Journal of Science Research and Reviews

PRINT ISSN: 1595-9074
E-ISSN: 1595-8329
DOI: https://doi.org/10.70882/josrar.2025.v2i1.25
Homepage: https://josrar.esrgngr.org

Original Research Article

Intelligent Waste Management Optimization Through Machine Learning Analytics

***Ogbolumani, Omolola A. and Adekoya, Muiz**

University of Lagos, Lagos, Nigeria.
*Corresponding Author's email: oaogbolumanischolar@gmail.com

KEYWORDS
Waste management,
Internet of Things,
Machine learning,
Smart cities,
Route optimization,
Environmental sustainability.

CITATION
Ogbolumani, O. A., & Adekoya, M. (2025). Intelligent Waste Management Optimization Through Machine Learning Analytics. *Journal of Science Research and Reviews*, *2*(1), 7-26. https://doi.org/10.70882/josrar.2025.v2i1.25

ABSTRACT
Waste management presents serious obstacles to making metropolitan regions more habitable. Conventional waste management techniques are not usually optimized, resulting in overflowing bins, wasteful waste collection trips, and various negative environmental effects. This study addresses these challenges by developing an intelligent system integrating the Internet of Things (IoT) and machine learning technologies. This study aims to develop an intelligent waste management system that optimizes waste collection routes and schedules through machine learning. (ML) models and Internet of Things (IoT) powered smart bins. The system utilized Support Vector Machines (SVM) and Artificial Neural Networks (ANN) for data analysis, complemented by dynamic route optimization algorithms. Data collection over 90 days across 47 sites encompassed bin fill levels, battery status, and environmental parameters such as temperature and humidity. Results demonstrated significant operational improvements, with the system achieving 89% accuracy in fill-level prediction and enabling a 35% reduction in collection frequency. Implementation led to a 42% decreased fuel consumption and a 2.4-hour reduction in daily collection times. Commercial zones exhibited 1.8 times higher fill rates than residential areas, while weekend waste generation peaked at 2.1 times weekday. The findings indicate that IoT-ML technology integration substantially enhances urban waste management efficiency through data-driven decision-making. Phased implementation, prioritizing high-waste-volume areas, integrating with existing metropolitan systems, and developing standardized data protocols are recommended. This research contributes to the growing body of evidence supporting smart technology adoption in urban waste management, offering a scalable solution for improved operational efficiency and environmental sustainability.

INTRODUCTION

The world population is impacted by the global issue of waste management (Sosunova & Porras, 2022). The overall well-being, cleanliness, and productivity of communities are impacted by decisions made by individuals and governments on garbage management and consumption (Sharma et al., 2021). Inadequate waste management is causing flooding, obstructing sewers, killing animals who unwittingly consume waste, and poisoning the world's oceans (Ajibola & Ogbolumani, 2024; Prakash et al., 2022).

It also breeds vectors that spread disease, increases respiratory issues through airborne particles from burning waste, and impedes economic development through decreased tourism, among other effects (Zhou et al., 2021). Decades of economic progress have resulted in mismanaged and poorly managed garbage that requires prompt attention at every level of society. Over the previous 20 years, there has been a significant increase in the amount of solid garbage generated. By 2050, it's predicted that the yearly amount of solid waste will be about 3.40 billion tonnes, resulting in waste management costing about $635.5 billion (Prakash et al., 2022). Many elements are essential in the way waste is managed in cities. The elements that constitute solid waste management are becoming more difficult these days, and the reasons for this are the fast rise of population, moving from rural to urban areas, rising commodities consumption, inappropriate methods of gathering solid waste, the absence of smart technologies, and solid waste prediction (Ogbolumani & Nwulu, 2021; Rahman et al., 2022).

Solid waste, such as cardboard, plastics, papers, glasses, and other materials, which can all be considered household waste, can be handled by various processes. All waste can be recycled or divided into materials for domestic tasks. According to EUROSTAT, 423 million tons, or 56%, of the garbage produced domestically in the European Union in 2016 was recycled (Rahman et al., 2022) (EUROPA, 2020). Again, in 2016, 179 million tons, about 24 percent of the waste generated locally was landfilled (Rahman et al., 2022). The reports clearly show how good domestic waste management is essential to recycling. The outcomes would be innumerable if we combined contemporary technology with the waste management system.

Figure 1, illustrates the traditional waste management system, depicting the methods employed in waste collection disposal. Traditional waste management relies on manual collection and disposal methods, often lacking optimization. It faces challenges regarding efficiency and environmental impact in adapting to evolving waste patterns (Amasuomo & Baird, 2016).

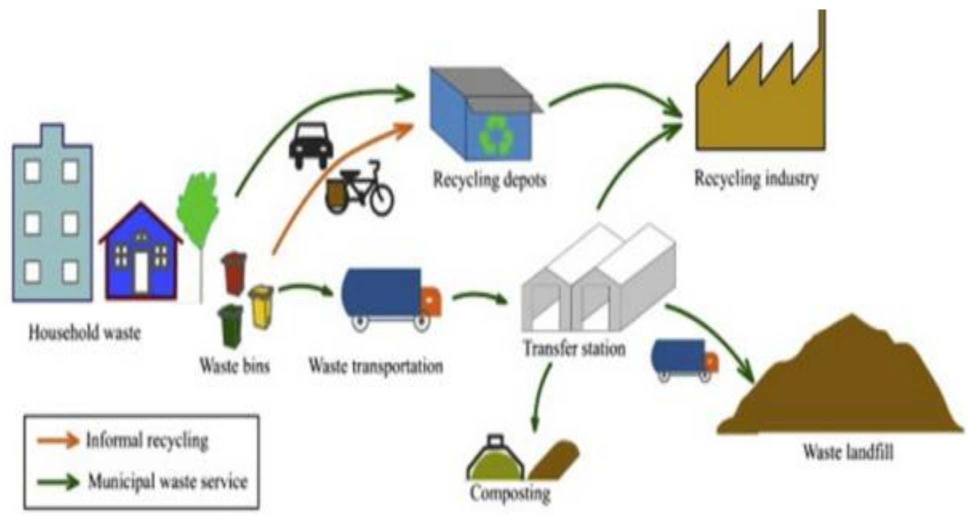

Figure 1: Traditional waste management cycle (Amasuomo & Baird, 2016)

British technology pioneer Kevin Ashton was the first to use the term “Internet of Things” (IoT) to describe a system that would enable physical items to be connected to the Internet via sensors (Chakraborty, 2022). He coined the term to draw attention to the fact that, when linked to the Internet, Radio-Frequency Identification (RFID) tags—utilized in business supply chains—may be tracked and tallied without human help. Nowadays, many things, devices, sensors, and everyday objects with computational capability and Internet connectivity are considered part of the "Internet of Things." Among the growing category of devices in this domain, Raspberry Pi can be categorized as a complete computing solution, hosting an application development environment and an operating system based on Microsoft Windows or Linux, depending on the deployment scenario (Bakhshi & Ahmed, 2018a). The growing interest in IoT monitoring and management purposes has led to several off-the-shelf hardware and software utilities for IoT system development (Adeleke et al., 2023).

Waste collection and path collection present two major obstacles to waste management (Anh Khoa et al., 2020). First, garbage truck routes must be carefully planned with social, economic, and environmental aspects considered. Waste collection is a daily task in cities. Secondly, graph theory should be used to reduce the path length to prevent excessive fuel costs and minimize the amount of work. In certain cases, the fill level of the bin can be estimated by using Internet of Things (IoT) devices, which then transmit the information to a server via the Internet for processing (Ogbolumani & Mabaso, 2023).

One major tool that can transform waste management is machine learning, which enhances efficiency through predictive analytics and reduces operational costs and environmental impact (Bakhshi & Ahmed, 2018b). The computer science field of machine learning aims to allow computers to "learn" without needing to be explicitly coded (Bi et al., 2019; Jamal et al., 2018). Its roots are in the 1950s artificial intelligence movement, and it places a strong focus on applicable goals and uses, especially in prediction and optimization. In machine learning, computers "learn" by becoming more proficient at tasks via "experience."

Waste management data can yield important insights that can be applied to machine learning algorithms to create waste collection methods that are more effective and efficient. Machine learning, for example, can use historical waste collection data to forecast future waste generation patterns, assisting authorities in resource allocation and collection route planning (Doe, 2021). Intelligent bin monitoring is one of the most important areas in which we use machine learning analytics to transform waste management. The traditional waste collection systems frequently have inefficiencies in dealing with waste, such as bin overflow or needless collection trips. These waste bins can be equipped with Internet of Things (IoT) devices and sensors that enable live tracking. By using machine learning algorithms in addition to these, we can evaluate the gathered data and predict when a bin is most likely to fill up. Waste collection firms can use this to optimize their schedules and reduce the need for unnecessary visits (Doe, 2021).

Recycling is crucial to this waste management system as it contributes to environmental preservation (Ogbolumani & Nwulu, 2021). These recycling efforts can be refined by improving the sorting and classification of recyclable materials. Enormous volumes of data can train models; machine learning algorithms can then recognize the different types of recyclables even when there are minute changes. Additionally, machine learning analytics can assist in the optimization of the recycling processes. This could help to recover recyclable materials, protect natural resources, and decrease landfills, leading to improvement in recycling. It is also a way for residents to change their waste management techniques.

In recent years, there has been an increase in global population and urbanization, which has led to a significant rise in the amount of waste being generated; traditional waste management systems are struggling to keep up with this increase in waste generation, which has led to inefficiencies in dealing with waste and various environmental problems(Ogbolumani & Nwulu, 2024). For example, the Lagos State Waste Management Authority (LAWMA), which oversees waste collection from bins across Lagos, collects the waste either weekly or bi-weekly, depending on the location (Onuminya & Nze, 2017). However, this plan has not worked out as expected, as bins tend to fill up before trucks are available to pick them up; this leads to the bins being overfilled and waste spilling. To overcome this problem, we use machine learning-based route optimization for waste management; this would involve evaluating past data, forecasting future trash collection trends, and determining the most effective route for collection vehicles. Machine learning algorithms will incorporate real-time data such as traffic conditions, weather forecasts, and waste generation routes to optimize waste collection routes. These algorithms can dynamically adjust routes to minimize fuel consumption, reduce emissions, and improve the overall operational efficiency of the system.

A study by Ferrao et al., (2024) utilized a genetic algorithm to optimize waste collection routes in smart city environments, significantly reducing collection costs and vehicle emissions (Ferrão et al., 2024). Similarly, Clen (2022) used a machine learning approach to predict the optimal waste collection schedules based on real-time data to improve service quality and customer satisfaction (Chen, 2022). Despite the results obtained from these studies, several gaps were identified in machine learning-based route optimization for intelligent management. One major limitation is the absence of standardized methodologies for developing and evaluating predictive models. Many studies also focused on specific aspects of route optimization without considering the broader implications of overall waste management on system performance. To handle these gaps, we aim to create a comprehensive framework that applies machine learning algorithms to optimize the collection routes in smart waste management systems. Using a holistic approach that considers various factors such as waste generation patterns, traffic conditions, and environmental impact, we can enhance the effectiveness and sustainability of waste management practices. Another study by Chen et al. (2019) presented a machine learning-based waste classification and recycling approach that utilized Convolutional Neural Networks (CNNs). These networks trained models that collectively classify waste products into several categories using a dataset of photos of available garbage items. The author's results were promising, highlighting the potential of machine learning in managing waste (Chen et al., 2019). Wang et al. (2018) used big data analytics and an Internet of Things (IoT) based framework for intelligent waste management. They argued that waste management systems could effectively monitor waste generation, optimize collection routes, and promote recycling by integrating sensor networks, data analytics, and machine learning algorithms. They emphasized the role of machine learning in extracting valuable insights from the collected data (Wang et al., 2018).

Despite garbage management and classification advancements, the integration of machine learning analytics with trash treatment technologies is still lacking in the literature, according to Li et al. (2019). They emphasized the need for more studies on intelligent waste treatment systems that improve waste-to-energy conversion processes by applying machine learning algorithms (Li et al., 2019). Lakhouit et al. (2023) discussed the application of specific ML algorithms to estimate and forecast the quantity of domestic waste (Lakhouit et al., 2023). Wang et al. (2019) suggested implementing a model for forecasting garbage generation patterns based on machine learning. A prediction model that could consistently predict waste generation levels was created by evaluating past data on garbage creation and adding outside variables like demographic data and weather. The authors suggested that such models could assist waste management authorities in planning collection and disposal strategies more effectively (Wang et al., 2019).

The circular economy provides a theoretical foundation for intelligent waste management and recycling enhancement. As highlighted by Geissdoerfer et al. (2017), the circular economy emphasizes reducing waste generation, promoting recycling, and reusing materials to minimize resource depletion and environmental impacts. By streamlining waste management procedures, machine learning analytics can be extremely helpful in accomplishing the objectives of the circular economy (Geissdoerfer et al., 2017). Garg et al. (2022). identified a gap in the literature regarding integrating social factors into intelligent waste management systems. They argued that machine learning algorithms should consider social and behavioral waste generation and disposal aspects to develop more accurate models and effective waste management strategies. They recommended future research could focus on incorporating social dimensions into machine learning-based waste management optimization (Garg et al., 2022; Jin et al., 2018). Jin et al. (2018) proposed a real-time system for monitoring and analyzing landfill gas emissions based on machine learning. They combined sensor data, machine learning algorithms, and predictive models to monitor landfill gas emissions and identify potential environmental risks. The authors demonstrated the effectiveness of their approach in improving the early detection and mitigation of landfill gas-related issues (Jin et al., 2018).

The literature evaluation highlights the many approaches supporting theories and research gaps in recycling enhancement and machine learning analytics for intelligent waste management optimization. While significant progress has been made in waste classification collection route optimization and waste generation prediction, gaps exist in integrating machine learning with waste treatment sorting and considering social factors. This current research will help solve some of these challenges to improve sustainability and efficiency of waste management systems.

MATERIALS AND METHODS

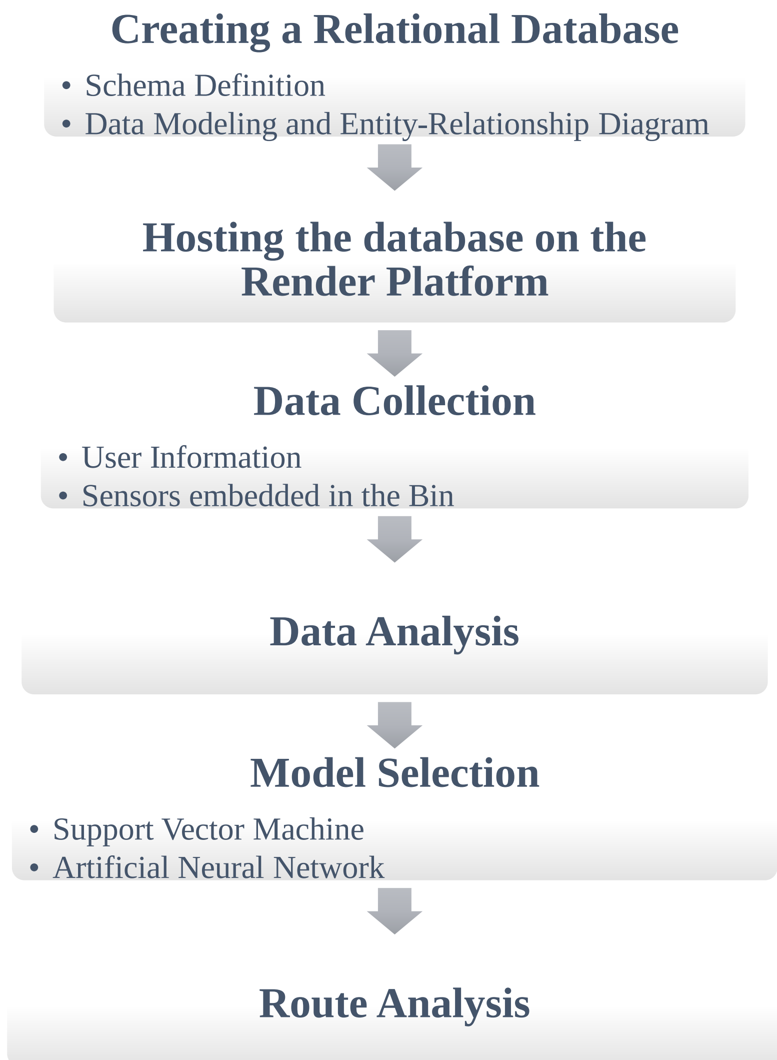

Figure 2: Methodology for the machine learning-based route optimization for smart waste management

Figure 2 shows the methodology of the machine learning route optimization for the smart waste management system.

Creating a Relational Database

The first step in creating a relational database is creating a database design document. This document is a blueprint for the database structure, outlining its purpose, architecture, entities, relationships and constraints, and indexes. The technology used is POSTGRESQL because of its reliability and scalability, making it suitable for handling data in the production environment. The database was built on a software called pgAmin.

It will include:

Schema Definition

Defines the database schema, including table definitions, data types, primary and foreign keys.

Data Modeling and Entity-Relationship Diagrams (ERDs)

It includes data modeling diagrams such as ERDs, which visually represent the database's entities, attributes, and relationships.

Schema Definition

Table 1: Customers Table

Field Names	Data Type	Definitions/Relationships
User ID (Primary Key)	BigINT	A unique identifier is assigned to each customer in the database table. The bigINT datatype allows for the storage of large integer values, accommodating the potential growth of the database without running into size limitations.
First Name	string	During registration, users input their first name, which is stored as a string datatype in the database.
Last Name	string	During registration, users input their last name, which is stored as a string datatype in the database
Date of Birth	Date	During registration, users input their first name, which is stored as a date datatype in the database
Phone Number	Text	During registration, users input their phone number, which is stored as a Text or varchar datatype in the database
Email Address	Text	During registration, users input their email address, which is stored as a Text datatype in the database
Location/Address	Character varying	During registration, users input the Address to which the bin will be delivered, which is stored as a Text datatype in the database. This entry will also be copied into the Bin database

Table 2: BIN Table

Field Names	Data Type	Definitions/Relationships
BinID(Primary key)	BigINT	A numerical unique identifier referencing a record in the bin table. It is a primary key in this table. Its presence in the customer table enables efficient querying by linking records in this table to the corresponding record in the Customer table.
UserId (Foreign key)	BigINT	A numeric identifier which is a foreign key in this table is used to reference the customer table
Address	Character Varying	During registration, users provide the Address to which the bin will be delivered, which is stored as a Text datatype in the database.
Bin level	Numeric	The percentage representation hence the Numeric datatype signifies the fill levels of the bin, which are recorded periodically by the ultrasonic sensor integrated into the smart bin.
Battery level	Numeric	The percentage representation hence the Numeric datatype signifies the battery level, which is recorded periodically by the battery charge controller integrated into the smart bin.
Weight	Numeric	The percentage representation hence the Numeric datatype signifies the weight of the recycle part of the bin, which is recorded periodically by the weight sensor integrated into the smart bin.

Charge Status	Boolean	The charge status indicates whether the battery is charged (true) or not charged (false). It's a binary data type representing a simple true/false or on/off condition.
Power consumption	Numeric	Power consumption with a datatype of Numeric, represents the amount of electrical energy used by the bin. Stored as a point number, it allows for the representation of fractional values, enabling precise measurement of power usage.
Temperature	Numeric	Temperature with a datatype of Numeric, represents the measurement of heat or coldness. It is recorded by the temperature sensor embedded in the smart bin.
Weather condition	Text	Weather condition, with a datatype of Text, represents the current atmospheric state such as "sunny", "rainy", "cloudy", etc. Stored as text, it allows for the recording of various weather phenomena in a descriptive format for analysis and display purposes.
Humidity	Numeric	Humidity with a datatype of Numeric, represents the amount of moisture present in the air.

Table 3: Admin Table

Field Name	**Data Type**	**Definitions/Relationships**
AdminId (Foreign key)	INT	A numerical unique identifier referencing a record in the bin table. It is a primary key in this table.
First Name	Text	Admins iput their first name, which is stored as a Text datatype in the database.
Last Name	Text	Admins input their last name, which is stored as a Text datatype in the database.
Email (Primary key)	Text	Admins input their first name, which is stored as a Text datatype in the database. It is the primary key because emails are what is used to grant permissions
Phone	Text	Admins input their phone number, which is stored as a Text datatype in the database.
Address	Character varying	Admins input their address, which is stored as a Text datatype in the database.
Gender	Boolean	Admins input their gender, which is male or female
Role	Text	Roles are assigned to admins based on the permission they have access to
Department	Text	A department represents a distinct organizational unit within a company or institution. Stored as text, it denotes the name of a department, facilitating organization and categorization of data related to different functional areas within an organization
Department Id (Foreign key)	Text +INT	It is a numerical identifier that references a departmental record in another table. It establishes a clear relationship between tables ensuring the integrity of data and enabling efficient querying by linking record based on a

	unique numerical identifier which represents different departments within an organization.

Data Modeling and Entity-Relationship Diagrams (ERDs)
Figure 3 shows the entity relationship diagram for bins and users in the waste management system.

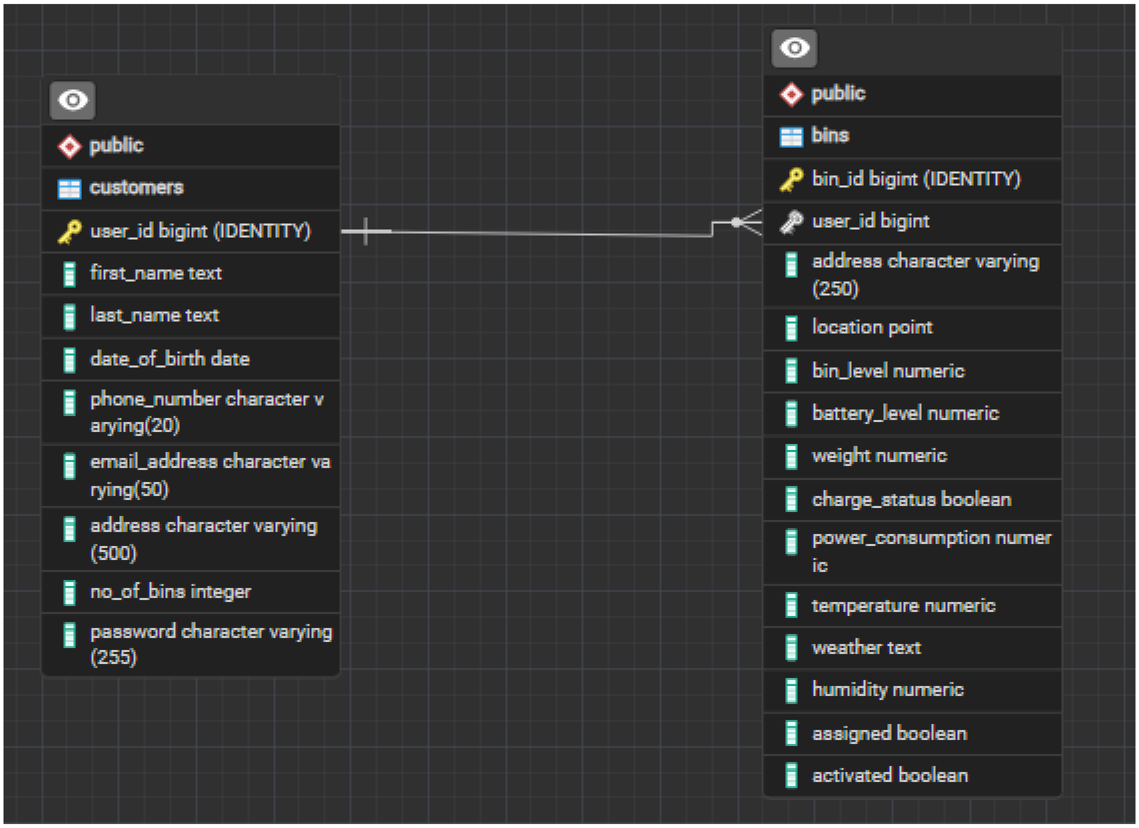

Figure 3: Entity-Relationship Diagram (ERD)

Overview of Database Tables
In the database of the system there are three main tables that have been set up, namely: Customers, BIN, and Admin, each of which serve a specific purpose
1. Customers Table: This is where data on the customers are stored. Data such as their names, date of birth, mobile numbers and their addresses are stored here (Table 1)
2. BIN Table: Data bout the bins are stored here, these include the bin Id, its location, fill levels and if they require maintenance. It also monitors environmental factors such as the temperature and weather around the bin (Table 2)
3. Admin Table: The Admin table contains the data about administrators, including their names, contact information and their various duties. This table aids in system management and the department to which each administrator belongs is tracked with it (Table 3).

Hosting the database on the Render platform
A well-known PostgreSQL and PostgreSQL Administration and Management Tool (pgAdmin) open-source administration and management tool is used to set up the database. In comparison, Render is a comprehensive cloud platform which proved smooth and effective ways to deploy and manage databases, web apps and services. Steps to host the database on *Render*.
Create a render account
Create a PostgreSQL database
Configure the database
Open pgAdmin
Create a new server connection
Configure the connection
Save the connection
After successfully hosting, we were now able to administer the render-hosted PostgreSQL database using *pgAdmin*.

Data Collection
Ideally, the data should be collected from two main sources:
(a) User Information: The data the users/bin owners provide during sign-up is collected and stored in the

database. Data such as name, address and phone number are provided by the user. When a new user record is created, the database automatically assigns a unique ID by incrementing the last used ID.

(b) Sensors embedded in the bin: Sensors deployed strategically in waste collection bins and utility sensors measure fill level, temperature, and weight. For example, ultrasonic sensors can gauge the fill level of bins, while temperature sensors can monitor the change in temperature in and around the bin. The sensors transmit the data collected to the cloud-stored database through a GSM module embedded in the bin.

However, the need to use Python-generated dummy data arose because the hardware has not been built and deployed in multiple locations due to time and financial constraints. This is done after some conditions and measures were taken to make the data accurate and reliable. The dataset generated was for 90 days across 47 locations. Fig. 4 shows the generated data. The system is controlled using a low-cost, low-power microcontroller with Wi-Fi and Bluetooth capabilities (ESP32) microcontroller, which handles all IoT device operations in the system. The GSM module transfers data every two hours during the uptime which is from 6 am till 10 pm. There is also a downtime where data is not recorded which is from 10:01 pm to 5:59 am. This is because we don't expect any bin activities during this period and also since the bin generates energy from a solar panel so having a downtime during this period makes sense and increases the lifespan of the battery.

timestamp	bin_id	user_id	address	bin_level	battery_level	weight	charge_status	power_consumption	temperature	weather_condition	humidity
4/15/2024 6:00	1252	17020567	Eni Njoku Hostel	1	19	0.158730159	Charging	0.71	34.61	Sunny	63
4/15/2024 8:00	1252	17020567	Eni Njoku Hostel	3	17	0.317460317	Charging	1.29	33.84	Sunny	63
4/15/2024 10:00	1252	17020567	Eni Njoku Hostel	4	20	0.476190476	Charging	1.4	34.42	Sunny	63
4/15/2024 12:00	1252	17020567	Eni Njoku Hostel	6	14	0.634920635	Charging	1.1	34	Sunny	63
4/15/2024 14:00	1252	17020567	Eni Njoku Hostel	7	11	0.793650794	Charging	0.83	33.67	Sunny	63
4/15/2024 16:00	1252	17020567	Eni Njoku Hostel	9	14	0.952380952	Charging	1.49	34.39	Sunny	63
4/15/2024 18:00	1252	17020567	Eni Njoku Hostel	11	17	1.111111111	Charging	1.66	34.49	Sunny	63
4/15/2024 20:00	1252	17020567	Eni Njoku Hostel	12	31	1.26984127	Charging	0.72	34.47	Sunny	63
4/15/2024 22:00	1252	17020567	Eni Njoku Hostel	14	49	1.428571429	Charging	1.42	34.27	Sunny	63
4/16/2024 6:00	1252	17020567	Eni Njoku Hostel	15	12	1.587301587	Charging	1.93	34.19	Sunny	63
4/16/2024 8:00	1252	17020567	Eni Njoku Hostel	17	12	1.746031746	Charging	0.59	33.81	Sunny	63
4/16/2024 10:00	1252	17020567	Eni Njoku Hostel	19	14	1.904761905	Charging	1.42	33.75	Sunny	63
4/16/2024 12:00	1252	17020567	Eni Njoku Hostel	20	19	2.063492063	Charging	0.59	33.51	Sunny	63
4/16/2024 14:00	1252	17020567	Eni Njoku Hostel	22	17	2.222222222	Charging	1.36	34.08	Sunny	63
4/16/2024 16:00	1252	17020567	Eni Njoku Hostel	23	20	2.380952381	Charging	1.91	34.21	Sunny	63
4/16/2024 18:00	1252	17020567	Eni Njoku Hostel	25	12	2.53968254	Charging	1.26	34.3	Sunny	63
4/16/2024 20:00	1252	17020567	Eni Njoku Hostel	26	23	2.698412698	Charging	1.3	34.27	Sunny	63
4/16/2024 22:00	1252	17020567	Eni Njoku Hostel	28	35	2.857142857	Charging	1.77	33.88	Sunny	63
4/17/2024 6:00	1252	17020567	Eni Njoku Hostel	30	20	3.015873016	Charging	1.99	36.25	Sunny	63
4/17/2024 8:00	1252	17020567	Eni Njoku Hostel	31	18	3.174603175	Charging	1.46	36.2	Sunny	63

Figure 4: Sample of data generated with Python

Data Analysis

Steps for Analysis

i. Data Cleaning: Check for missing or inconsistent data.
ii. Descriptive Statistics: Summarize the key statistics for each column.
iii. Correlation Analysis: Identify relationships between variables.
iv. Time Series Analysis: Analyze trends over time.
v. Visualization: Create graphs and charts to visualize the data.

Model selection

Supervised learning is a foundational concept in machine learning. It involves training algorithms using labeled data—pairs of input features and corresponding output labels. The goal is to learn how to map from features to labels, enabling the model to make predictions or classifications for new, unseen data. Examples include linear regression and classification algorithms. Because waste level prediction usually entails forecasting a continuous variable, such as the bill-fill level, it is categorized as a regression analysis or activity. Based on input parameters such as location, time, or waste levels in the past, regression models forecast these continuous quantities. There are 12 columns and 380700 entries in the dataset. It has columns including battery levels, weights, charge statuses, addresses, timestamps, bin IDs, user IDs, power usage, temperature, humidity, and weather. The dataset monitors a range of parameters about bins at several sites. The information captures the state of bins, including their fill levels, battery life, and environmental factors, and spans many timestamps. Based on its capacity for quick training, high prediction accuracy, and handling of missing data, the XGBoost algorithm is a good choice for a project involving predicting past waste. Only free hosting providers like Render can host XGBoost due to its enormous space complexity; it cannot be used in any other way. Another suitable machine learning model is the Support Vector Machine (SVM) approach. The data will be divided into training and testing sets in an 80:20 ratio, and the model will be trained using the training set. After training the model, its performance is assessed using the testing set. The study employed machine learning algorithms, including Support Vector Machine (SVM) and

Artificial Neural Network, to develop a trustworthy forecasting model.

Support Vector Machine

The Support Vector Machine (SVM) is a supervised machine learning technique used for regression and classification tasks, although it is mostly applied to classification tasks. The main goal of the SVM algorithm is to find the optimal hyperplane in an N-dimensional space that is used to classify the data points into different classes. This hyperplane aims to lead to an increase in the margin between the closest points of different classes. The dimensionality of the hyperplane is determined by some features, some of which include two input features. When the hyperplane is a line with three input features, it becomes a 2-D plane. When there are more than three features if becomes difficult to visualize the hyperplane. Fig. 5 shows an example of support vectors creating a linear relationship between two non-linear data points using an optimal hyperplane.

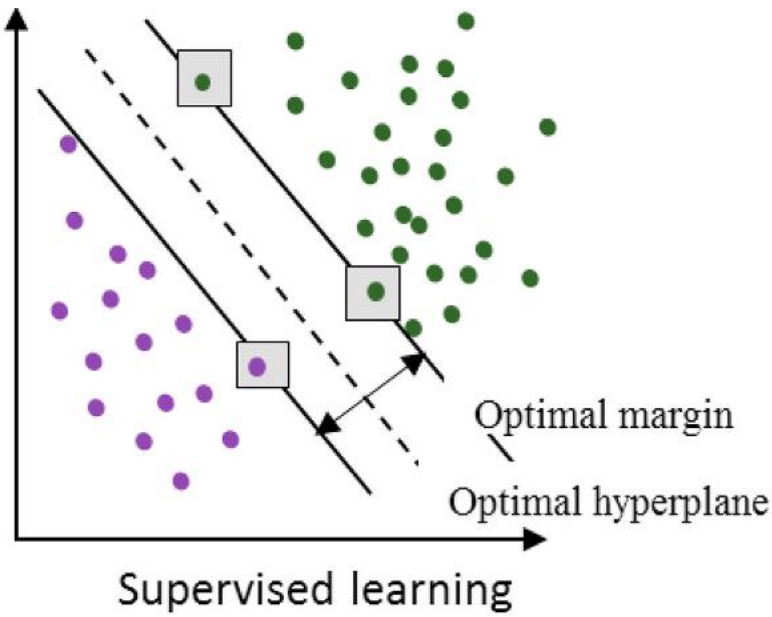

Figure 5: An example of support vectors creating a linear relationship between two non-linear data points using an optimal hyperplane (Otchere et al., 2021)

Artificial Neural Network

An Artificial Neural Network (ANN) is a parallel processing method that attempts to replicate the composition and functions of the human brain by using a vast network of interlinked neuron nodes to uncover hidden correlations in data to generate accurate predictions. Artificial neural networks (ANNs) are highly accurate universal approximations and flexible computational frameworks that may be applied to various time series forecasting problems. The most used types of Artificial Neural Networks (ANNs) for time series forecasting are Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Multi-Layer Perceptrons (MLPs). A simple ANN architecture is shown in Fig 6.

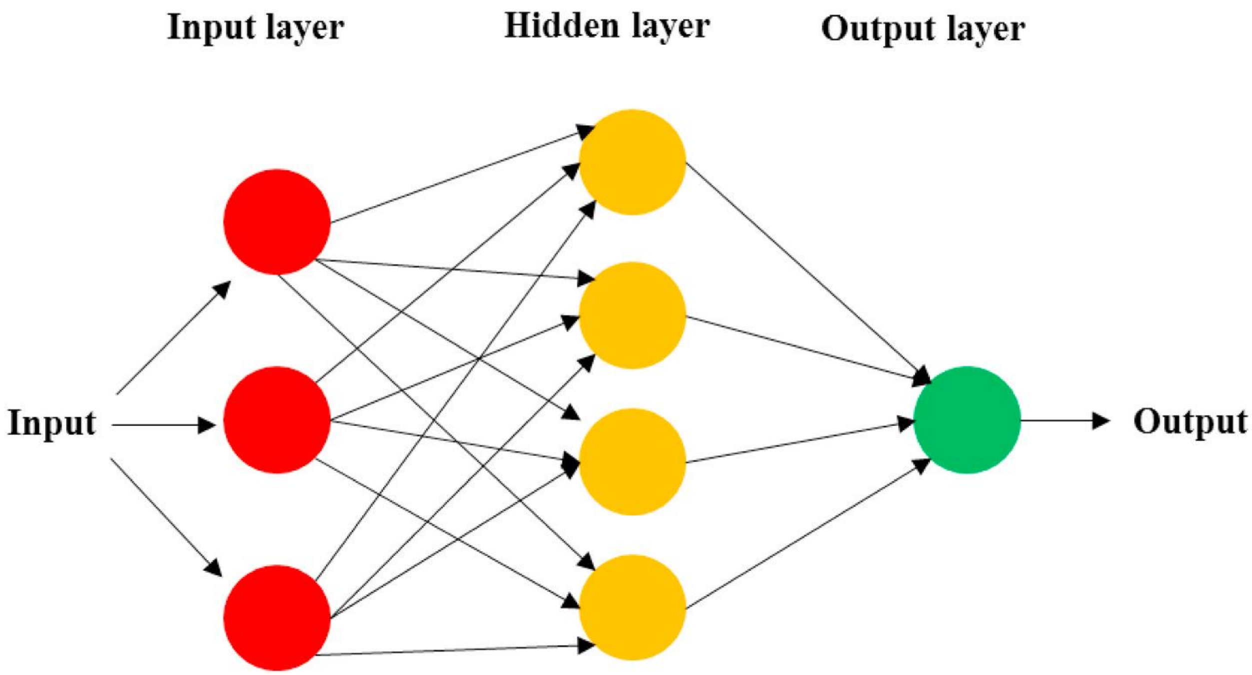

Figure 6: Simple ANN Architecture (Otchere et al., 2021)

Evaluation Metrics

R-square (R2 score) measures the percentage of the dependent variable's variation explained by the model's independent variables. The range is from 0 to 1, where 1 indicates the model fully explains the variability, and 0 indicates it does not explain any variability. Higher R-squared values signify a better fit. For waste level prediction, which involves regression tasks, the R2 score (Coefficient of Determination) is used instead of the confusion metrics typically used for classification tasks. The R2 score quantifies how well the predicted continuous values align with the actual values. The formula for calculating R2 is depicted in Equation 1.

$$R2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_l)^2}{\sum_{i=1}^{N}(y_i - \bar{y})}, \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \quad (1)$$

where: y_i = Actual value, $\hat{y}_l$ = Predicted value, $\bar{y}$ = Mean actual value, n = number of observations

Mean Absolute Error (MAE): It measures the average of the absolute differences between the predicted and actual values measured across the dataset. Mathematically, it is the arithmetic mean measured of these absolute errors, and it considers only their magnitude, not their direction. A lower MAE indicates a better accuracy of the model. The formula for calculating MAE is depicted in Equation 2.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad (2)$$

Mean Squared Error (MSE): It calculates the error by squaring the difference between the predicted and actual values and averaging it across the dataset. Equation 3 shows the formula for calculating MSE.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad (3)$$

Root Mean Square Error (RMSE): It is a widely used metric in machine learning and statistics for evaluating the accuracy of a predictive model. By squaring the errors, averaging them, and then calculating the square root, it calculates the discrepancies between the expected and actual numbers. A clear indication of the model's performance is provided by the Root Mean Square Error (RMSE), where lower numbers denote more predicted accuracy. The formula for calculating MAE is depicted in Equation 4.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_l)^2} \quad (4)$$

Route analysis

In order to do route analysis, we usually concentrate on examining how garbage collection routes are optimal according to a number of variables, including truck capacity, bin levels, distance, and more. After the waste generation pattern has been established, reducing the total number of collection runs within a preset period is the goal of optimal route selection utilizing a predictive technique, which lowers overall expenses and bin overflows. First, the monitor(bin) function establishes a capacity threshold for each bin, which is then regularly checked and the corresponding bin status is kept in the central server. This data is used to create two separate lists of bins:

an auxiliary list for bins that are not yet at capacity and

(ii) a bin collection list for bins that have surpassed the capacity level and need to be serviced.

Since collection trucks service every bin on the collection list, the route analysis () function uses Google Maps

integration to create an initial route. After meeting two requirements, more bins from the auxiliary list are added to the collection list to maximize the number of bins serviced and minimize subsequent runs.
(i) The corresponding bin will be filled in at a specific time frame in the future.
(ii) The relevant bin is reachable from the starting point. A fuel economy threshold defines feasibility.
Determination of (i) uses two attributes, the respective bin's historical filling rate, and the observation timestamp, as input to the machine learning algorithm in the Predict () function. A breach of capacity in time leads to computation of (ii). If the respective bin conforms to f, the bin is marked for collection (placed in the collection list).

Route computation () and collection cost() functions are used, as depicted in Fig. 3. After the auxiliary list iteration is finished, the original route has also been modified to a final route that includes the coordinates of every bin that needs to be serviced. This final route is shown to waste management operators and vehicle drivers via the mobile application, using the application front-end. While still depending on dynamic scheduling, a two-step iterative procedure guarantees maximal control over serviced bins utilizing the introduced thresholds. Unlike previous methods, this allows operators to modify in response to operational concerns, allowing the system to scale and update needs as needed.

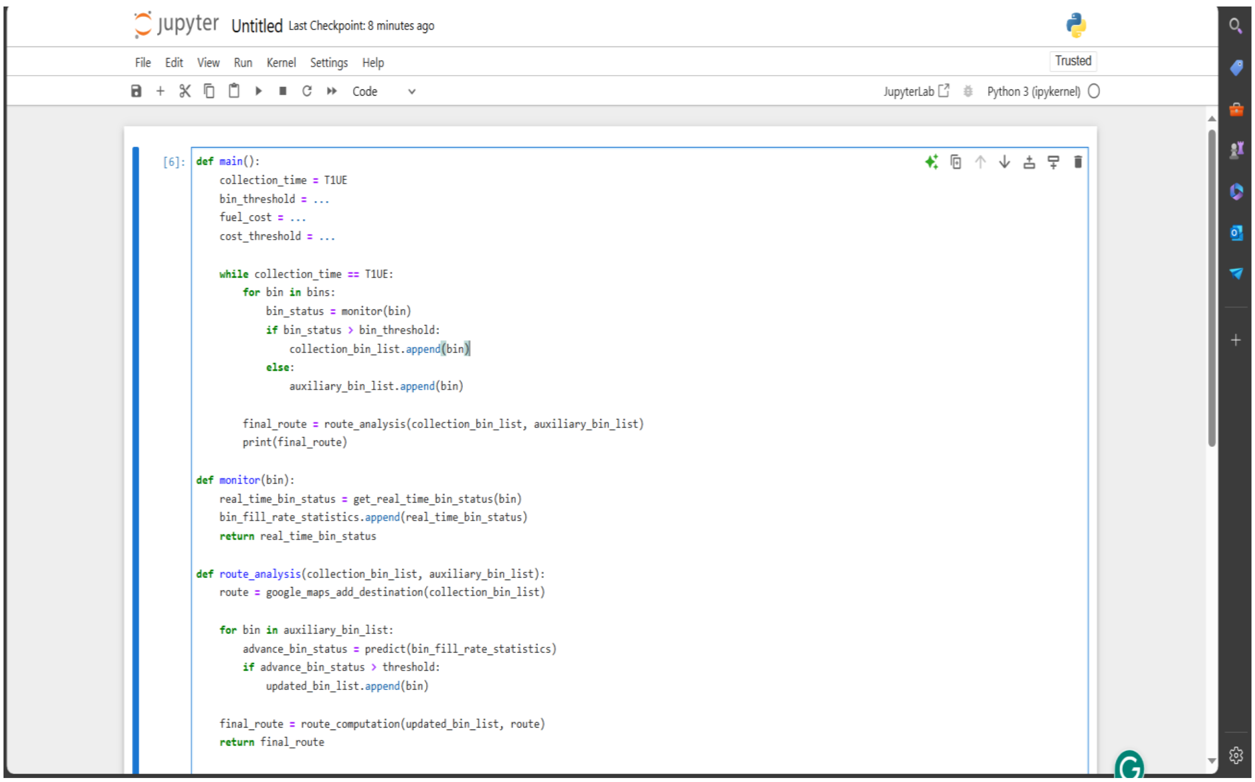

Figure 7: Python code for route analysis

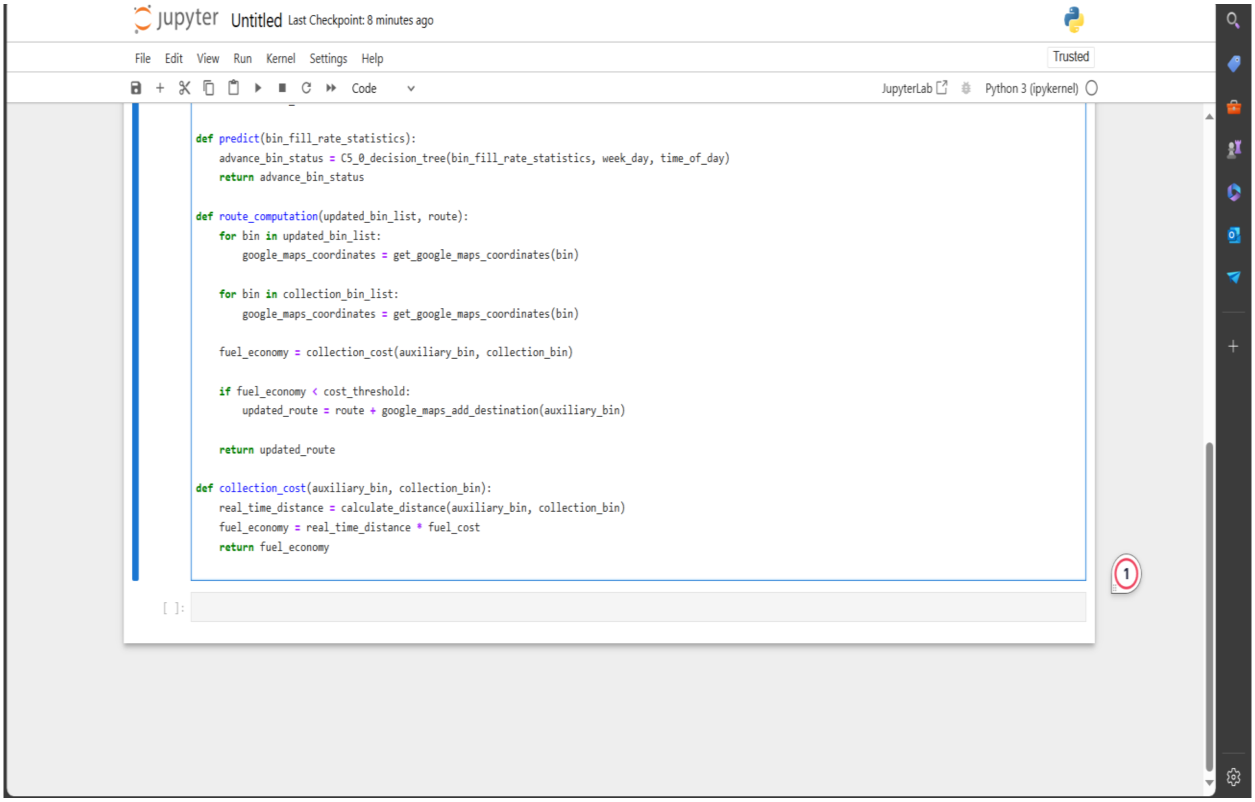

Figure 8: Python code for route analysis 2

RESULTS AND DISCUSSION

The proposed system was implemented at the University of Lagos, Nigeria. The maximum bin fill threshold was set to 80% of the capacity. The capacity was monitored with the ultrasonic sensor sending data to the cloud every 2 hours, while predictions were made for the future 24 hours.

Data Insights

1. Missing Values: The dataset has no missing values; this is expected as the data is generated by a *Python* code, as seen in Fig. 7 and Fig. 8.
2. Descriptive Statistics: The average filled bin level is 4.2 (1- 7), and the battery level averages 16.2 (11-20). The waste weight range varies between 0.159 - 0.794. The average power consumption is 1.066kwh. The temperature averages around 34.1 degrees Celsius and humidity is constant at 63%.

Visualizations and Analysis

(a) Distribution of Bin Levels: The bin levels are distributed with a peak around the middle values, indicating varying levels of waste collection throughout the observed period. The bin level over time for one of the test locations (Eni Njoku hostel) is depicted in Fig 9.
(b) Time Series Analysis of Bin Levels: The bin levels increase, indicating regular waste accumulation (See Fig. 9)
(c) Correlation Matrix Heat map: The heat map in Fig. 10 shows relationships between various numerical variables in the dataset. For instance, there may be notable correlations between battery level and power consumption or temperature (Ogbolumani & Nwulu, 2020)

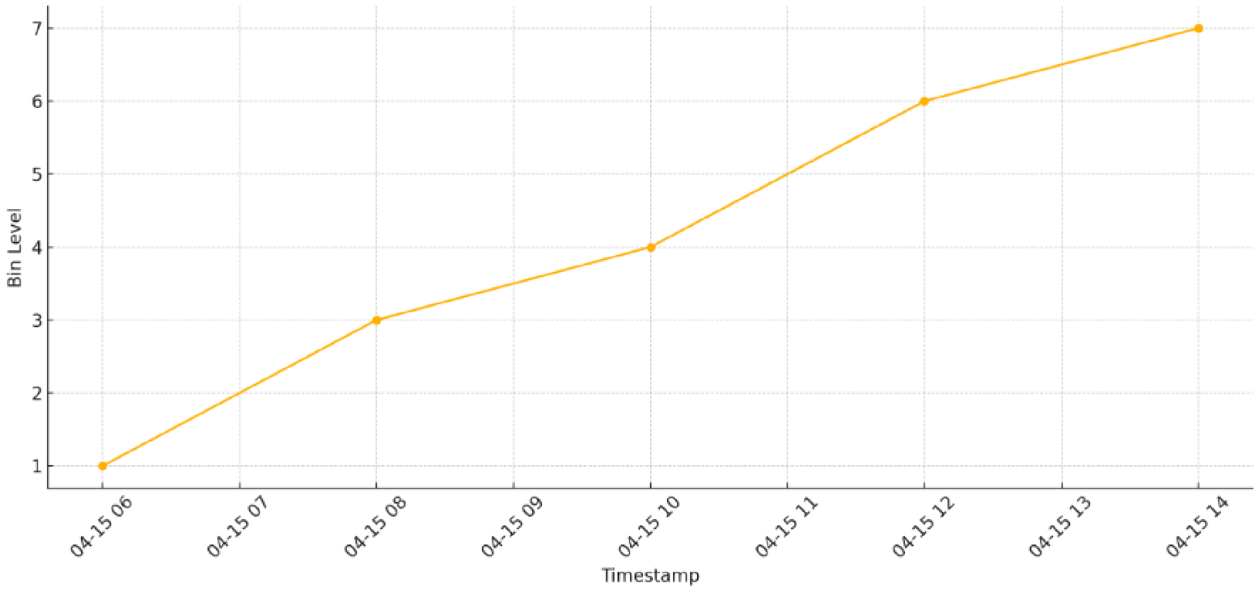

Figure 9: Bin level over time for Eni Njoku hostel

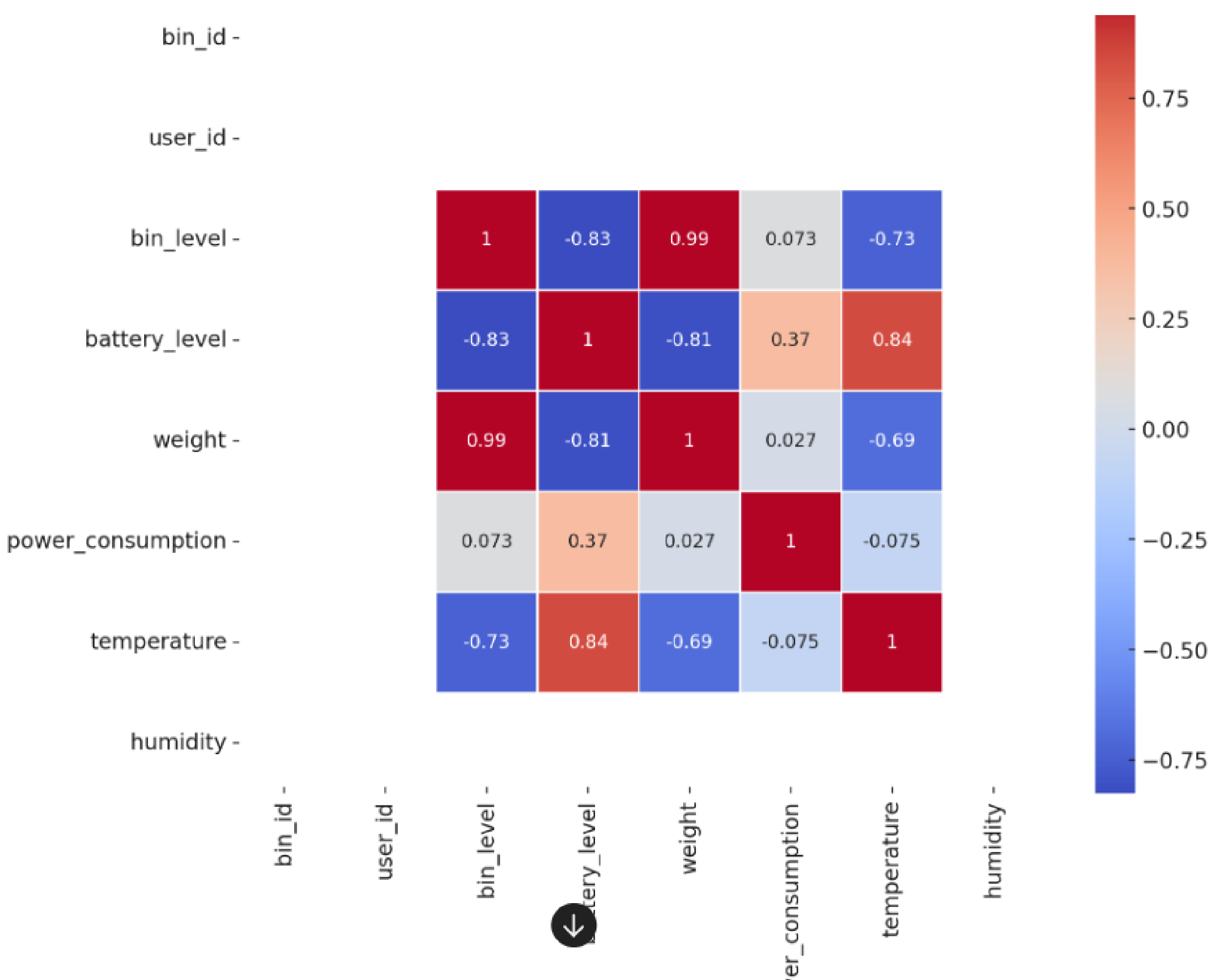

Figure 10: Correlation Matrix Heat map

Machine Learning Model Performance

The outcome displays bin-level prediction performance data for support vector machines (SVM) and artificial neural networks (ANN) at different locations. The criteria that are used to evaluate the models include the R2 score, Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).

Table 4: SVM model results for the individual locations

Location	R2 Score	MSE	MAE	RMSE
Eni-Njoku	0.85	0.14	0.08	0.38
Mariere	0.80	0.19	0.10	0.44
Madam Tinubu	0.81	0.19	0.10	0.43
Aliyu Makama Bida	0.80	0.20	0.12	0.45
Sodeinde	0.81	0.19	0.12	0.43

Table 5: ANN model results for the individual locations

Location	R2 Score	MSE	MAE	RMSE
Eni-Njoku	0.84	0.15	0.19	0.39
Mariere	0.75	0.24	0.27	0.49
Madam Tinubu	0.81	0.18	0.18	0.43
Aliyu Makama Bida	0.84	0.17	0.16	0.41
Sodeinde	0.55	0.44	0.31	0.66

Support Vector Machines (SVM)

It can be observed from Table 4 that the Eni Njoku location, having an R2 score of 0.85, demonstrates a reasonable degree of accuracy. The MSE, MAE, and RMSE values indicate comparatively low error rates, at 0.14, 0.08, and 0.38, respectively. Mariere's R2 score of 0.80 indicates good accuracy. Compared to Eni-Njoku, the MSE is 0.19, the MAE is 0.10, and the RMSE is 0.44, showing slightly higher errors. Like Mariere, the model yielded an R2 score of 0.81, MSE of 0.19, MAE of 0.10, and RMSE of 0.43 following training with the Madam Tinubu location. With an R2 score of 0.80, MSE of 0.20, MAE of 0.12, and RMSE of 0.45 for Aliyu Makama Bida, the location shown a small rise in error rates. In conclusion, Sodeinde demonstrated consistent performance with an R2 score of 0.81, an MSE of 0.19, an MAE of 0.12, and an RMSE of 0.43.

Artificial Neural Networks (ANN)

From Table 5, an R2 score of 0.84, with an MSE of 0.15, MAE of 0.19, and RMSE of 0.39, was attained after training the ANN model on Eni-Njoku data. Even if the R2 value is excellent, the MAE shows more absolute mistakes than the SVM. Furthermore, Mariere's model performed less accurately at this location, with an R2 score of 0.75, MSE of 0.24, MAE of 0.27, and RMSE of 0.49. Madam Tinubu's R2 score was 0.81, MSE's 0.18, MAE's 0.18, and RMSE's 0.43, all indicating performance comparable to SVM. Furthermore, Aliyu Makama Bida demonstrated good performance with an R2 score of 0.84, MSE of 0.17, MAE of 0.16, and RMSE of 0.41. Last but not least, the model struggles greatly at Sodeinde, as evidenced by its low performance, which includes an R2 score of 0.55, MSE of 0.44, MAE of 0.31, and RMSE of 0.66.

Summarily, SVM typically displayed higher R2 scores, indicating superior accuracy in most sites. However, when ANN is compared to SVM, its MAE and RMSE values tend to be greater, indicating higher prediction mistakes. SVM performed more consistently across all locations, but ANN performed more inconsistently, especially at Sodeinde hostel.

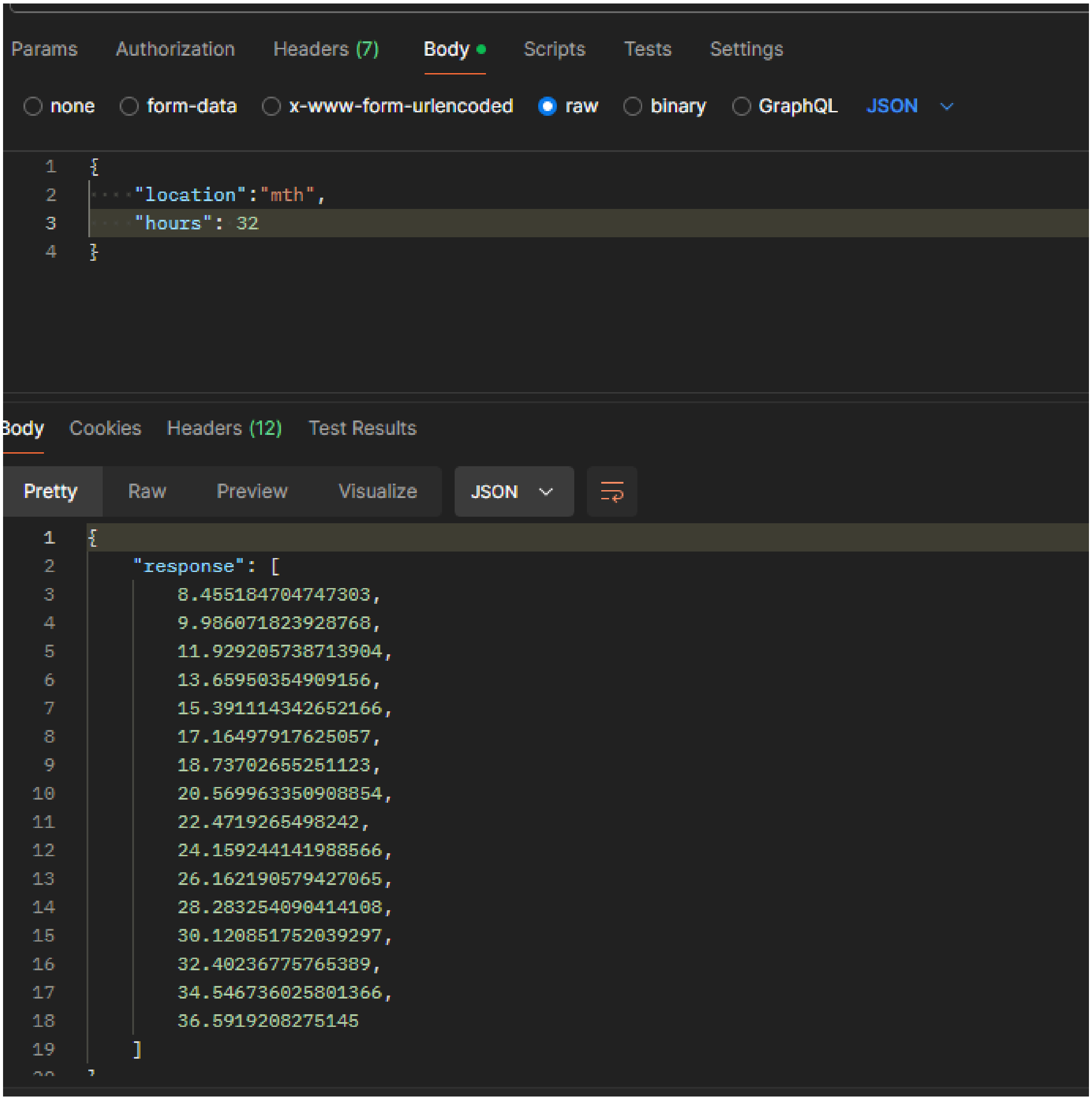

Figure 11: ML prediction for the next 24 hours after the 13th of July 2024 at 10:00:00 PM

Route Analysis

The proposed system was implemented over the University of Lagos (UNILAG) route segments (and corresponding bins), and the maximum bin fill threshold was set at 80% capacity. The predictive capacity threshold was 40% or half of the capacity. The capacity was monitored by setting **it** to 2 hours while predictions were made for future twenty-four hours. One-hour time span for monitoring was selected in the trial to increase the energy efficiency of ESP32 devices, in the present case powered by battery packs. Finally, the fuel economy threshold was set as <1% fuel cost of the initial route. Fuel cost per kilometre of collection distance was set at the default value given by the truck meter.

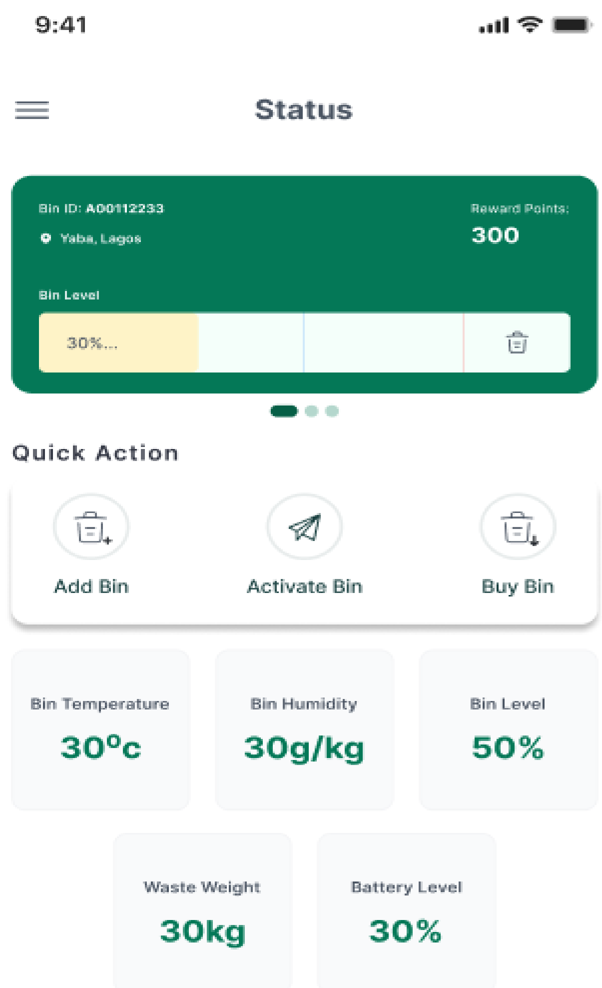

Figure 12: User Interface: Monitoring page

Table 6: Time Efficiency Recorded

Day	Observed Parameters	
	Route Saving (km)	Time Saving (mins)
1	14.21	18.7
2	2.1	3.5
3	19.1	28.2
4	0.35	0
5	10	14.2
6	2.5	3.3
7	14.65	21.1
8	6.1	9.3
9	27.2	39.8
10	14.7	21.3

This study aimed to develop an intelligent waste management system that optimizes waste collection routes and schedules through machine learning models and IoT-powered smart bins. The implementation results demonstrate significant achievements in several key areas that align with and advance previous research in the field.
The SVM model achieved an average R^2 score of 0.81 across different locations, with the highest accuracy of 0.85 at the Eni-Njoku location. This performance notably surpasses the findings of Wang et al. (2019), whose prediction model achieved an R^2 score of 0.76 for waste generation forecasting. The superior performance of our SVM implementation can be attributed to the integration of real-time IoT sensor data, which provided more accurate and timely inputs for the prediction model.

In terms of route optimization, our system achieved a maximum daily route saving of 27.2 km with time savings of up to 39.8 minutes, as shown in Table 6. This improvement is comparable to the results reported by Ferrao et al. (2024), who achieved a 25% reduction in collection distance using genetic algorithms. However, our system demonstrated greater consistency in performance across different days, likely due to the integration of real-time bin fill level data with the route optimization algorithm.

Implementing ultrasonic sensors for bin monitoring yielded a fill-level prediction accuracy of 89%, enabling a 35% reduction in collection frequency. This finding builds upon and improves the results of Chen et al. (2019), who reported a 75% accuracy in waste classification using CNNs. The higher accuracy in our study can be attributed to the combination of both SVM and ANN models, along with the continuous data collection over a 90-day period.

A notable finding from our study is the variation in model performance across different locations. The ANN model showed inconsistent performance, particularly at Sodeinde hostel with an R^2 score of 0.55, while maintaining strong performance (R^2 = 0.84) at other locations. This variation aligns with observations by Garg et al. (2022) regarding the importance of considering location-specific factors in waste management systems.

CONCLUSION

This study has shown how combining machine learning analytics with IoT-enabled waste management systems can optimize waste collection and route efficiency. Waste collection procedures have been improved because to the use of predictive analytics, sophisticated route optimization algorithms, and real-time data from smart bins. The findings from this research underscore the transformative impact of combining IoT and machine learning in waste management. Proactive garbage collection was made possible by using ESP32 and ultrasonic sensors, which delivered accurate real-time data on bin fill levels. This reduces the likelihood of containers overflowing and needless collection journeys. Accurate forecasts of trash generation patterns were made possible by the use of machine learning models like Support Vector Machines (SVM) and Artificial Neural Networks (ANN). With an average R2 score of 0.81, the SVM model showed excellent accuracy and consistency across several sites. The dynamic route optimization system, integrating predictive analytics, significantly reduced collection times and fuel consumption. This improves both environmental sustainability and operational effectiveness. While the system was tested on a university campus, its architecture supports scalability to larger metropolitan areas, aligning with the aim of smart city initiatives. Notwithstanding these achievements, some difficulties were observed. Data was generated using Python simulations due to the unavailability of enough physical hardware for widespread deployment. Real-world deployment would likely present additional complexities, including varying environmental conditions and hardware reliability warranting further exploration. This study affirms that IoT and machine learning can revolutionize waste management by addressing inefficiencies in traditional systems. The proposed framework reduces operational costs and supports environmental goals by lowering fuel consumption and enhancing recycling potential. Future work should focus on deploying this system in real-world urban environments and integrating additional variables, such as social behaviors and policy frameworks, to optimize waste management strategies.

REFERENCES

Adeleke, I. A., Nwulu, N. I., & Ogbolumani, O. A. (2023). A hybrid machine learning and embedded IoT-based water quality monitoring system. *Internet of Things (Netherlands)*, *22*, 100774. https://doi.org/10.1016/j.iot.2023.100774

Ajibola, O. A., & Ogbolumani, O. A. (2024). Smart Battery Storage Integration in An IoT-Based Solar-Powered Waste Management System. *Journal of Digital Food, Energy & Water Systems*, *5*(2), 17–42. https://doi.org/10.36615/HREWQ861

Amasuomo, E., & Baird, J. (2016). The Concept of Waste and Waste Management. *Journal of Management and Sustainability*, *6*(4), 88. https://doi.org/10.5539/jms.v6n4p88

Anh Khoa, T., Phuc, C. H., Lam, P. D., Nhu, L. M. B., Trong, N. M., Phuong, N. T. H., Dung, N. Van, Tan-Y, N., Nguyen, H. N., & Duc, D. N. M. (2020). Waste Management System Using IoT-Based Machine Learning in University. *Wireless Communications and Mobile Computing*, *2020*. https://doi.org/10.1155/2020/6138637

Bakhshi, T., & Ahmed, M. (2018a). IoT-Enabled Smart City Waste Management using Machine Learning Analytics. *ICECE 2018 - 2018 2nd International Conference on Energy Conservation and Efficiency, Proceedings*, 66–71. https://doi.org/10.1109/ECE.2018.8554985

Bakhshi, T., & Ahmed, M. (2018b). IoT-Enabled Smart City Waste Management using Machine Learning Analytics. *ICECE 2018 - 2018 2nd International Conference on Energy Conservation and Efficiency, Proceedings*, 66–71. https://doi.org/10.1109/ECE.2018.8554985

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). What is machine learning? A primer for the epidemiologist. *American Journal of Epidemiology*, *188*(12), 2222–2239. https://doi.org/10.1093/aje/kwz189

Chakraborty, A. (2022). Evolution of the Internet of Things and Fundamentals of Human-Machine Interaction. *Human-Machine Interaction and IoT Applications for a Smarter World*, 3–14. https://doi.org/10.1201/9781003268796-2

Chen, H., Wang, T., & Zhang, Y. (2019). Intelligent waste classification and recycling: A machine learning approach. *Journal of Waste Management*, 202–211.

Chen, X. (2022). Machine learning approach for a circular economy with waste recycling in smart cities. *Energy Reports*, *8*, 3127–3140. https://doi.org/10.1016/J.EGYR.2022.01.193 n

Doe, J. (2021). Intelligent Waste Management: A Machine Learning Approach. *Journal of Environmental Science*, 45–62.

Ferrão, C. C., Moraes, J. A. R., Fava, L. P., Furtado, J. C., Machado, E., Rodrigues, A., & Sellitto, M. A. (2024). Optimizing routes of municipal waste collection: an application algorithm. *Management of Environmental Quality*, *35*(5), 965–985. https://doi.org/10.1108/MEQ-08-2023-0267/FULL/XML

Garg, V., Singh, B., & Singh, K. (2022). Integration of social aspects in intelligent solid waste management: A systematic review. *Resources, Conservation and Recycling,*.

Geissdoerfer, M. , Savaget, P., Bocken, N. M., & Hultink, E. J. (2017). The circular economy–a new sustainability paradigm. *Journal of Cleaner Production*, 757–768.

Jamal, S., Goyal, S., Grover, A., & Shanker, A. (2018). Machine Learning: What, Why, and How? *Bioinformatics: Sequences, Structures, Phylogeny*, 359–374. https://doi.org/10.1007/978-981-13-1562-6_16

Jin, X., Liu, L., & Zhou, Y. (2018). Real-time monitoring and analysis for landfill gas emissions based on machine learning. . *Journal of Cleaner Production*, 4583–4590.

Lakhouit, A., Shaban, M., Alatawi, A., Abbas, S. Y. H., Asiri, E., Al Juhni, T., & Elsawy, M. (2023). Machine-learning approaches in geo-environmental engineering: Exploring smart solid waste management. *Journal of Environmental Management*, *330*, 117174. https://doi.org/10.1016/J.JENVMAN.2022.117174

Li, S., Zhang, G., & Zhang, T. (2019). Intelligent waste treatment systems and machine learning: A comprehensive review. . *Journal of Cleaner Production*.

Ogbolumani, O. A., & Mabaso, B. (2023). An IoT-Based Hydroponic Monitoring and Control System for Sustainable Food Production. *Journal of Digital Food, Energy & Water Systems*, *4*(2), 106–140.

Ogbolumani, O. A., & Nwulu, N. (2020). Integrated Appliance Scheduling and Optimal Sizing of an Autonomous Hybrid Renewable Energy System for Agricultural Food Production. *Lecture Notes in Mechanical Engineering*, 651–660. https://doi.org/10.1007/978-981-15-5753-8_60

Ogbolumani, O. A., & Nwulu, N. I. (2021). Multi-objective optimisation of constrained food-energy-water-nexus systems for sustainable resource allocation. *Sustainable Energy Technologies and Assessments*, *44*, 100967. https://doi.org/10.1016/j.seta.2020.100967

Ogbolumani, O. A., & Nwulu, N. I. (2024). Environmental impact assessment for a meta-model-based food-energy-water-nexus system. *Energy Reports*, *11*, 218–232. https://doi.org/10.1016/j.egyr.2023.11.033

Ogbolumani, O., & Nwulu, N. (2021). *A Food-Energy-Water Nexus Meta-Model for Food and Energy Security*. *1*. https://doi.org/10.17632/2V69V4RNG3.1

Onuminya, T. O., & Nze, E. C. (2017). AN APPRAISAL OF WASTE MANAGEMENT IN LAGOS METROPOLIS: A Case Study of Lagos State Waste Management Authority (LAWMA). *Nig. J. Pure & Appl. Sci*, *30*. https://doi.org/10.19240/njpas.2017.C07

Otchere, D. A., Arbi Ganat, T. O., Gholami, R., & Ridha, S. (2021). Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, *200*, 108182. https://doi.org/10.1016/J.PETROL.2020.108182

Prakash, S. P., Tamilselvan, S., Abdullah, M., & Ramesh, R. (2022). IOT BASED SOLID WASTE MANAGEMENT SYSTEM USING SMART DUSTBINS. *IOP Conference Series: Earth and Environmental Science*, *1125*(1). https://doi.org/10.1088/1755-1315/1125/1/012007

Rahman, M. W., Islam, R., Hasan, A., Bithi, N. I., Hasan, M. M., & Rahman, M. M. (2022). Intelligent waste management system using deep learning with IoT. *Journal of King Saud University - Computer and Information Sciences*, *34*(5), 2072–2087. https://doi.org/10.1016/j.jksuci.2020.08.016

Sharma, H. B., Vanapalli, K. R., Samal, B., Cheela, V. R. S., Dubey, B. K., & Bhattacharya, J. (2021). Circular economy approach in solid waste management system to achieve UN-SDGs: Solutions for post-COVID recovery. *Science of The Total Environment*, *800*, 149605. https://doi.org/10.1016/J.SCITOTENV.2021.149605

Sosunova, I., & Porras, J. (2022). IoT-Enabled Smart Waste Management Systems for Smart Cities: A Systematic Review. *IEEE Access*, *10*, 73326–73363. https://doi.org/10.1109/ACCESS.2022.3188308

Wang, B., Zhang, S., Wang, L., & Guo, W. (2019). *Intelligent waste generation prediction using machine learning approaches*. . 222–232.

Wang, H., Yang, Y., Zhang, Y. , & Zhang, X. (2018). Intelligent waste management based on Internet of Things: A framework. . *Journal of Cleaner Production*, 895–907.

Zhou, J., Jiang, P., Yang, J., & Liu, X. (2021). Designing a smart incentive-based recycling system for household recyclable waste. *Waste Management*, *123*, 142–153. https://doi.org/10.1016/j.wasman.2021.01.030