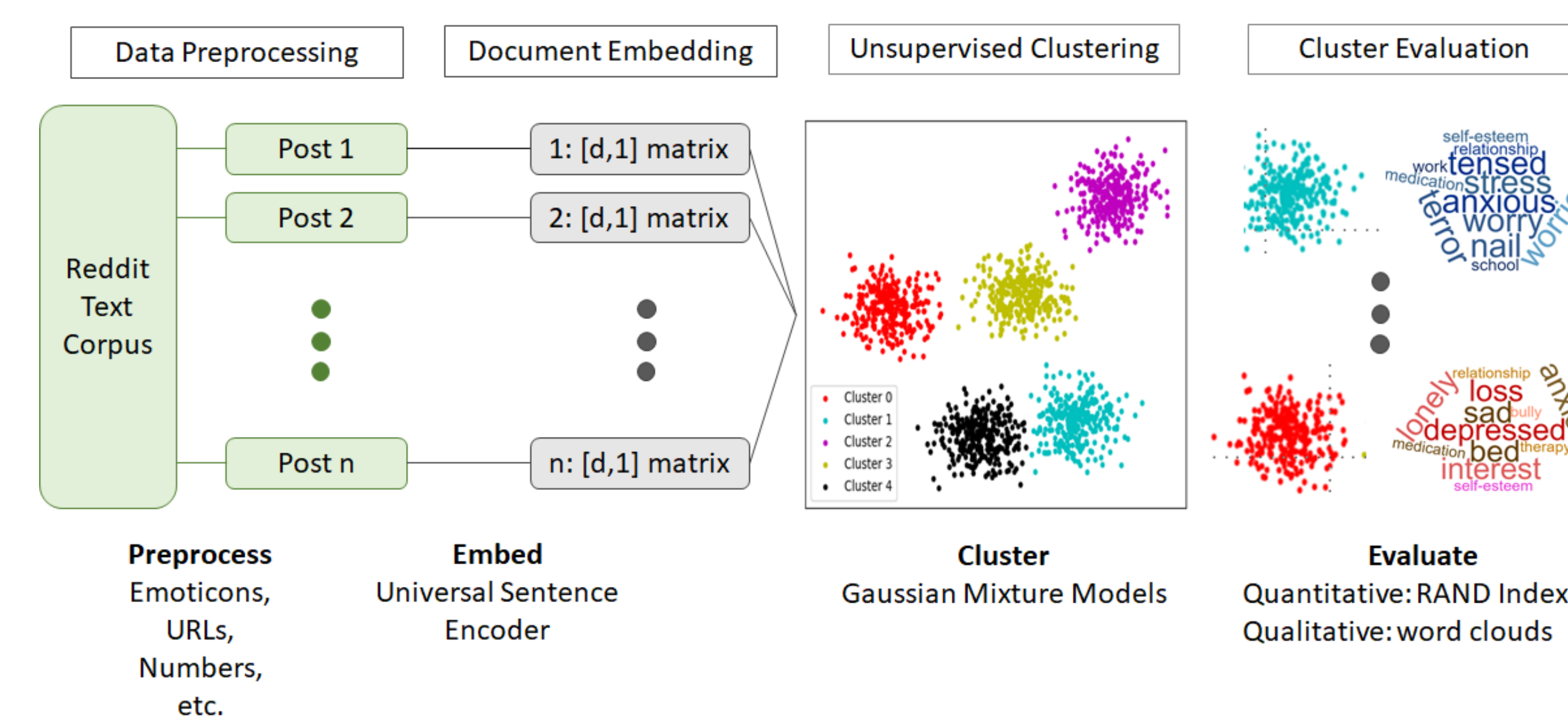

Embeddings and Comorbidities

- Subreddits reflect diagnostic comorbidities
- Carry semantic meaning
- Subreddit labels are noisy and overlapping
- Current DSM-5 categories can also be noisy and overlapping
- Is it possible that they are noisy and overlapping in the same way?
- If so, can we do better?

Motivation

- Psychological disorders -- leading cause of disability worldwide (~ **$1 trillion** loss to the global economy)
- Computational analysis of narrative texts using natural language processing (**NLP**) and other artificial intelligence (**AI**) techniques applied to texts related to patients could help study these at a large scale

- **Clinical notes:** Contain only professional impressions at point of care
- **Social media (SM) platforms:** Could provide a lens into the patient's life; they are ubiquitous, pervasive, and recorded at the moment
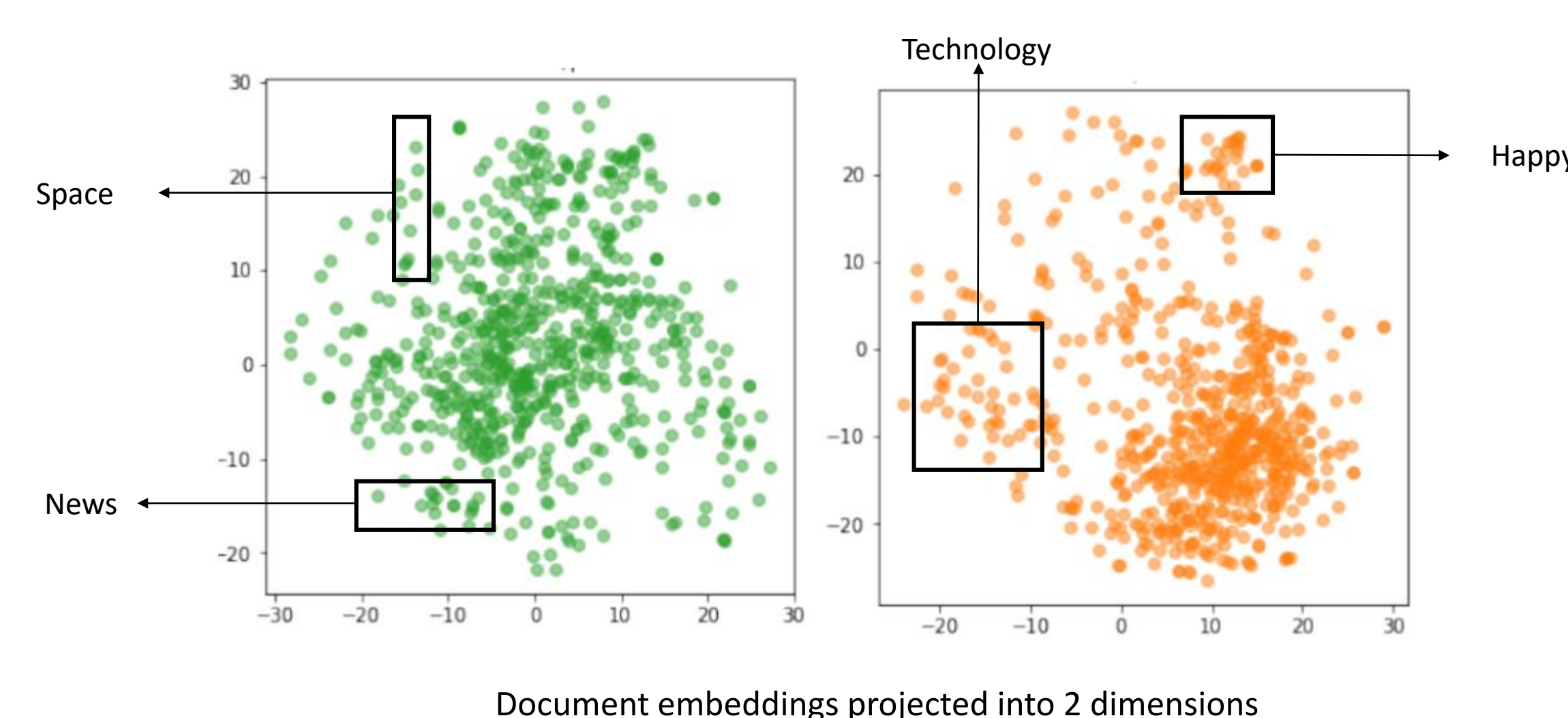
Our overall methodology

Prior Work

- **Focus:** automated classification of social media posts for predictive analysis
- **Approach:** Multi-class classification of relevant posts to the appropriate (user-defined) mental health disorder category
- **Method:** Supervised machine learning with either subreddits or heuristically extracted features as training labels

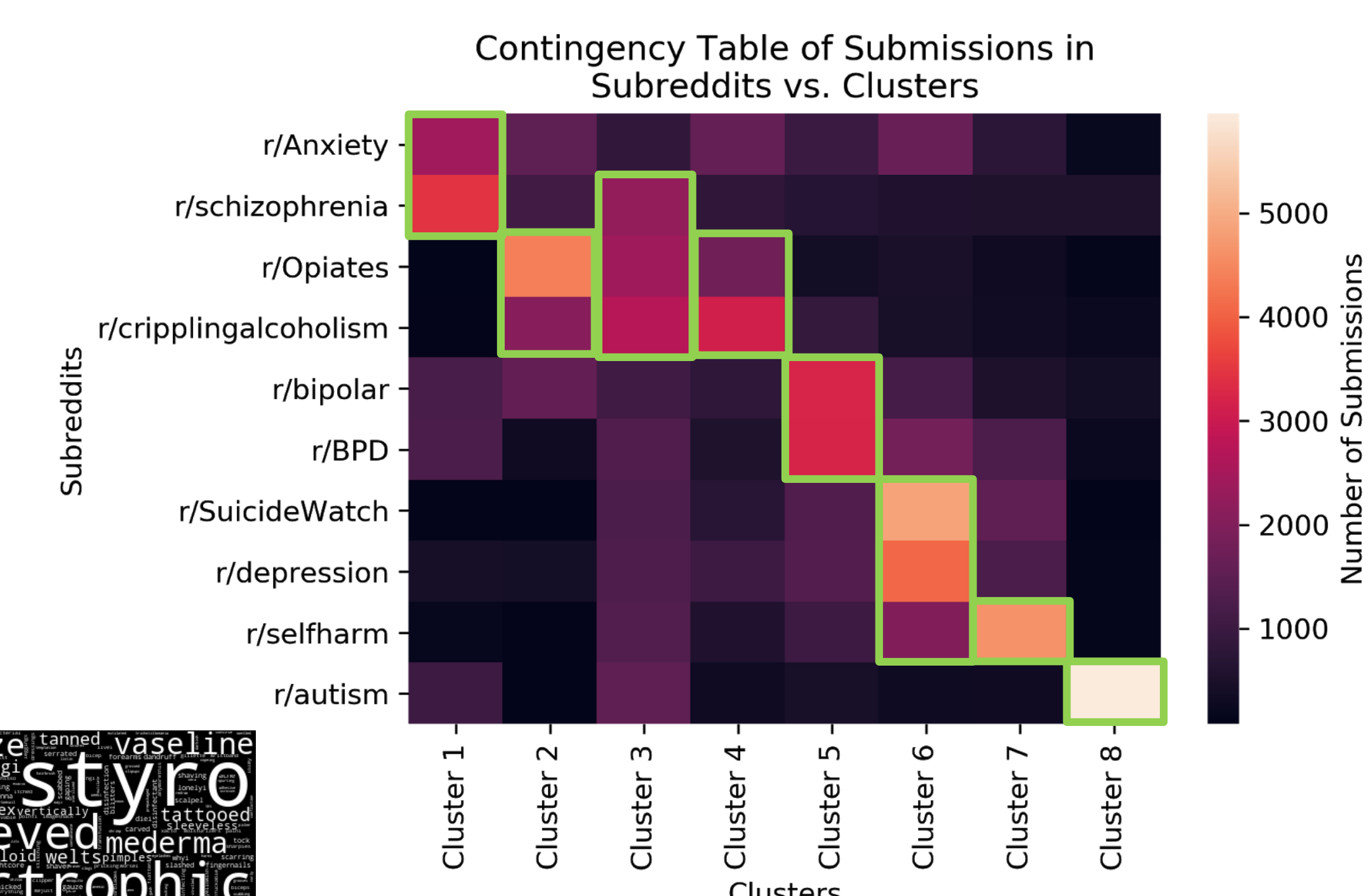[1] Gkotsis et al. (2017) Nature Scientific Reports

Word Embeddings, Document Labelling, Supervised Learning

- Word embeddings capture semantics of words as points in vector space

- Can be used as features to train a supervised model for document classification with a small number of ground truth labels

Document embeddings projected into 2 dimensions

Results

- Meaningful, reproducible clusters
- RAND index = 0.79
- (Green): >2000 posts
- Cluster 4 ~ crippling alcoholism
- Cluster 7 ~ self harm

Bigger Vision

Leverage NLP/AI for automated web screening:

- For early diagnosis
- For timely intervention and clinical referrals
- To reach a wider, underserved population
- To augment clinical decision making
- Potential healthcare cost savings