L&T_EduTech

Garbage Bin Level Prediction and Collection

Problem Statement 2 (PS-2)

Problem Description

➢ Cities are exploring new ways to manage and treat municipal solid waste in order to improve public health and the overall livability of urban areas.
➢ By using AI algorithms to predict, when garbage bins will need to be emptied, waste collection vehicles can be operated more efficiently
➢ When garbage bins are full, collection services can be optimized to ensure that waste is collected before it becomes a public health or sanitation issue
➢ The process of trash/garbage bin level prediction typically involves collecting data from sensors or cameras installed on the garbage bins and using this data to train a machine learning model
➢ The model can then be used to predict the fill level of a garbage bin based on its current sensor readings or images
➢ AI-based trash bin level prediction can generate a large amount of data that can be used to
improve waste collection and management
➢ This is to ensure that waste is managed in an efficient, sustainable, and environmentally friendly manner
➢ And also, this data can be used to optimize routes, plan for future waste management needs,
and identify areas where improvements can be made

Requirement Specification

➢ Use the trash bin levels dataset (dataset attached) with different features like date, time, bin level, Bin ID, week no, fill level (in liters), total(liters), fill percentage, location, latitude, longitude, temperature (in ºc) and battery level
➢ Develop appropriate ML models for predicting filled status of the bin level binary classifications such true /false for different areas and localities

Judging Metrics

➢ Trash bin Level Classification -Precision, Recall, Accuracy, F1-score and confusion matrix

AIM OF THE PROJECT

The aim of this project is to predict Garbage Bin Level and Collect using machine learning algorithms.

UNIQUENESS/DIFFERENTIATOR TRIED OUT IN THE PROJECT

A variety of machine learning models were employed, including Logistic Regression, Random Forest, and K-Nearest Neighbors (KNN).

INFERENCE

The analysis revealed remarkable performance across multiple models:

Model Performance:

- **Logistic Regression**:
  - Accuracy: 1.0
  - Mean Squared Error: 0.0
  - Cross-Validation Accuracy: 1.0

- **Random Forest**:
  - Accuracy: 1.0
  - Mean Squared Error: 0.0
  - Cross-Validation Accuracy: 1.0

- **K-Nearest Neighbors (KNN)**:
  - Highest Accuracy achieved with k = 4, k = 5, k = 12, k = 16, k = 17, k = 18, k = 19, k = 20 (all with Accuracy: 1.0)
  - Cross-Validation Accuracy with 4 neighbors: 0.9999094612947035

- **Custom Linear Regression**:
  - Mean Squared Error: 0.08782254413761884
  - Accuracy Score: 0.9121774558623812
  - Cross-Validation Accuracy: NaN

Metrics:

- **F1 Score**: 1.0

- **Recall**: 1.0

Inference:

- Logistic Regression and Random Forest models achieved perfect accuracy.

- KNN models with various values of k showed high accuracy, with the highest being 1.0.

- Custom Linear Regression performed well with a relatively low mean squared error.

- All models demonstrated excellent F1 scores and recall, indicating strong predictive capabilities.

Github link:

- **https://github.com/Karthik0827/Garbage-levelGarbage_Bin_Level_Preciction_and_Collection**