Glass Half-full:

Using Machine Learning to keep clean water flowing in Tanzania

Rashid Baset
Alphonso Woodbury

Statement of Work

RBAW Consulting will aid the Tanzanian government in analyzing waterpoints to improve their function and water availability.

Steps:

- Explore and understand problem and dataset
- Create a classification model
- Analyze model and provide recommendations

An oasis, a challenge...

Developing nation

African Great Lakes

56M people

More than half travel more than 1 Km to water

Challenge accepted!

Significant investment in water infrastructure since the 1950s

88% of funding is from external donors

Urban communities benefiting most

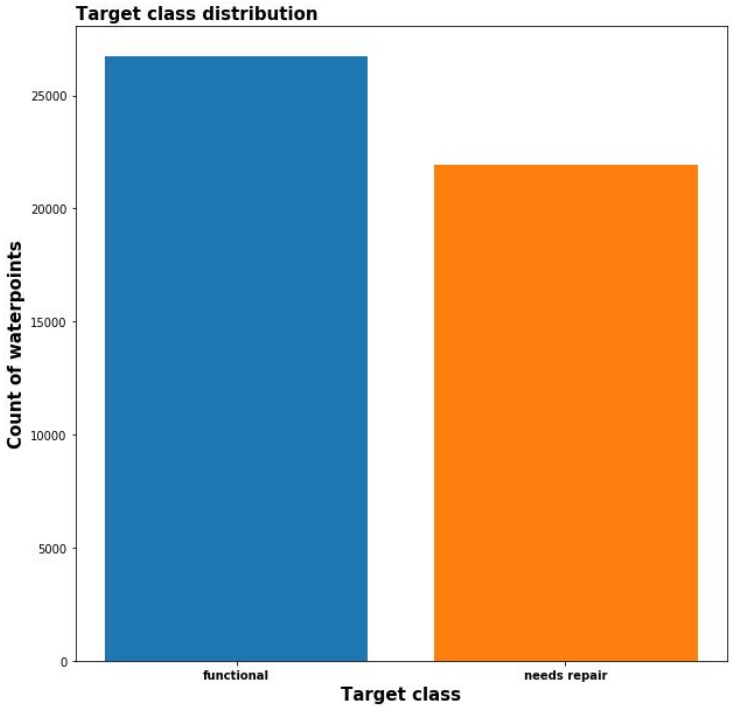
Not all have been successful

What we want to discover

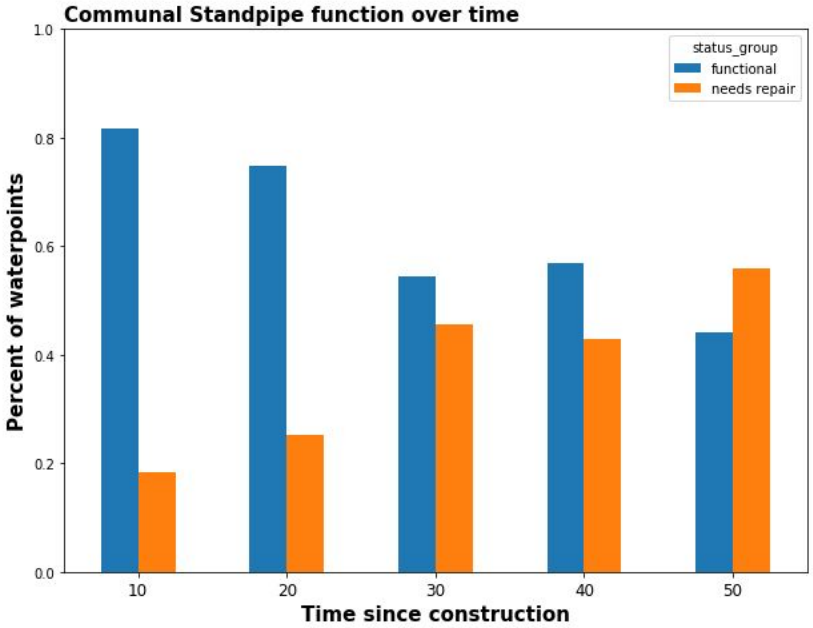
1. Which waterpoint features ensure waterpoint success?
2. Who should fund and install waterpoints?
3. How well can we identify (predict) waterpoints that need repair?

Data and Assumptions

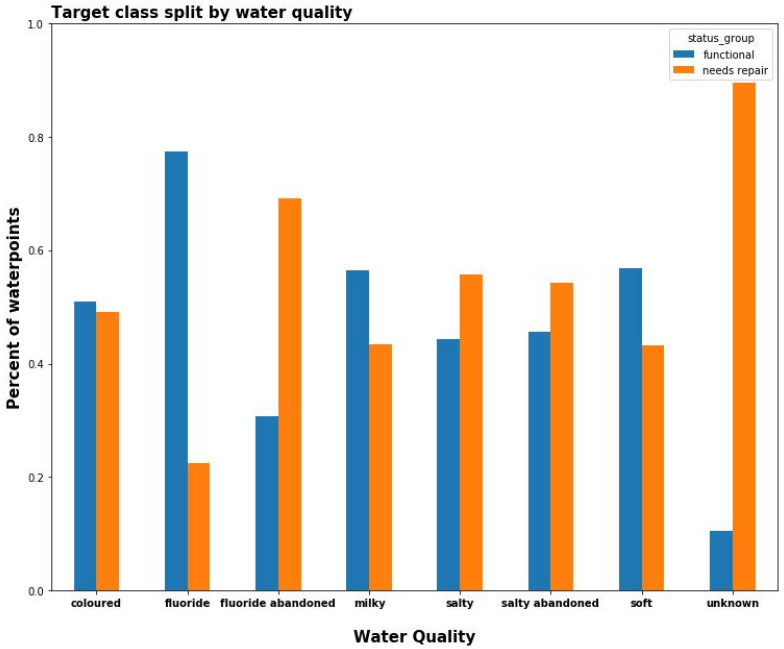
- DrivenData competition
- Specific to Tanzanian waterpoints
- Data collected from March 2011 to September 2011



Waterpoint function declines over time



Waterpoints with fluoride may be most effective?



Water Point Mapping - Data Entry Form

Date of record_____ Recording Organisation_____

Region_____ Village Longitude_____
District_____ Village Latitude_____
Water_____ Village registration no _____
Village_____ Village Photo ID _____
Sub village_____

WP name_____ GPS waypoint Number_____
Scheme name_____ WP Longitude_____
Funder_____ WP Longitude_____
Installer_____ Elevation_____
Year constructed_____ WP Photo ID_____

Source Type
- O Shallow well
- O Hand-drilled tube well
- O Machine-drilled borehole
- O Dam
- O River/Lake
- O Spring
- O Rainwater Harvesting
- O Other…

_____

Extraction system
- O None
- O Mono
- O Cemo
- O Climax
- O KSB
- O Submersible
- O Gravity
- O Afridev
- O Nira/Tanira
- O SWN 80
- O India mark II
- O Walimi
- O Windmill
- O Other…

_____

Water point type
- O Hand pump
- O Communal standpipe
- O Communal standpipe multiple
- O Cattle trough
- O Dam
- O Improved Spring
- O Windmill
- O Other

_____

Status
- O Functional
- O Non Functional

Main hardware problem
- O None
- O Source damaged
- O Pump broken
- O Pump Stolen
- O Engine broken
- O Engine stolen
- O Tank out of use
- O Pipe broken
- O Tap poorly sited
- O Tap broken
- O Under construction

Water quantity
- O Enough
- O Insufficient
- O Seasonal
- O Dry

Water quality
- O Soft
- O Milky
- O Coloured
- O Salty
- O Salty abandoned
- O Fluoride
- O Fluoride abandoned
- O Other

Breakdown date _____

Other reason WP not functional _____

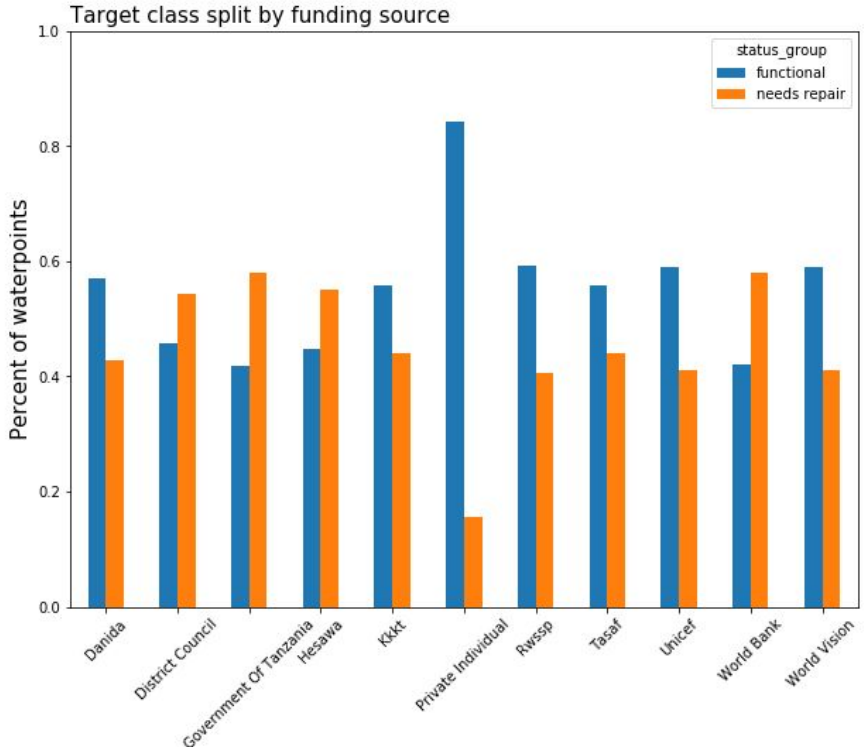
Scheme Ownership
- O VWC
- O WUG
- O WUA
- O Company
- O Trust
- O Water Board
- O Parastatal
- O Private individual
- O Other…

WP Management
- O VWC
- O WUG
- O WUA
- O Company
- O Trust
- O Water Board
- O Parastatal
- O Private operator
- O Other…

Water Payment
- O Pay per bucket
- O Pay monthly
- O Pay annually
- O Pay when scheme fails
- O Never pay
- O Other

Public meeting about income and expenditure?
- O Yes
- O No
- O Don't know

Waterpoint function varies by funders



Top 10 funders (by waterpoint count) and their respective functional classes. Is private funding the way?

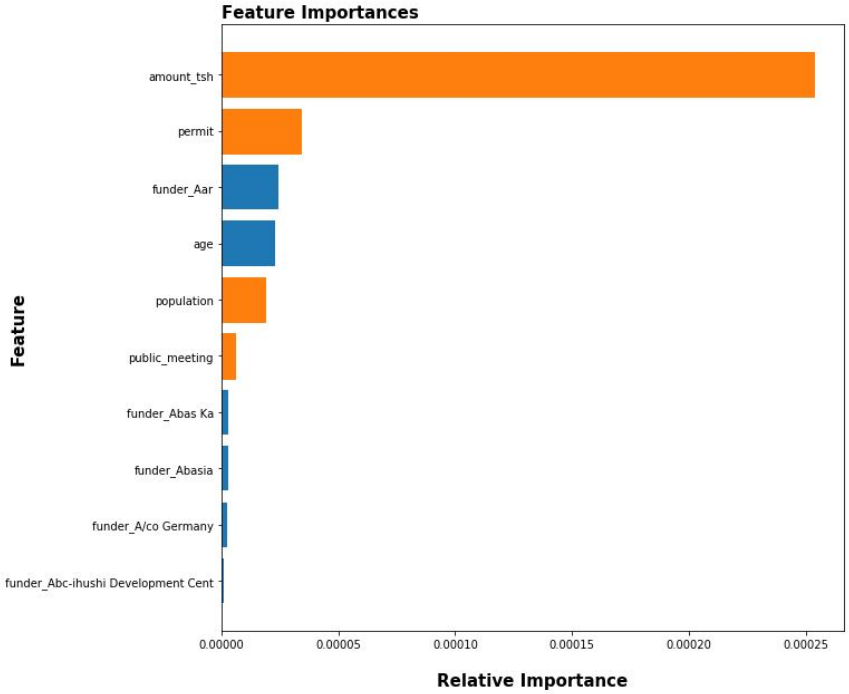
Model Results

- Our best model was able to correctly classify waterpoints that needed repair **83%** of the time.

- This means maintenance efforts can be coordinated to be tactical and cost-effective

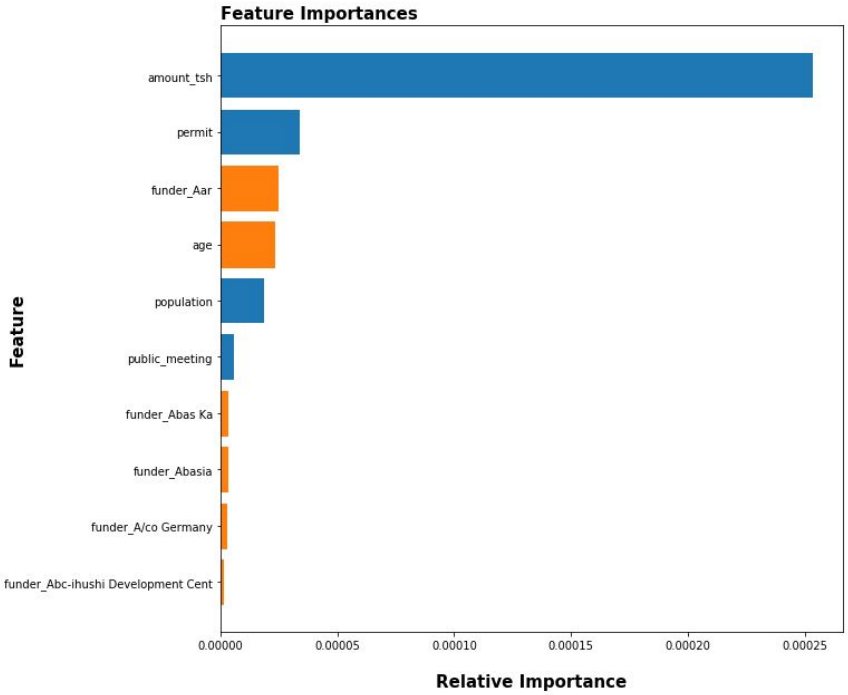
Recommendations

- Waterpoint **total static head**
- **Permittance**
- **Population**
- **Public meetings**



Recommendations

- Partnering with **Funders**



Limitations & Further Research

- The majority of observations had 0 total static head; we would like to know if this is expected or missing data
- We would also like data on waterpoint installation and maintenance cost
- Can we use remote-linked meters to collect data on quality, amount, etc., or is this a ‘pipe’ dream?

Thank You