

WPI

DepreST-CAT: Retrospective Smartphone Call and Text Logs Collected During the COVID-19 Pandemic to Screen for Mental Illnesses

ML Tlachac, Bryant University

**Ricardo Flores, Miranda Reisch, Katie Houskeeper,
Elke Rundensteiner**, Worcester Polytechnic Institute

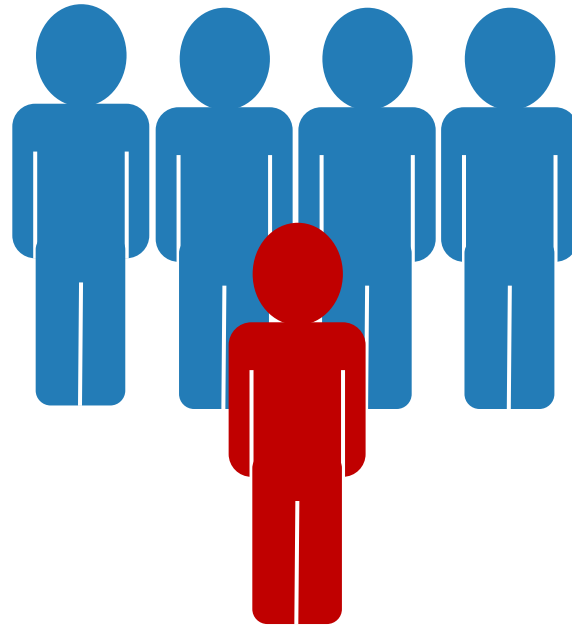
ACM IMWUT 6(2) & UbiComp 2022

Mental Illnesses are Prevalent and Costly

Annually, more than

1 in 5

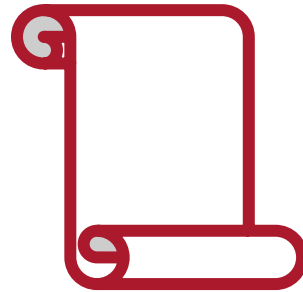
U.S. **adults** experience
mental illnesses



Mental Illness Screening



Interview



Survey



App

DepreST- Call And Text logs (DepreST-CAT)

- Goal: A dataset of mobile communication logs for passive and unbiased mental illness screening
- Data: Retrospective call and text logs with depression and anxiety screening labels
- Population: 369 crowdsourced participants from Prolific
- Collection: Approximately a year into COVID-19
- Analysis: Classification of log time series

DepreST Crowd-Sourced Data Collection

Study Info

EMV Study Overview

Study Overview

Goal: The goal of this project is to build an AI tool to predict mental health. The more data you share, the more useful our tool will be.

Procedure: This survey will only take around **4 minutes** to complete. You will be asked to answer survey questions, record samples of your voice, and share phone data, such as text logs and messages.

Privacy: All information you give will be stored **anonymously** on a secure server and will not be tied to you.

Voluntary/Risk: You can share as much or little data as you would like and you may stop the study at any time.

- WPI 2020 -

I AGREE

Upload Phone Logs

EMV Phone Data

Accepting the permissions will allow the

Allow **Mental Health Detection...** to access your contacts?

Allow

Deny

NEXT

Depression Survey

EMV Survey 1

Over the last 2 weeks, how often have you been bothered by any of the following problems?

0 - Not At all
1 - Several Days
2 - More than Half the Days
3 - Nearly Every Day

1. Little interest or pleasure in doing things

0 1 2 3

2. Feeling down, depressed or hopeless

0 1 2 3

3. Trouble falling asleep, staying asleep, or sleeping too much

0 1 2 3

4. Feeling tired or having little energy

0 1 2 3

Swipe down to scroll

SUBMIT

Anxiety Survey

EMV Survey 2

Over the last 2 weeks, how often have you been bothered by the following problems?

0 - Not At all
1 - Several Days
2 - More than Half the Days
3 - Nearly Every Day

1. Feeling nervous, anxious or on edge

0 1 2 3

2. Not being able to stop or control worrying

0 1 2 3

3. Worrying too much about different things

0 1 2 3

4. Trouble relaxing

0 1 2 3

SUBMIT

Demographics

EMV Demographics

1. What is your gender?

Man Woman Other

2. What is your age?

18-23
 24-39
 40-55
 56+
 Prefer not to answer

3. Are you a student?

Yes, I am a undergrad student
 Yes, I am a graduate student
 Yes, I am a student (Other)
 No, I am not a student
 Prefer not to answer

4. Have you received treatment for depression?

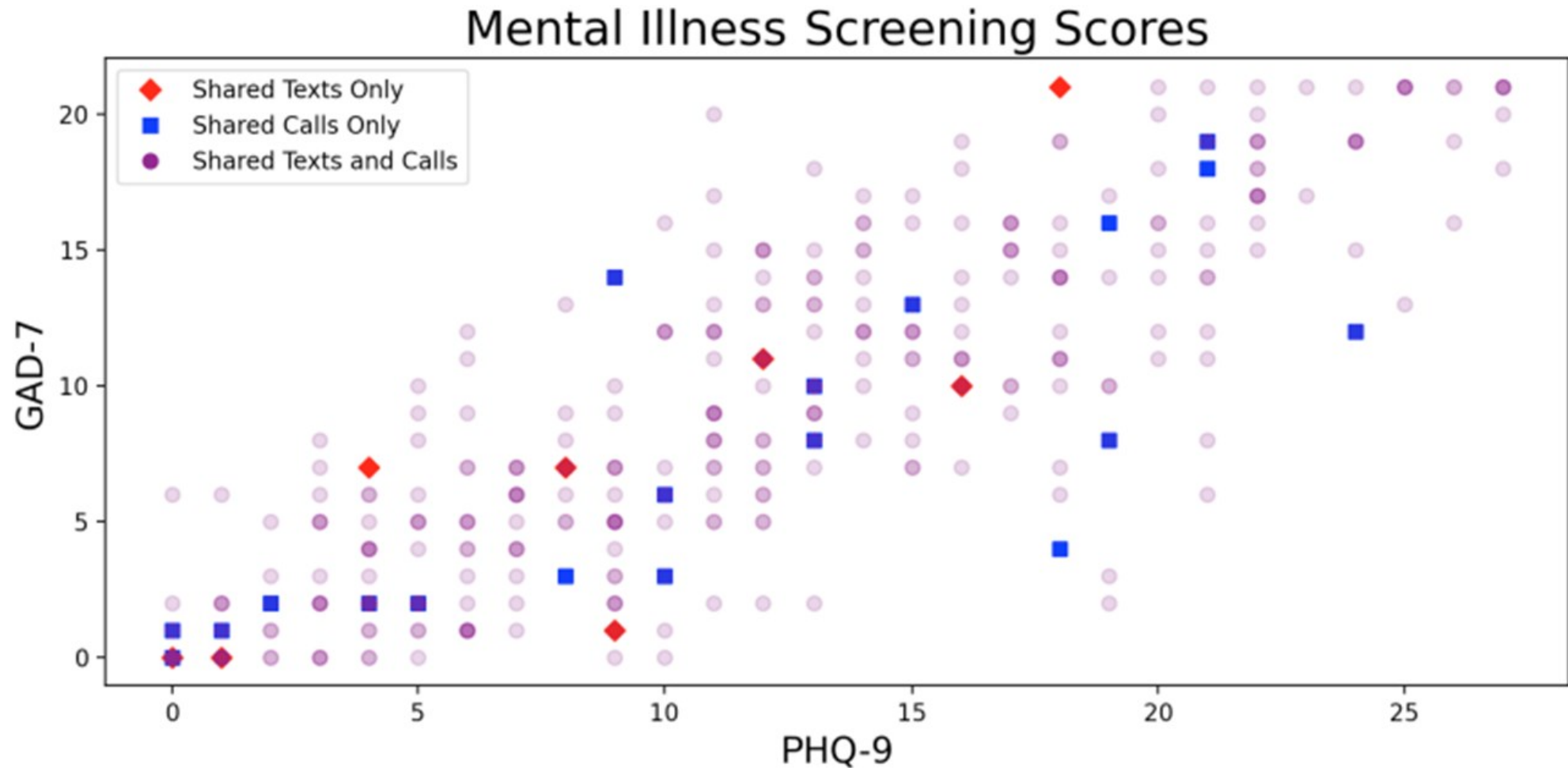
Yes No Prefer not to answer

5. How do you identify yourself?
Choose all that apply

SUBMIT

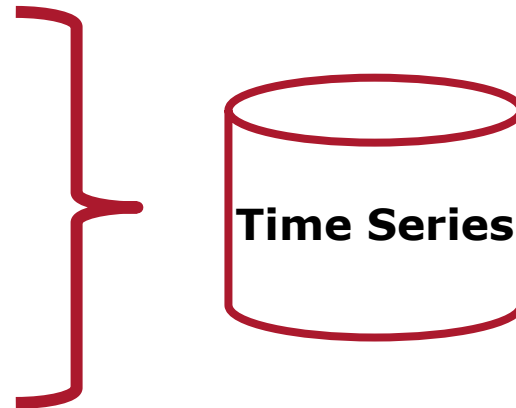
Mental illness labels for the data

369 DepreST-CAT Participants

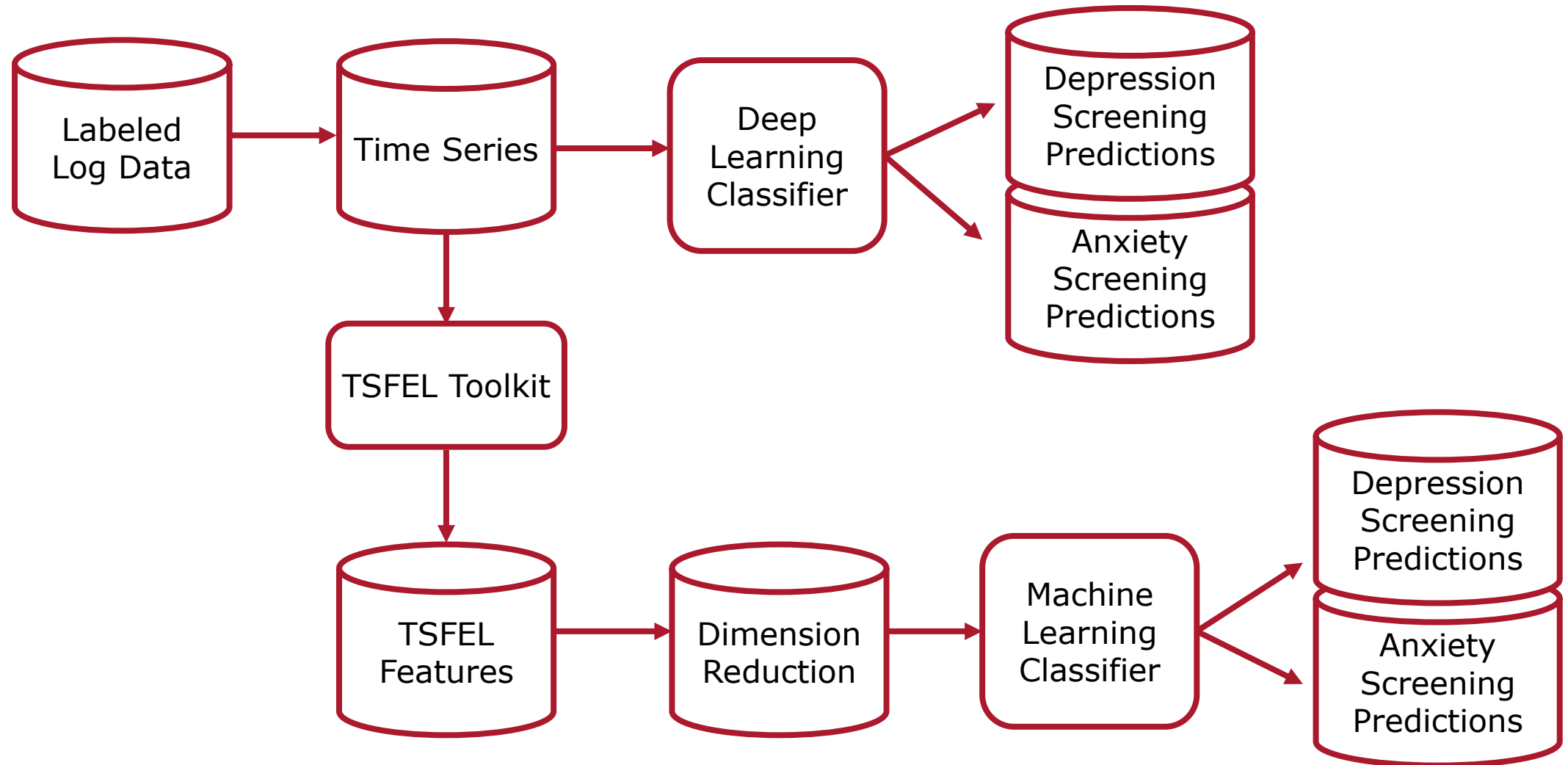


696,608 Logs Shared

- 143,280 incoming calls logs
- 98,247 outgoing call logs
- 368,807 incoming text logs
- 76,118 outgoing text logs
- 5,186 calendar logs
- 4,970 contact logs

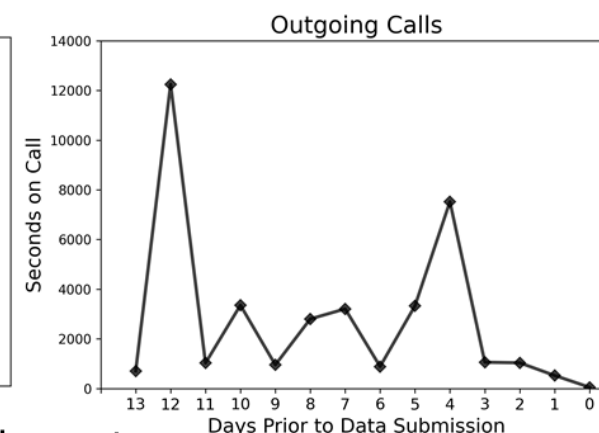
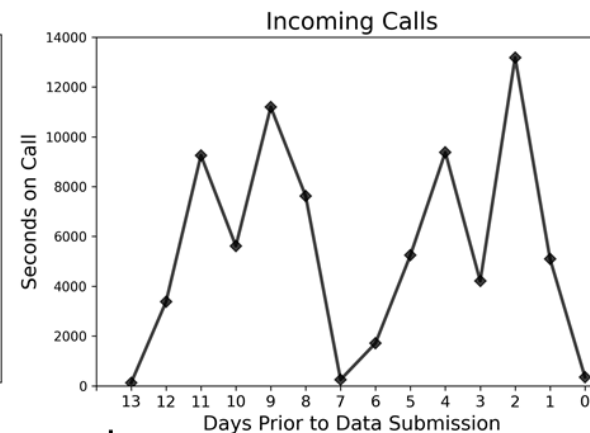
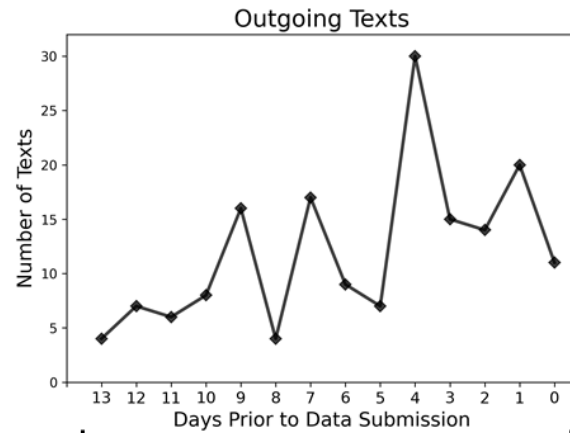
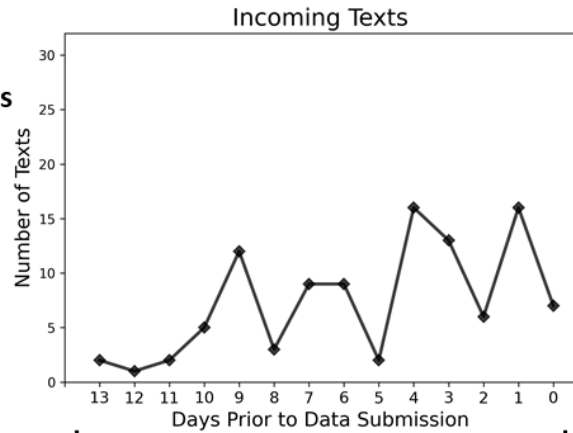


Methodology: Screening with DepreST-CAT Logs



Data Preparation for Machine Learning

Step 1:
Create
Time Series



Step 2:
Extract
Features

[|energy|, ..., min, ..., ZCR]
[1119, ..., 1, ..., 0]

[|energy|, ..., min, ..., ZCR]
[2698, ..., 4, ..., 0]

[|energy|, ..., min, ..., ZCR]
[6.5(10^8), ..., 119, ..., 0]

[|energy|, ..., min, ..., ZCR]
[2.5(10^8), ..., 50, ..., 0]

Step 3:
Concatenate
Features

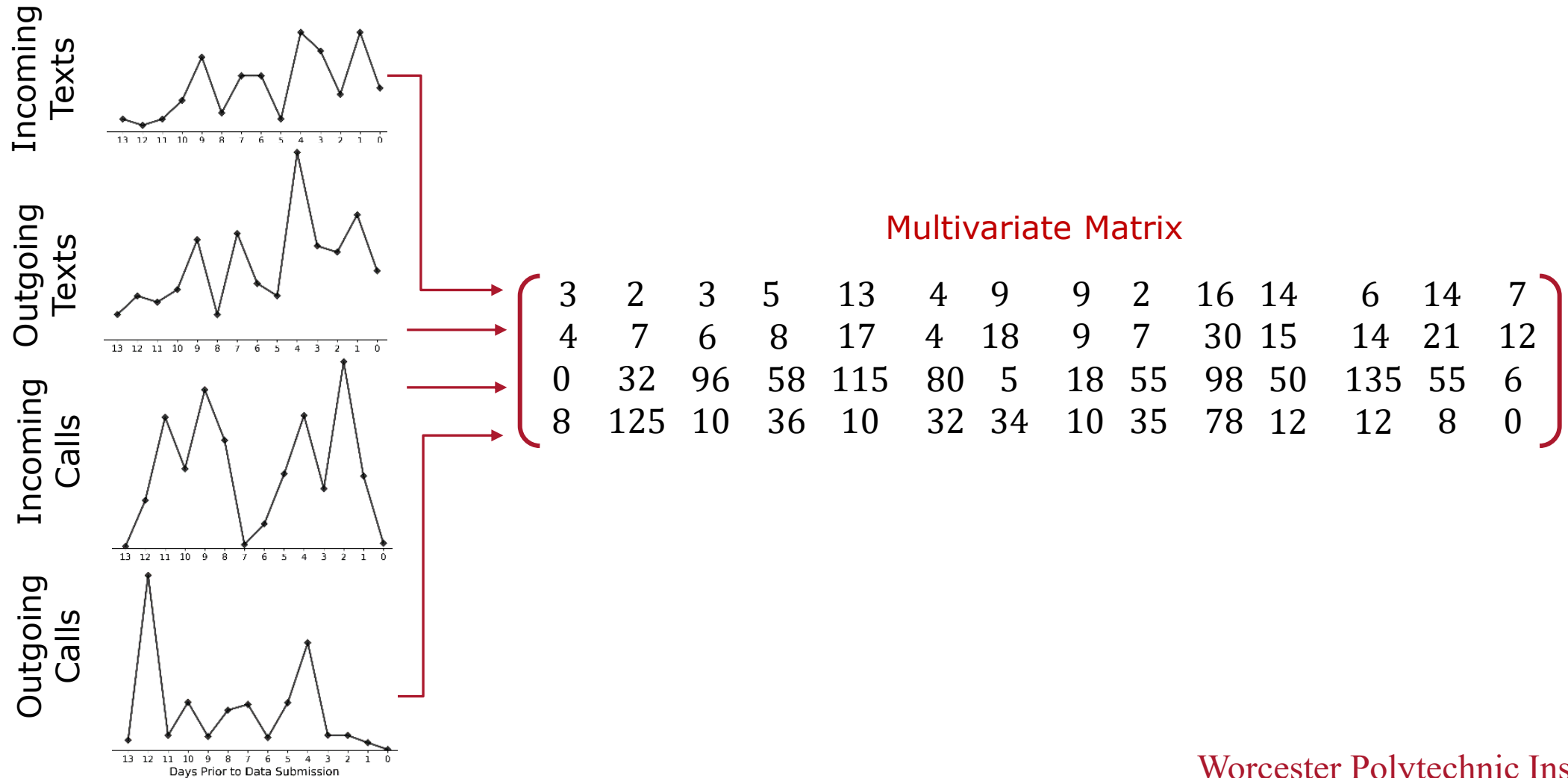
[1119, ..., 1, ..., 0,

2698, ..., 4, ..., 0,

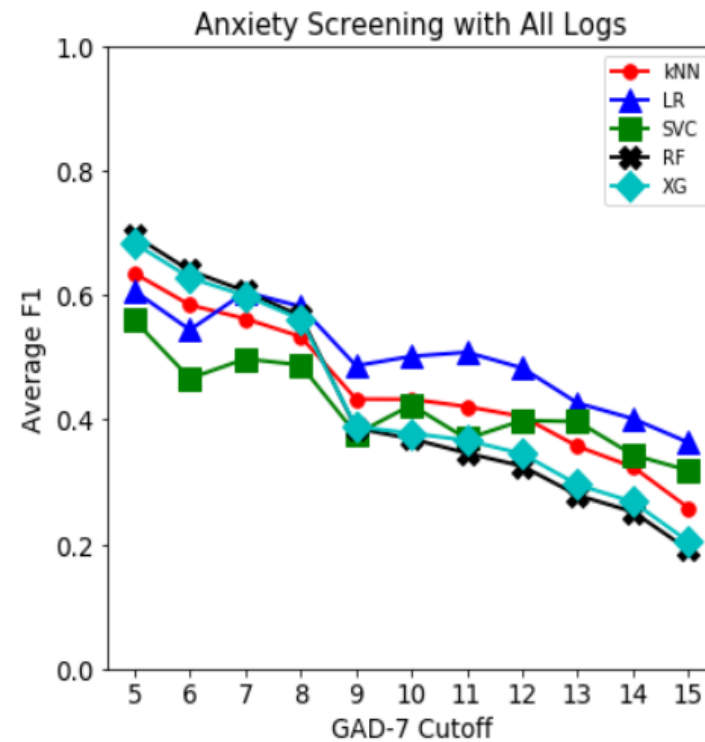
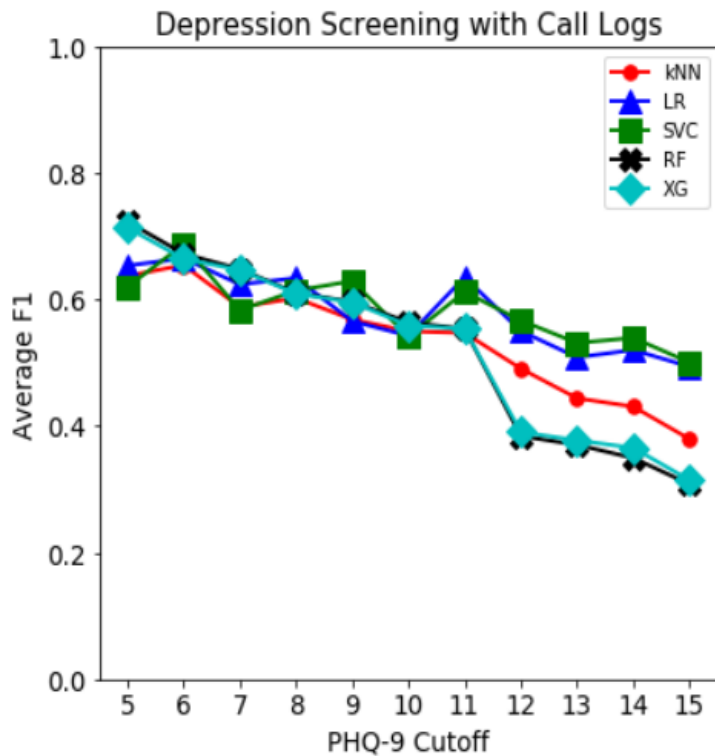
6.5(10^8), ..., 119, ..., 0,

2.5(10^8), ..., 50, ..., 0]

Data Preparation for Deep Learning

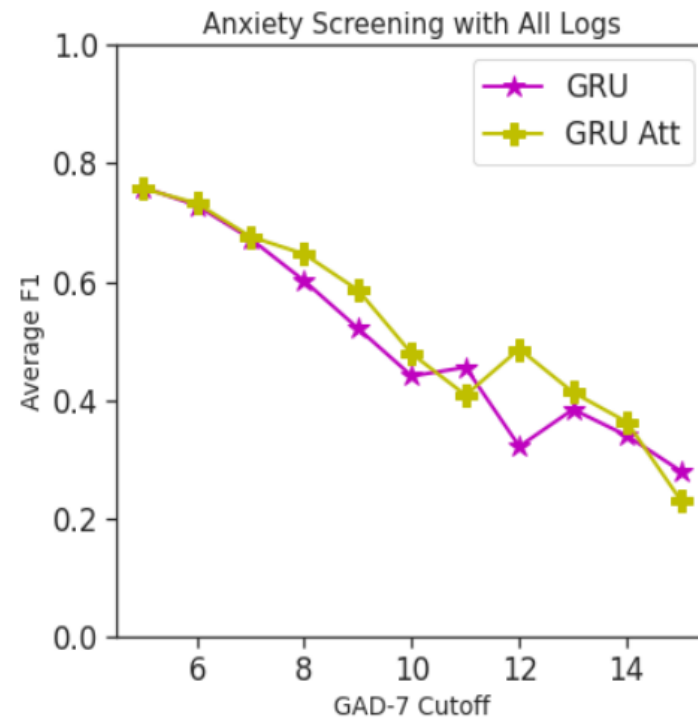
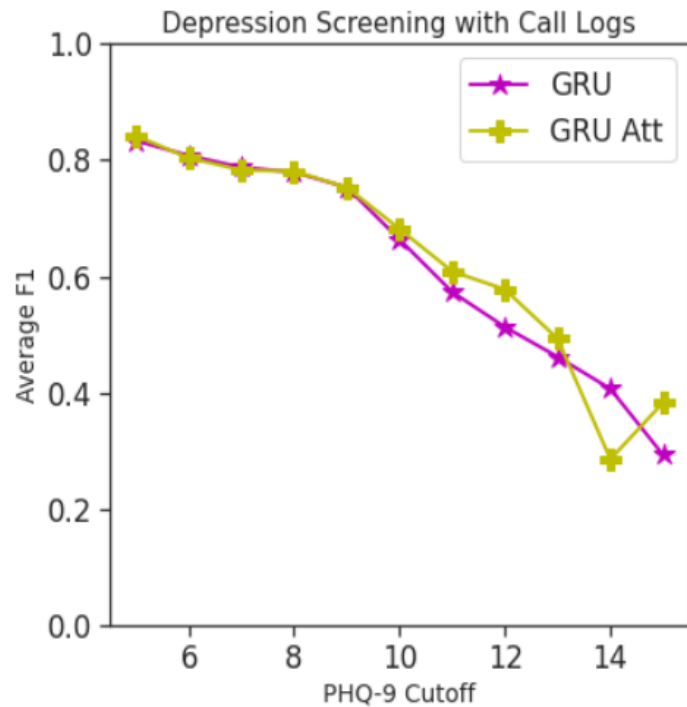


Machine Learning Results: Mental Illness Screening



- Models only need
 - two weeks of data
 - one principal component
- Equally good at mild depression and mild anxiety screening
- Better at depression screening than anxiety screening for higher cutoffs

Deep Learning Results: Mental Illness Screening



- Models only need two weeks of data
- Slightly better at depression screening than anxiety screening
- Deep learning better for lower cutoffs than machine learning

DepreST-CAT Log Availability



DepreST-CAT

The DepreST Call and Text (DepreST-CAT) dataset was collected between December 2020 and April 2021 by Miranda Reisch, ML Tlachac, and Prof. Rundensteiner. The DepreST-CAT dataset contains retrospective call and text logs labeled with demographics, PHQ-9 depression screening scores, and GAD-7 anxiety screening scores from over 369 Prolific crowd-sourced participants.

- Paper: <https://dl.acm.org/doi/10.1145/3534596>
- Data: <https://github.com/mltlachac/DepreST-CAT>

Thank You

This work was supported by:

- US Dep. of Education P200A180088
- AFRI Grant 1023720NSF
- Fulbright Foreign Student Program
- Chilean National Agency for Research and Development Scholarship Program
- DraftKings Fellowship
- Data Science Department at WPI

Results were obtained in part using an HPC from NSF MRI grant DMS-1337943 to WPI

