

Final Project

Kushal Ismael and Matt Kosko

6/23/2021

Introduction

Problem Statement

Target

[Put image here]

Plan

Outline

Contribution

- ▶ This data set does not have any published papers relevant to it (Redmond 2009)
 - ▶ So any analysis with this particular data set is new
- ▶ There is a large literature on the use of machine learning in so-called “predictive policing” (Hardyns and Rummens 2018)
- ▶ Many predictive policing systems use a single method ML method, like
- ▶ We have tried many different methods and compared them

Missing Data

- ▶ Data are “missing not at random” (Rubin 1976)
- ▶ Impute data using *MICE* algorithm (Van Buuren 2018)
 - ▶ Implemented as part of `MultipeImputer`
- ▶ Check analysis as part of robustness check

Preprocessing

- ▶ Filter features based on correlation

Solver

Hyperparameter Tuning

Results

References

Hardyns, Wim, and Anneleen Rummen. 2018. "Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges." *European Journal on Criminal Policy and Research* 24 (3). Springer: 201–18.

Redmond, Michael. 2009. "Communities and Crime." <https://archive.ics.uci.edu/ml/datasets/Communities+and+Crime>.

Rubin, Donald B. 1976. "Inference and Missing Data." *Biometrika* 63 (3). Oxford University Press: 581–92.

Van Buuren, Stef. 2018. *Flexible Imputation of Missing Data*. CRC press.