



Get Pumped Up! Sprint 2

Fatemeh Ebrahimi, Kateryna Ponomarova, James Donahue

Data Circle

May 21, 2025

Feature Engineering

A Dark Art

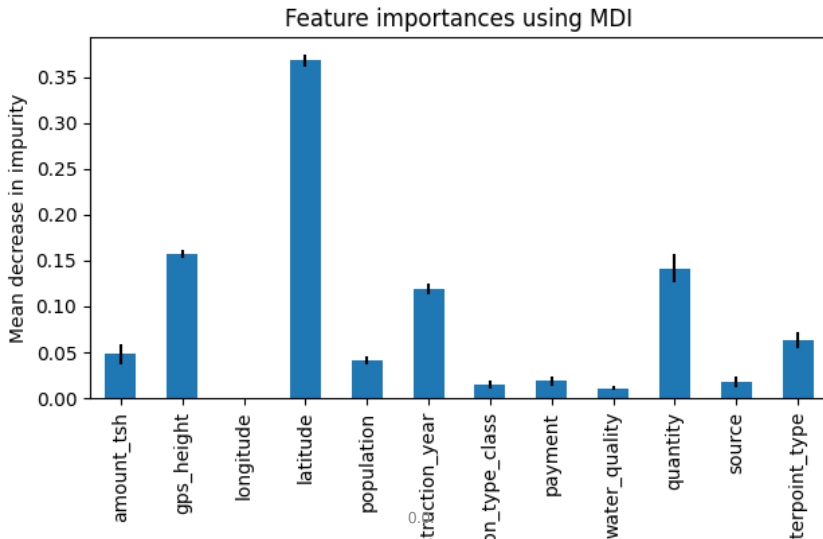
- Amount TSA - still there
- Population - binned
- Interactions - possible
- Geospatial analysis - impossible

Feature Importance

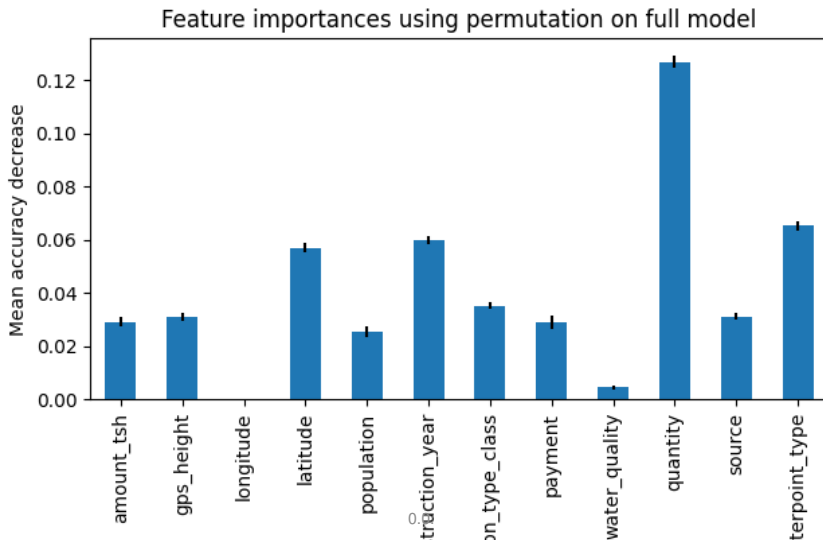
Some are more equal than others

- Latitude but not longitude
- Quantity, GPS height
- Not water quality, little on population

Feature Importance: MDI



Feature Importance: Permutations



Modeling

Too Logit to quit

- Performed reasonably well, compared to others we tried, and easy to interpret
- Still not predicting well with the minority class

Class	Precision	Recall	F1-score
functional	0.71	0.91	0.79
functional needs repair	0.62	0.03	0.06
non functional	0.78	0.60	0.68

Thank you!

Random Graphic

Might look familiar?

