Tanzania Water Pump Classification

Using classification models to determine water pump functionality.

Alex Beat

Business Overview

Problem: Predict which Tanzanian water pumps are functional, which need repairs, and which don't work at all.

Business Value: A smart understanding of which waterpoints will fail can improve maintenance operations and ensure that clean, potable water is available to communities across Tanzania. This understanding includes predictability based on location and types of pumps, for faster assessment and follow through.

Method

Tested three different algorithm models:

Ran several models to find best performance of classification.

Looked at most important features involved in classification from model evaluation.

Followed up with visuals to help explain each variables effect on class.

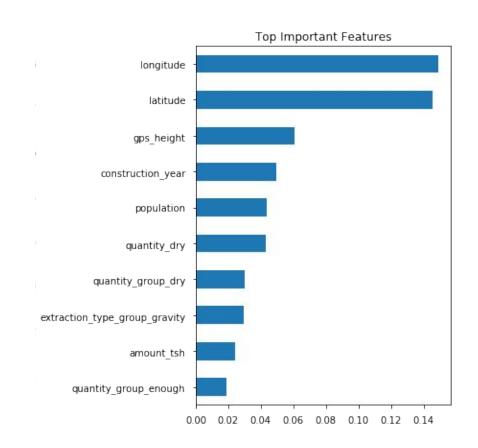
- Focused mainly on location, population and gravity well pump technology.

Important Features To prioritize.

RandomForest Model Accuracy: 0.79

Chosen features:

- Longitude & Latitude
- Population
- Extraction type: Gravity pump



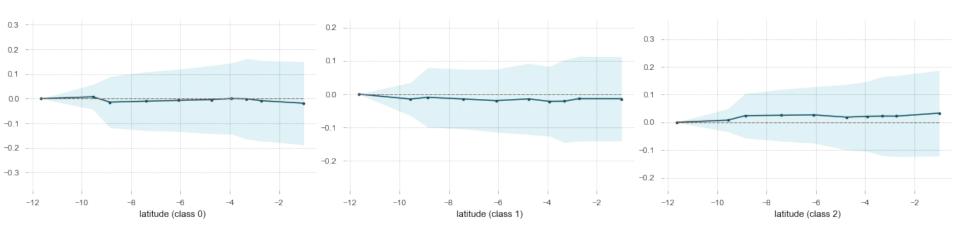


PDP for feature "latitude"

Number of unique grid points: 10

Nonfunctional pumps likely to be located:

- Latitudes between -9 to -1.
- Below 34 degree longitude and above 37 degrees longitude



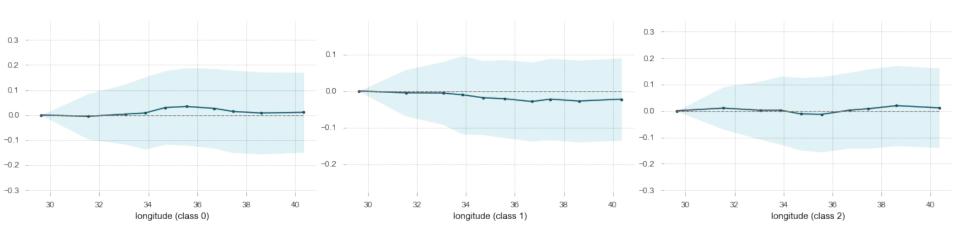


PDP for feature "longitude"

Number of unique grid points: 10

Nonfunctional pumps likely to be located:

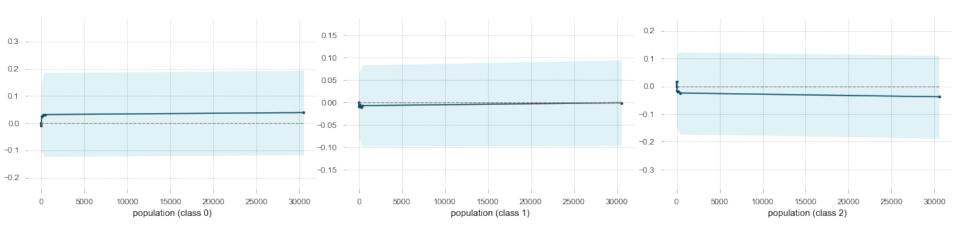
- Below 34 degree longitude and above 37 degrees longitude





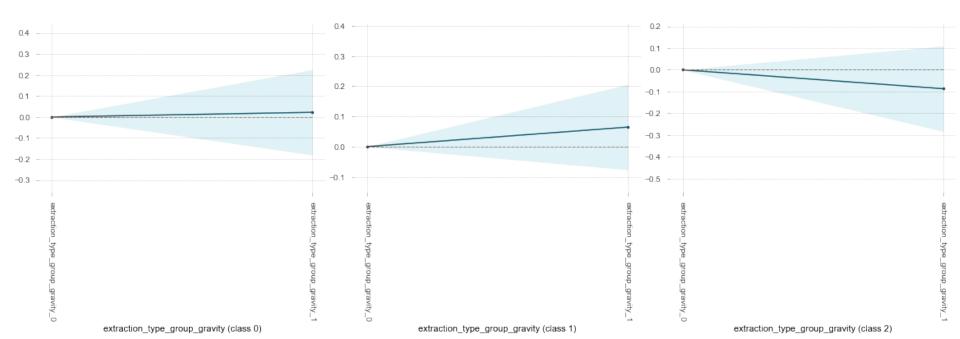
PDP for feature "population"

- Nonfunctional pumps likely to be located near smaller population concentration or zero population.
- Focus on those pumps first for replacement if not dry.
- Repairs: Pumps near 0 population and pumps near high population with high use.



Gravity extraction pumps

- High chance of needing repair
- Most likely functional.
- Check on those as priority.
- Build different type, more resilient pumps.



Summary & Recommendations

- Look for nonfunctional pumps:
 - Between latitudes -9 to -1.
 - Below 34 degree longitude and above 37 degrees longitude
- Check on gravity type pumps as priority.

High chance of needing repair but still functional.

Build different type, more resilient.

2

Most nonfunctional pumps in remote location - if needed for use, replace.

Most in need of repairs: Pumps near 0 population and pumps near high population with high use.

Future work

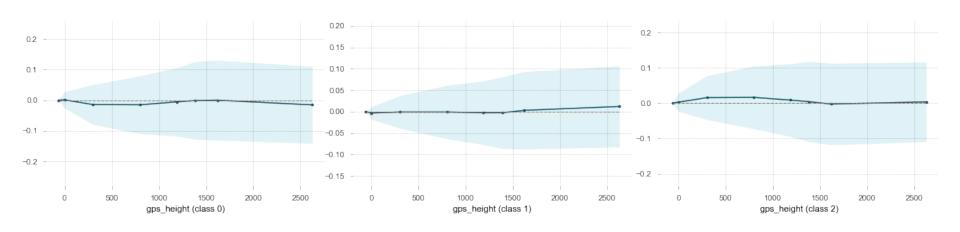
- Fix filler data in longitude, latitude for better accuracy of geo coordinates for nonfunctional pumps. (Same for population and amount of total static head.)
- Do the latitude and longitude degrees where most broken water pumps are found have to do with lack of proper authorities and management (business or political)?
- Compare that to data on public meetings held for different districts for better political insight.

Thank you.



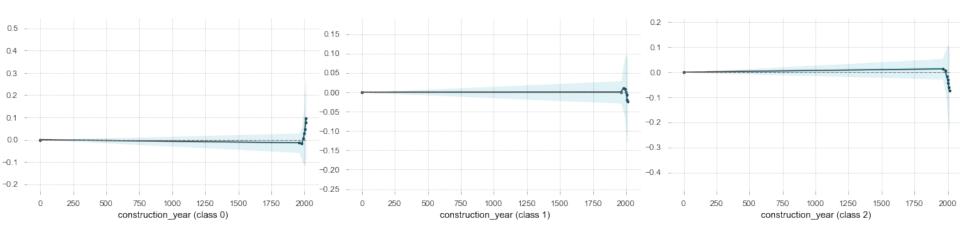
Appendix: Further Visuals of Features - GPS Height Sea Level

PDP for feature "gps_height"



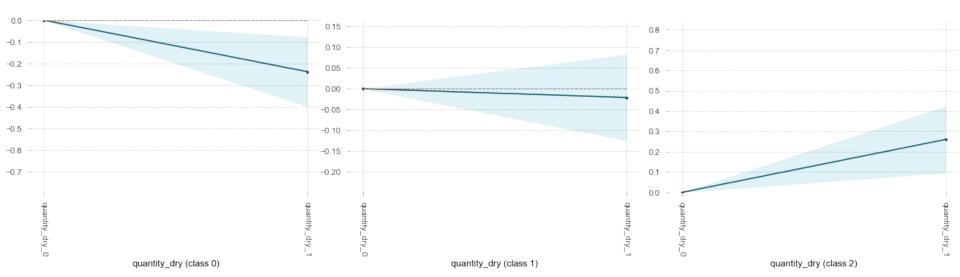


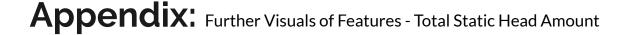
PDP for feature "construction_year"



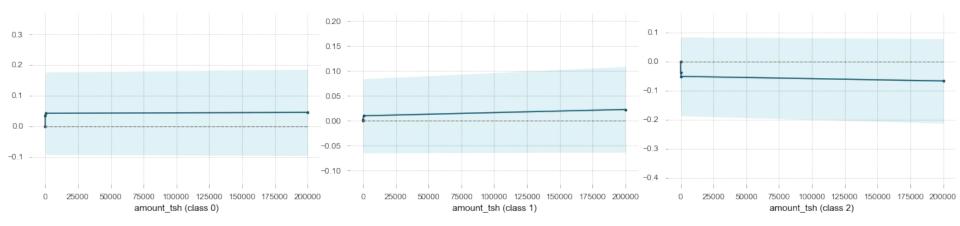
Appendix: Further Visuals of Features - Dry Wells

PDP for feature "quantity_dry"





PDP for feature "amount_tsh"



Appendix: Further Visuals of Features - Enough Water

PDP for feature "quantity_group_enough"

