Machine Learning
Regression Model

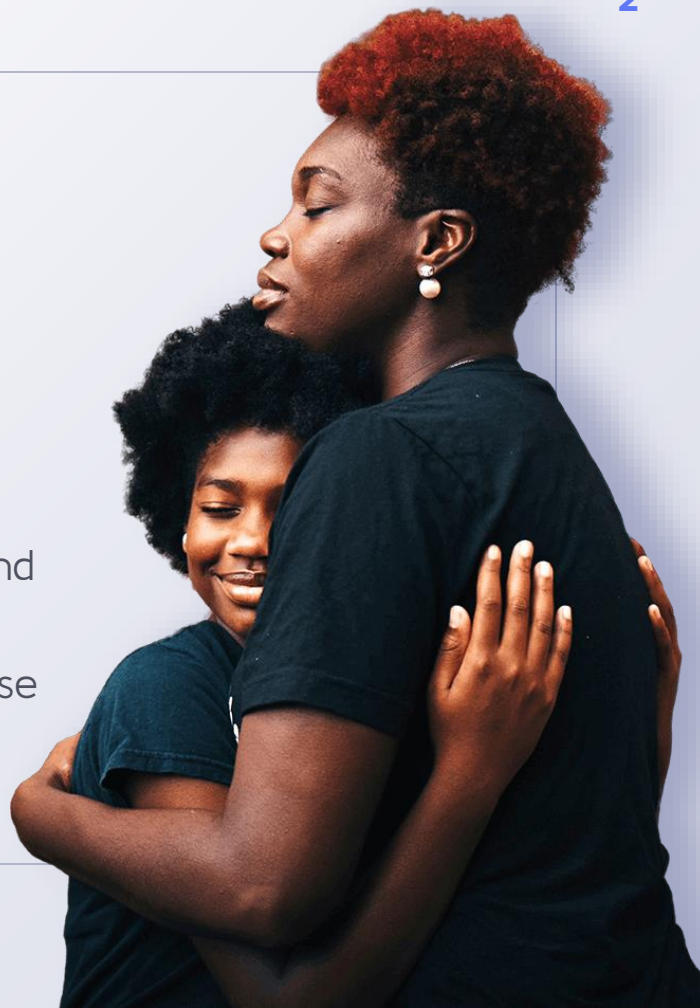
Mental Illness in Communities Estimator

September 2020

Francis Morales

Project Objective

Estimate the percentaje of population with mental illnesses in communities to help local governments and health organizations raise awareness and allocate resources in mental care to build better lives for those affected.



Mental Illness in the U.S.

What is a mental illness?

Mental illness is defined as having a diagnosable mental, behavioral, or emotional disorder, other than a developmental or substance use disorder

1 in **5** U.S. adults experience mental illness (19%)

1 in **6** U.S. youth experience mental illness (6–17 years)

90% of people who die by suicide have experienced symptoms of a mental health condition

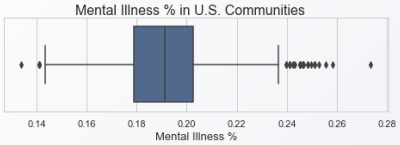
11 Years

The average delay between symptom onset and treatment

Sources:
https://nami.org/About-Mental-Illness
https://www.samhsa.gov/
https://www.psychiatry.org/

DATASET

- **820** counties from 48 states (26% counties) – excluding Hawaii and Alaska
- **30** different features: economic, demographics, family structure, education and weather
- **3** data sources: API & Data Download



United States Census Bureau

NOAA NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION

SAMHSA Substance Abuse and Mental Health Services Administration

METHODOLOGY & MODEL OVERVIEW

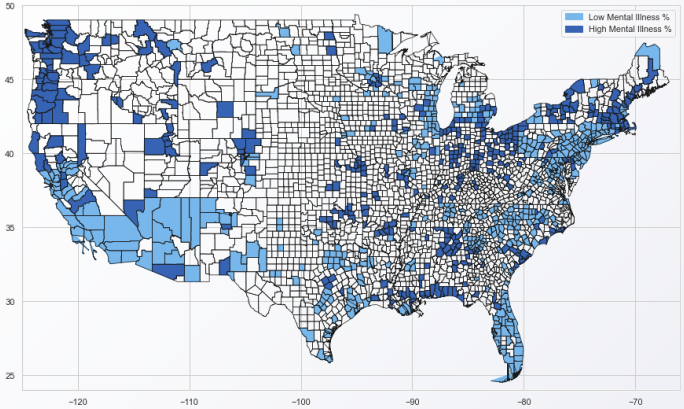
Data Sourcing & Scrubbing ⟩⟩⟩ Exploratory Data Analysis ⟩⟩⟩ Model Training & Evaluation ⟩⟩⟩ Delivery of Results

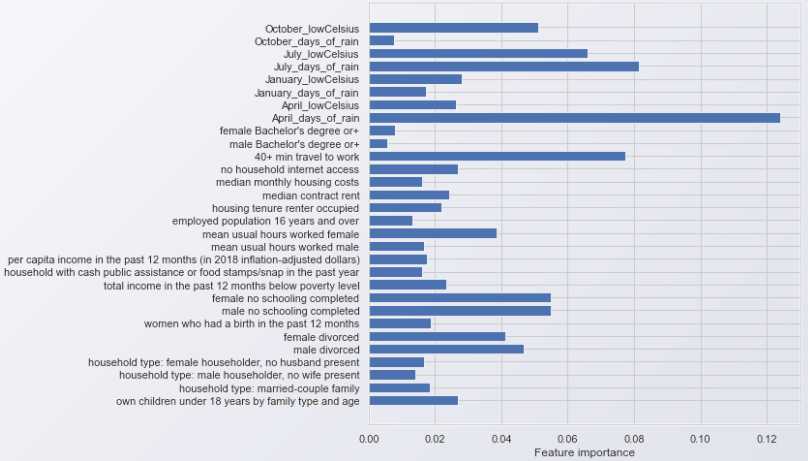
Model Selection: Support Vector Regressor Model with a RMSE of 1.54% and R2 of 0.445

Regression Model	RMSE	R2
Baseline	0.021	-0.001
SVR	0.0154	0.445
Random Forest Regr.	0.0167	0.348
XGBRegressor	0.0156	0.432
Deep Neural Network	0.017	0.27

DATA INSIGHTS

MENTAL ILLNESS IN COUNTIES

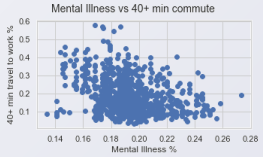


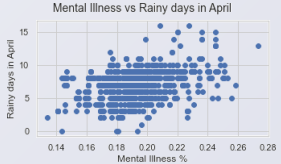


Counties with **higher** rates of long **commutes** show **less** % of **mental illness**

Counties with **higher** number of **rainy days** in January, April and October show **higher** % of **mental illness**

Counties were women work **less hours** show **higher** % of **mental illness**





RECOMMENDATIONS / NEXT STEPS

- Tune model parameters to improve RMSE and R2 before moving model to production
- Review dataset to eliminate noise from unnecessary variables
- Explore further insights from independent variables

Thanks!

Any questions?

You can find me at https://www.linkedin.com/in/francis-morales-89b22b66/