Faulty Water Pumps Prediction

Mahdi KRIAA

Objective

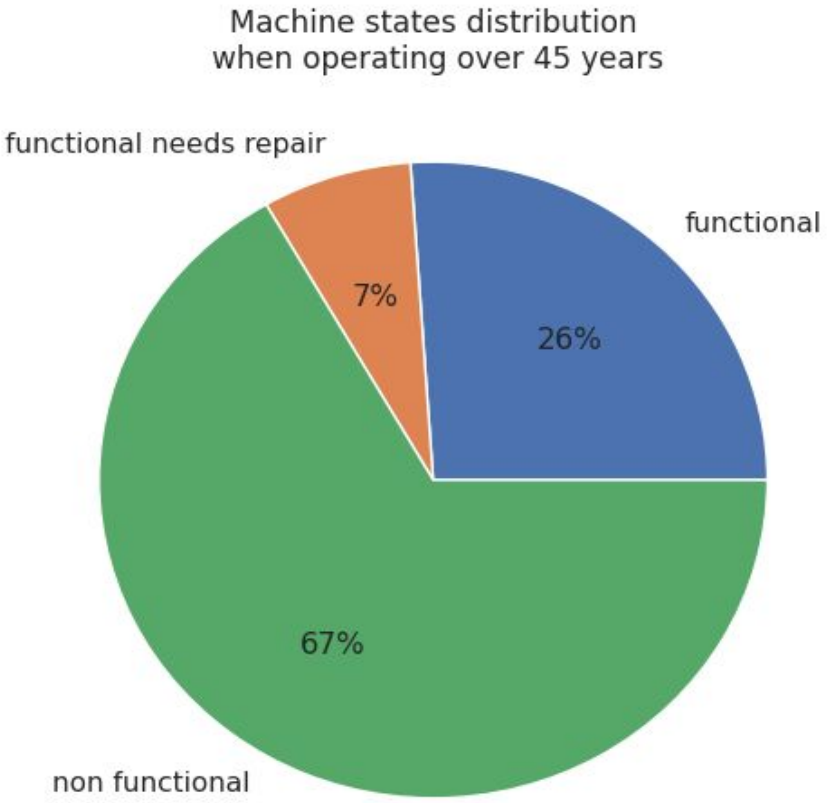
This project aims to leverage machine learning techniques to predict the operational status of water pumps in Tanzania. The goal is to identify faulty water pumps before they fail, allowing for timely maintenance and minimizing downtime. Water supply in Tanzania is critical, and pump failures can have significant impacts on local communities. By predicting these failures, we can ensure a more reliable water supply and improve the quality of life for these communities.

Data Description

The data for this project comes from Kaggle and it is originally comes from the Taarifa waterpoints dashboard, which aggregates data from the Tanzania Ministry of Water. It contains informations about what kind of pump is operating, when was installed, where it is located and how it is managed. This dataset contains 59400 rows and 41 columns including the target variable.

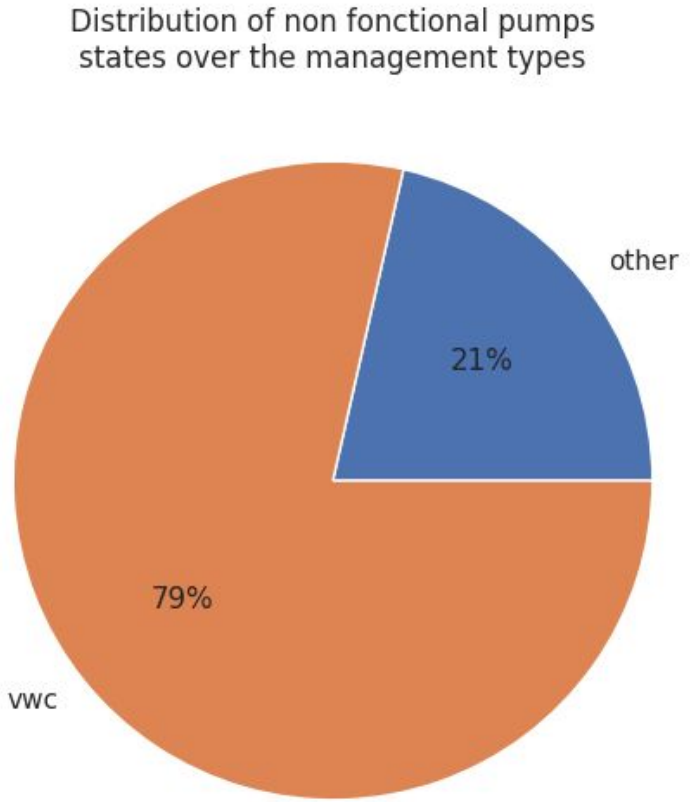
Data Explanation

The majority of pumps states is non functional when the pump is operating over 45 years so there is high risk of failure over this operating period.



Data Explanation

The majority of non functional states are vwc managed. So it's most likely that there is a problem with this type of management.



Model Strengths and Limitations

The model has a good performance on detecting states of non functionality, but it is weak face to states where the water pump is functional but need repair so using it to detect these states is not reliable.

Recommandations

- When the type of the water pump is "other" then the risk of failure is high. so this type must be dealt with.
- There is a few samples when the water pump is functional but needs repair, hence more data about this state must be provided to improve the model performance face to it.