Team 5

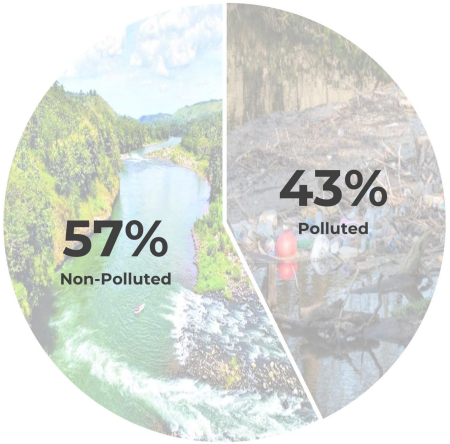
Brodeth
Van Jersey Paolo
Advanced System Administration

Pascual
Ken Leonard
Data Science

Engr. Ji Han Gang
Adviser

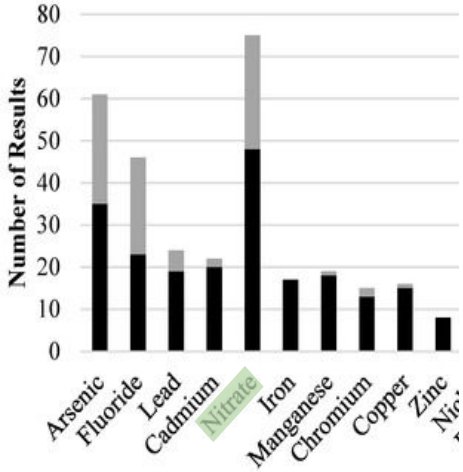
The Problem and Client

43% of principal rivers in the Philippines are polluted due to domestic, industrial and agricultural waste



from WEPA report

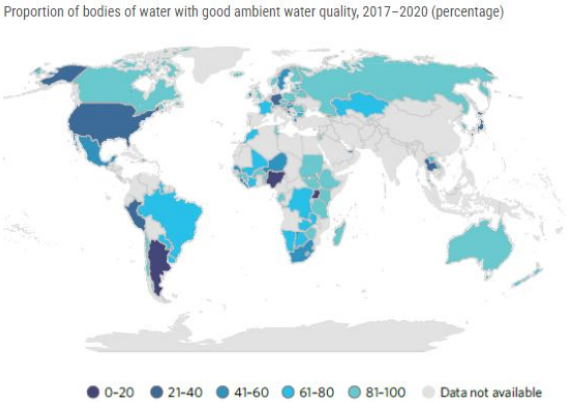
Water contamination due to chemicals is a growing problem.



from https://www.mdpi.com/2071-1050/16/16/7107

Lack of data poses a risk to people within areas where water quality of freshwater is unknown, including hikers and campers.

There is a lack of affordable, user-friendly technologies to allow users to quantitatively assess and monitor local water quality independently



from https://unstats.un.org/sdgs/report/2023/goal-06/

Client Details

Danielle Fortich Dolom

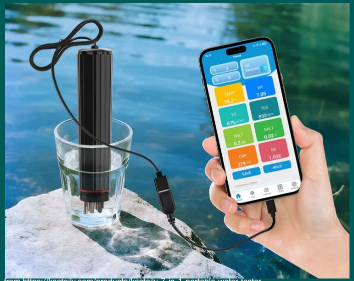
Casual Hiker
Hikes through mountainous regions and terrain 3-5 times a year
Part of a created team of hikers.

Existing Solutions & Gaps

SIR-based Smartphone Colorimetry

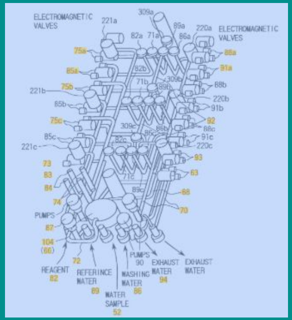
from https://pubs.acs.org/doi/10.1021/acsomega.8b00625

Kactoily 7-in-1 Water Tester



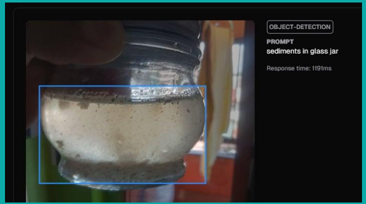
from https://kactoily.com/products/kactoily-7-in-1-portable-water-tester

Hitachi Patent on Water Quality Monitoring Tool



from https://patents.google.com/patent/US6444172B2/en

Moondream 2 AI VLM



WaterScope Testing Platform

from https://www.nature.com/articles/s41545-024-00357-y

Design of a Machine Learning-based Water Quality Classification Tool for On-site Colorimetric Analysis

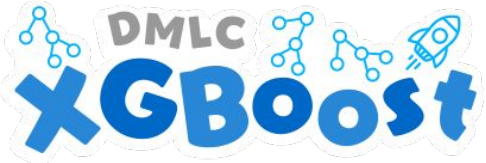
Objectives

1. Develop a web-based application that:
    a. Processes and classifies water sample quality with focus on chemical contamination and water usability using machine learning algorithms.
    b. Outputs water classification and analytical interpretations based on colorimetry data gathered
    c. Saves and organizes the output for each specified testing location.
2. Test and evaluate the system's accuracy

Potential Alternative Designs

LightGBM

- More efficient in larger datasets



XGBoost

- Fast training time

CatBoost

- Yields high performance even in small datasets

General System Architecture

Input

Water Sample
+
Chemical Reagent

↓

Substances in water
(Nitrate, Ammonia, etc.) react with
Chemical Reagent

↓

Water color changes

↓

Colors are converted to
numbers via
Colorimetric Reading

↓

Numeric concentration values

→

Process

Numeric
concentration values

↓

Machine Learning
Model

↓

Processing

→

Output

Water Classification

↓

Store record into
database

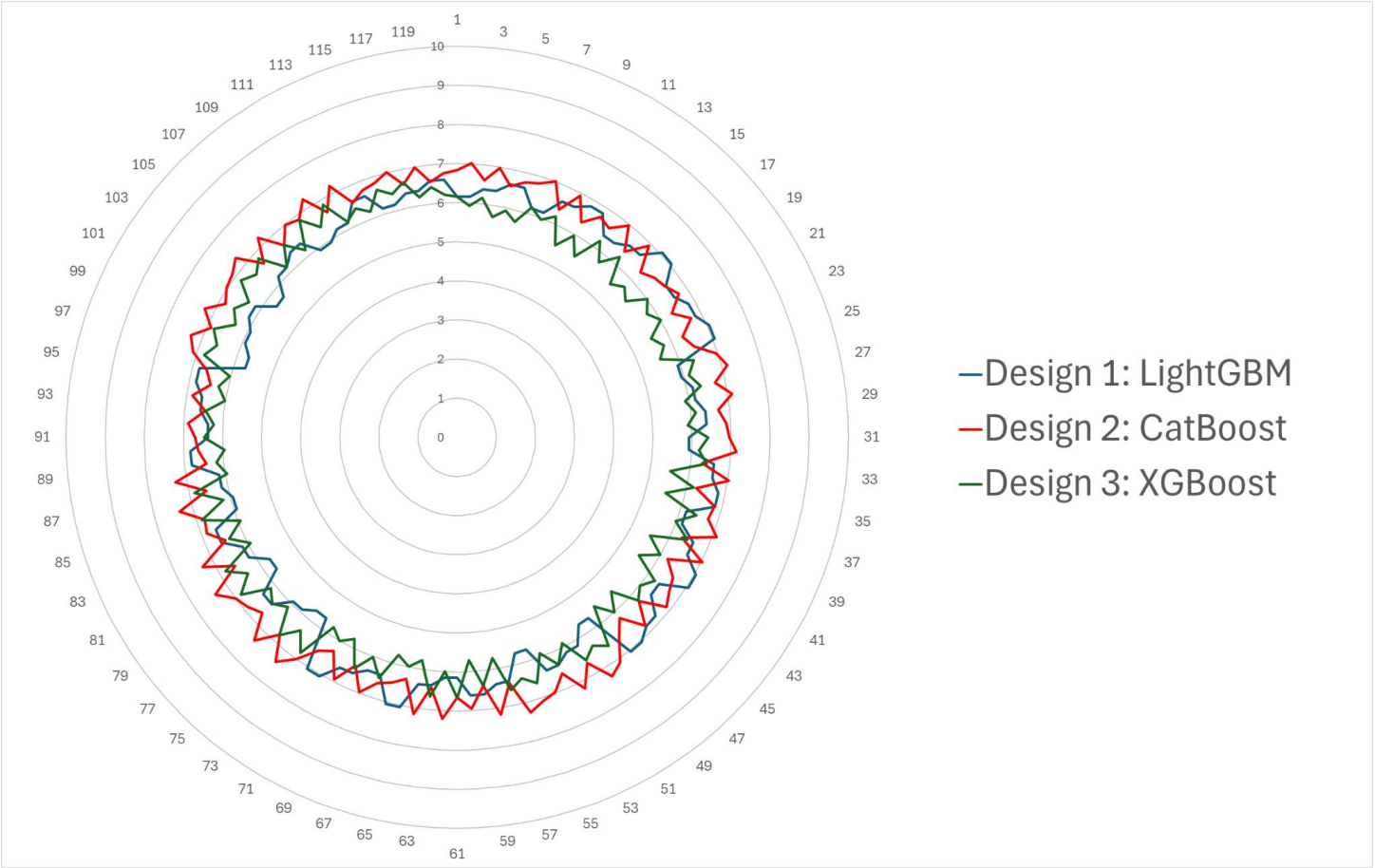
Design Constraints

Safety	Misclassification Rate
Performance	Inference Time
Manufacturability	Training Time
Compatibility	Maintainability Index Score
Efficiency	Storage Consumption

Risks and Action Plan

Identified Risk	Action Plan
Lack of local data, which can affect accuracy of trained model upon deployment	Gather data, and reuse the data for periodical model training
Data drift	Regular model evaluation; Periodical model training

Appendix A: Sensitivity Analysis



Appendix B: Proof of Concept

```
Please input the following water colorimetry concentration values:

Ammonia (mg/l): 90
pH (ph units): 5
Nitrate (mg/l): 90
C:\Users\Leon\miniconda3\envs\CPE312_Pascual\Lib\site-packages\sklea
  warnings.warn(

Predicted water class: C4

Prediction saved to predictions_log.csv
Current file size: 182 bytes
```

Ammonia (mg/l)	pH (ph units)	Nitrate (mg/l)	Timestamp	Predicted_Class
0.05	8	2	2025-11-24T10:22:57	A3
2	9	7	2025-11-24T10:23:52	C3
90	5	90	2025-12-06T11:08:30	C4

Appendix C: Concept UI

Water Quality Classifier - Log Viewer

Choose a CSV file

Drag and drop file here
Limit 200MB per file • CSV

Browse files

predictions_log.csv 145.0B

Data Overview

Rows
2

Columns
5

Data Preview

	Ammonia (mg/l)	pH (ph units)	Nitrate (mg/l)	Timestamp	Predicted_Class
0	0.05	7.8	2	2025-12-07T21:28:59	A1
1	0.05	5	7	2025-12-07T21:29:46	D4

Column Statistics

	Ammonia (mg/l)	pH (ph units)	Nitrate (mg/l)
count	2	2	2
mean	0.05	6.4	4.5
std	0	1.9799	3.5355
min	0.05	5	2

Appendix D: Gantt Chart

TEAM 5

Pascual, Ken Brodeth, Van Jersey Paolo

Project star **Fri, 8/1/2025**

Display wee **11**

GANTT CHART

TASK	ASSIGNED TO	PROGRESS	START	END
Client Identification and Topic Formulation				
Preliminary Investigation		100%	8/1/25	8/15/25
Review of Related Literature		90%	8/15/25	11/10/25
Client Consultation 1		90%	10/10/25	10/15/25
Client Consultation 2		90%	11/1/25	11/10/25
Development of Alternative Designs				
Data Gathering / Dataset		100%	10/13/25	10/30/25
Data Cleaning		100%	10/30/25	11/5/25
Design Draft		100%	11/5/25	11/24/25
Design Evaluation and Metrics		100%	11/24/25	11/26/25
Tradeoff and Sensitivity Analysis		100%	11/26/25	12/1/25
Documentation and Presentation				
Chapter 1		100%	8/16/25	11/17/25
Chapter 2		100%	11/10/25	12/6/25
Chapter 3		80%	11/16/25	12/6/25
Prelim Presentation		100%	9/1/25	9/5/25
Final Presentation		100%	11/10/25	12/7/25