Predicting the Status Of Water Wells

Tanzanian Minister of Water and Flatiron Partnership

Data and Key Processes

The data contains 59,400 rows and 40 features

The project is broke into two parts

1. Actionable BI for Key Stakeholders
2. POC for Predictive Modelling

Installers with non Functional Wells

DWE	6347
Government	1034
RWE	765
Central government	450
DANIDA	425

We want to fix broken wells quickly. Therefore we generated, the instsallers with the top ‘non functional wells’

The top are listed here for quick communication and preparation for repairs program

Funders with non functional wells

Government Of Tanzania	4663
Danida	1242
Hesawa	1034
World Bank	707
Kkkt	498

- Reach out to these funders to get funding for repairing the wells
- Critical source of revenue and there must be engaged to allocate resources to repair

Predictive modelling

- Target variables are status of well

- Models used are
  - Random Forest – 37%
  - Xgboost – 74%

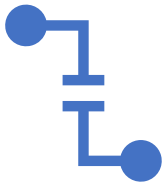
Is this model ready for deployment?

- At 74%, with 80% being the top scoring results, I would consider this model ready for deployment

- However, before deployment, further measures need to be taken to acoid wasting resources on a model that can be improved

Next Steps

Dashboard for installers to locate and fix wells



Deploy model if threshold of 80% is reached

Improve ML score to at least 80% through following measures

Feature engineering

Perameter tuning