TANZANIA-WATER- PUMP

SarahM

Business Problem:

The Tanzania Government is trying to improve their water pump maintenance operations in order to ensure that clean, potable water is available to communities across Tanzania. In order to accomplish this, the Non government stakeholder wants to be able to better predict which pumps will fail, and to better identify pumps that need repair and what factors need to be considered in the future. This will help in improving maintenance efficiency and water access.

Study Area

- Tanzania is an East African country with about 59,678,000 in population. It is located just south of the Equator. Tanzania is a developing country that still has struggles with providing clean water to its entire population. There are many water points already established in the country, but some are in great need of repair while others have failed altogether.

Data used:

- The data set has different variables that describes the pump status (the target variable),The geographical location of the pump, what kind of pump is operating, date of installation, etc. The data has 59,400 entries of individual pumps recorded from the year 2011 to 2013.

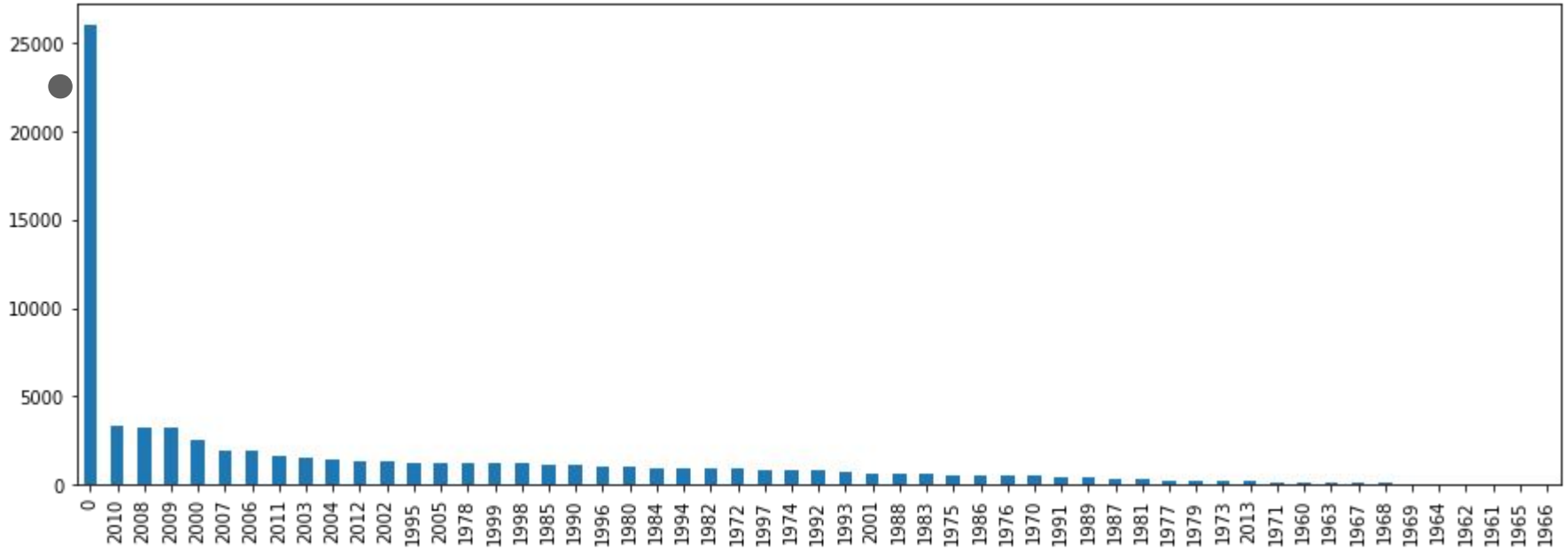
Methodology:

- Understanding the data
- Preparing the data for the model.
- Build different predictive classification models for the functionality of the wells.
- Compare results from each Model.

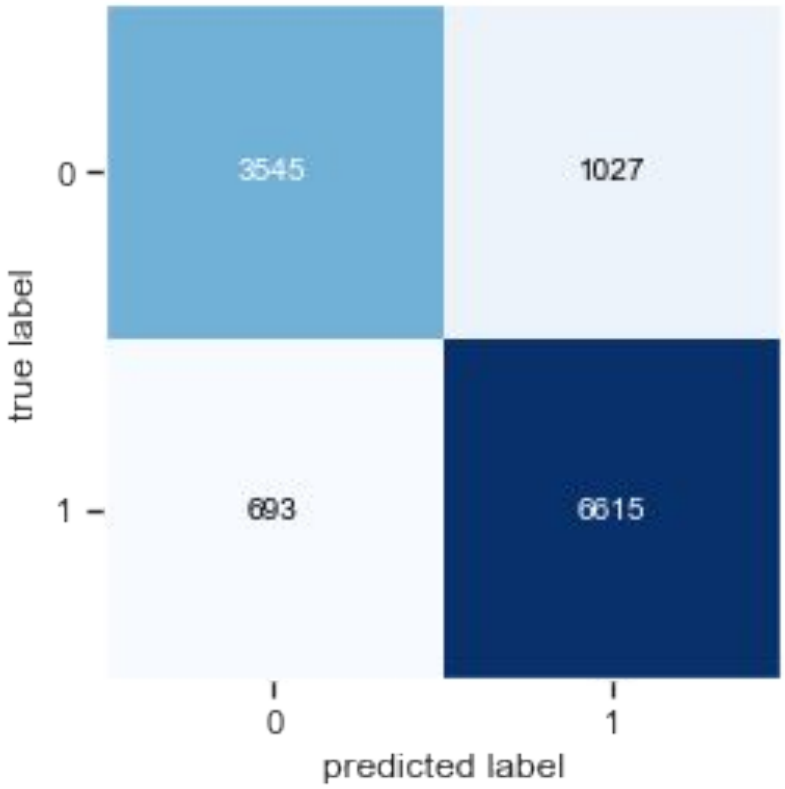
Visualization

This is the Year the waterpoint was constructed.

- Years 1960-2013



Results:



From the Random forest model prediction.From the confusion matrix results 6615 water points were correctly classified as functional while 693 water points were wrongly classified as non-functional but were actually functional water points.Also, 3545 water points were correctly classified as non-functional while 1027 water points were wrongly classified as functional but were actually non-functional.

Findings:

1. We did several models were we compared and the best mode was found to be Random Forest classifier with an accuracy of 85%.
2. There were dried 4272 wells that dried up but with good water quality.
3. Finding clean water sources is a problem in Tanzania.
4. 2226 wells making up 7% of the total wells have enough and soft water but needs repair. The Tanzanian Government should invest in repairing them otherwise they will be nonfunctional.
5. 8035 wells have enough, good quality water but are nonfunctional. Thus the need to find ways to pump the water from these wells.

Recommendation:

- I would recommend that the Non Government to continue to making recordings of the wells so as to continue monitoring them since the data provided was only recorded mostly 2011 and 2013. This would provide more accurate data for model improvement.
- Different regions have different factors like climate, rainfall seasons etc. so thus different modes should build for each region.

Thank You!!