

## Project Plan

### Product-based Capstone Project

River pollution due to the accumulation of waste has become an escalating environmental issue worldwide. This pollution not only degrades water quality but also endangers aquatic ecosystems, disrupts biodiversity, and poses severe health risks to communities living near affected rivers. Common pollutants, such as plastics, industrial waste, and household trash, accumulate in rivers, obstructing water flow, spreading toxins, and affecting both wildlife and human populations that rely on these rivers for water, food, and recreation. Despite growing awareness, efforts to monitor and manage river pollution often face challenges due to the lack of community engagement and insufficient, real-time data.

There is a critical need for a user-friendly, accessible platform that empowers local communities to report and monitor pollution levels in rivers. Currently, the absence of such tools limits the public's ability to contribute to environmental protection efforts. As a result, river cleanups remain sporadic, under-prioritized, and reactive rather than preventive. By bridging this gap with a community-driven system, it's possible to enable a more proactive approach to river conservation. Such a system could empower local users to document pollution, assist authorities in prioritizing cleanup activities, and educate the public on the impact of waste accumulation.

**Team ID** : C242-PR601

**Team Member** :

1. (ML) M279B4KX4492 – Wilma Nur Fatimah – Universitas Negeri Malang - [Active]
2. (ML) M010B4KY4129 – Shawn Michael Dayanti Intong – Universitas Indonesia - [Active]
3. (CC) C119B4KY1484 – Ferdinand Zulvan Lindan – Institut Teknologi Sumatera - [Active]
4. (CC) C179B4KY2517 – Moch Ihsan Ramadhani – Universitas Ahmad Dahlan - [Active]
5. (CC) C123B4KY0419 – Alvian Fahrul Fanani – Politeknik Elektronika Negeri Surabaya - [Active]
6. (MD) A73B4KY0205 – Ahmad Ferli – Institut Teknologi dan Bisnis PalComTech - [Inactive]

# Project Plan

## Product-based Capstone Project

### Final Selected Themes:

Sustainable Futures: Nurturing harmony between humanity and the environment ▾

### Title of the Project:

Sapu (Sahabat Peduli Sungai)

### Executive Summary/Abstract:

The cleanliness of rivers has been a major issue in large cities. Not only do dirty rivers serve as breeding grounds for harmful microorganisms, but they are also unpleasant to look at, detracting from the aesthetics of an urban environment. One of the main challenges is the low public awareness around keeping rivers clean, resulting in a near-constant need for cleanup by responsible authorities. However, without anyone directly monitoring the rivers in real time, ensuring cleanliness becomes difficult. Reports on river cleanliness rarely reach the relevant authorities promptly, leading to delays in cleanup.

A direct solution is needed. With the advancement of machine learning and artificial intelligence, this problem could be addressed through a mobile application. This app would allow citizens to report river pollution by uploading images as evidence. These images would be validated by an image classification model and, along with the user's location, would be stored in a database. An interactive map highlighting polluted river locations based on report frequency would then help authorities prioritize and respond more effectively to areas in need of immediate cleanup.

### How did your team come up with this project?

This project idea came to mind when one of our team members realized that the river near his house is always dirty and almost never cleaned. Doing the clean-up alone would seem impossible because of the huge amount of waste. Not knowing what to do and to whom he should report directly to, this becomes a problem that needs special attention and an immediate solution.

# Project Plan

## Product-based Capstone Project

Project Scope & Deliverables:

### WEEK 1

TASKS	9	10	11	12	13	14	15
ML: Data Gathering	✓	✓	✓				
ML: Data Cleaning				✓	✓	✓	✓
CC: Backend Creating	✓	✓	✓	✓	✓	✓	✓

### WEEK 2

TASKS	16	17	18	19	20	21	22
ML: Data Preparation	✓	✓					
ML: Model Selection			✓	✓			
ML: Model Training and Evaluation					✓	✓	✓
CC: Backend Creating	✓	✓	✓				
CC: Backend Testing				✓	✓	✓	✓

### WEEK 3

TASKS	23	24	25	26	27	28	29
ML: Error Analysis	✓	✓	✓				
ML: Retrain Model				✓	✓	✓	✓
CC: Backend Documentation	✓	✓					
CC: Calculating Cloud Infrastructure Costs			✓	✓			
CC: Cloud Infrastructure Setup					✓	✓	✓

## Project Plan

### Product-based Capstone Project

#### WEEK 4

TASKS	30	1	2	3	4	5	6
ML: Model Deployment	✓	✓	✓	✓	✓	✓	✓
CC: Cloud Deployment	✓	✓	✓				
CC: Cloud Deployment Testing				✓	✓	✓	✓

#### Project Schedule:

MILESTONE	WEEK 1 NOV 9-15	WEEK 2 NOV 16-22	WEEK 3 NOV 23-29	WEEK 4 NOV 30-DEC 6
ML: Data Gathering and Cleaning CC: Backend Creating	✓ ✓			
ML: Model Training CC: Backend Creating & Testing		✓ ✓		
ML: Error Analysis CC: Cloud Infrastructure Setup			✓ ✓	
ML: Model Deployment CC: Cloud Deployment & Testing				✓ ✓

## Project Plan

### Product-based Capstone Project

**Based on your team's knowledge, what tools/IDE/Library and resources that your team will use to solve the problem?**

- ML: VSCode, Google Colab, Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn, Tensorflow
- CC: VSCode, Postman, Git, GitHub, Cloud Run, CloudStorage

**Based on your knowledge and explorations, what will your team need support for?**

- Mentorship in environmental science and pollution indicators
- Access to high-quality datasets specific to river waste types
- Resources for cloud deployment (Google Cloud credits or alternatives)

**Based on your knowledge and explorations, tell us the Machine Learning Part of your Capstone!**

A deep learning model will be trained and fine-tuned using tensorflow which classify dirty river or clean river images to validate the truthness of report evidence. To do this, a representative dataset must be created which comprises images taken from internet and phone cameras. To simulate low quality images some augmentation will be done like image blurring and lighting variation.

**Based on your knowledge and explorations, tell us the Cloud/Web/Frontend/Backend Part of your capstone?**

We will build the backend for the application using Node.js with the Express framework to manage API requests and data interactions. User report data and river pollution information will be stored in Cloud SQL as the main database, so that the data is stored neatly and safely. Photos uploaded by users will be stored in Cloud Storage. We will run both the backend and machine learning models on Google Cloud Run, so the application can handle many requests efficiently without the need for complex server maintenance.

This project was initiated based on the observation that river pollution is a persistent issue in urban areas, often going unnoticed and unreported. According to various environmental studies, the cleanliness of rivers has a significant impact on public health and urban aesthetics. The common problem of individuals being unsure about how to report pollution effectively inspired us to create this project. Through the application we are developing, we aim to raise awareness about river cleanliness and enable the community to contribute actively by reporting polluted rivers. Our goal is to support local authorities in maintaining cleaner rivers and ultimately create a healthier environment for everyone.

## Project Plan

### Product-based Capstone Project

**Based on your team's planning, is there any identifiable potential Risk or Issue related to your project?**

- For the machine learning part, the identifiable risks and issues are low model performance and failure of deployment. In order to address this, data used to train the model is ensured to be very clean by checking carefully each image for potential defects. Also, a careful selection and design of models to ensure balance so that it is not too large and has fast inference time but still having a great performance.
- For the cloud computing part, identifiable risks and issues include potential data privacy breaches and unexpected cloud cost overruns. To address data privacy, strict access controls and encryption will be implemented, ensuring sensitive information is protected. In terms of cost management, we will monitor usage closely and set budget alerts to avoid exceeding limits. Additionally, the scalability of cloud resources is essential; therefore, Google Cloud Run and Cloud SQL will be optimized for efficient scaling and performance. These precautions will help maintain a secure, cost-effective, and scalable cloud infrastructure for our project.

**Any other notes/remarks we should consider on your team's application**

The project aims to bridge technology and community action, empowering individuals to actively monitor river pollution. The app will be designed with user accessibility in mind, ensuring a seamless experience for a variety of skill levels. Our main goal is to make a positive contribution to environmental conservation through an innovative, community-centered approach.