Into the Wild: Transitioning from Recognizing Mood in Clinical Interactions to Personal Conversations for Individuals with Bipolar Disorder

Katie Matton[1], Melvin G McInnis[2], Emily Mower Provost
Computer Science and Engineering[1], Psychiatry[2], University of Michigan, Ann Arbor, Michigan.

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

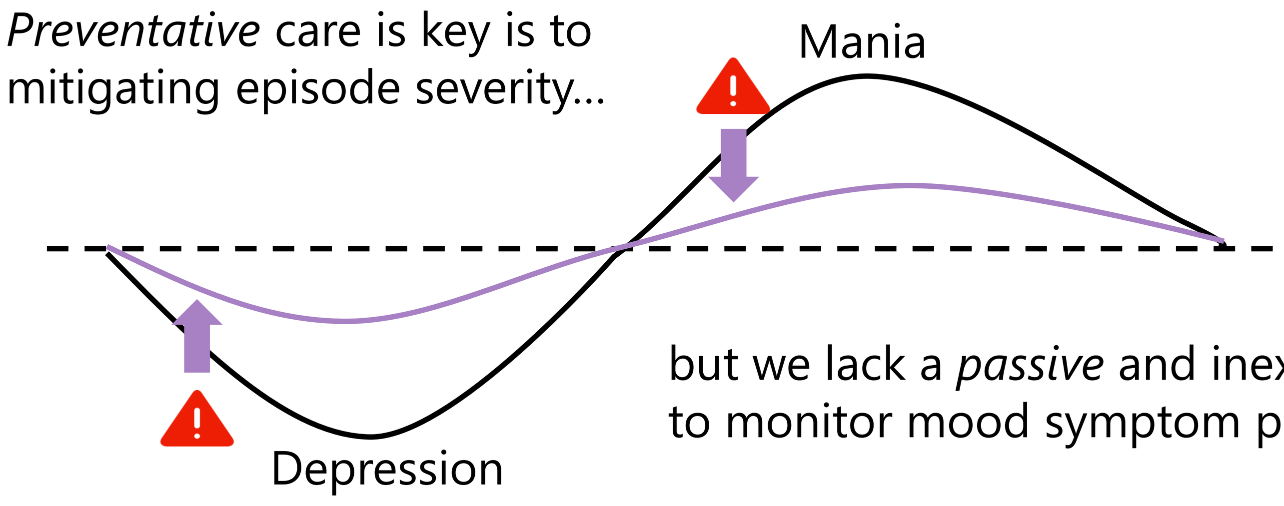
Introduction

Problem Statement

Bipolar Disorder (BD) is a severe and chronic mental illness characterized by mood transitions into **episodes of mania and depression**

Consequences can be devastating: suicide rate is 10-15%[1]

Preventative care is key is to mitigating episode severity…



but we lack a *passive* and inexpensive way to monitor mood symptom progression

Automatic speech-based monitoring is promising, but most research has used data from structured conversations – not "in-the-wild" speech

Objectives

Investigate how **interaction context** (clinical vs "in-the-wild") **influences the utility of speech and language features** for mood detection

Develop method to **detect mood severity from "in-the-wild" speech**

Data

PRIORI Dataset

Smartphone conversations from 51 individuals with BD collected over a period of 6-12 months

Transcripts obtained with Automatic Speech Recognition (ASR) model

Assessment calls: weekly clinical interviews to assess mood severity

Personal calls: everyday, "in-the-wild" calls (only use from *day of* assessment)

Subset of data used in this study:

	Train (35 subjects)	**Test** (12 subjects)
Assessment	131 / 155 / 286	67 / 100 / 167
Personal	88 / 127 / 215	50 / 87 / 137

Euthymic (asymptomatic) | Depressed

Methods

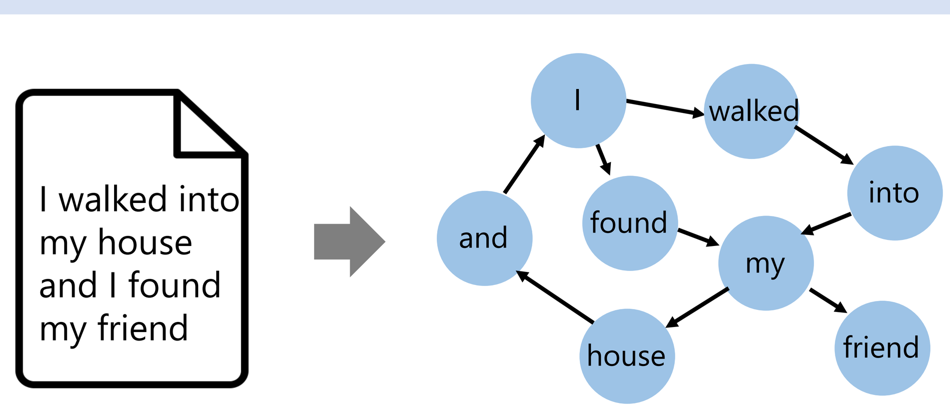
Feature Extraction

Linguistic Style

Complexity + Verbosity
- Word count, syllable count, etc.

Syntax
- Part-of-speech, verb tenses

Graph Analysis[2]
- Node count, loop count, etc.



Transcript — Speech Graph

Semantic content

LIWC Psychological Categories[3]
- Emotions, biological processes, etc.

TF-IDF unigrams & bigrams
- 20,000+ from full PRIORI dataset

Speech and Non-verbal

Speech Intelligibility: ASR confidence

Non-verbal Exp: Laughter, Noise

Speaker Timing

Speaking Duration
- Words, phones, pauses, call

Speaking rate
- Per-second timing of words, phones, pauses; per-minute timing of segments

Speaking Quantity
- Counts of words, phones, pauses, etc.

Data Modeling

Linear regression because of limited dataset size + desire interpretability

Feature Selection

- Pearson Correlation Coefficient (PCC) to filter and rank features
- Select from ranked list with nested Leave-One-Subject-Out (LOSO) Cross-Validation (CV)

Evaluate with PCC + compute across 12 test subjects w/ LOSO CV

Results

Depression Severity Detection

	Assessment PCC	Personal PCC
All Features	**.64 ± .12**	**.32 ± .25**
Speaker Timing	.63 ± .15	-
Linguistic Style	.30 ± .38	-
LIWC Pysch. Categories	.22 ± .35	.29 ± .37
TF-IDF	.46 ± .22	-
Speech Intelligibility	-	.15 ± .40

'-' indicates that applying feature elimination step resulted in empty feature set for at least one training fold

Feature Analysis

What features are useful for clinical assessment data?

Speaker timing features gain utility from interview structure

- Total duration: 0.42 PCC (assess) vs. 0.11 PCC (personal)
- Segment Count: 0.37 PCC (assess) vs 0.07 PCC (personal)

Easier to identify keywords because conversation is focused on mood

Feature	β	Feature	β
yes	2.3 ± .49	people	.84 ± .16
good	-1.14 ± .35	bad	.61 ± .18
normal	-1.12 ± .28	hand	.60 ± .21
yeah	.93 ± .14	nope	-.56 ± .15
really bad	.90 ± .10	every day	.42 ± .33

Features selected for TF-IDF only model trained on assessment

What features are useful for "in-the-wild" data?

Significant Features	Assessment PCC	Personal PCC
negative emotion	**.25**	**.37**
laughter*	-.04	**.32**
ASR conf. med	-.07	**-.32**
anger	.04	**.31**
ASR conf. mean	-.08	**-.31**
anxiety	**.23**	**.30**
death	.12	**.30**

*ASR model typically output "laughter" when crying occurred

Conclusion

Utility of speech features depends on interaction context

- Timing and TF-IDF gain utility from clinical interview structure
- Emotional distress and personal concern are useful for personal data

Detect mood severity from "in-the-wild" data, demonstrating the potential for passive, smartphone-based monitoring of BD

Acknowledgement

This work was supported by the National Science Foundation (CAREER-1651740), NIMH R34MH100404, and the Heinz C Prechter Bipolar Research Fund and the Richard Tam Foundation at the University of Michigan.

References: [1] Hawton, Keith, et al., *The Journal of clinical psychiatry* (2005).[2] Mota et al., PLoS One 7(4), (2012) [3] Pennebaker et. al, LIWC2007 Manual (2007)