IBM SB AI Summer Certification Program 2025

From Learner to Builder: Become an AI Agent Architect

Presented By: **SoloBot**

Group Members:

1. **Sagar Kumar**

College: Marwari College, Ranchi, Jharkhand

Report

Title**: Development of an AI-Powered Multi-Agent System for Monitoring Marine Pollution**

Introduction

Marine pollution is a severe and growing threat to global marine ecosystems, biodiversity, and coastal communities. Traditional methods for detecting and monitoring this pollution are often slow, expensive, and lack the real-time capabilities needed for effective intervention. This leads to significant delays in cleanup efforts, resulting in more extensive environmental damage. This project aims to address this critical challenge by developing a multi-agent AI system that leverages advanced technologies to autonomously detect and alert about marine pollution.

Problem Statement

The primary challenge addressed by this project is the lack of a scalable, non-invasive, and efficient system for real-time marine pollution monitoring. Existing approaches are often manual, resource-intensive, and provide delayed information, hindering rapid response to events like oil spills and plastic accumulation. This gap leads to prolonged environmental damage and increased cleanup complexities.

Objective

The main objective of this project is to design, develop, and evaluate a multi-agent AI system capable of predicting and detecting marine pollution based on various environmental indicators and data sources. The system aims to improve early detection, facilitate timely intervention, and ultimately contribute to the conservation and sustainable use of oceans, seas, and marine resources, directly supporting the United Nations Sustainable Development Goal 14: Life Below Water.

Solution Overview

Our solution is a multi-agent AI system designed to provide continuous, autonomous monitoring of marine environments. It comprises three specialized agents that work collaboratively:

- **Data Ingestion Agent**: This agent is responsible for collecting raw data from diverse sources, including publicly available satellite imagery (e.g., from Copernicus) and environmental sensor data (e.g., water quality

parameters). It securely stores this data in **IBM Cloud Object Storage**.

- **Pollution Analysis Agent**: This agent retrieves data from Cloud Object Storage and uses advanced machine learning models to identify pollution. It employs a trained RandomForestClassifier model to analyse water quality parameters and could be extended with computer vision models for satellite image analysis (e.g., detecting oil slicks or plastic patches). This model is deployed as a public API endpoint using **IBM Watson Machine Learning**.
- **Alert and Reporting Agent**: This agent acts as the system's communicator. It periodically calls the deployed Pollution Analysis API with new data. Based on the prediction (e.g., water is "non-potable" indicating pollution), it generates immediate alerts and reports, which can be sent via email or displayed on a dashboard.

Key Features of the Solution:

- **Real-time Detection**: Enables rapid identification of pollution events as they occur.
- **Multi-Source Data Integration**: Combines different types of data (sensor, satellite) for comprehensive monitoring.
- **Autonomous Operation**: Reduces reliance on manual inspection and human intervention.

- **Scalable and Cloud-Native**: Built on **IBM Cloud** services (Cloud Object Storage, Watson Studio, Watson Machine Learning), ensuring scalability and reliability.
- **Actionable Alerts**: Provides timely notifications to facilitate quick response and mitigation efforts.

Technical Implementation

The project was developed using Python and several IBM Cloud services:

- **Python Libraries**: pandas for data manipulation, scikit-learn for machine learning model development, requests for API interactions, and ibm-boto3 for Cloud Object Storage integration.
- **IBM Cloud Services**:
    - **IBM Cloud Object Storage**: Used as a scalable and secure repository for raw and processed data.
    - **IBM Watson Studio**: Provided the integrated environment for data loading, model training, and evaluation.
    - **IBM Watson Machine Learning**: Utilized for deploying the trained RandomForestClassifier model as a REST API endpoint, making it accessible for real-time predictions.
- **Multi-Agent Communication**: Agents communicate via API calls, where the Data Ingestion Agent populates Cloud Object Storage, the Pollution Analysis Agent consumes

from it and exposes an API, and the Alert Agent consumes from this API.

Results

The developed system successfully demonstrated its core functionality:

- The **Data Ingestion Agent** successfully uploaded a water_potability.csv dataset (serving as a proxy for real-world marine sensor data) to IBM Cloud Object Storage.
- The **Pollution Analysis Agent** trained a RandomForestClassifier model on this data, achieving an accuracy of approximately 67% in predicting water potability. This model was then successfully stored and deployed as a public API endpoint on IBM Watson Machine Learning, providing a promising baseline for initial water quality assessment.
- The **Alert and Reporting Agent** successfully called the deployed API with sample data. When provided with data indicating poor water quality, it correctly predicted "NON-POTABLE" (0) and generated a corresponding alert, demonstrating the system's ability to identify and report potential pollution.

Conclusion

This project successfully developed a multi-agent AI system for marine pollution monitoring, leveraging IBM Cloud services. By providing an autonomous, scalable, and real-time detection

solution, this system has the potential to significantly improve environmental response efforts and contribute to the health of marine ecosystems.

Future Work

- **Integration of Satellite Imagery Analysis**: Incorporate computer vision models (e.g., using TensorFlow or PyTorch) to analyze satellite images for visual pollution cues like oil slicks or large plastic accumulations.
- **Real-time Sensor Data Streams**: Integrate with actual IoT marine sensors for live data ingestion.
- **Advanced Alerting**: Implement more sophisticated notification channels (e.g., SMS, dedicated mobile app notifications) and integrate with environmental response platforms.
- **Interactive Dashboard Deployment**: Deploy the developed Streamlit dashboard to a public cloud platform (e.g., IBM Cloud Code Engine) to provide a user-friendly interface for real-time visualization of pollution data and alerts.
- **Model Improvement**: Continuously retrain and refine the machine learning models with larger and more diverse datasets to improve accuracy and detection capabilities.