PREDICTING THE OPERATIONAL STATUS OF WATER ACCESS POINT IN TANZANIA

INTRODUCTION

Access to clean and functional water sources is a fundamental necessity, yet many regions in Tanzania struggle to provide reliable water points for their population of over 67 million. Despite numerous water points being established across the country, a significant number are either non-functional or in need of repair. This situation hampers access to clean water, leading to health challenges, time lost in fetching water, and increased burdens on communities, particularly women and children

INTRO CONTS….

This project seeks to address this issue by building a machine learning model to classify the operational status of water points. By analyzing data related to water point features such as pump type, installation details, and geographic characteristics, the project aims to predict whether a water point is functional, functional but requiring repair, or non-functional. Such insights can empower stakeholders to take targeted and proactive actions to improve water access.

PROBLEM STATEMENT

- In Tanzania, many water points fail to provide reliable access to clean water, impacting millions of people and burdening rural communities. As a data scientist, my goal is to develop a machine learning model to predict the operational status of water points—functional, in need of repair, or non-functional. This will enable stakeholders to prioritize maintenance, allocate resources effectively, and address systemic issues, ensuring sustainable access to clean water for the population.

OBJECTIVES

1. Predict Water Point Functionality: Develop a machine learning models to classify water points as functional, functional but in need of repair, or non-functional.

2. Evaluate Model Performance: Assess the accuracy, precision, recall, and overall effectiveness of the models to ensure reliable predictions.

OBJECTIVES CONTS...

3

Identify Failure Patterns: Analyze the data to uncover factors and trends contributing to water point failures.

4

Optimize Resource Allocation: Provide actionable insights for stakeholders to prioritize maintenance and repair efforts effectively.

DATA

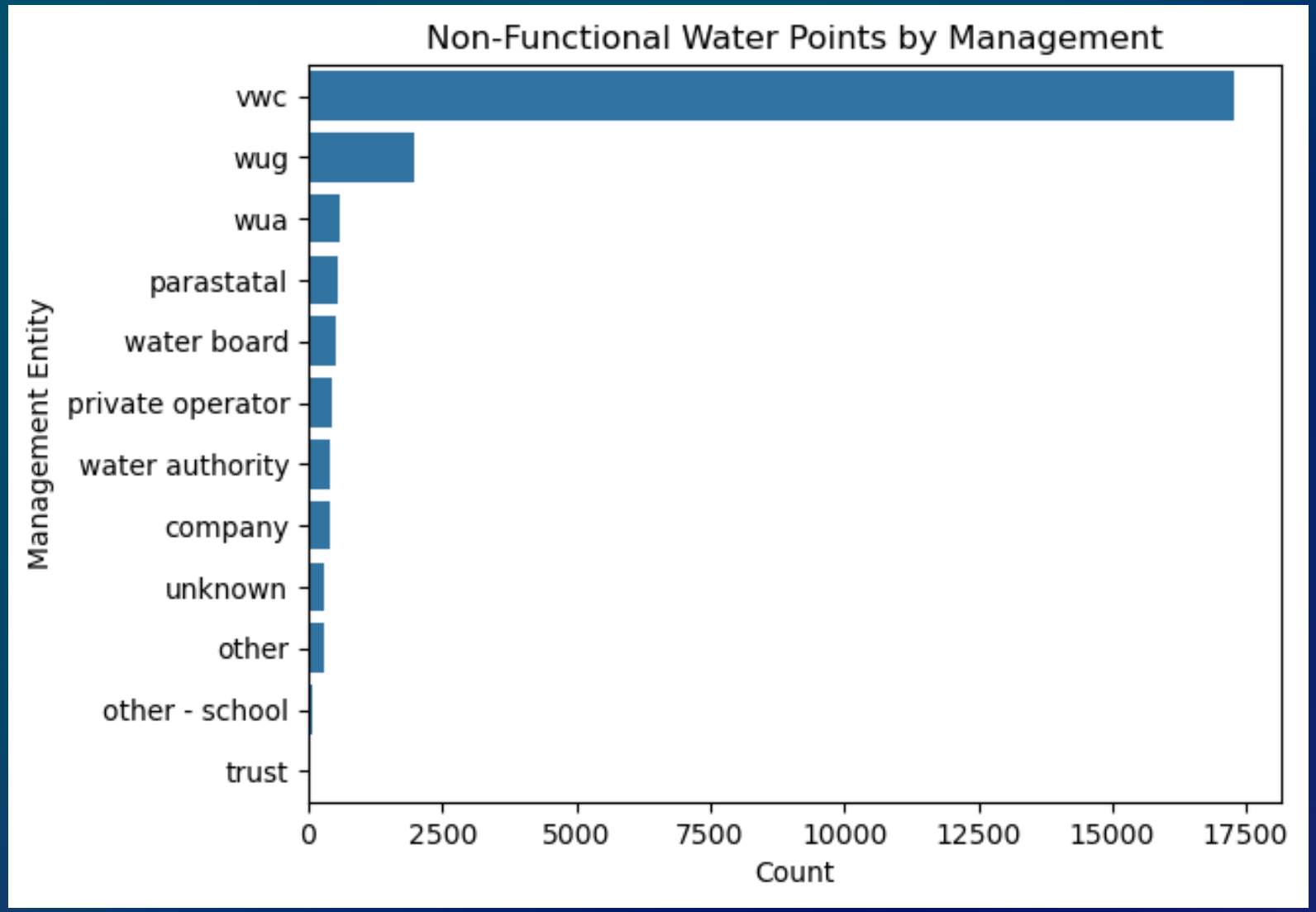
Source: The dataset was sourced from Taarifa and the Tanzanian Ministry of Water and is available on DrivenData. https://www.drivendata.org/competitions/7/pump-it-up-data-mining-the-water-table/page/23/

Features: The dataset includes attributes such as amount_tsh, funder, gps_height, installer, longitude, latitude, region, management, permit, construction_year, and more.

Target Variable: status_group (categorical: functional, functional needs repair, non-functional).

Methods

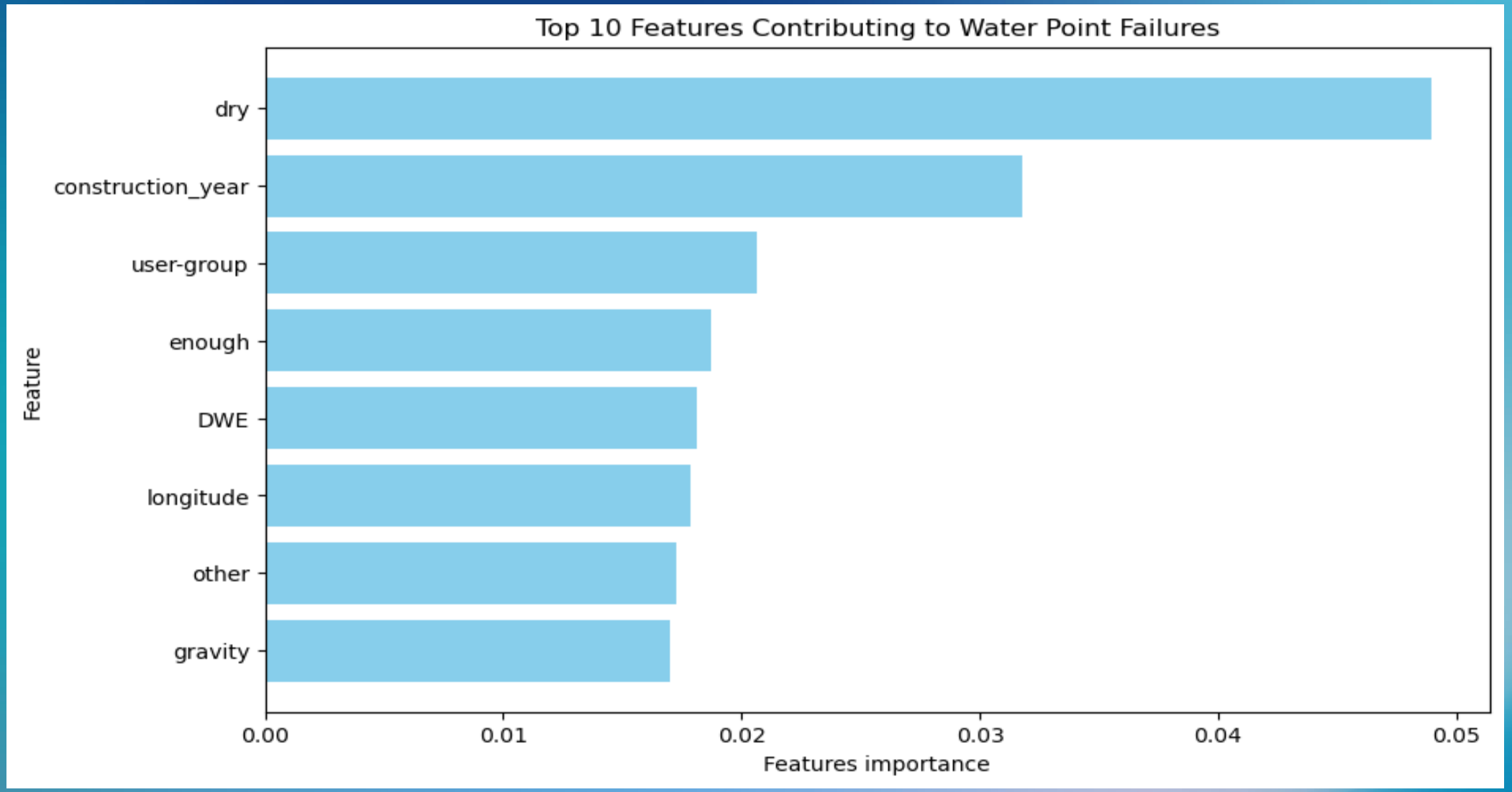
- Data preprocessing (e.g., handling missing values, feature scaling).
- Models used (e.g., Random Forest, Decision Tree, Logistic Regression).
- Evaluation metrics (e.g., accuracy, AUC, precision).


Non-Functional Water Points by Management
Management Entity
vwc
wug
wua
parastatal
water board
private operator
water authority
company
unknown
other
other - school
trust
0
2500
5000
7500
10000
12500
15000
17500
Count

OBSERVATIONS

VWC (Village Water Committees): This entity accounts for the majority of non-functional water points, indicating potential challenges in maintenance or resource allocation.Investigate the capacity, training, and resources available to VWC. Provide targeted training or support to improve their management effectiveness.

WUG (Water User Groups):The second-highest number of non-functional water points is managed by WUG.Evaluate their management processes and identify gaps that lead to failure.


Top 10 Features Contributing to Water Point Failures
dry
construction_year
user-group
enough
DWE
longitude
other
gravity
Feature
0.00
0.01
0.02
0.03
0.04
0.05
Features importance

OBSERVATIONS

Dry : This appears to be the most influential feature in predicting water point failures. It could indicate whether the water source has dried up. Prioritize investigating and restoring water points marked as "dry."

construction_year : Older water points might be more prone to failure . Focus maintenance on older water points or consider replacement for very old installations.

user-group: This might reflect how the water point is managed or shared among users . Assess the impact of user-group dynamics on functionality and repair needs.

Actionable Recommendations

Focus on older water points and areas with dry sources.
Investigate user management practices (user-group) and funder performance (DWE).
Focus on Underperforming Managers : Prioritize support for VWC and WUG to reduce the number of non-functional water points. Conduct capacity-building workshops and allocate additional resources.

THANK YOU