

Communities and Crime: Using an Ensemble of ML Algorithms to Predict Crime

By Jake Blancher

?



The Data

1,994

Instances

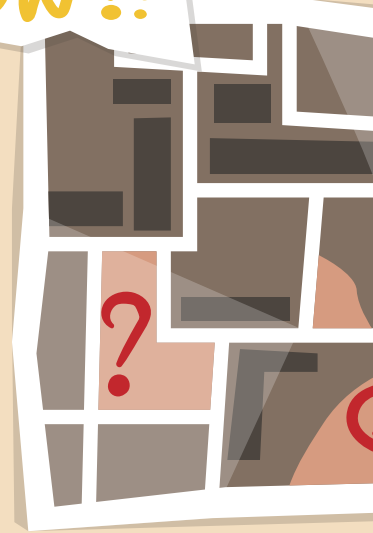
127

Features

1

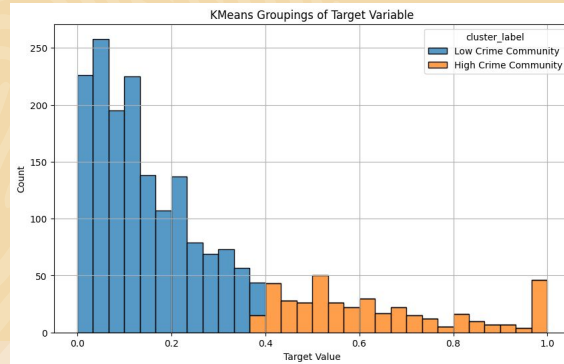
Target Value (violent crimes per capita), continuous and normalized

HOW?!



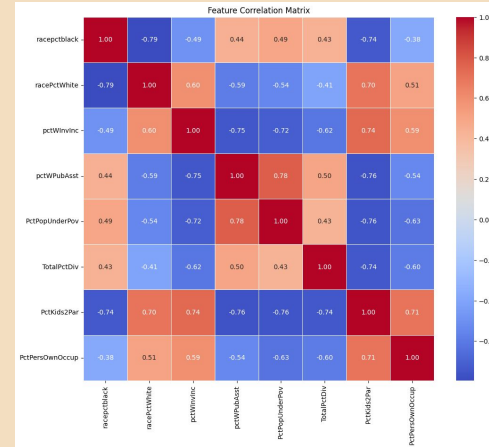
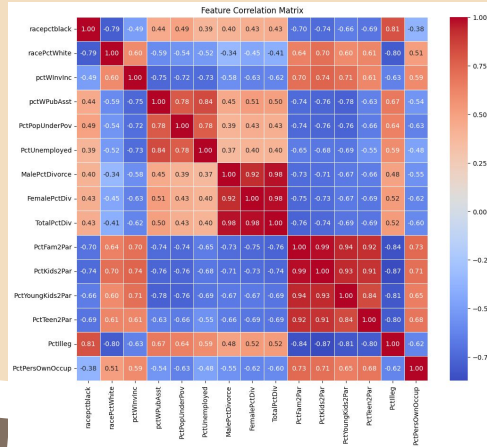
Binning

- Used K-means with $k=2$ to separate into high and low crime
- Frequently ran into low recall, needed to weight class 1



Feature Selection

- Didn't use PCA, need simplicity for GUI
- Removed features with missing data or ones that measure non-numeric data (like state or city name)
- Ran correlation between each feature and the target value, removing all those with an r value between -0.5 and 0.5
- Ran pairwise correlation, eliminating less robust multi collinear features (eg male divorce rate, female divorce rate, and total divorce rate)



11:15 pm

Classification Algorithms

01

Logistic Regression

Liblinear solver, 4:1 weight

02

Support Vector Classifier (SVC)

You can describe the topic of the section here

03

Neural Network

1 layer, sigmoid activation,
binary cross entropy loss, 1001
total parameters

04

Deep Neural Network (DNN)

3 hidden layers, ReLu between
them, sigmoid output layer,
43521 parameters

Ensemble

Logistic Regression

86.0% accurate

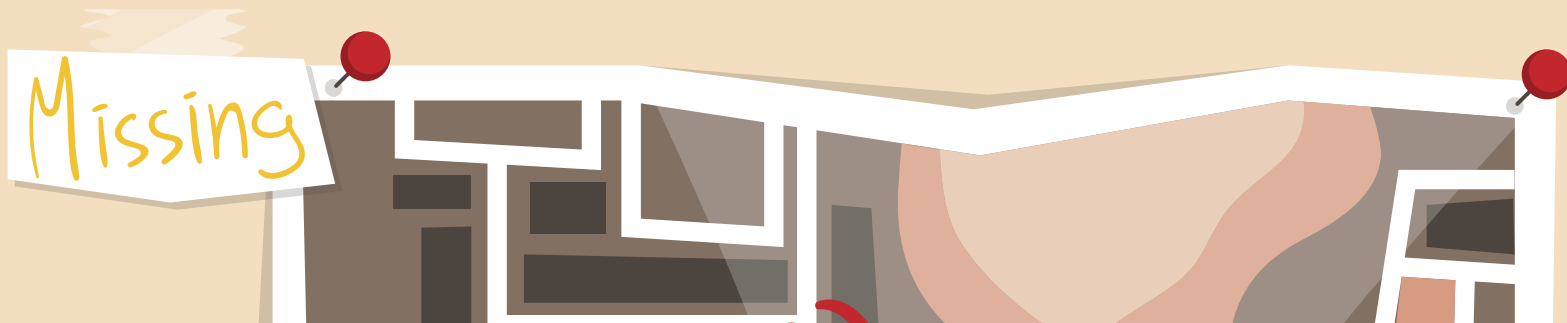
SVC

84.0% accurate

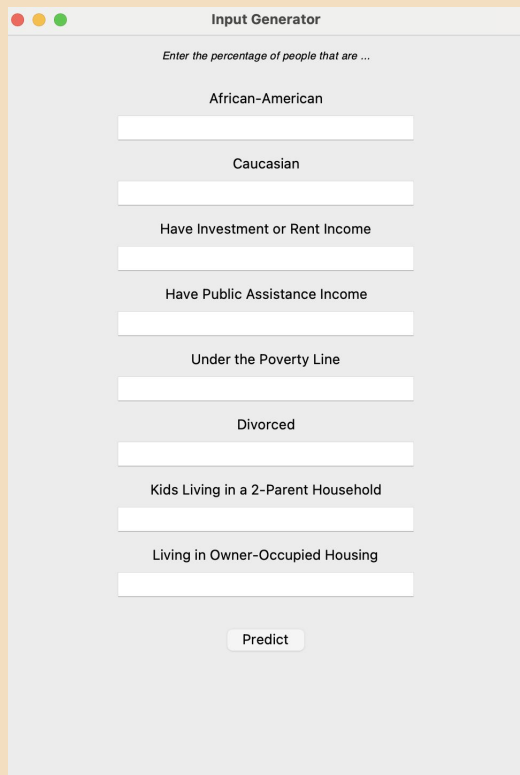
DNN

86.7% accurate

88.5% accurate, almost 2% better than our best model, DNN



Graphical User Interface (Tkinter)



The image shows a Tkinter window titled "Input Generator". It contains a series of input fields for demographic and socioeconomic data, each preceded by a label. The labels are: "Enter the percentage of people that are ...", "African-American", "Caucasian", "Have Investment or Rent Income", "Have Public Assistance Income", "Under the Poverty Line", "Divorced", "Kids Living in a 2-Parent Household", and "Living in Owner-Occupied Housing". Each label is followed by a white rectangular input field. At the bottom of the window is a "Predict" button.

Input Generator

Enter the percentage of people that are ...

African-American

Caucasian

Have Investment or Rent Income

Have Public Assistance Income

Under the Poverty Line

Divorced

Kids Living in a 2-Parent Household

Living in Owner-Occupied Housing

Predict



Discussion - Over-Policing



73.5%

Of investigatory stops target
Black Chicagoans

Despite representing only 8%
of the neighborhood

3.3X

Times more likely to have their
cars searched by police

As compared to White
residents

Cherrone, Heather. "Chicago Police More Likely to Use Force against Black Chicagoans: Watchdog." WTTW News, 1 Mar. 2022,
news.wttw.com/2022/03/01/chicago-police-more-likely-use-force-against-black-chicagoans-watchdog.