

# Classification for Water Pumps in Tanzania

Max Zimmerman

# Pump it Up: Data Mining the Water Table

HOSTED BY DRIVEN DATA



Can you predict which water pumps are faulty?

Using data from [Taarifa](#) and the [Tanzanian Ministry of Water](#), can you predict which pumps are functional, which need some repairs, and which don't work at all? This is an intermediate-level practice competition. Predict one of these three classes based on a number of variables about what kind of pump is operating, when it was installed, and how it is managed. A smart understanding of which waterpoints will fail can improve maintenance operations and ensure that clean, potable water is available to communities across Tanzania.

---

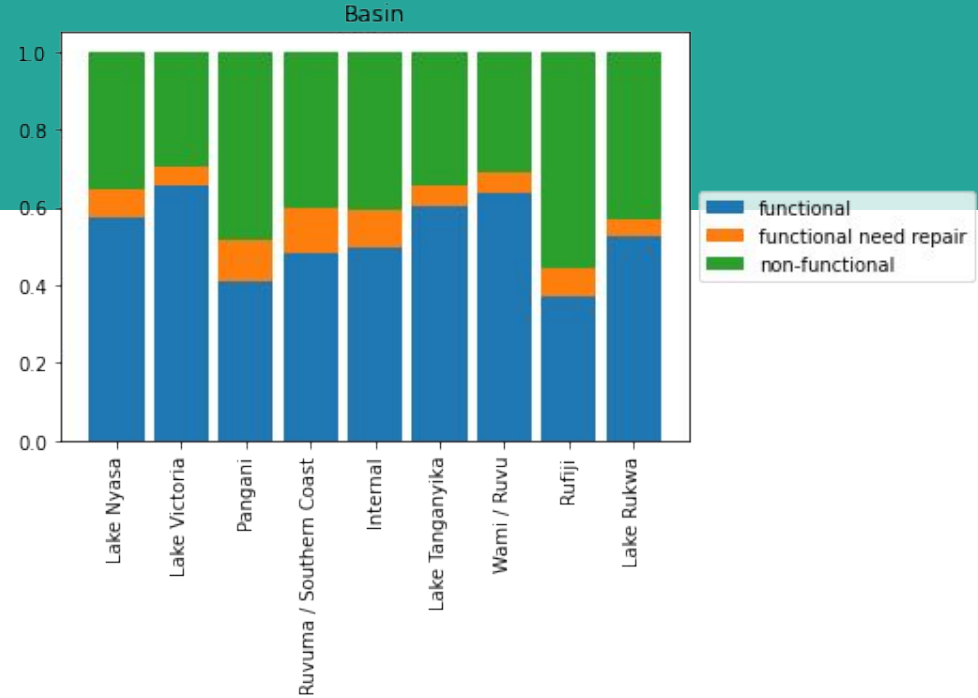
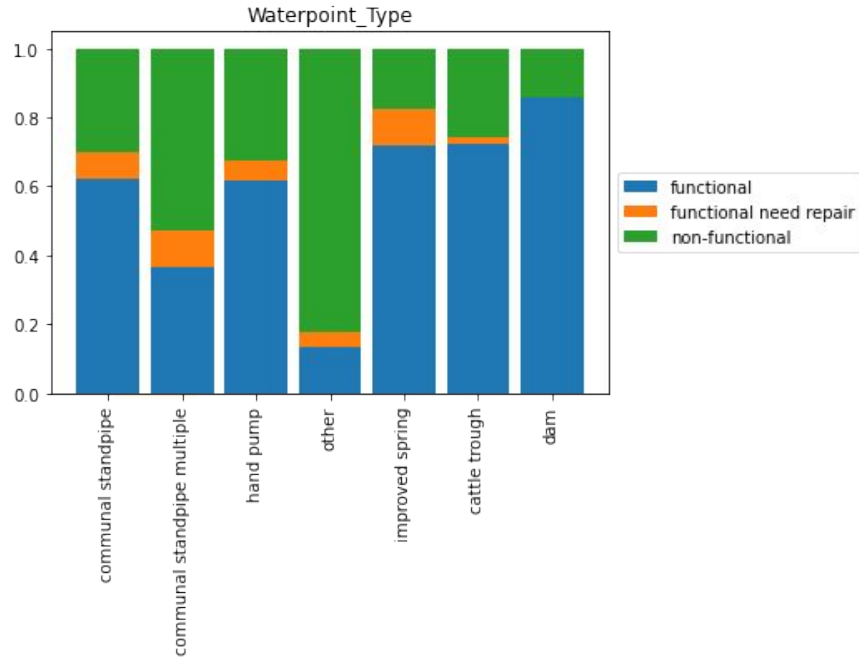
# Main Goals



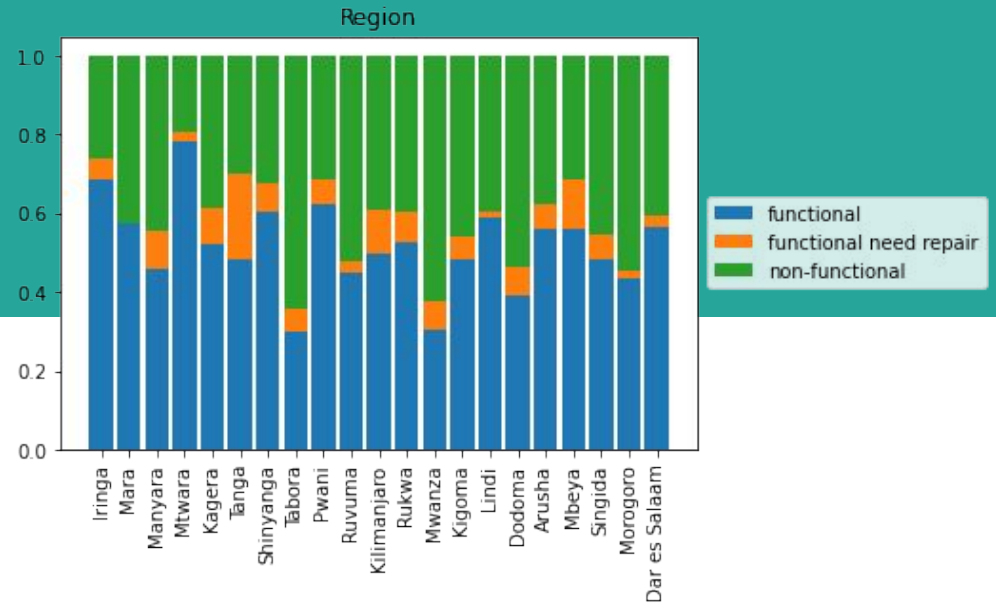
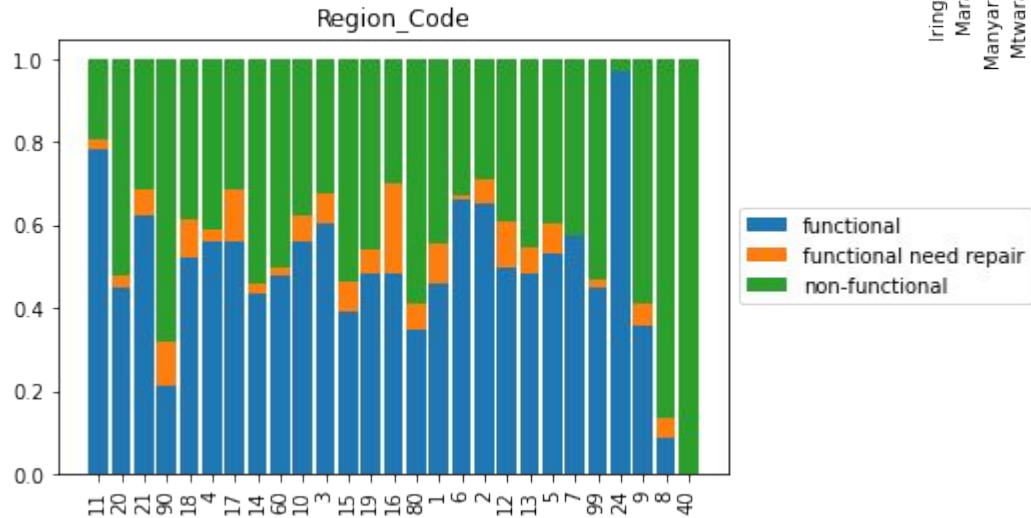
# Accuracy!



# Feature Selection



# Continued..



# Final List of Features

```
['basin',  
 'tsh_0',  
 'gov_funding',  
 'installer',  
 'region',  
 'region_code',  
 'district_code',  
 'population',  
 'public_meeting',  
 'scheme_management',  
 'permit',  
 'construction_year',  
 'extraction_type',  
 'management',  
 'management_group',  
 'payment',  
 'water_quality',  
 'waterpoint_type',  
 'quantity',  
 'source',  
 'source_class',  
 'payment']
```



## Submissions

BEST

0.8031

CURRENT RANK

1993

# COMPETITORS

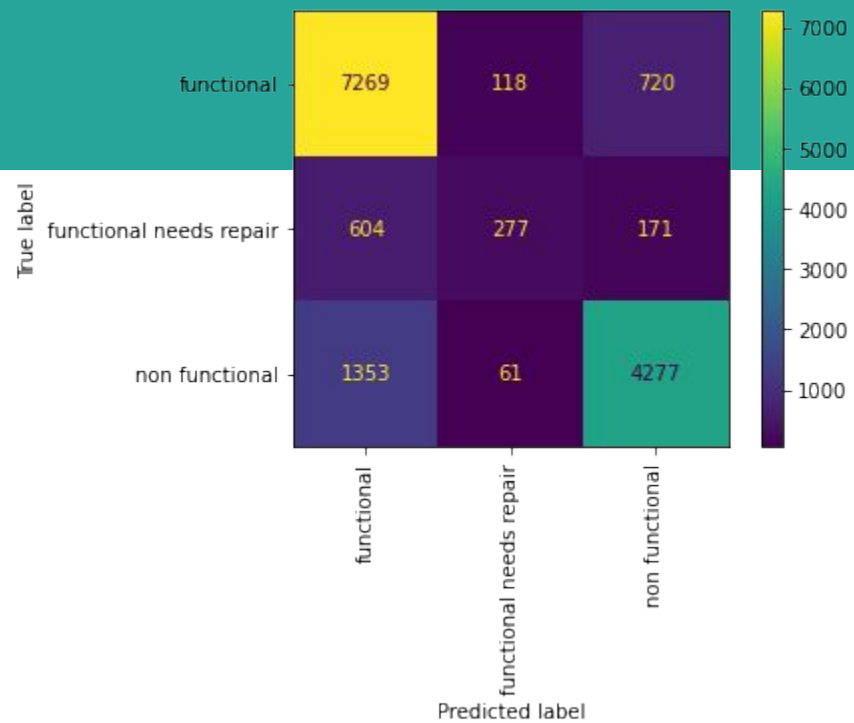
10553

SUBS. MADE

0 of 3

Diving a Little Deeper....

---



functional

functional needs repair

non functional

# Future Research

Tweek model to reduce the worst kind of errors

More research on the main factors that contribute to conditions of water pumps

# Recommendations

- Use the model to make predictions, but don't solely rely on it
- Be aware of the type of errors that are most common
- Always be on the lookout for ways to gather more data/characteristics - for the most part, more information will almost always increase accuracy throughout machine learning models

# Thank You!

