TANZANIA WATER WELLS PREDICTIONS

ABOUT THE PROJECT

This study aims to forecast the functional condition of Tanzanian water wells using machine learning classification algorithms. Functional, non-functional, and functional but in need of repair are the several status groups for classification. The idea is to increase access to water throughout Tanzania by forecasting a well's operational state.

DATA PROCESSING

The data set used for training our model was from the data driven website <https://www.drivendata.org/competitions/7/pump-it-up-data-mining-the-water-table/page/23/>

The data went through various cleaning steps which included filling in null values and dropping unnecessary columns.

After encoding our data using binary encoding, scaling it using standard scaler and resampling it using SMOTE , we split it and
created various models , ultimately choosing model 7 which was our best performing model.

MODELLING SETBACKS AND SOLUTIONS

CLASS IMBALANCES

There were class imbalances which greatly affected the scores of the functional needs repair class. The solution would be collecting more data to balance the data

MISSING INFORMATION

Missing information which made the cleaning process very tedious

TUNING

Fine tuning our model to improve its accuracy

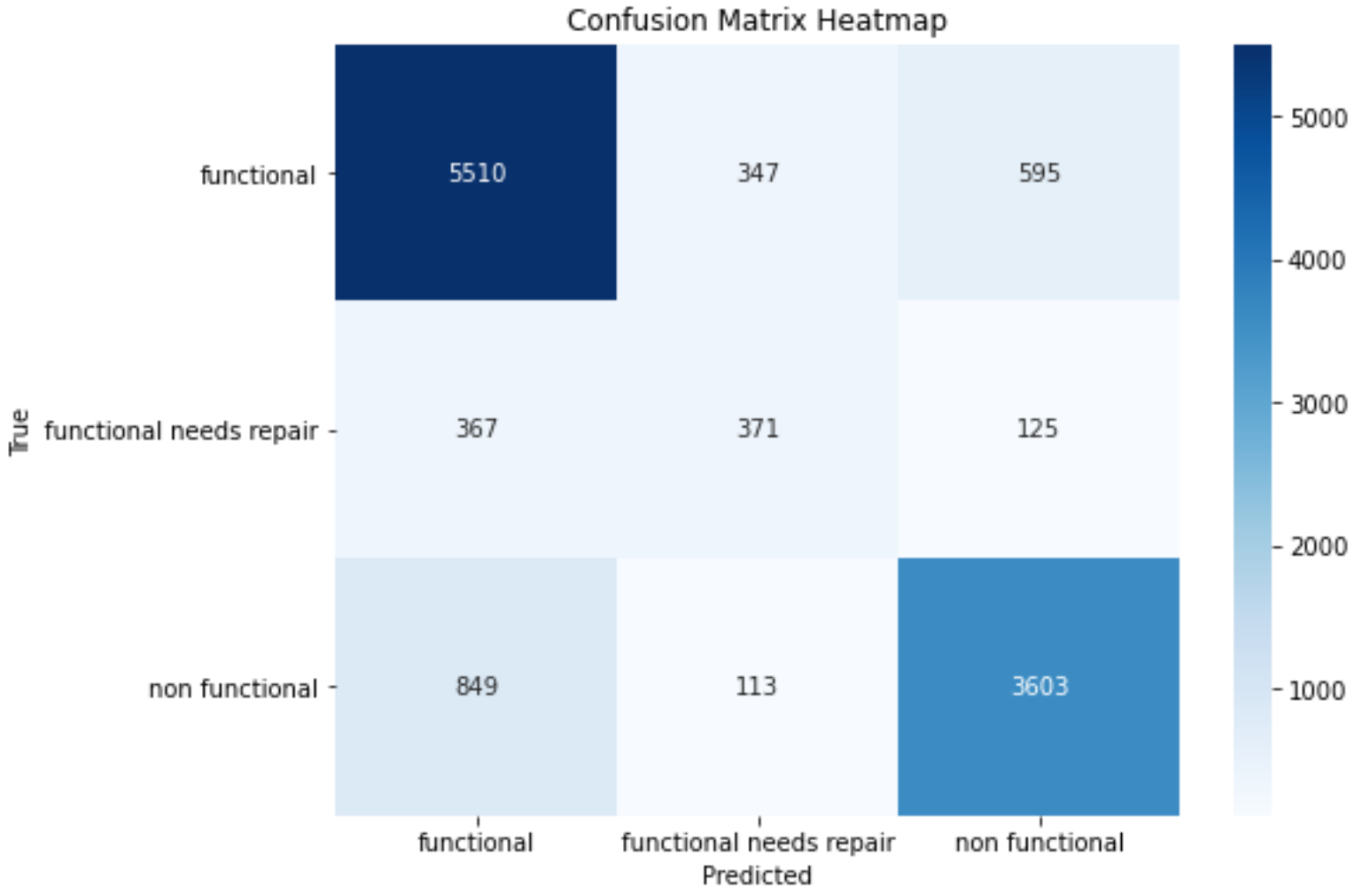
MODELLING

Experiment with more modelling techniques

BEST PERORMING MODEL

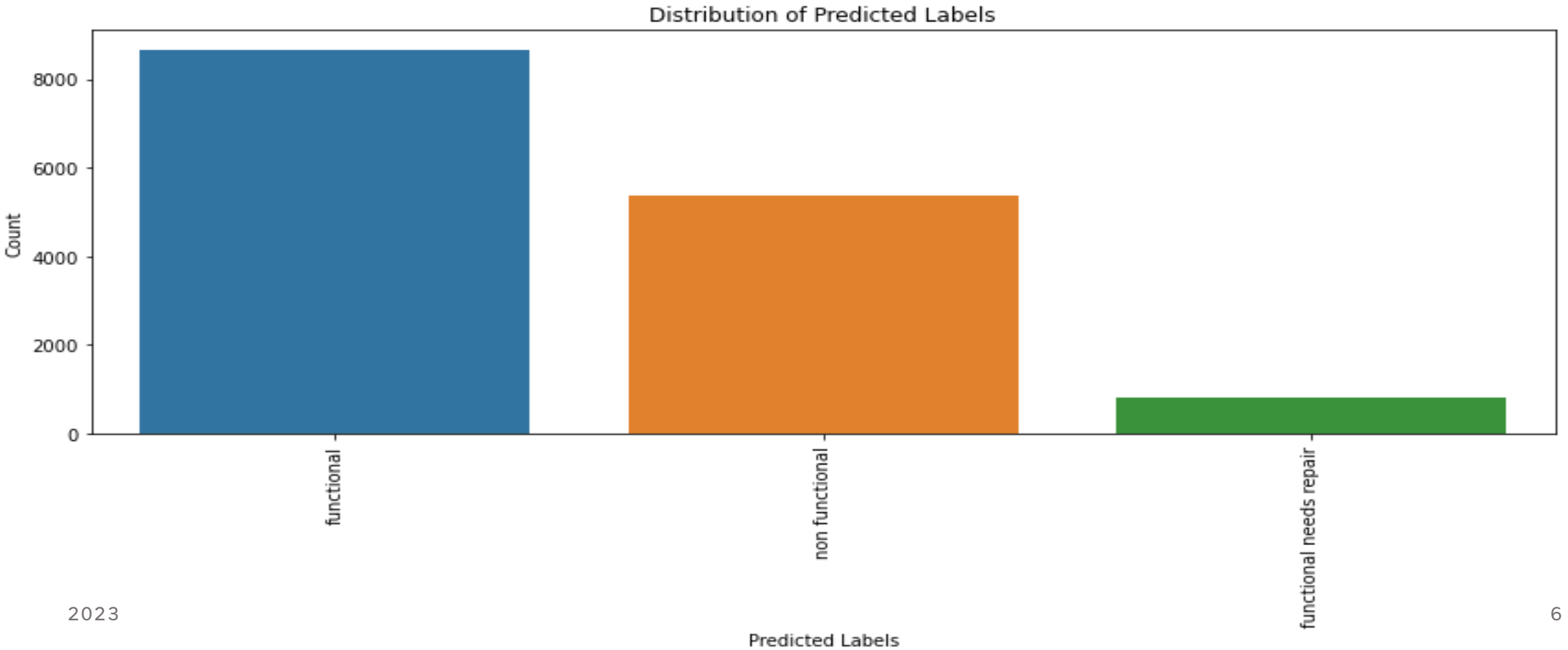
It correctly predicted 5510 wells as functional , 371 wells as functional needs repair and correctly predicted (3603) wells as nonfunctional

The best performing model was the voting classifier model which had an accuracy of 80%

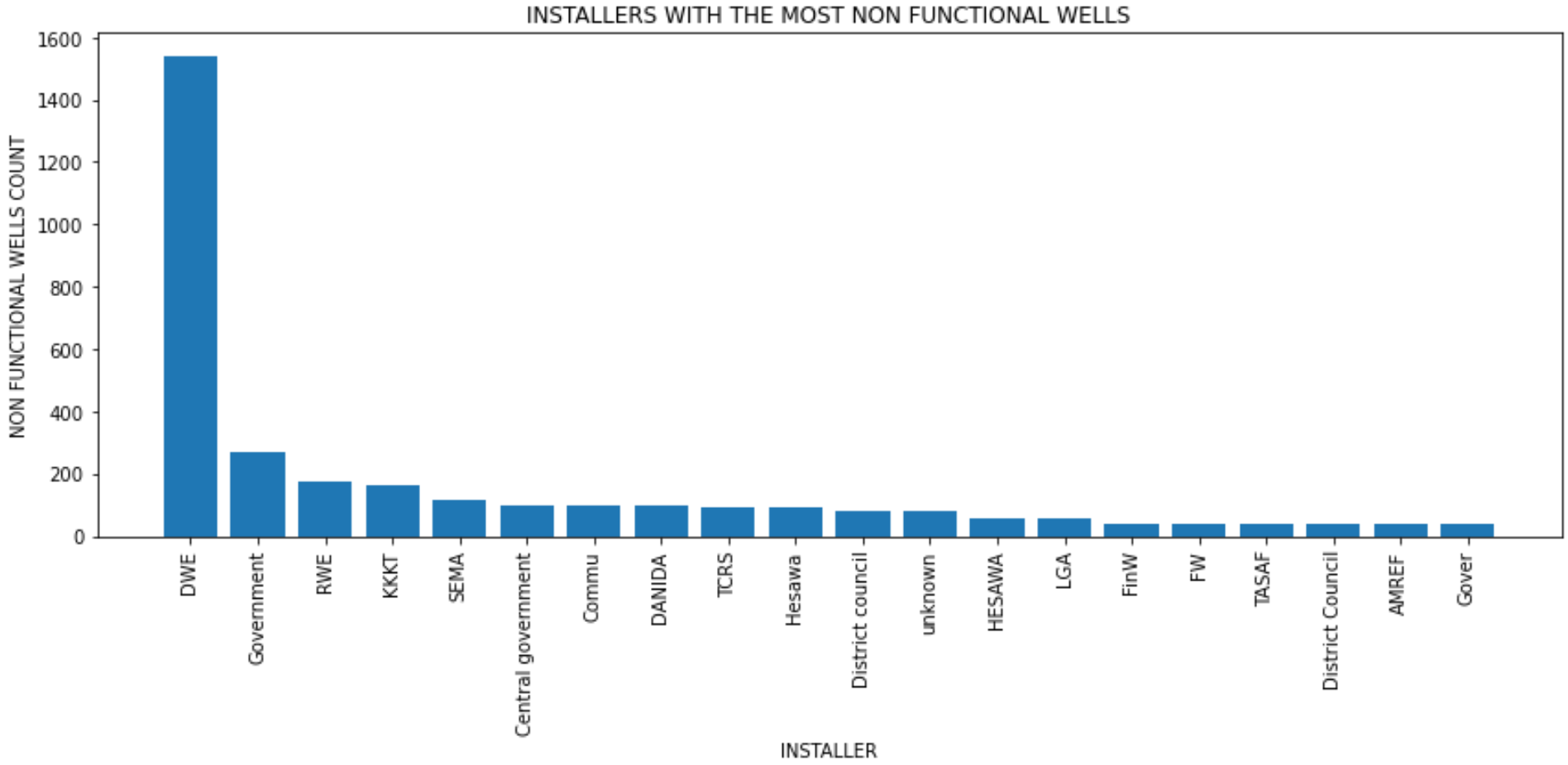


DISTRIBUTION OF THE WELLS ACCORDING TO PREDICTIONS MADE ON OUR TEST DATA.

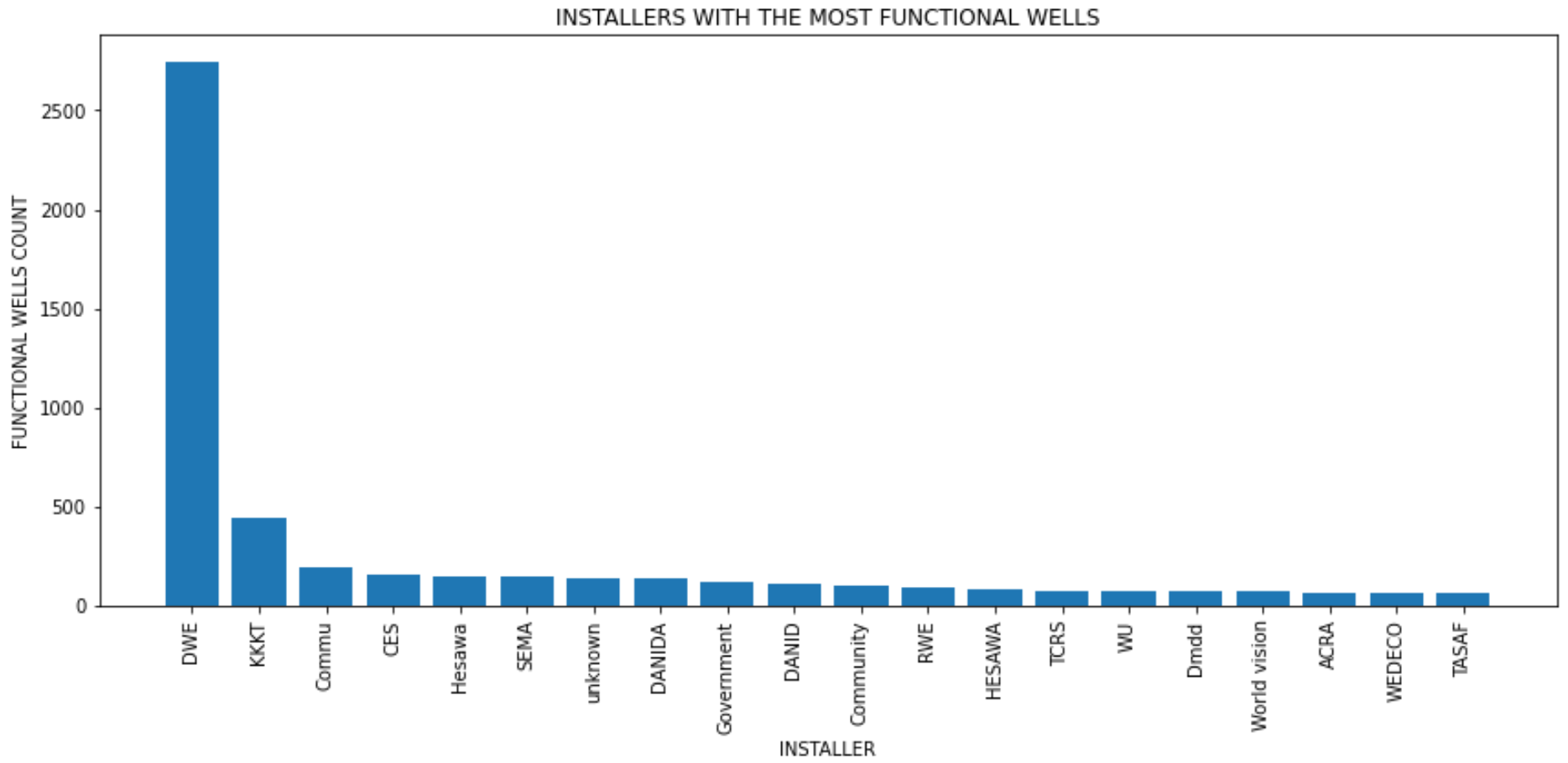
functional 8730
non functional 5363
functional needs repair 757



INSTALLERS WITH THE HIGHEST NUMBER OF NON FUNCTIONAL WELLS



INSTALLERS WITH THE HIGHEST NUMBER OF FUNCTIONAL WELLS



To analyze the effectiveness of installers we see that

1. Despite having the highest number of nonfunctional wells , DWE also has the highest number of functional wells. It is essential to investigate the causes of the high number of non-functional wells and consider improvements to installation and maintenance practices.

2. The Government installer has a relatively high number of non-functional wells (272) compared to their functional wells (116).

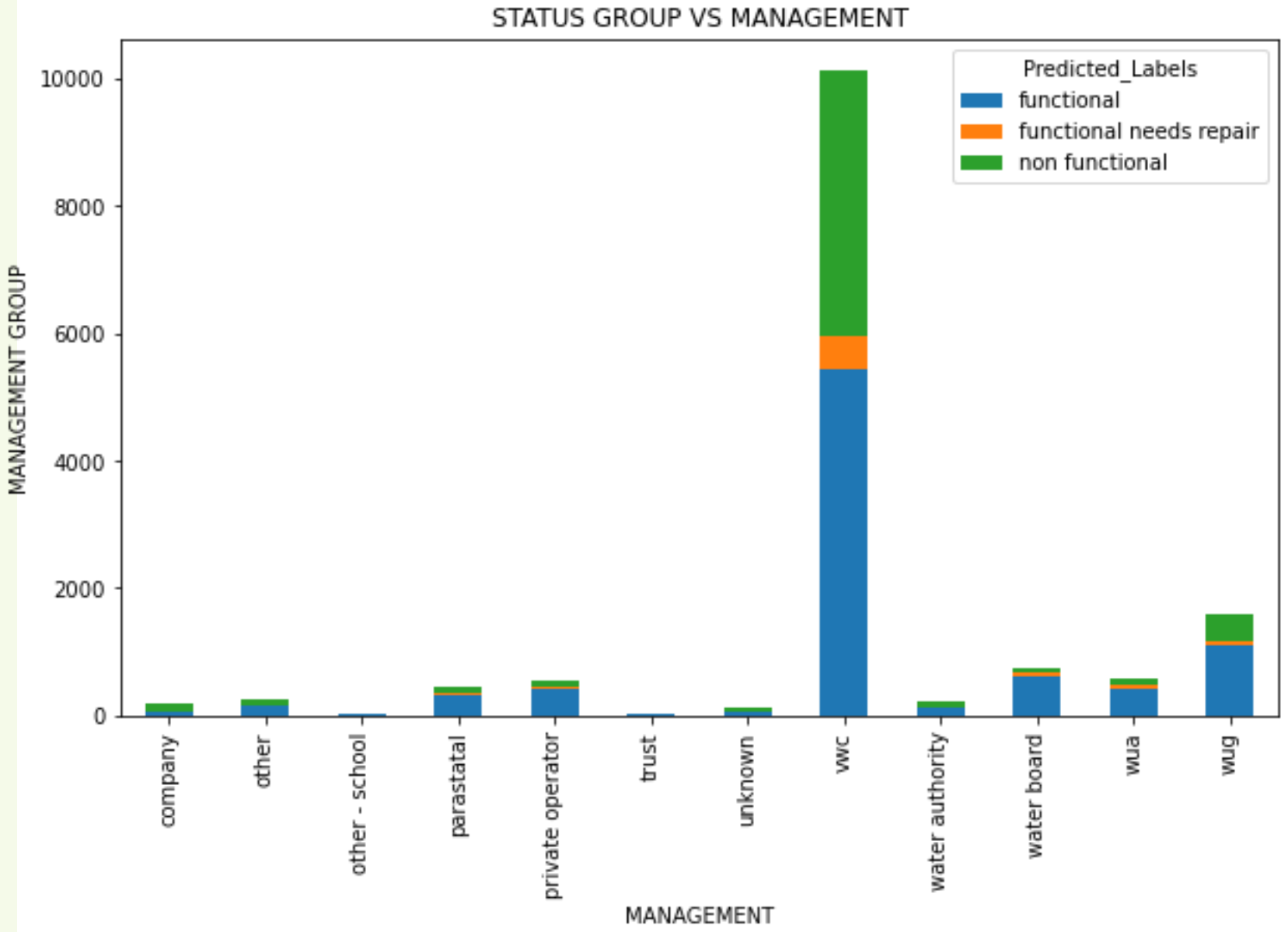
3. RWE also has a relatively high number of non-functional wells (176) compared to functional ones (88). This indicates room for improvement in their well installations.

4. KKKT has a relatively high count of functional wells(437), compared to the non functional (160) which is relatively okay.

5 .Hesawa has more functional wells(78) compared to the non-functional ones(57), suggesting their effectiveness in providing reliable water sources.

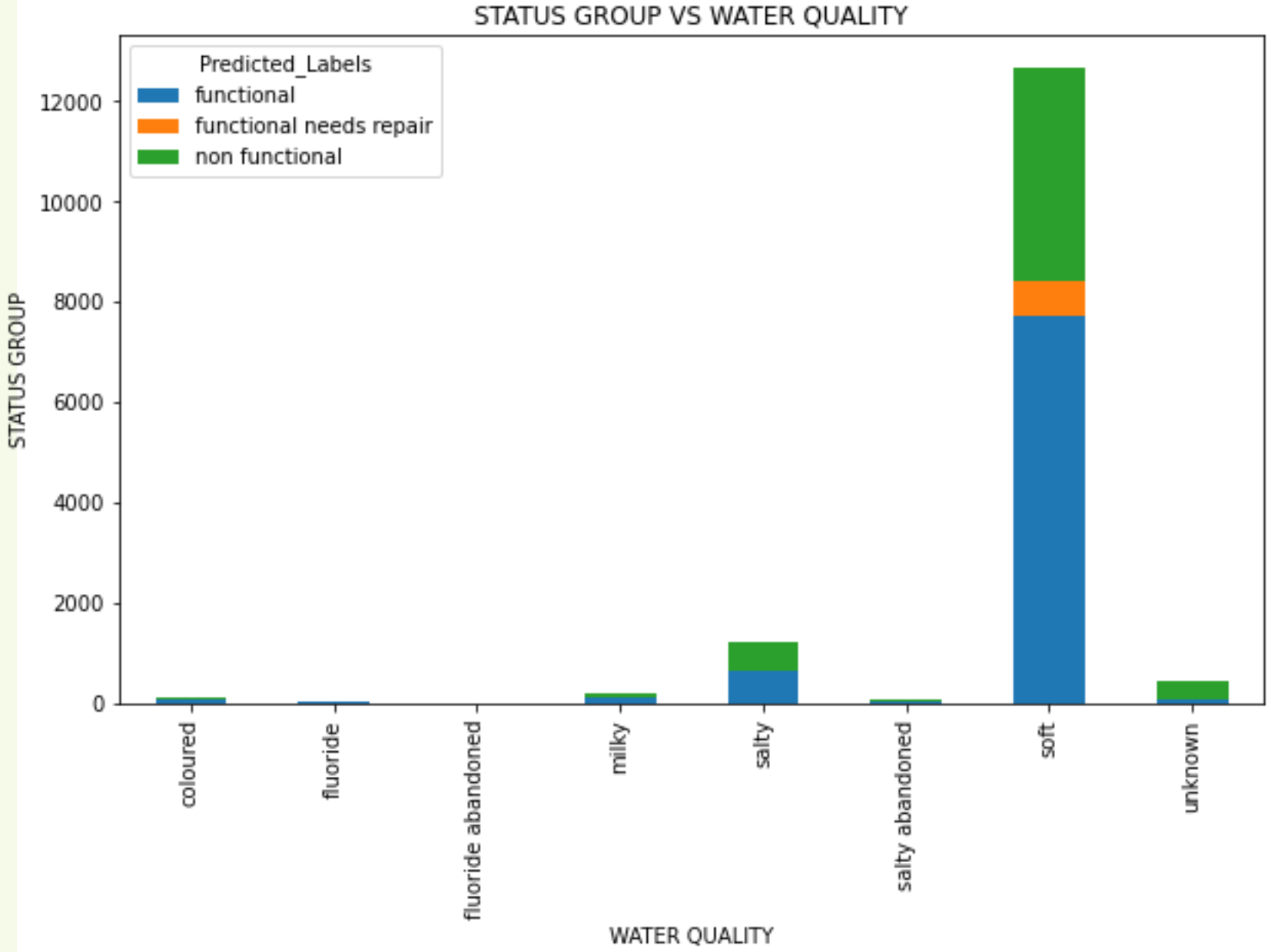
MANAGEMENT VS FUNCTIONALITY

VWC has the highest number of wells under its management followed by WUG. Trusts and schools have the lowest amount of wells under their management



WATER QUALITY VS FUNCTIONALITY

Most wells are soft water wells. Followed by salty water wells. Unknown wells are mostly non functional.



RECOMENDATIONS

1. Quality Assurance:
Implement quality control systems to ensure that wells are constructed and maintained in compliance with accepted standards.

2. Transparency and Accountability:
Promote transparency in funding allocation and project management. This can help reduce corruption and mismanagement

3. Evaluation and monitoring:
Set up methods for monitoring and evaluating well operation, water quality, and user satisfaction. This information can be used to identify issues and come up with ways of solving them through informed decision making.

4. Data Collection Infrastructure:
Install data collecting and reporting infrastructure to enable real-time monitoring and problem solving.

SUMMARY

Access to safe and clean drinking water is not only a basic human right, but also an important aspect of the well-being and health of the public. We are taking a key step towards securing a happier and more secure future for our communities by addressing challenges associated with water resource mismanagement and inadequate installation practices.

THANK YOU

JOAN NJOROGE
Joanpriscillanjoroge@gmail.com

https://github.com/PriscillaJoan