Pump it Up!

Using AI to help improving the water supply in Tanzania

Data Science Bootcamp 2020

This is the presentation of my Capstone Project of the Data Science Bootcamp from February until May 2020 at neuefische GmbH, Hamburg

01

Pump it Up!
About the project and the goals

02

Tanzania – Water supply
Facts and insights

03

How AI could help
How to participate in improving
access to potable water

04

Recommendations & Conclusion
What was learned and how can
you take advantage of it

05

Future Work
What is left?

01 Pump It Up!

"When the well's dry, we know the worth of water."
Benjamin Franklin

- Human Right to water and sanitation since 2010
- drinking water and sanitation are essential to the realization of all human rights
- Still around 2.2 billion people worldwide lack regular access to clean water

Development AId®, Tanzania

Using Machine Learning Methods to improve the use of development money and funds and hence provide more people with access to clean and potable water

02 TANZANIA

Population

- About 56 million
- Approx 70% rural
- 44% below 15 years
- 125 ethnics

Geography

- 30% of TZA is national parks
- highest Mountan on african continent: Kibo summit of Mount Kilimanjaro
- Zanzibar

Access to water

- Only 61% with access to imprroved drinking water sources
- 86% in Urban Mainland
- Only 49% in Rural Mainland
- Zanzibar 98%

Water & Sanitation

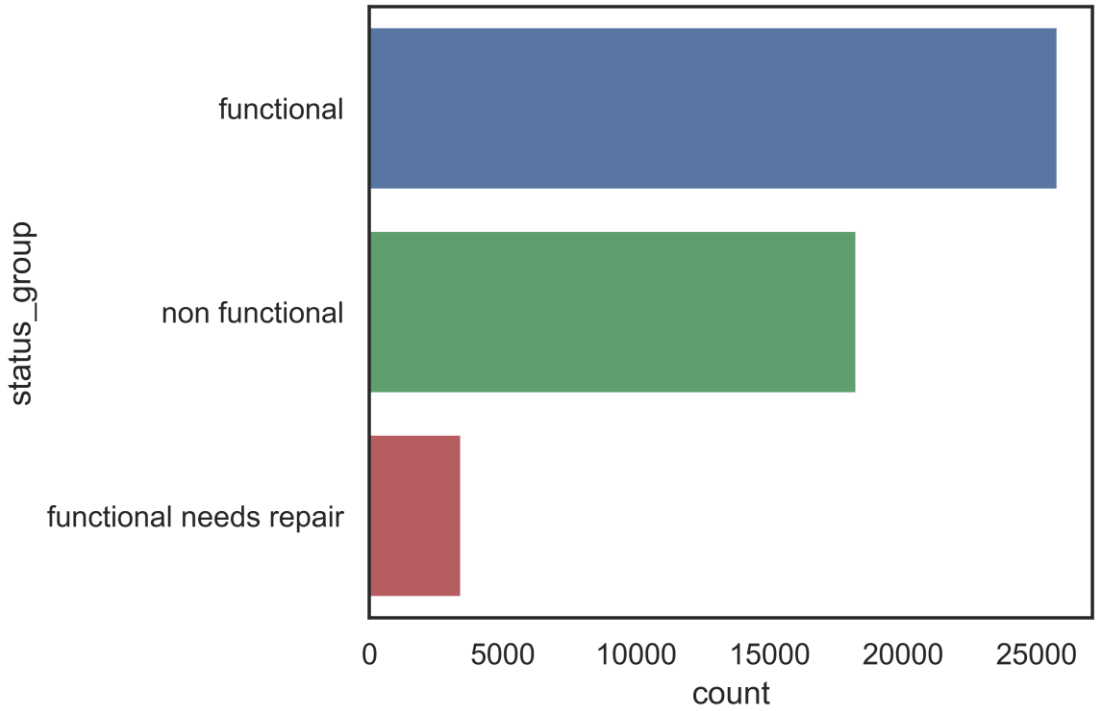
- Only 19% of have improved unshared toilet facilities
- 46% of schools lack functioning water supply

03 How could AI help?

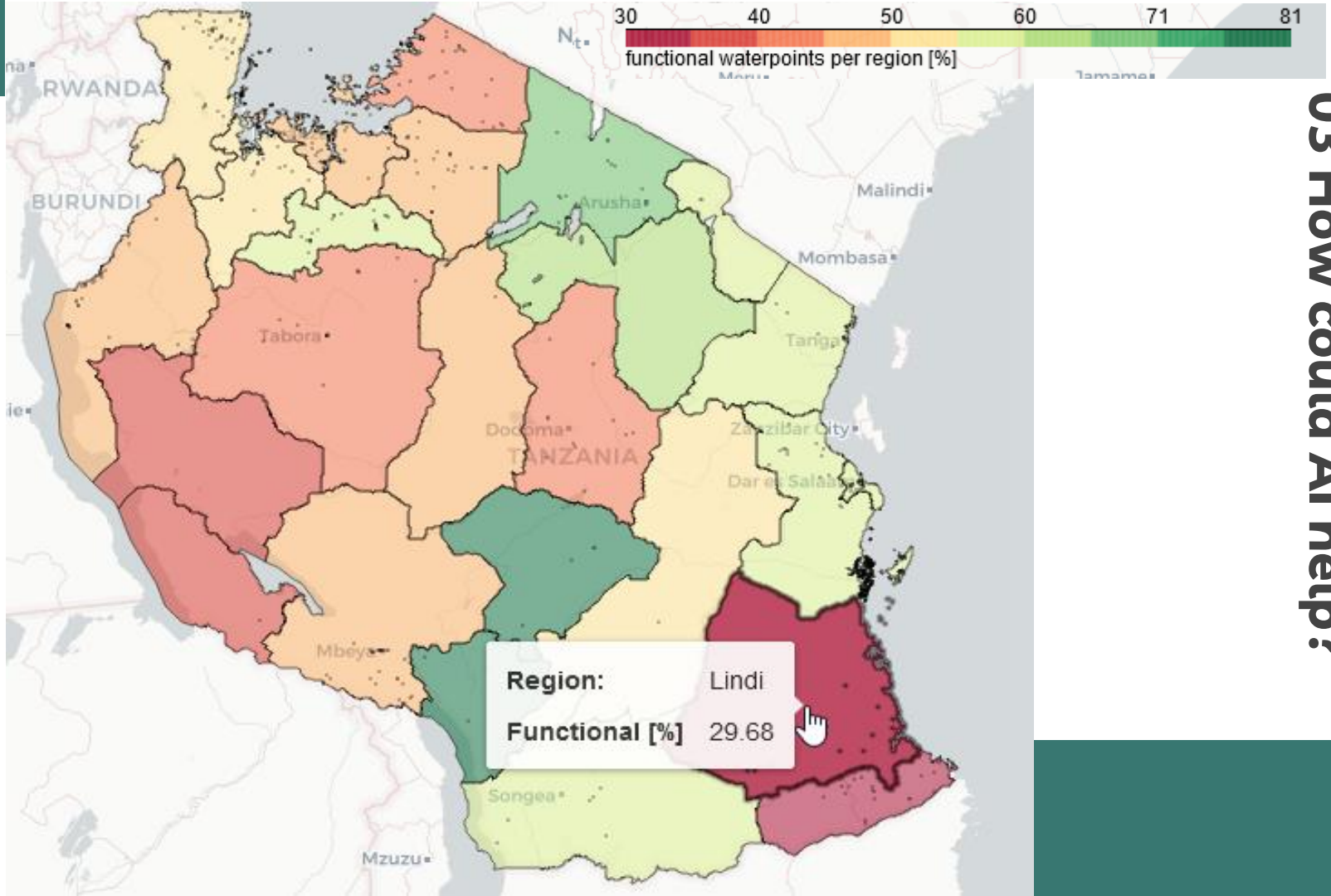
Ensure availability of clean water by predicting the functionality of water points in Tanzania

- Improving maintenance and repair regarding time and money need
- Highlighting where further fundings are needed the most
- About the data:
  - 59 400 waterpoints
  - 41 features (uncleaned)
  - Target: Status of a waterpoint
    - Functional
    - Functional – needs repair
    - Non functional

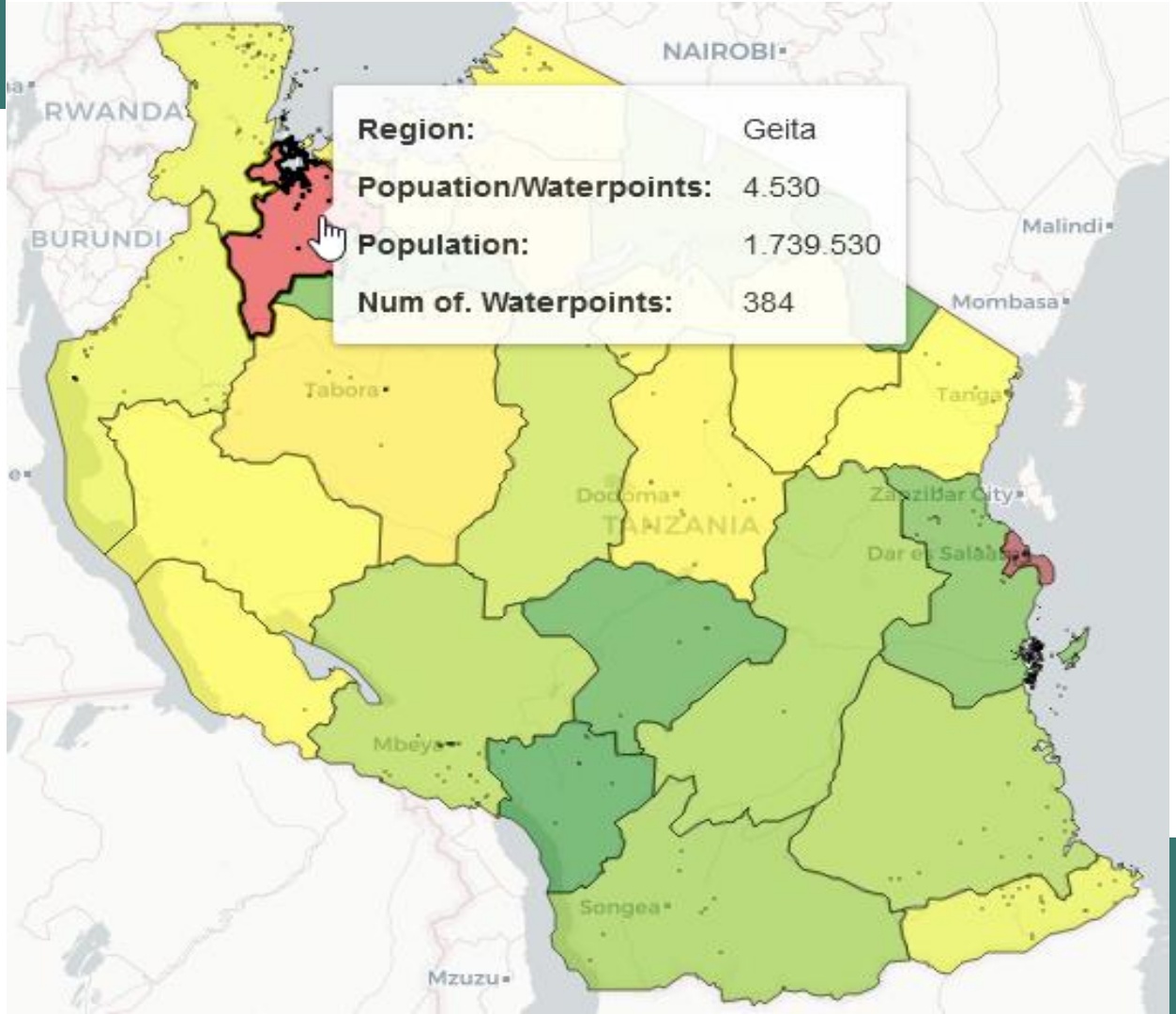
03 How could AI help?



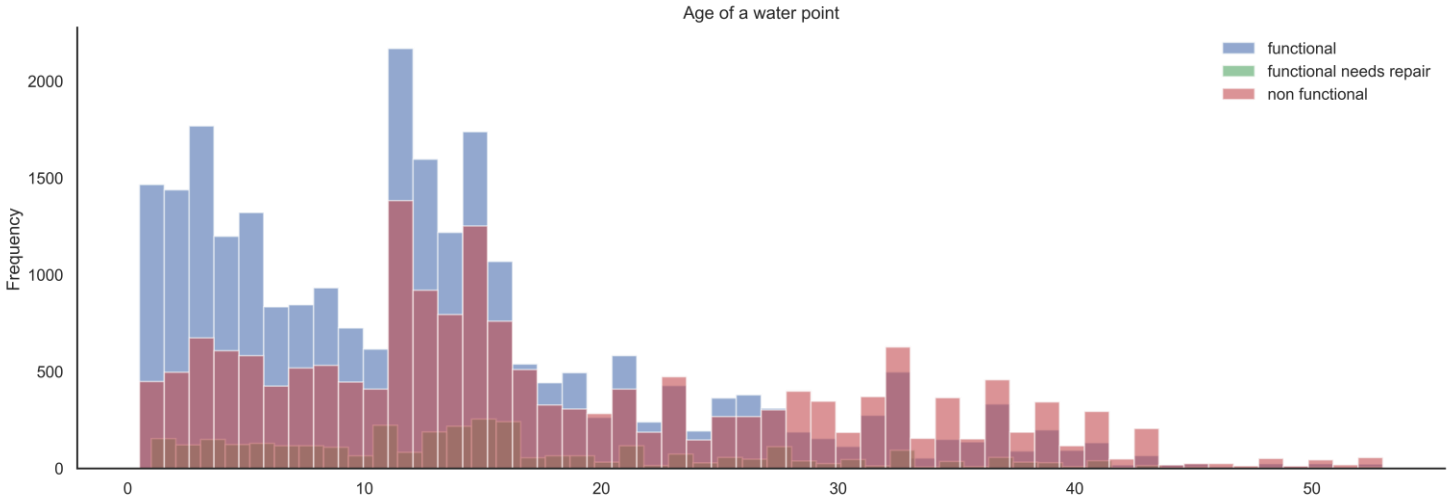
03 How could AI help?



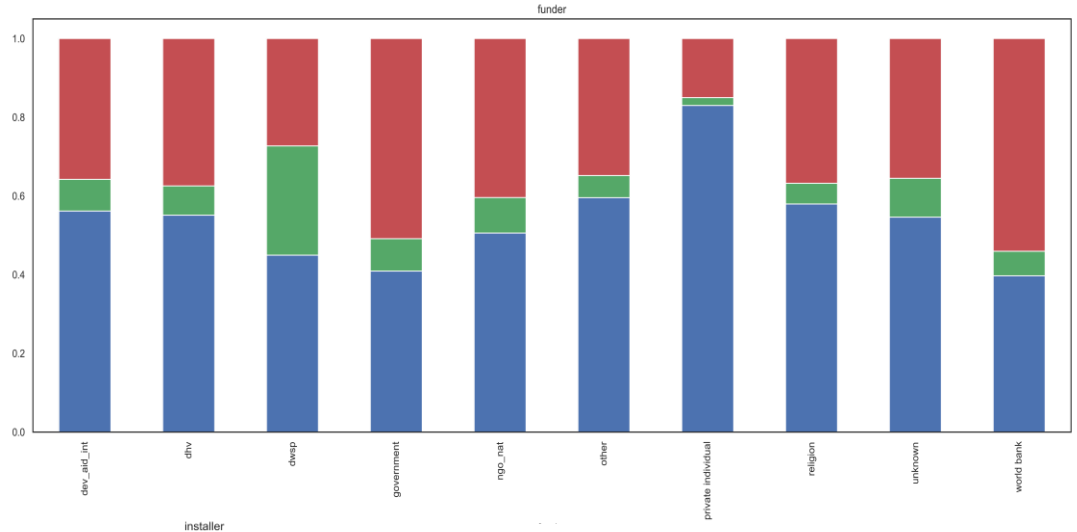
03 How could AI help?

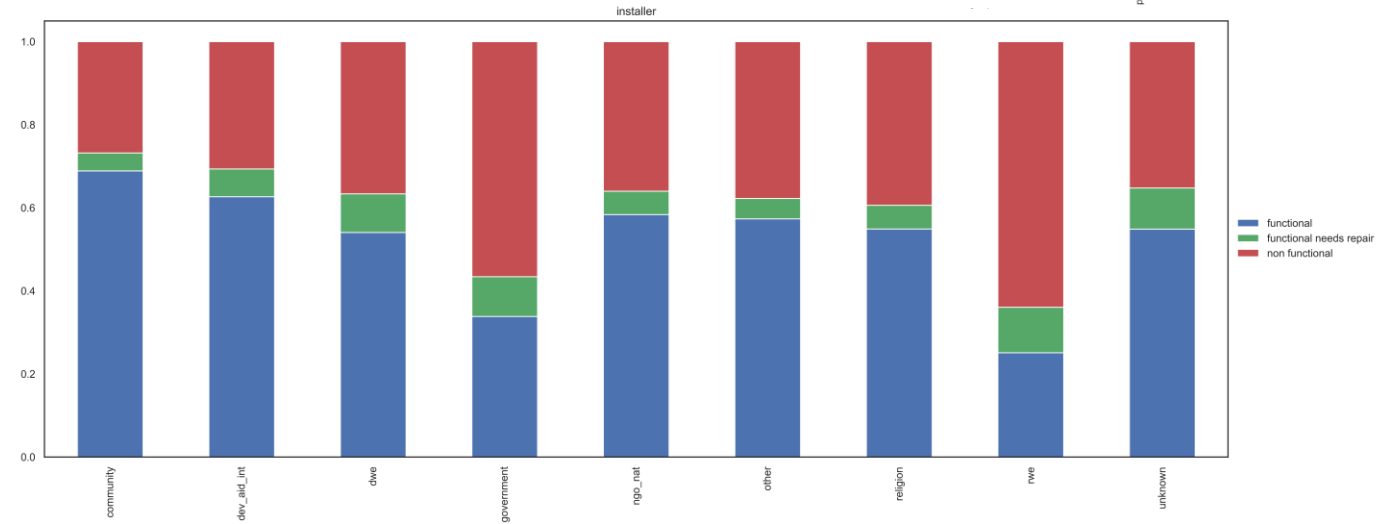


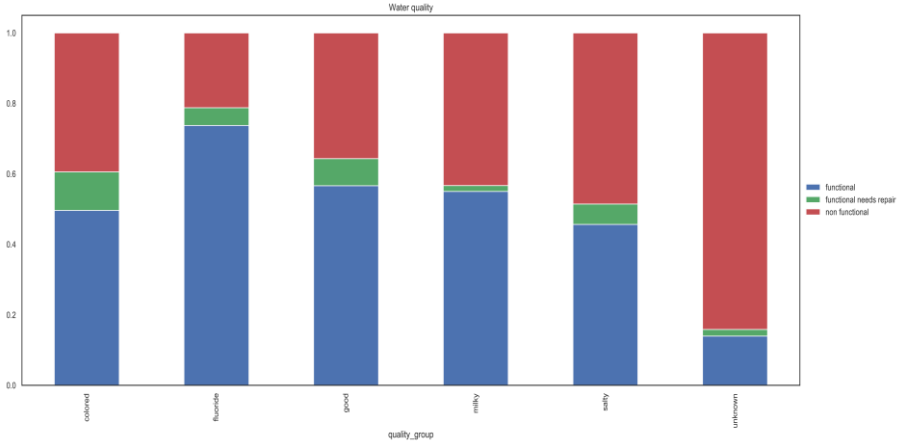
How could AI help?

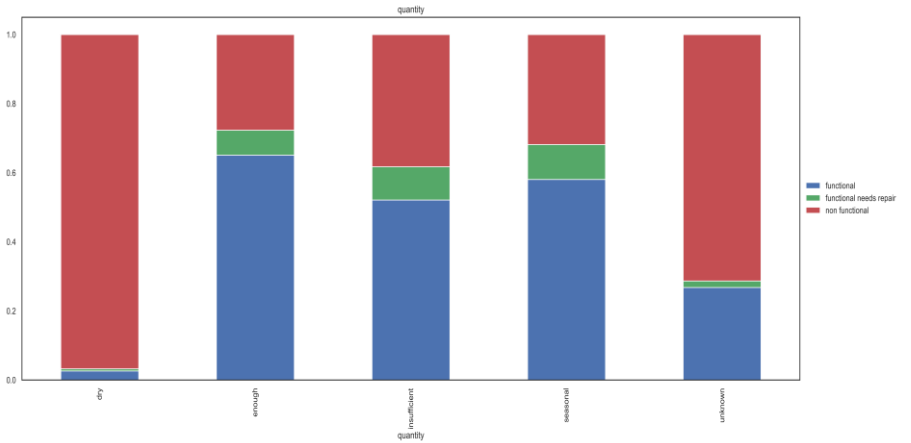


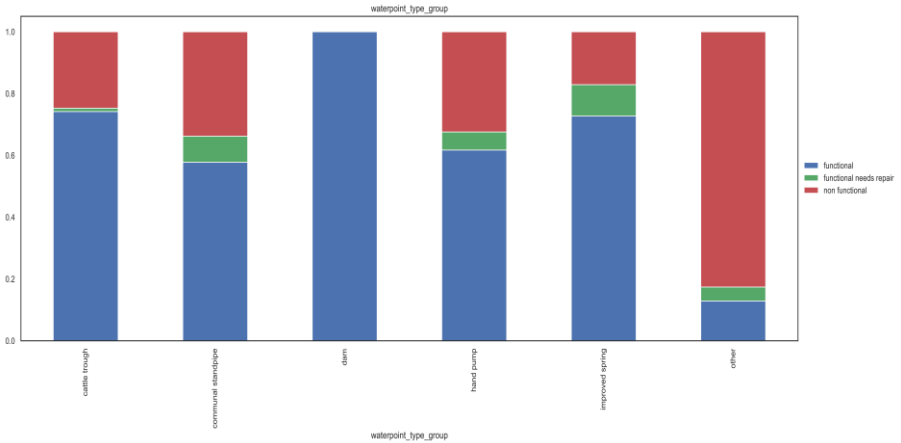
03 How could AI help?



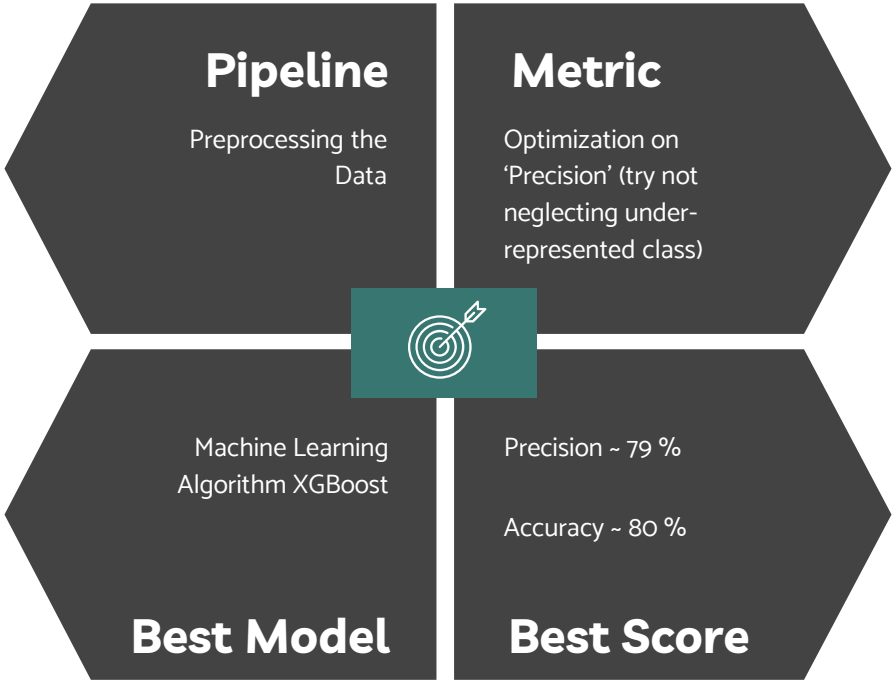



Water quality
1.0
0.8
0.6
0.4
0.2
0.0
colored
fluoride
good
milky
salty
unknown
quality_group
functional
functional needs repair
non functional

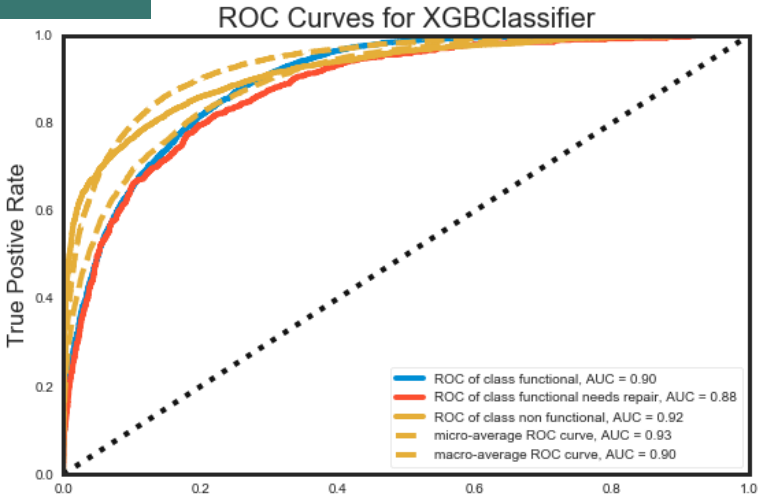

quantity
1.0
0.8
0.6
0.4
0.2
0.0
dry
enough
insufficient
seasonal
unknown
quantity
functional
functional needs repair
non functional

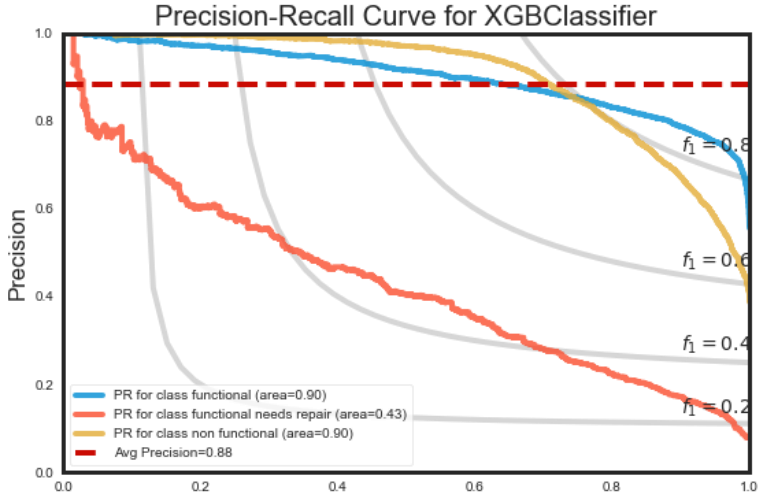

waterpoint_type_group
1.0
0.8
0.6
0.4
0.2
0.0
cattle trough
communal standpipe
dam
hand pump
improved spring
other
waterpoint_type_group
functional
functional needs repair
non functional

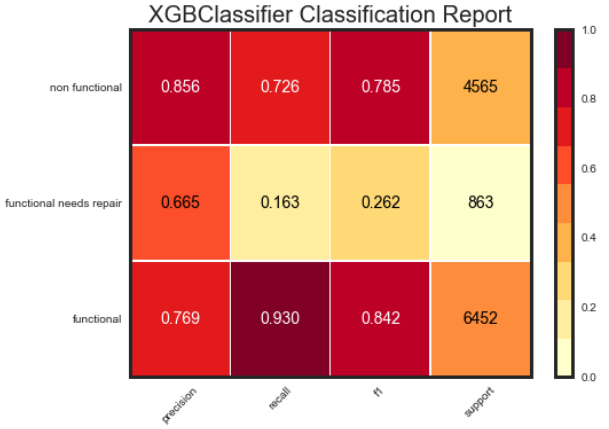
Prediction Model

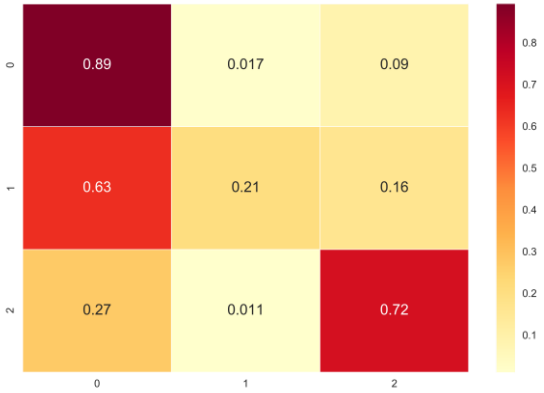


Prediction Model: XGB









Recommen-dation & Conclusion

- Fund a project in **Lindi ,Mtwara, Katavi.** A high percentage of non-functional water points leads to more people lacking access to potable water

- Do **not** fund in **Njombe or Iringa**. They both have a high percentage of working water points and a good water point to people ratio

➢ **Following these recommendations will help as much people as possible**

Tanzania is just the beginning…

Help making the world a better place!

Recommen-dation & Conclusion

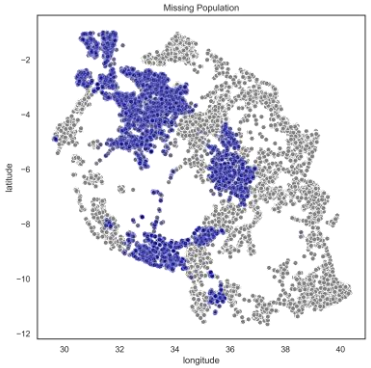
- Use **Development AId** after funding or when operating a water point.

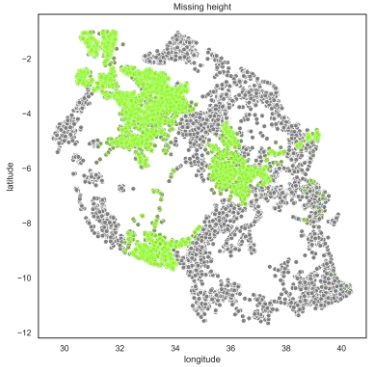
➢ **Following this recommendation will ensure a efficient use of time in money regarding the maintenance of your well**

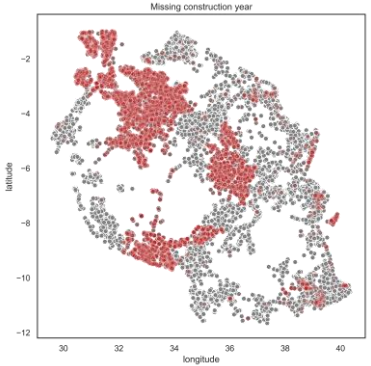
Tanzania is just the beginning…

Help making the world a better place!

Future Work 1 – Project and Data







- Improve the data especially for some regions
- Establish a better structured recording of existing and future waterpoints

Future Work 2 - Modeling

- Further Developments:
  - Additional improvement of models
  - Stacking of the best models to optimize the output
  - Use Neural Network work prediction
  - Develop more sophisticated strategies for imputing missing values

THANK YOU FOR YOUR ATTENTION

Any questions?

Follow the project at github: