

Predicting Water Pump Functionality in Tanzania

Author: Janine Makorre





Business Context

Why This Matters

- Tanzania faces a water crisis, with many of its 57 million people lacking clean water.
- The Ministry of Water needs to maintain thousands of water pumps but struggles to identify which ones are broken or at risk.
- Our project uses data science to predict pump functionality, helping prioritize repairs and improve water access.



Business Problem and Key Questions

The Challenge

Engineers lack a reliable way to know which pumps are non-functional or likely to fail, delaying repairs.

Key Questions

- *Which pumps are functional, non-functional, or need repairs?*
- *What factors (e.g., water quantity, pump age) predict pump failure?*
- *How can we prioritize maintenance to ensure clean water access?*

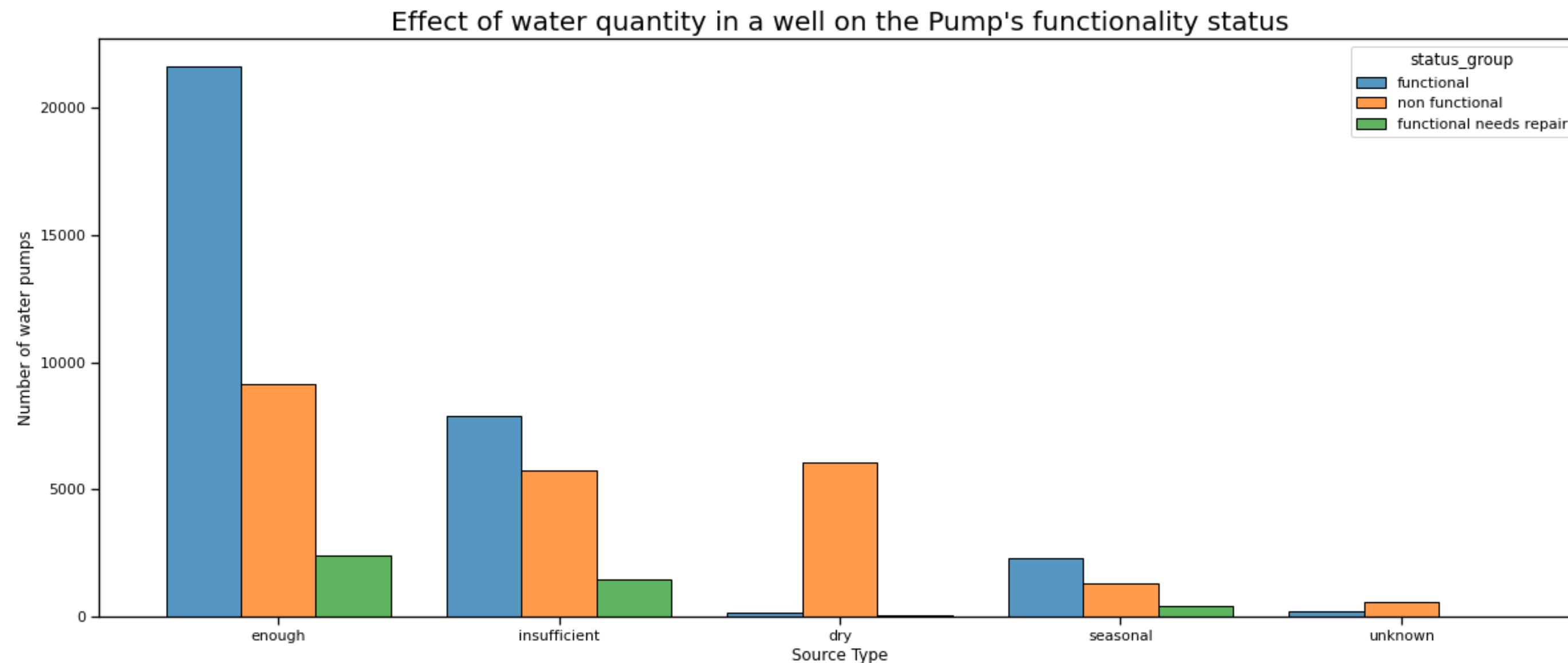
Dataset Overview

- Sourced from Taarifa and the Tanzanian Ministry of Water.
- 59,400 training records, 14,850 test records, with 40 features (e.g., water quality, pump age, location).
- Labels: Functional (54%), Non-Functional (38%), Needs Repair (7%).
- Visual: Table or chart showing data breakdown (e.g., pie chart of label distribution)

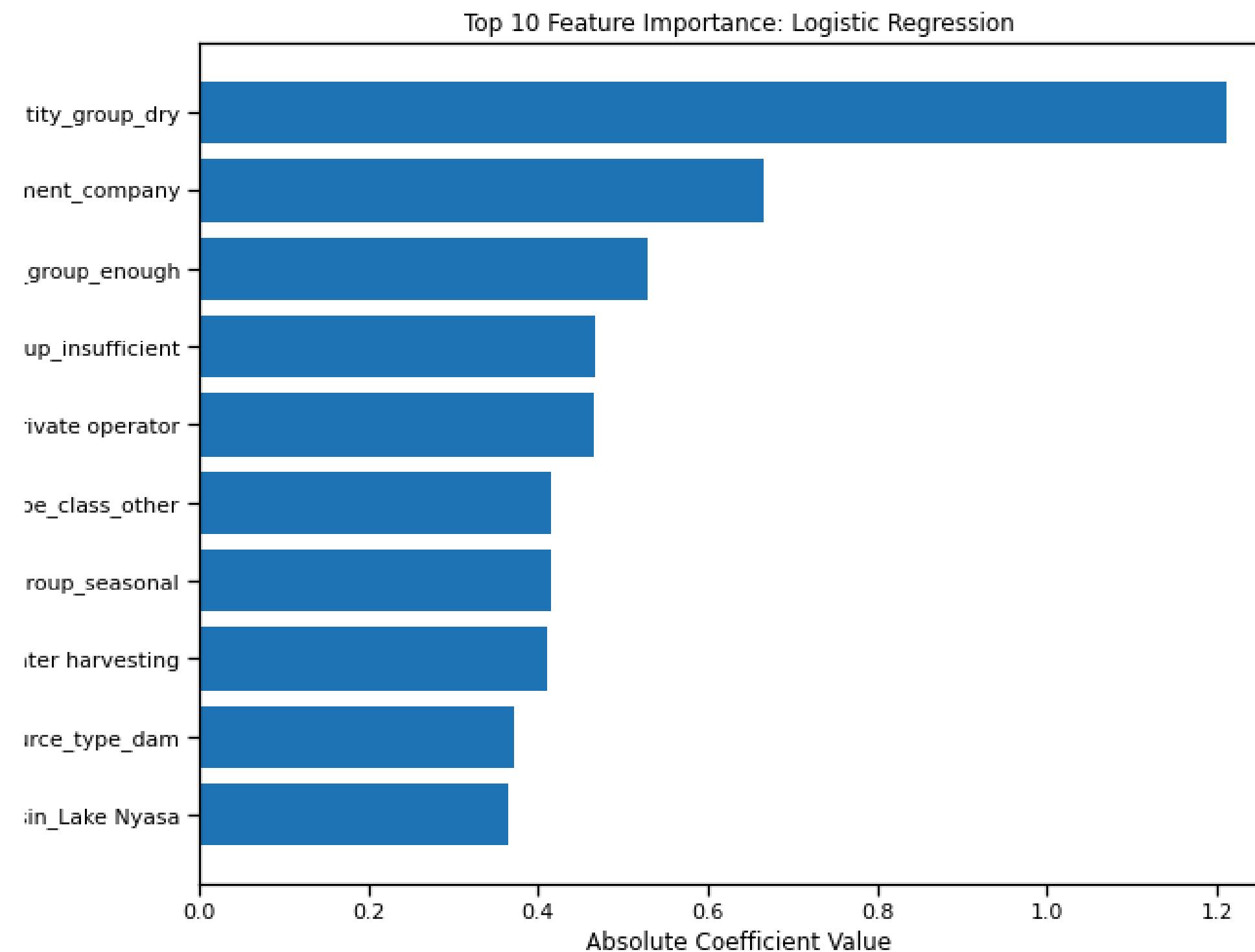
Our Results:



1. Dry wells strongly predict non-functional pumps

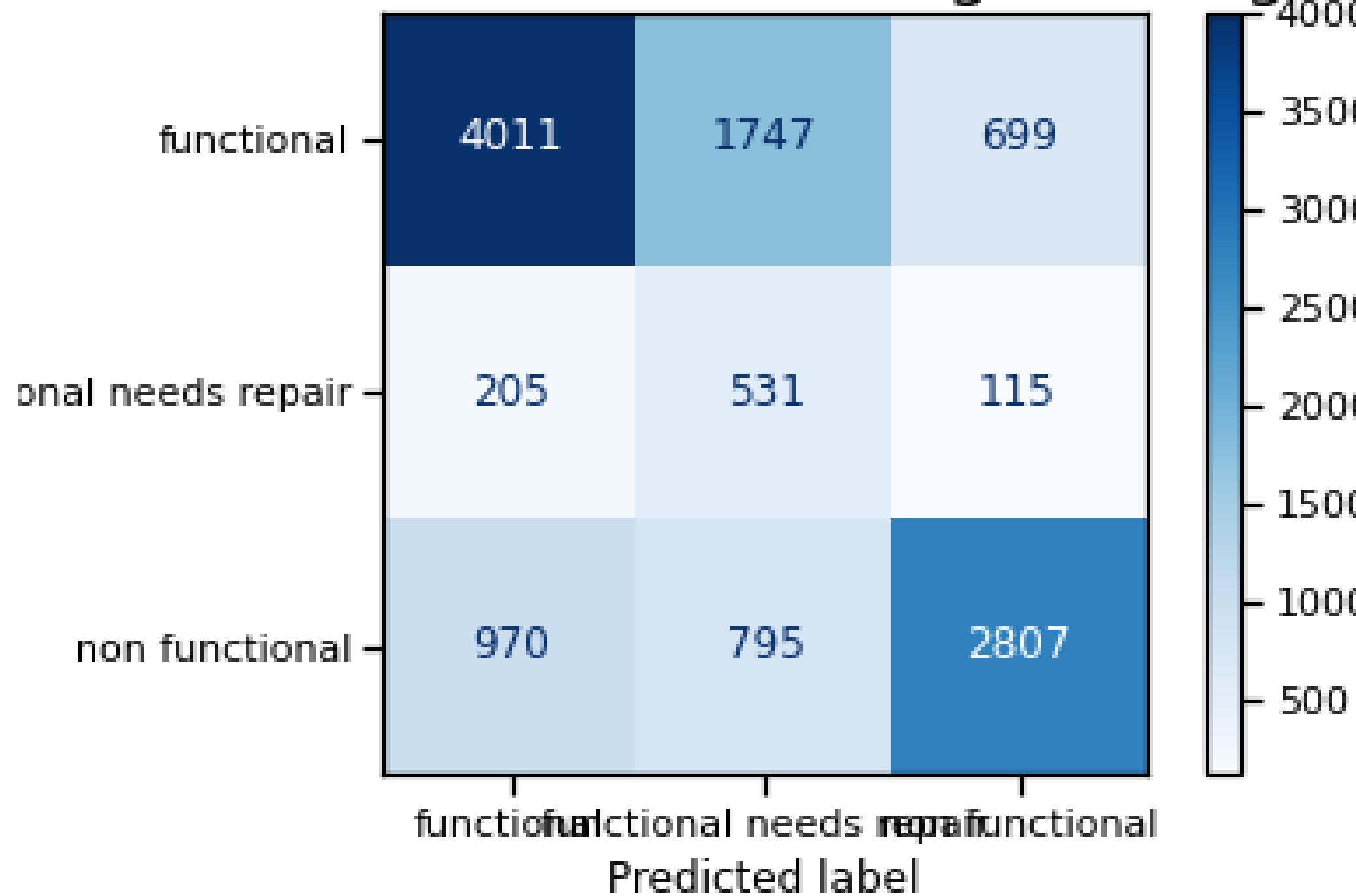


2. Older pumps and those without permits are more likely to fail



3. Non-functional pumps are identified well, but “needs repair” is harder to predict

Confusion Matrix: Tuned Logistic Regression



Recommendations

- Prioritize Repairs: Focus on pumps in dry wells and those over 20 years old.
- Review Processes: Investigate permit and public meeting practices, as they impact functionality.
- Improve Data: Collect more detailed data (e.g., maintenance history) to boost predictions.

Next Steps

- Test advanced models (e.g., Random Forest) to improve prediction accuracy.
- Collaborate with the Ministry to integrate predictions into maintenance workflows.
- Expand data collection to include real-time pump status updates.

THANK YOU

Any Questions?