Cloud Computing for Data Analysis Final Project

University of North Carolina at Charlotte

STUDENT NAMES:

Rishi Vamshi Athinarap
Venkata Sai Shalini Ganni
Siva Naga Rutwik Reddy Chintha
Darshini Chalumuri

Project Name:

Business scenario overview

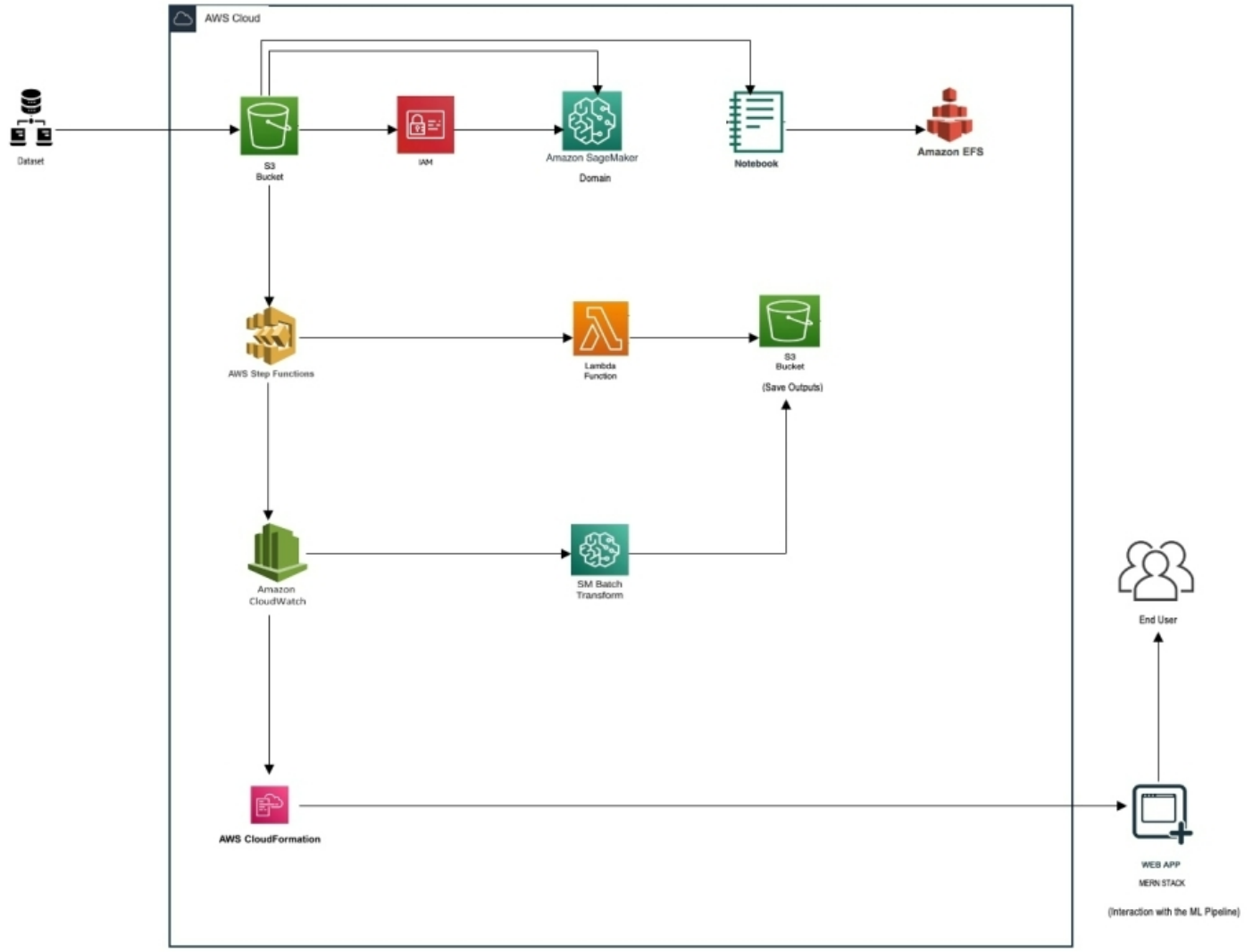
Problem Statement

- The project addresses the critical environmental problem of plastic bag pollution, which poses a grave threat to marine life, disrupts ecosystems, and contributes to the larger global crisis of plastic waste.
- Traditional methods for detecting and collecting these pollutants are inadequate, often missing small or submerged plastic bags that can cause significant harm.
- By improving detection, the project facilitates more effective and efficient clean-up operations, directly contributing to the mitigation of pollution, the protection of wildlife, and the promotion of ecological sustainability.

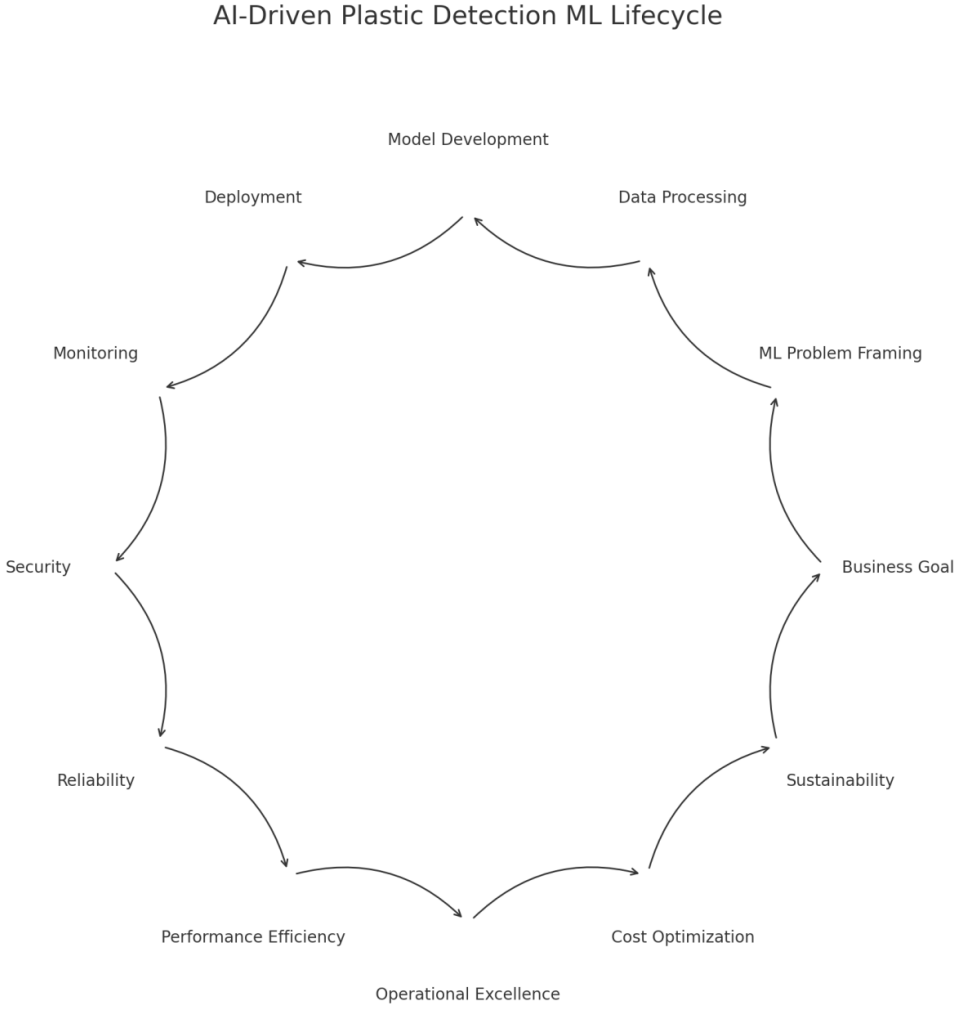
Solution overview

- We developed an AI-driven platform designed to identify plastic bag pollutants in various environments.
- This platform integrates a user-friendly interface for uploading images, ranging from satellite data to high-resolution fieldwork photographs.
- Utilizing AWS technologies like Sagemaker, S3, and Lambda, the platform processes these images in batches daily, optimizing computational resources.
- The project aims to provide a scalable, cost-effective solution for environmental researchers, volunteer cleanup crews, and other stakeholders in marine and terrestrial ecosystem conservation.

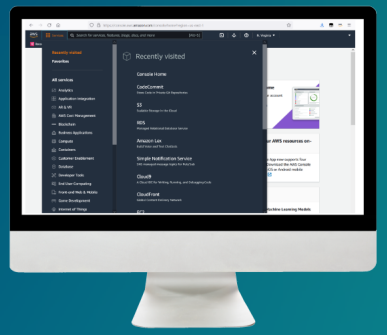
Architecture diagram of the solution



Machine Learning Lifecyle



Demo



- Project Repo link –
- https://github.com/rishivamshi/CCDA-group11

Lessons learned

- Challenges that we overcame: We overcame the challenge of accurately detecting plastic bags in varied and complex environmental images by enhancing our machine learning models with a more diverse training dataset and advanced algorithms.
- List resources that we found helpful: AWS documentation and the Google Open Images Dataset were invaluable in training our models, alongside machine learning communities like Towards Data Science for practical insights.
- New skills that we used: We developed proficiency in AWS Sagemaker for model training and deployment, and gained new insights into optimizing machine learning models for better performance and accuracy.
- Next steps: The next steps involve scaling the solution to handle larger datasets, expanding the types of detectable pollutants, and collaborating with environmental organizations to deploy our platform for real-world impact.

Thank you