PREDICTING WATER WELL CONDITIONS IN TANZANIA

Presented By : Evans Oyugi

OVERVIEW

1. Business Understanding
2. Data understanding
3. Data prepararation
4. Modelling and Evaluation
5. External Validation
6. Conclusion
7. Recommendation

I.BUSINESS UNDERSTANDING

Problem statement

Lack of clean and potable water is a major issue in communities across Tanzania. The Tanzania Ministry of Water has installed several water wells. The aim is to improve maintenance operations and ensure that clean and portable water is available to communities across Tanzania

PROJECT GOAL

The goal of this project is to build a predictive model that can accurately predict the condition of water wells in Tanzania based on the variables provided in the data.

OBJECTIVES

Main Objective

To predict the condition of water wells in Tanzania to ensure that clean and portable water is available to communities across Tanzania.

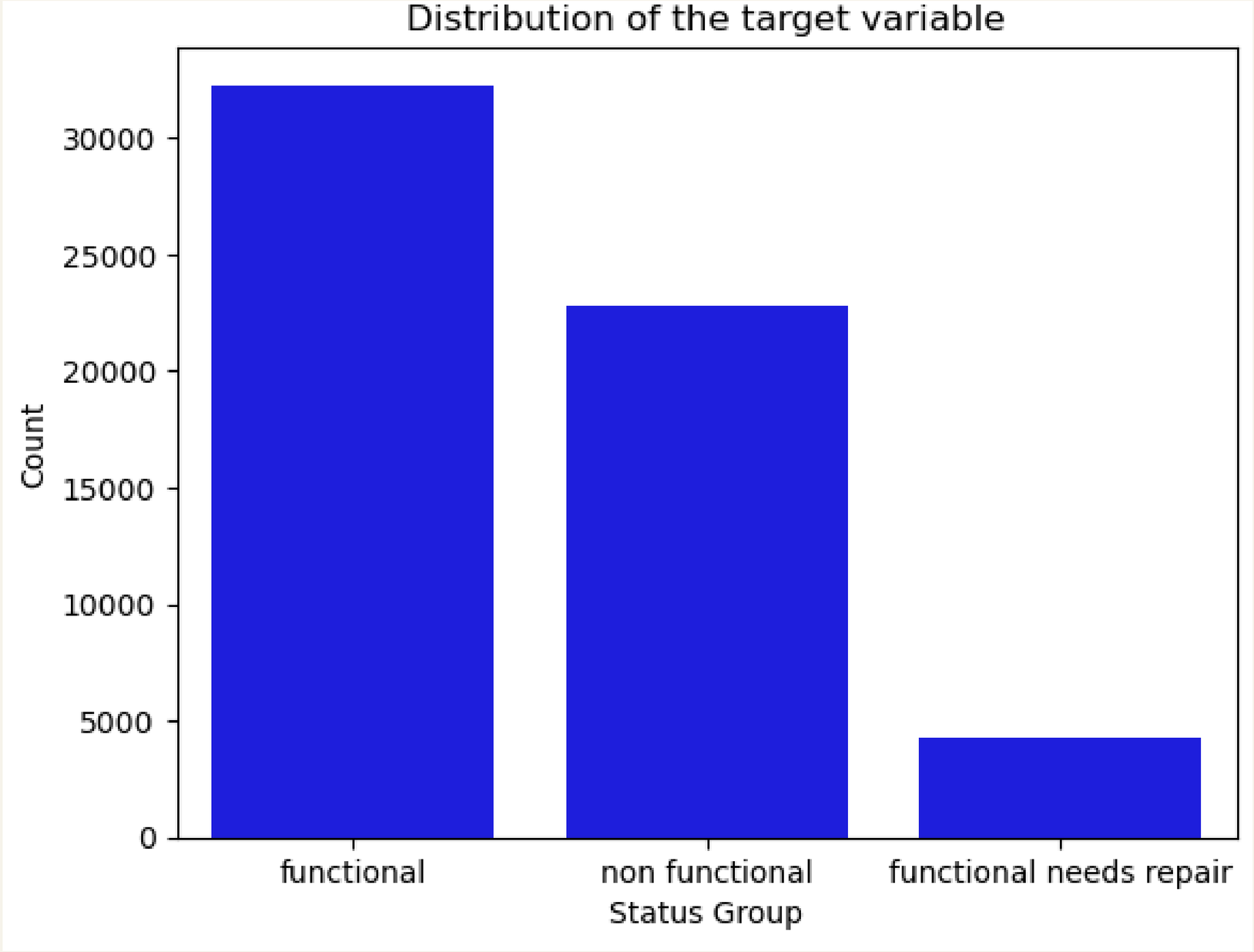
Specific Objective

I. To understand the problem statement and the goal of the project 2. To identify the variables that can impact the functionality of water wells 3. To determine the target variable (functional, need repairs, or non-functional)

ANALYSIS

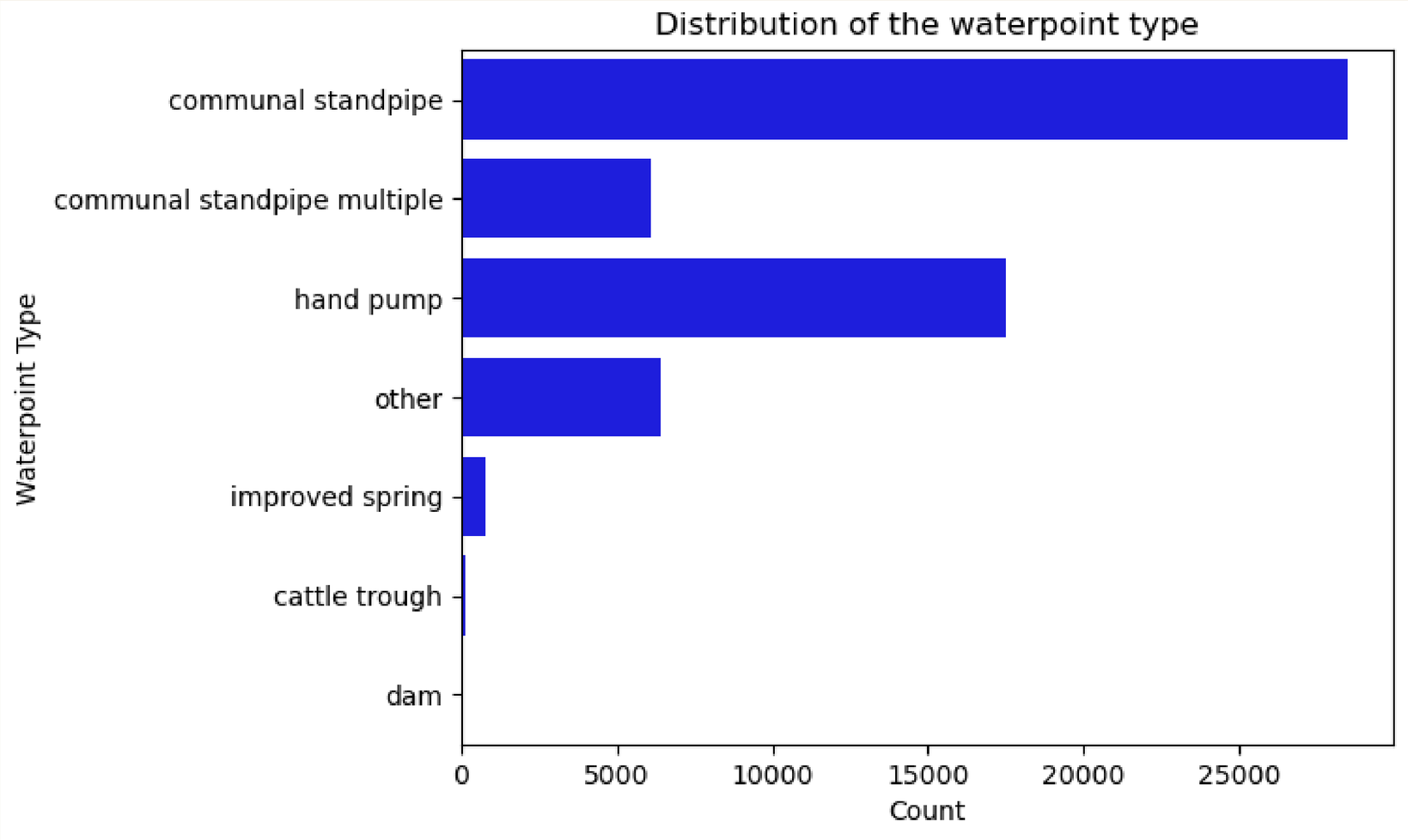

PEXELS
Best free stock photos in one place. Learn more
Search for free photos...
Top Photographers
Free Stock Photos
DELL

Status group



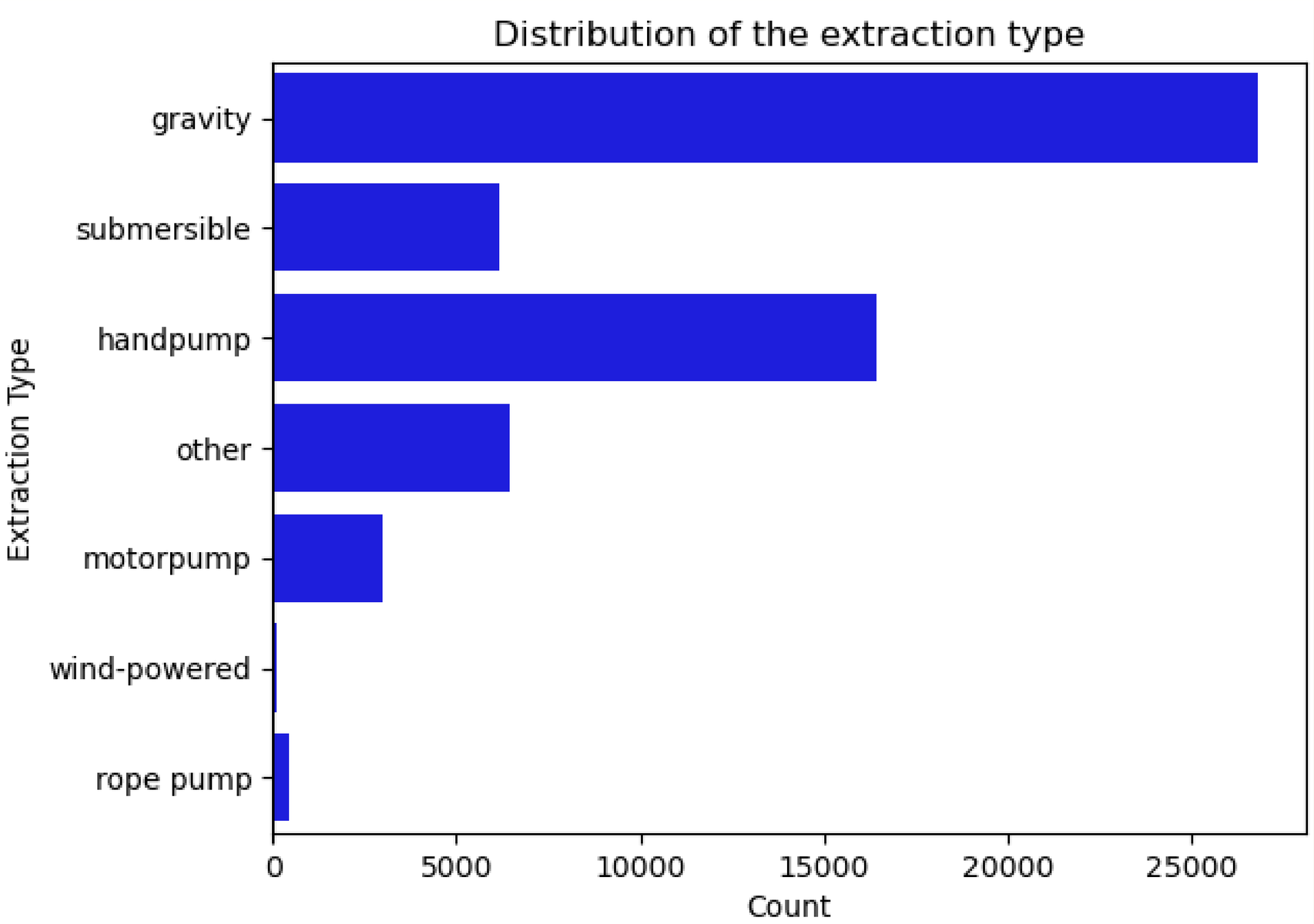
THE MAJORITY CLASS IS THE FUNCTIONAL CLASS WHILE THE MINORITY IS THE FUNCTIONAL NEEDS REPAIR CLASS

Waterpoint Type



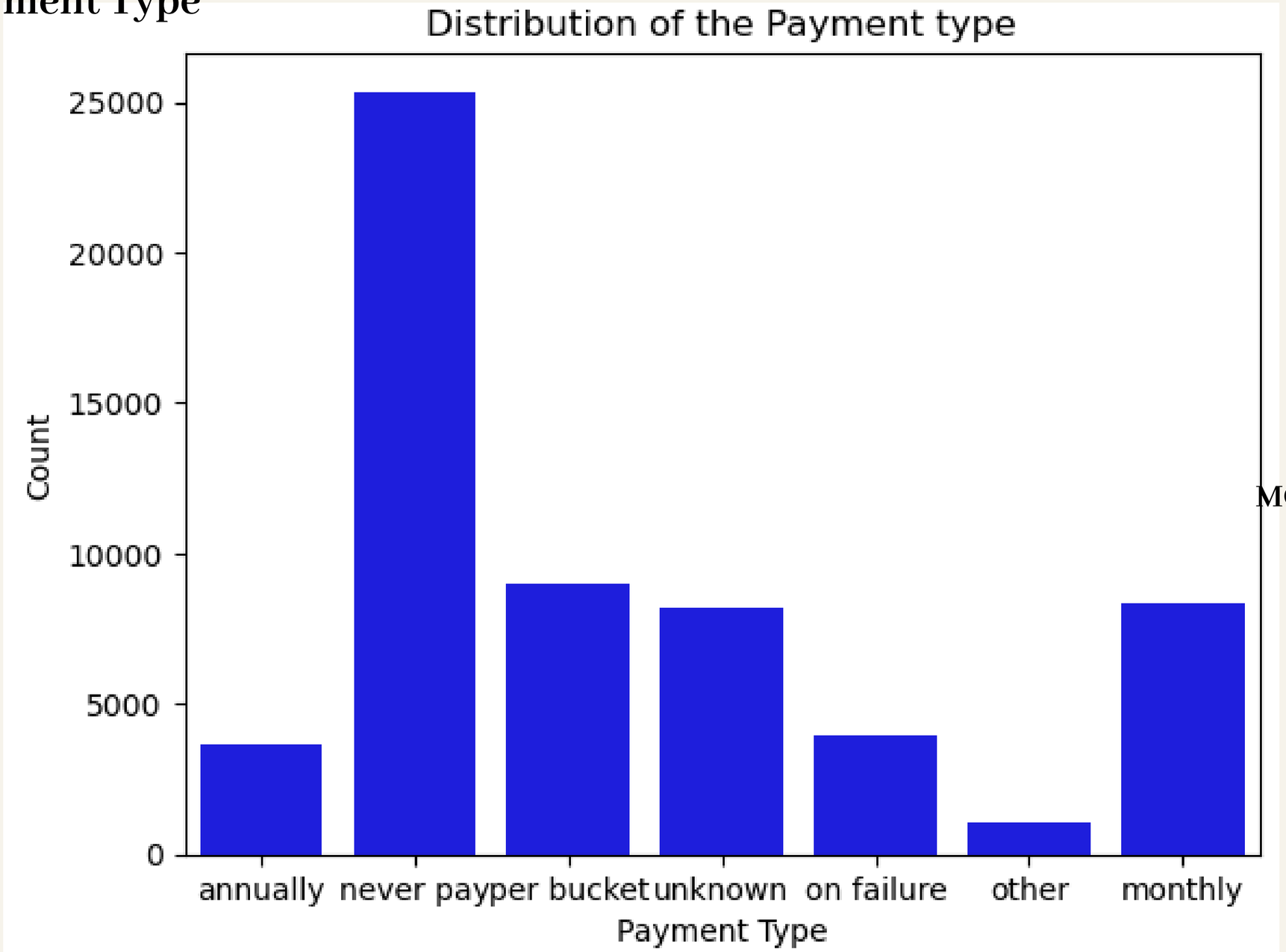
THE TYPE OF WATERPOINT FOR MOST WELLS IS THE COMMUNAL STANDPIPE FOLLOWED BY HANDPUMP

Extraction Type



THE EXTRACTION TYPE FOR MOST WELLS IS THROUGH GRAVITY

Payment Type



MOST WATERPOINTS ARE NEVER PAID FOR

3.DATA PREPATION

- Handling missing values
- Encoding and Categorical Variable
- Feature selection
- Scalling & Handling Class Imbalance

4.MODELING & EVALUATION

THE BEST MODEL (RANDOM FOREST) WAS SELECTED FROM FIVE CLASSIFIER MODELS THIS MODEL WAS TUNED TO IMPROVE ITS PERFOMANCE THE ACCURACY OF THE MODEL IS 0.8168, WHICH MEANS THAT THE MODEL CORRECTLY PREDICTS THE STATUS GROUP WITH AN ACCURACY OF 81.68%

EXTERNAL VALIDATION

SUBMISSION OF THE TEST PREDICTIONS MADE BY THE MODEL TO THE "PUMP IT UP: DATA MINING THE WATER TABLE" COMPETITION HOSTED BY DRIVENDATA:

- THE CLASSIFICATION RATE FOR THE SUBMISSION IS 0.8210 WHICH MEANS IT WORKS WELL ON BOTH SEEN AND UNSEEN DATA TO PREDICT THE WATER WELLS CONDITION

CONCLUSION

THE MODEL COULD BE FURTHER IMPROVED BY INCORPORATING MORE DATA ESPECIALLY FOR THE FUNCTIONAL NEEDS REPAIR CLASS TO HANDLE IMBALANCE FOR THE CLASSES

RECOMMENDATION

- THE TANZANIA MINISTRY OF WATER SHOULD INVEST IN BETTER WATERPOINT TYPES SUCH COMMUNAL STANDPIPES AND HAND PUMPS
- THE TANZANIA MINISTRY OF WATER SHOULD ENSURE THAT THE EXTRACTION TYPE FOR THE WELLS IS MOSTLY THROUGH GRAVITY AND HANDPUMP
- THE TANZANIA MINISTRY OF WATER SHOULD ENSURE THAT THE GPS HEIGHT(ALTITUDE OF THE WELL) FOR MOST WATERPOINTS IS HIGH ENOUGH
- THE TANZANIA MINISTRY OF WATER SHOULD ALSO ENSURE THAT THE PEOPLE USING THE WATERPOINTS PAY EITHER MONTHLY, ANNUALY OR PER BUCKET TO ENSURE THAT THE WELLS ARE WELL MAINTAINED

MORINGA SCHOOL | 2024

THANK YOU

Presented By : Evans Oyugi