BY DIANA TERRY AWUOR ALOO

SUSTAINING TANZANIA'S WATER WELLS WITH MACHINE LEARNING

Harnessing Machine Learning to Predict and Prevent Well Failures

Overview

- Rural Tanzanian communities depend on water wells for clean water
- Many wells break down or become non-functional.
- The project applies data-driven insights to predict and prevent failures.
- Outcome: More efficient maintenance, resource use, and community well-being

Business Understanding

Stakeholders:

- Ministry of Water & Irrigation, NGOs (e.g., WaterAid, UNICEF), donors, and local maintenance teams.

Problem:

- Many wells fall into disrepair due to age, environment, or poor construction.
- Maintenance funds are limited — predicting failures can save time and resource

NGO'S

Donors

Water Aid

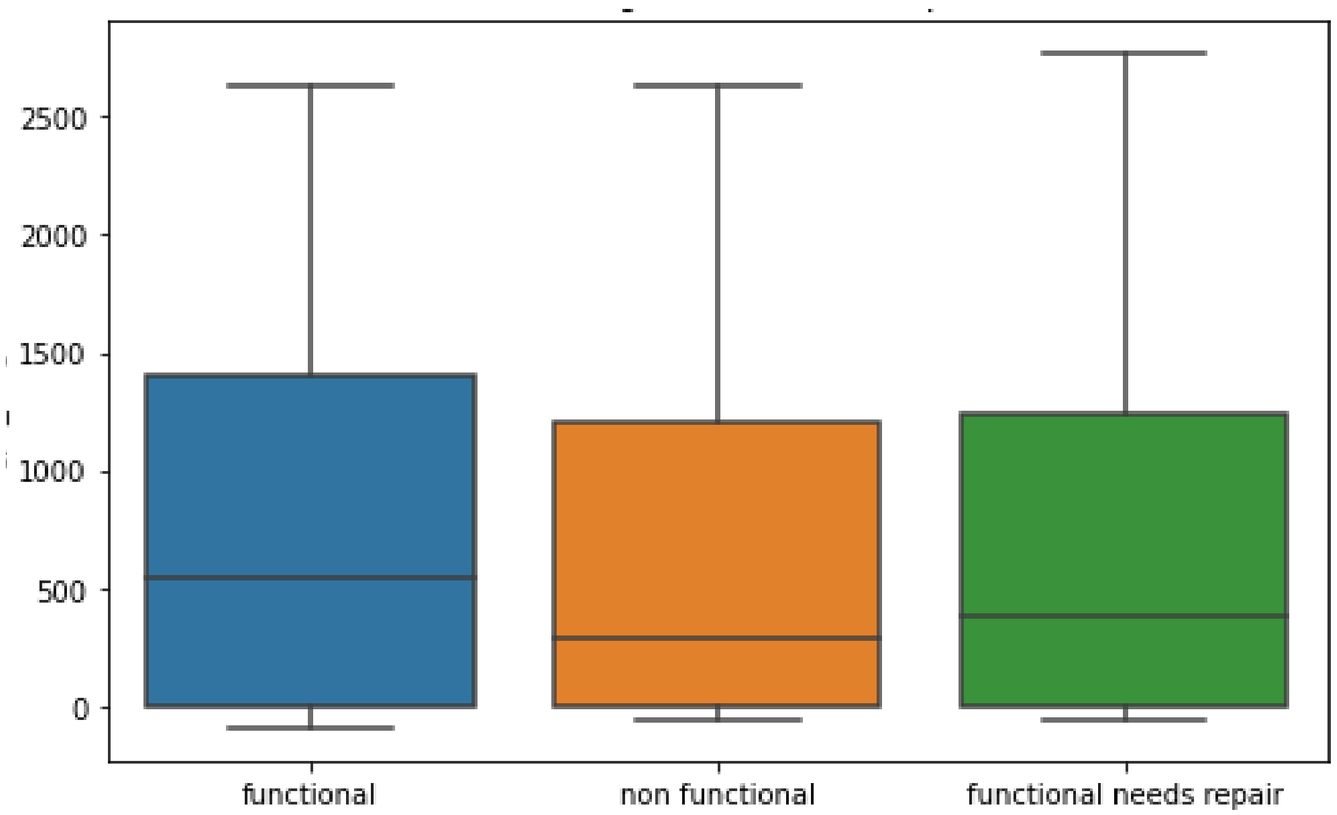
Ministry

Project Objectives & Business Objectives:

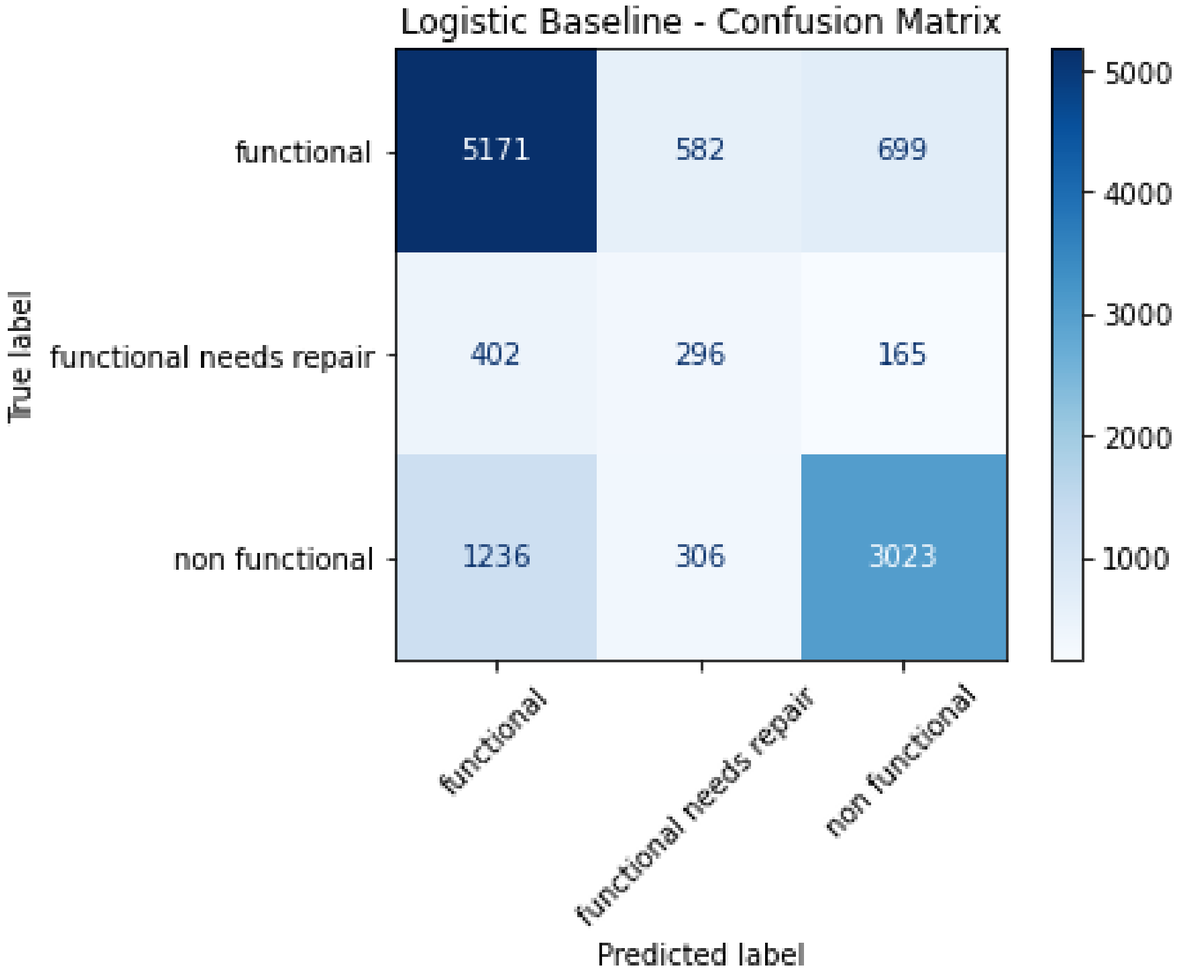
- Identify wells most likely to fail before they do.
- Support better maintenance scheduling and budget allocation.
- Reveal key factors behind well failures (e.g., installer, pump type, region).
- Build an interpretable classification model to predict well status.
- Use data to uncover trends that can guide sustainable water management.
- Deliver actionable insights and a model that stakeholders can use in planning.

Data Understanding

- Data includes well characteristics: location, installer, pump type, construction year, water source, usage level.
- Target variable:
  - Functional
  - Functional needs repair
  - Non-functional
- Covers multiple Tanzanian regions and community setups.



Bivariate that explains the data sources, outlines what the chart illustrates, and provides guidance on interpreting the visualizations to inform decision-making.

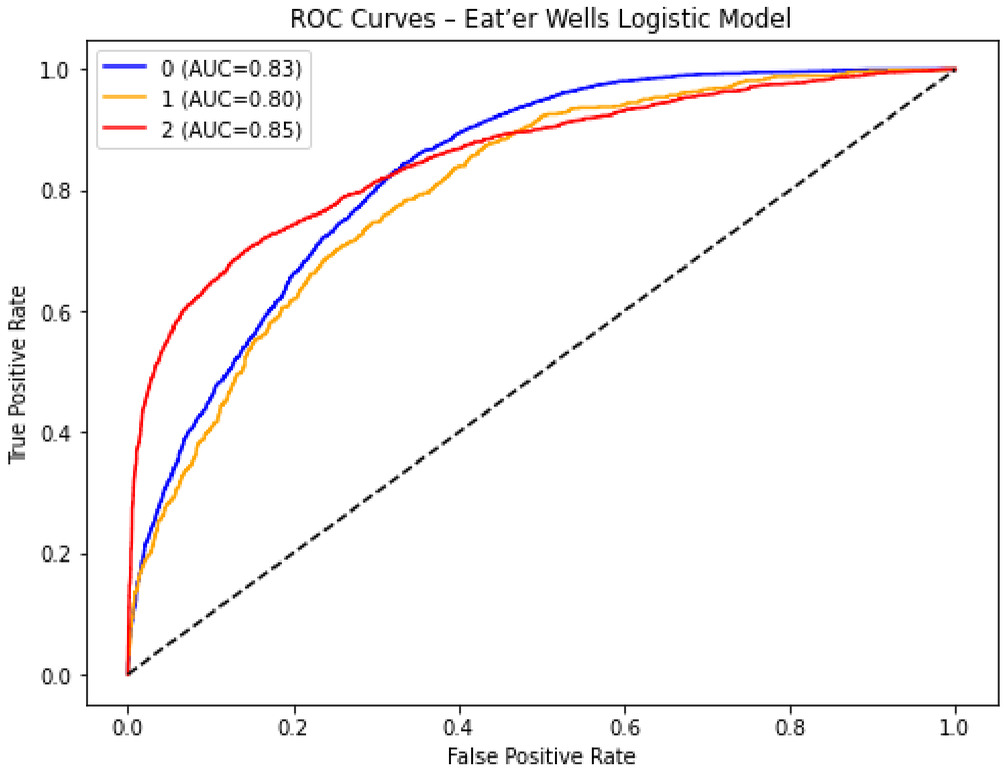


Modeling Approach

- Used classification models to group wells based on operational status.
- Models tested: Logistic Regression, Decision Tree, and possibly Ensemble models.
- Classification chosen because it helps categorize wells by risk — high, medium, low.

Model Evaluation

- Evaluated on unseen data to test real-world reliability.
- Metrics like accuracy show how well predictions match actual outcomes.
- Example: “Our model correctly predicts 8 out of 10 wells’ status.”
- Indicates strong potential for practical field deployment.

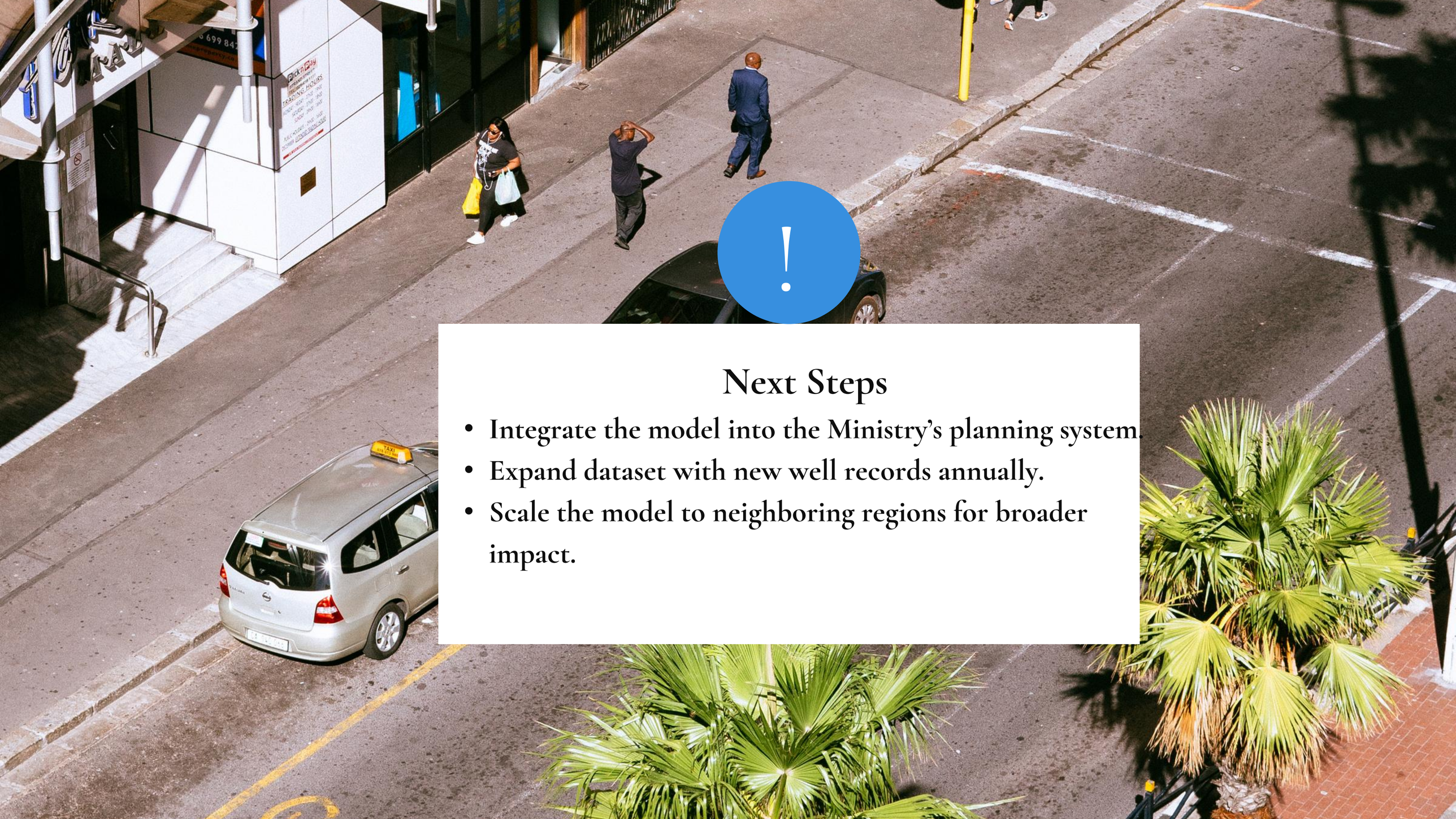


Key Insights

- Age of well and pump type strongly affect functionality.
- Installer quality and region also influence performance.
- Wells with less frequent maintenance show faster decline.

Recommendations

- Prioritize maintenance on wells predicted to fail.
- Standardize successful pump types and installation practices.
- Train field teams on data collection to keep model performance strong.
- Develop a dashboard to visualize and track well health.



Next Steps

- Integrate the model into the Ministry's planning system.
- Expand dataset with new well records annually.
- Scale the model to neighboring regions for broader impact.


THANK YOU!