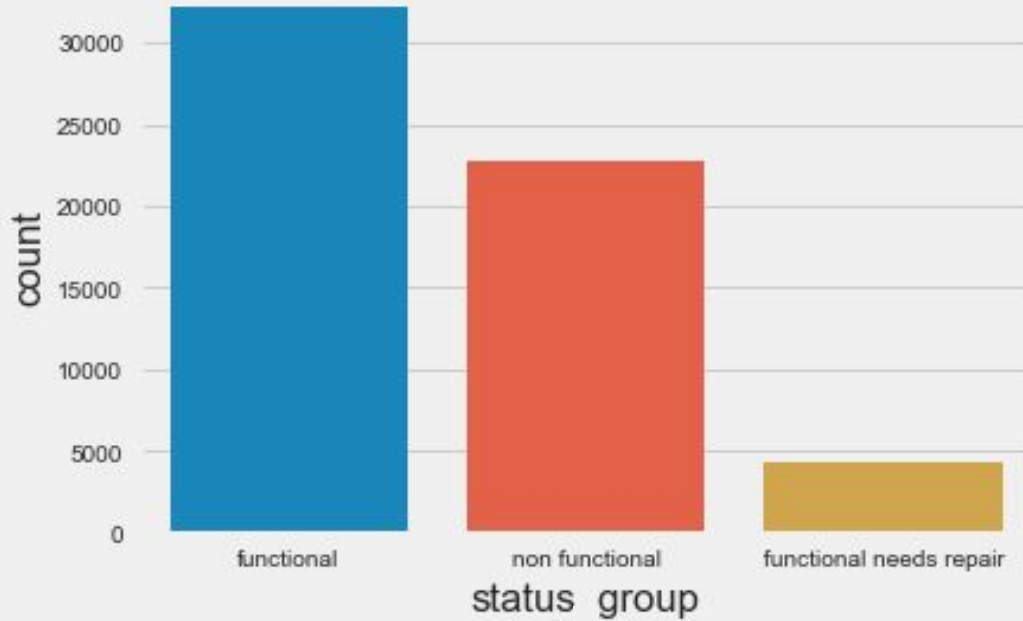

Tanzanian Water Supply

— Taarifa Water Table Classification —

Jeremy Betz
Metis
May 2017

Taarifa

<https://github.com/taarifa>



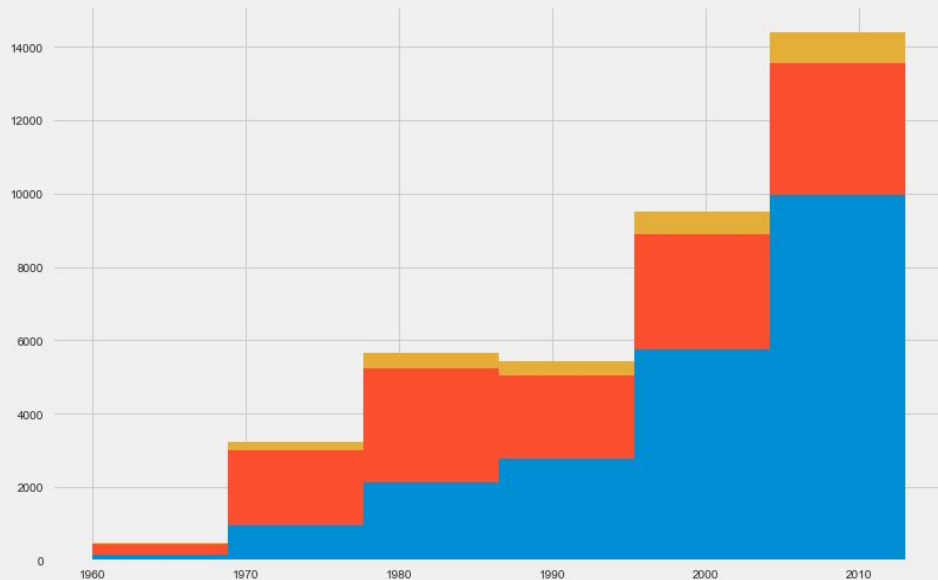
Data Cleaning and Feature Selection

Numerical

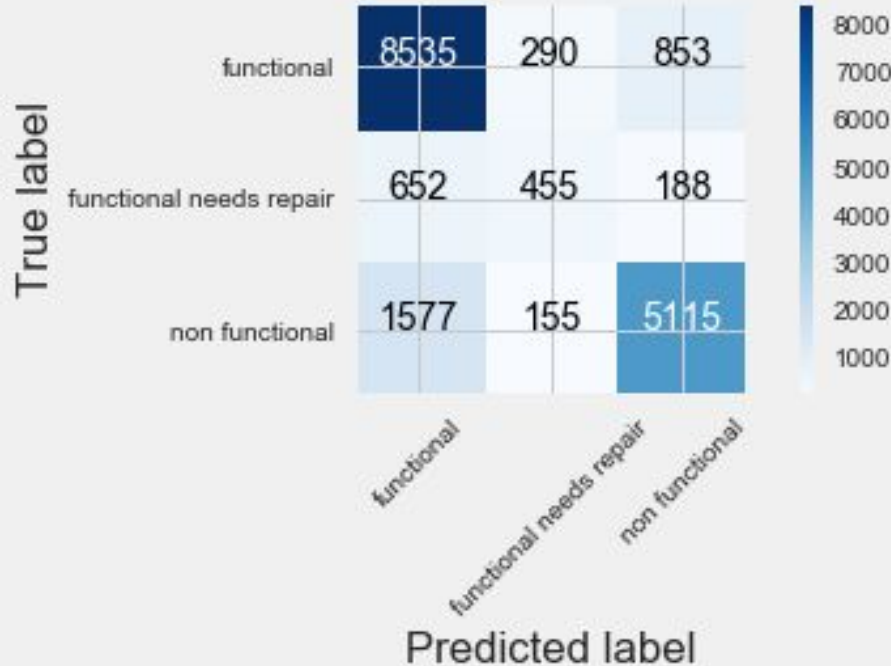
- Construction Year
- Latitude, Longitude
- Elevation, Depth

Categorical

- Water Basin
- Design
- Management

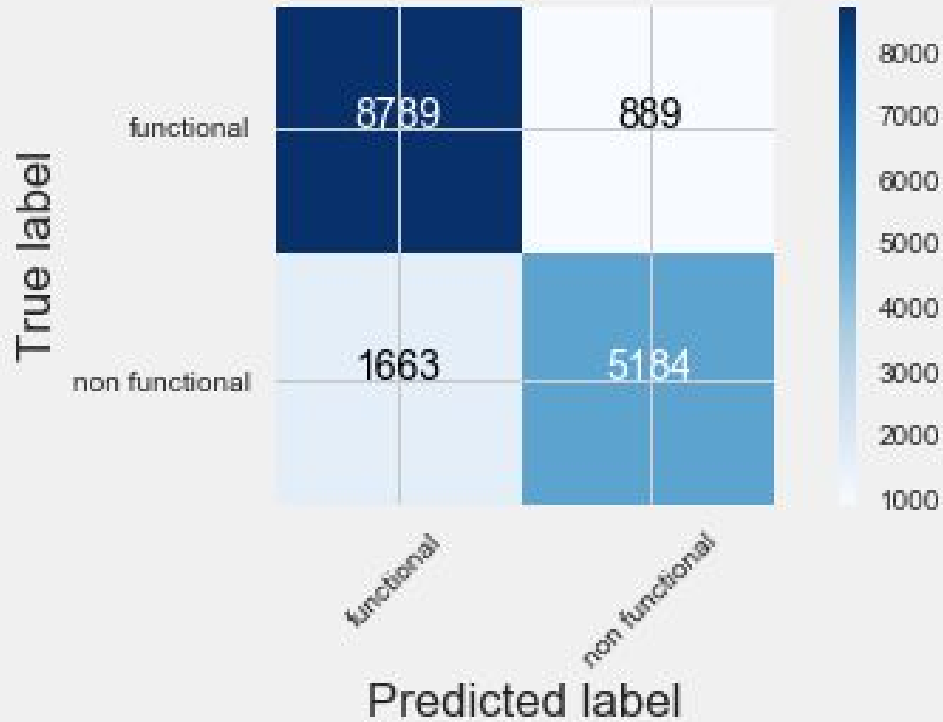


Random Forest Confusion matrix



- All 59,400 Water Sources
- 80% Accuracy
- 75% Recall of *non functional*
- 35% Recall of *needs repair*

Random Forest Confusion matrix



- Removed 4,400 *needs repair*
- 85% Accuracy
- 76% Recall of *non functional*

Forward

- More Complex Models
- Heatmap of Water Sources and Status

Tanzanian Water Supply

— Taarifa Water Table Classification —

Jeremy Betz
Metis
May 2017