
Predicting Mental Health

(and identifying key features)



Sonia Sharapova • 05.20.2024

Motivation

- Current mental health crisis
- Increase in mental health issues among adolescents



Impacts

- Early detection
- Prevention
- Tailored support and care
- Supportive diagnostic
- Better mental health resources



Datasets

Datasets

Borealis

- University Student Mental Health
- Collected 2022
- 1192 Participants

EXTRA:

OpenICPR

- Social Media and Mental Health
- Collected 2022
- 581 Participants

Kaggle

- Social Media and Mental Health
- Mental Health in Tech

Work Done

Data Processing



-
- Student mental health survey
 - Preprocessing:
 - Dropped extraneous fields (answer specifications)
 - Removing NaN rows
 - Dropped invalid responses
 - Even scaling

Responses of interest:

- Mental Health Metrics
-

Borealis

- **Preprocessing:** Categorize mental health factors
 - Drop irrelevant features
 - Binary Classification for specific symptoms
 - Survey: count responses: higher response = more relation
-

Classification Models

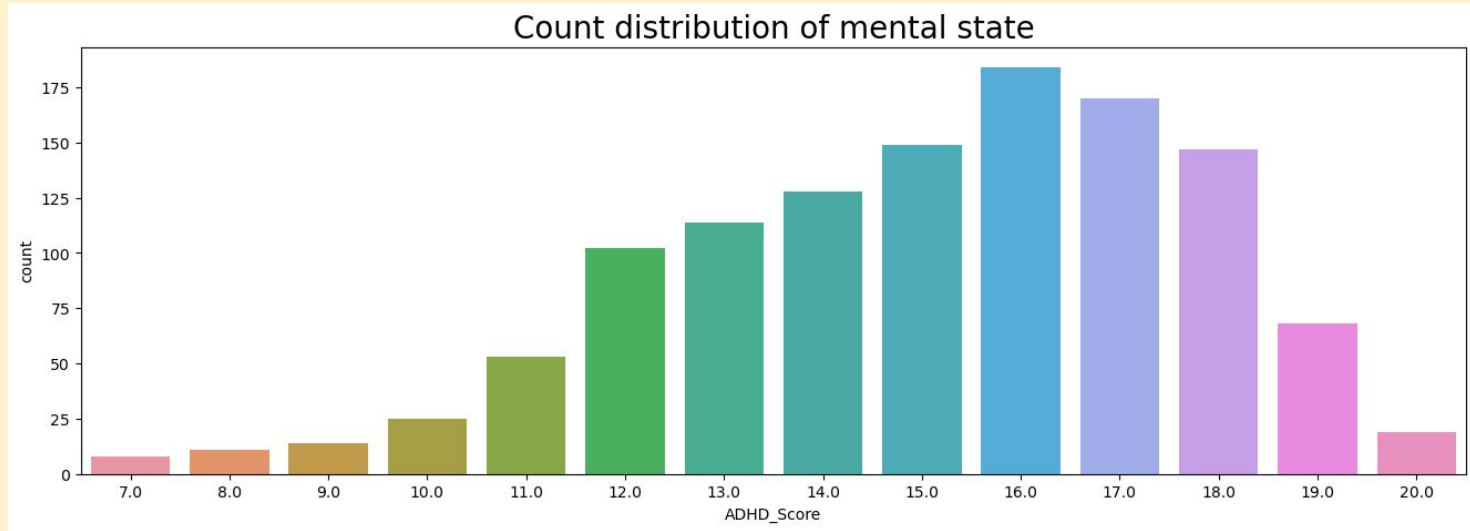
- KNN
- SVM (Linear/Radial)
- Gaussian
- Decision Tree
- Random Forest
- Neural Network (MLP Classifier)
- AdaBoost

Evaluation Metrics

- F1 Score, ROC-AUC, Accuracy
 - ADHD
 - Anxiety
 - Depression
 - General
-

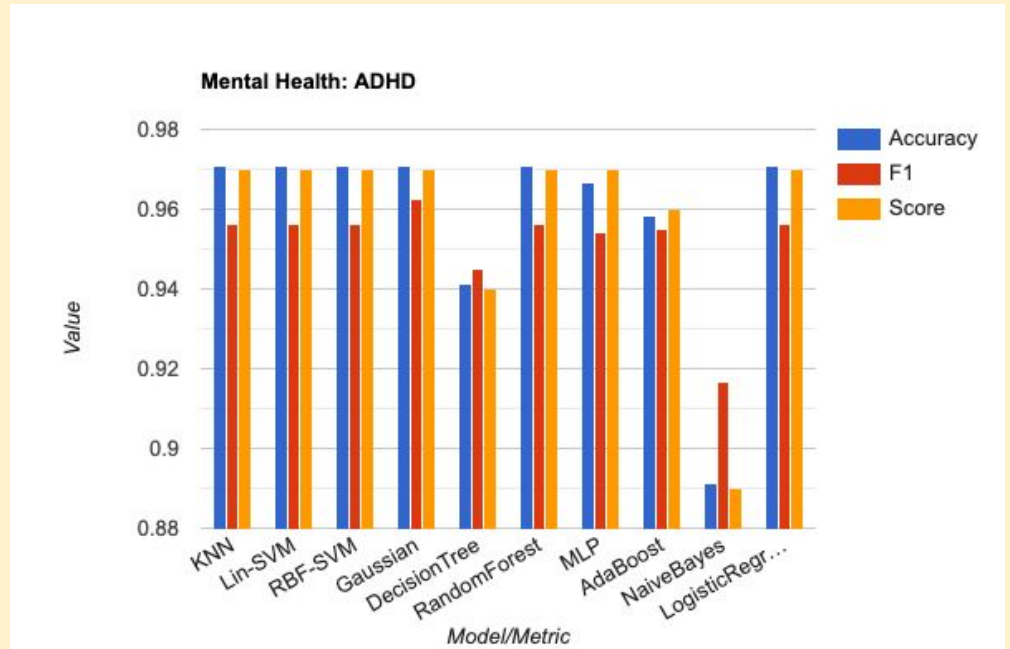
ADHD

- Markers of concentration issues



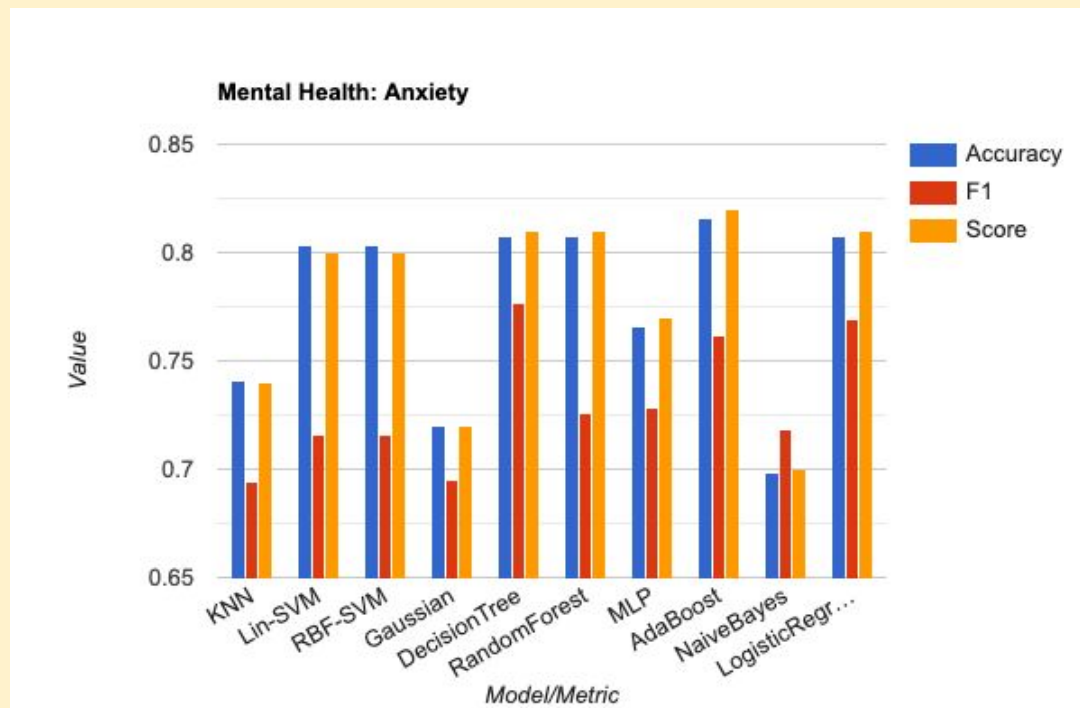
ADHD

- Best: Gaussian Process
- Overall score: 97%



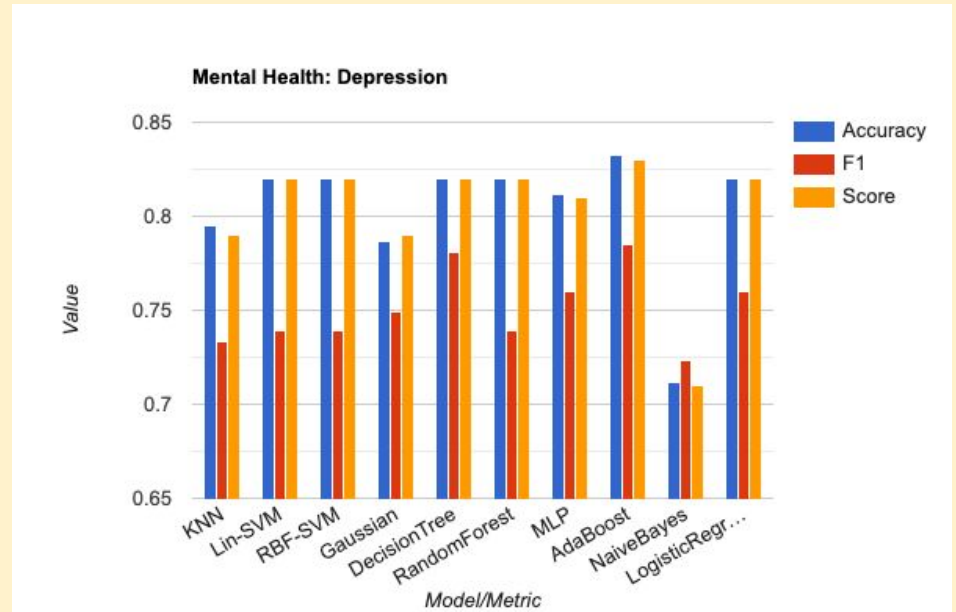
Anxiety

- Best: AdaBoost
- Overall Score: 83%



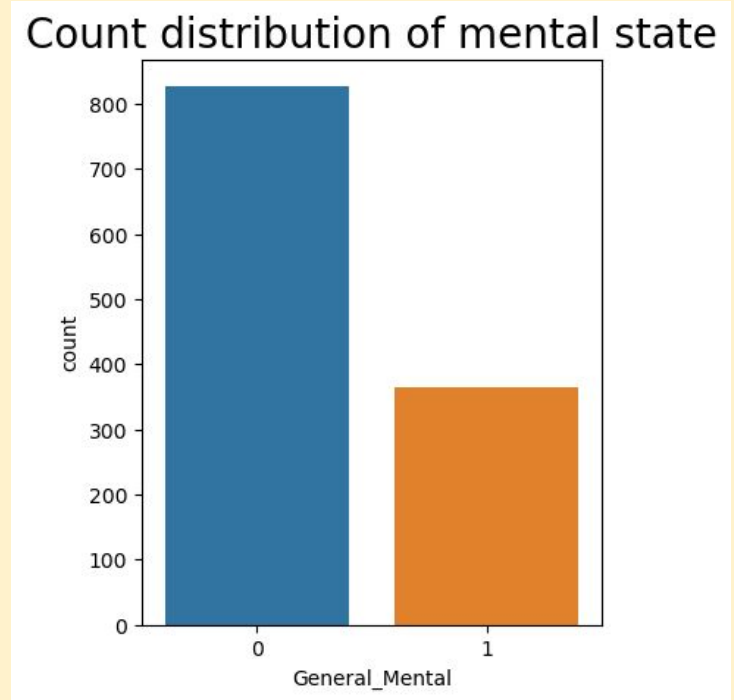
Depression

- Best: AdaBoost
- Overall score 83%



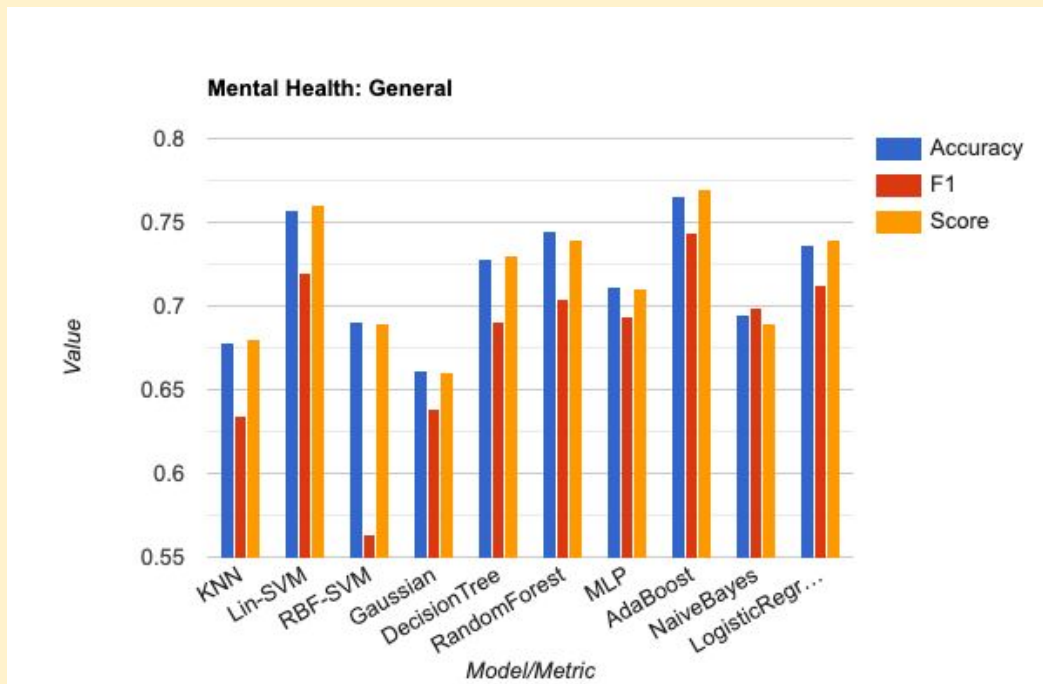
General Mental Health

- Markers of mental health
- Methods:
 - Sum over responses, take mean.
Above the mean = 1, below mean = 0
 - Mine: 1 if present, 0 if not



General Mental Health

- Best: AdaBoost
- Overall score 77%



Top Features

Rank	ADHD	Anxiety	Depression	General
1	Disability	Disability	Disability	CATS_6
2	Age	Mild exercise	Age	Hours_sleep
3	Mild exercise	SSS_6	Hours_sleep	Mild exercise
4	Hobbies_Time_6	Ethnicity	Hobbies_Time_6	SSS_11
5	Hobbies_Imp_2	Age	Year_credits	Age
6	CATS_5	Hours_sleep	Moderate exercise	Moderate exercise
7	Mindfulness_freq	Moderate exercise	Mild exercise	CATS_4
8	CATS_2	Hobbies_Time_4	Living	SSS_8
9	CATS_8	CATS_7	CATS_7	Hobbies_Imp_4
10	CATS_6	Program	CATS_4	CATS_2

Challenges

Finding Data

- Categorical Variables
 - Biased towards university students
 - Cleaning survey data
 - Mental health is very nuanced
-

Future Work

Further Work

- Have a more precise metric for mental issue at hand
 - Compare with more previous works and papers
 - Collect more data on social media usage to see its correlation with mental health
-

Thank You!!

OPENICPSR

- Mental health survey, Feb. 20, 2022
- Preprocessing:
 - Dropped timing fields (recorded time of click)
 - Replacing NaN fields with 0
 - Dummies for categorical variables

Responses of interest:

- Depression within the last 12 months
 - Depression diagnosis
 - Feeling down, depressed, or hopeless.
-