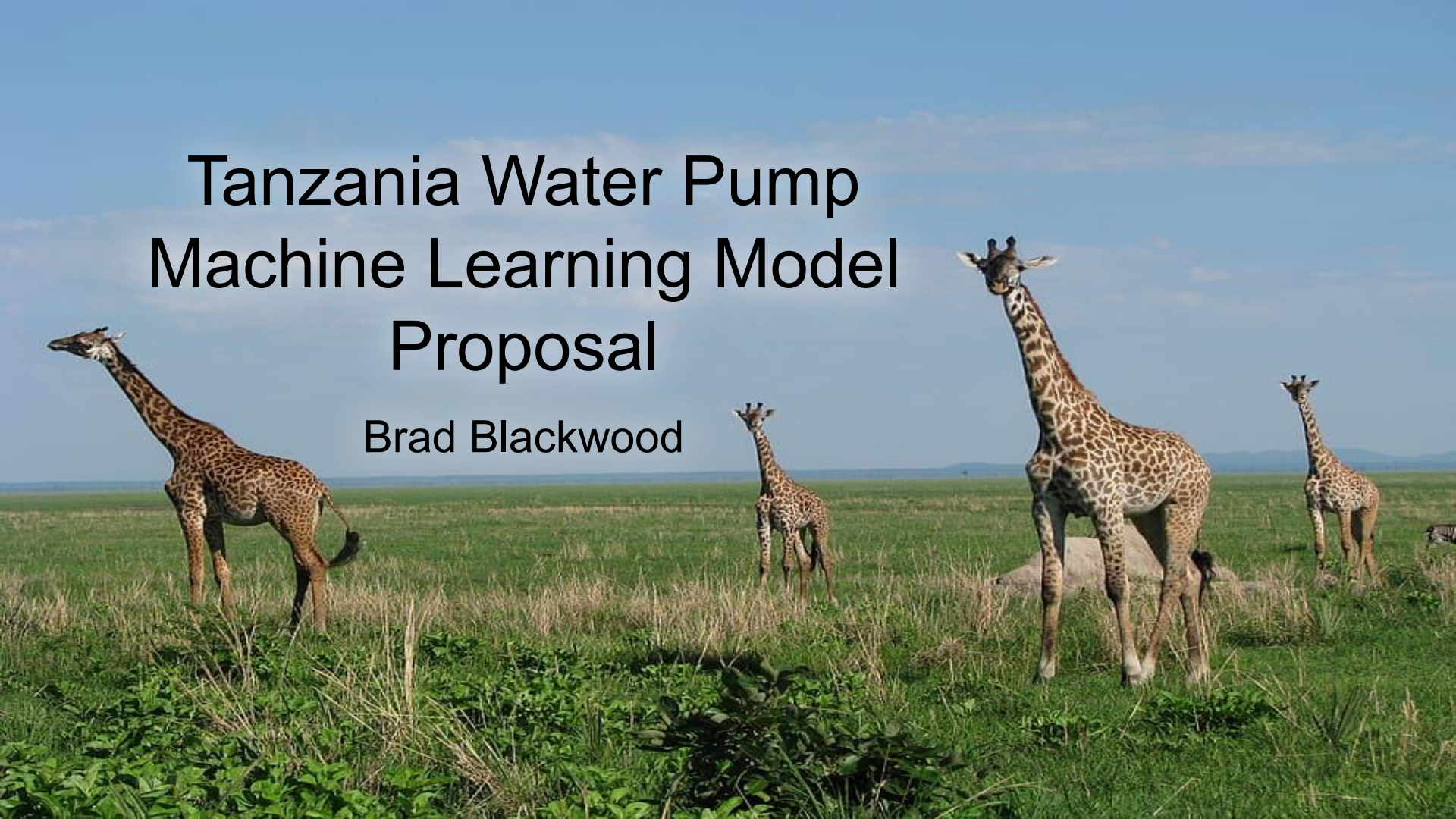


Tanzania Water Pump Machine Learning Model Proposal

Brad Blackwood

Overview:

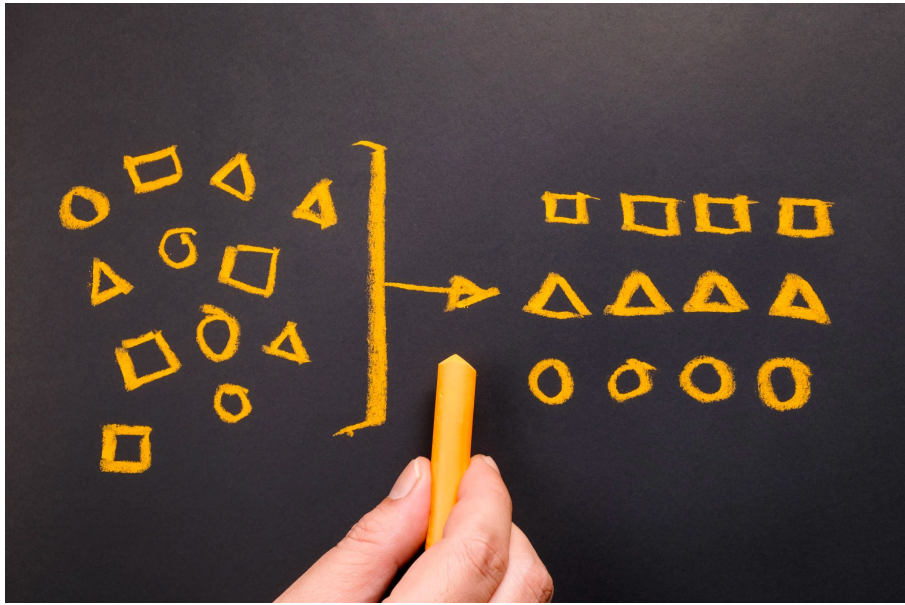
1. Problem
2. Data
3. Analysis
4. Conclusion

Problem:

Water pumps are in need of repair in Tanzania.

Solution:

Use Machine Learning to classify pumps that require attention.



Data:

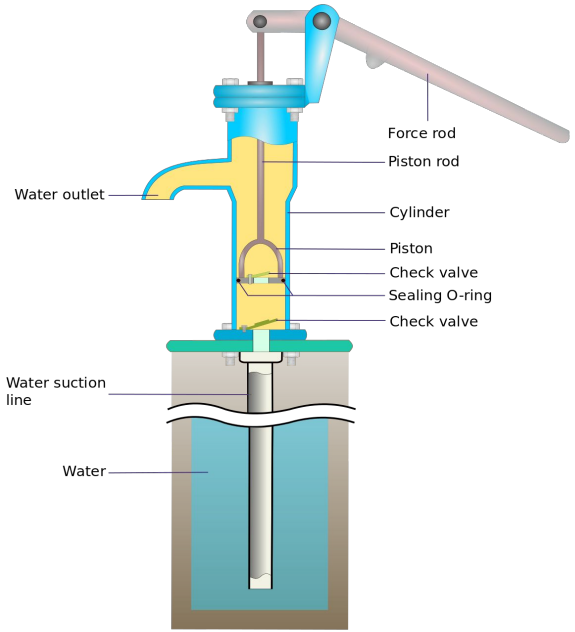
Provided by:

[Pump it Up: Data Mining the Water Table - Competition]

- Taarifa
- Tanzanian Ministry of Water

Shape:

- 59,400 water pumps
- 40 features



Data:

Features:

- Water source
- Location
- Pump type
- Management
- ect.

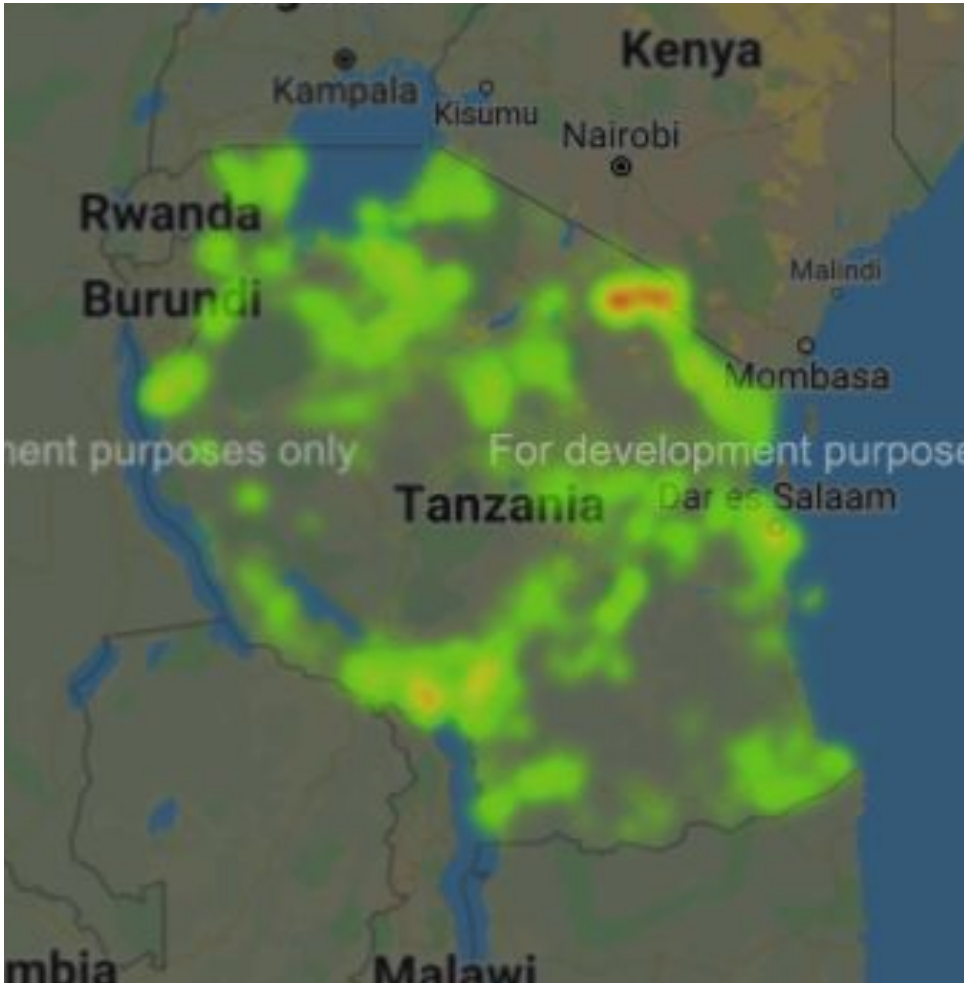
Target:

- Functional: 54.3%
- **Non functional/Need repair: 45.7%**

Modeling Objectives:

Identifying non functional pumps is **most** important.

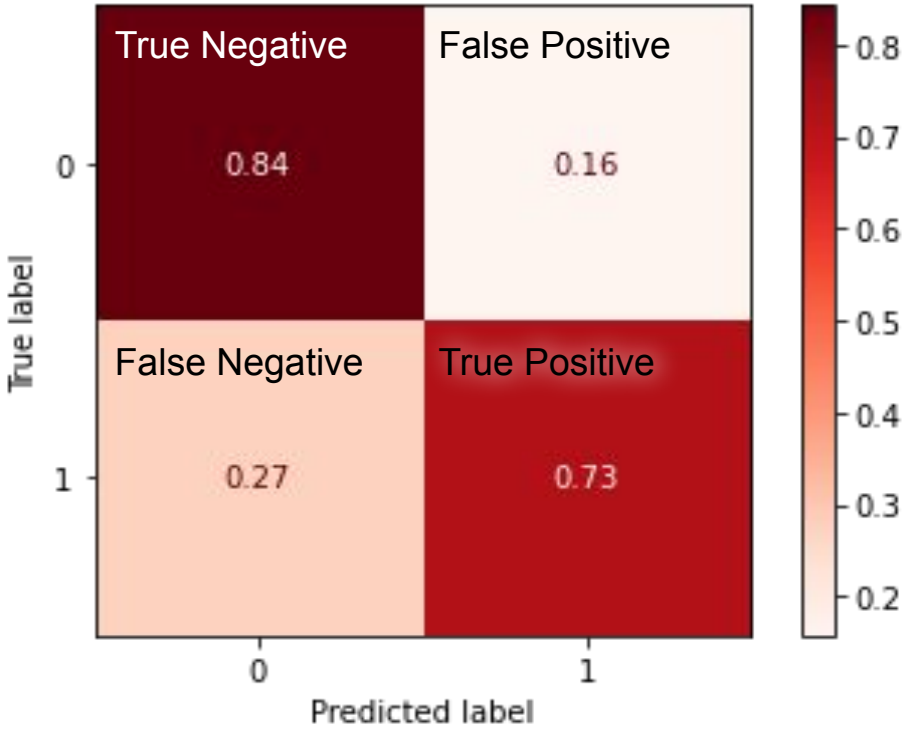
Minimizing the amount of prediction error is also important.



Analysis:

Random Forest Classifier

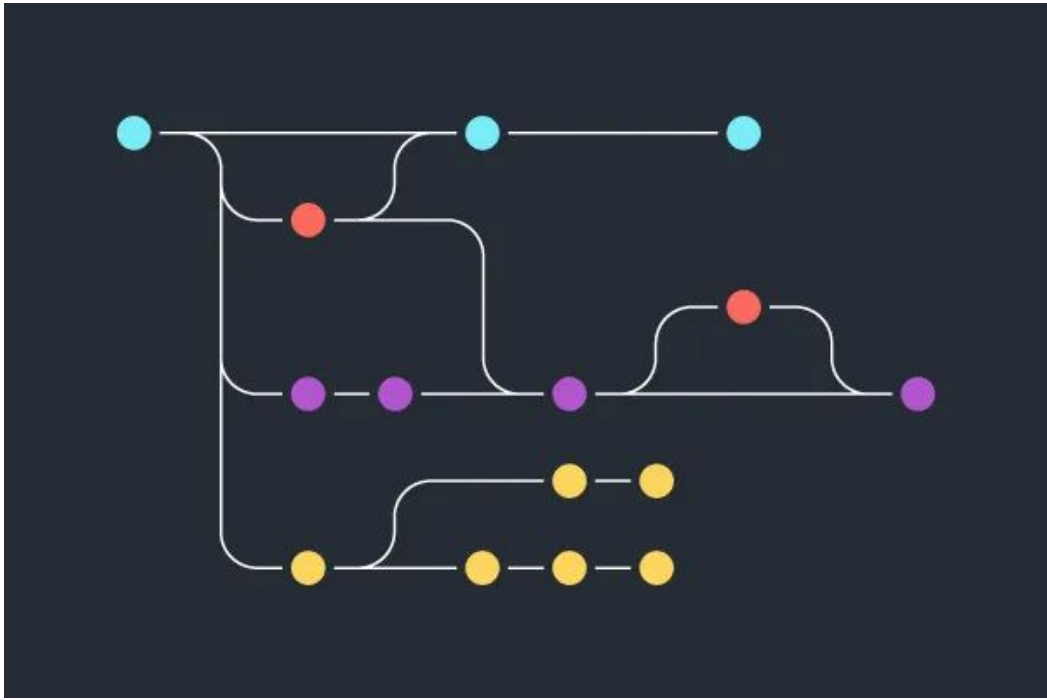
- Recall = 73% (optimized)
- Precision = 80%



Analysis:

The model **doesn’t** detect 27% of faulty pumps. (minimized)

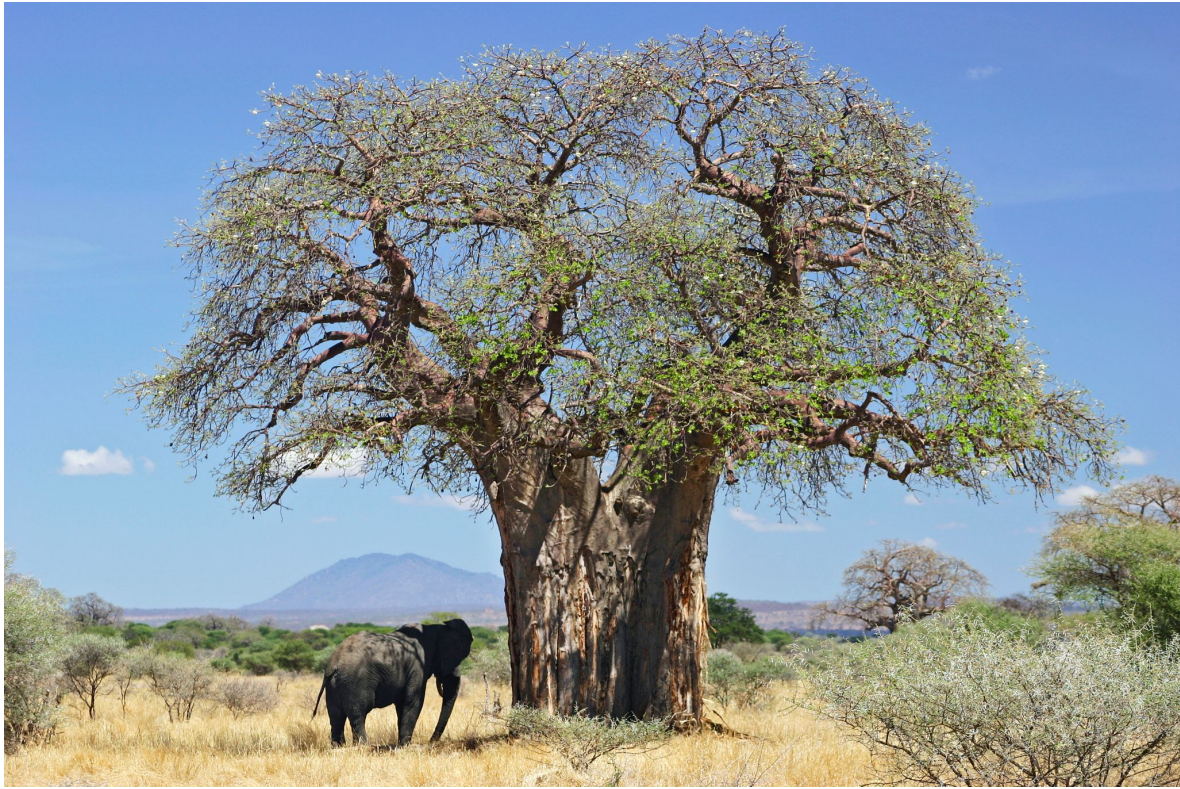
4 out of 5 faulty pumps the model predicts is correct.



Conclusion:

Recommendations

- Deploy the model (80% resource efficiency)
- Use alternative strategies to identify remaining faulty pumps.



Future Improvements:

Gain insight into data relationships.

What are the contributing factors?

Identifying types of pump failures.

Questions:

Gmail: brad.blackwood.03@gmail.com

LinkedIn: https://www.linkedin.com/in/blackwoodbs/

Github: https://github.com/blackwoodbs

Medium: https://medium.com/@brad.blackwood.03