



## APPLICATION OF MACHINE LEARNING IN WASTE MANAGEMENT: AN INTRODUCTORY APPROACH

Uchenna Chinyere Onyemauche<sup>1</sup>, Jonathan Ugwu. Okwor<sup>2</sup>,  
and Charity Elochukwu Mbanusi<sup>2</sup>, Blessing Ulunma Osondu<sup>3</sup>

<sup>1</sup>Department of Computer Science, Federal University of Technology, Owerri, Imo State, Nigeria.

Email: [osigwe.uchenna@yahoo.com](mailto:osigwe.uchenna@yahoo.com), [uchenna.onyemauche@futo.edu.ng](mailto:uchenna.onyemauche@futo.edu.ng)

<sup>2</sup>Department of Computer Education, Federal College of Education (Technical), Umunze, Anambra State, Nigeria.

Email: [chary4sam@gmail.com](mailto:chary4sam@gmail.com); [jonathanokwor@gmail.com](mailto:jonathanokwor@gmail.com)

<sup>3</sup>Department of Health Information System, Imo State College of Health Technology and Management Sciences, Amaigbo, Nigeria.

Email: [blessingosonndu51@yahoo.com](mailto:blessingosonndu51@yahoo.com)

### Cite this article:

Uchenna C. O., Jonathan U. O., Charity E. M., Blessing U. O. (2024), Application of Machine Learning in Waste Management: An Introductory Approach. *Advanced Journal of Science, Technology and Engineering* 4(2), 36-46. DOI: 10.52589/AJSTE-RSWSCYFZ

### Manuscript History

Received: 12 May 2024

Accepted: 4 Jul 2024

Published: 19 Jul 2024

### Copyright © 2024 The Author(s).

This is an Open Access article distributed under the terms of Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), which permits anyone to share, use, reproduce and redistribute in any medium, provided the original author and source are credited.

**ABSTRACT:** *Waste Management is a daily task in urban areas, which needs a huge amount of labour resources and this affects natural, budgetary, efficiency and social aspect of our cities. Manual sorting of garbage is a difficult process that is expensive and that is why scientists create and study automated sorting methods that increase the efficiency of the recycling process. Most recently, there has been a drift in combining waste that is prime scheme with low cost IoT architectural design on a test board. However, the results from all these past approaches and techniques are still not clear and cannot be applied in real systems, such as in cities and campuses. This work introduces the design of a micro controller that is single, low cost, straight forward with an ultrasound sensor which can measure the filling height of a garbage Trash Bin and send information using LORA Technology. A novel IoT based Machine Learning method in combination with Genetic Algorithm to predict the probability of collecting waste in real environment based on historical data were used in this study. This is combined with a microcontroller system designed with a sensor module for measuring the height that is the fillings levels of each trash bin. The system can optimize the collection of waste with the shortest path by using genetic algorithm. Python was used for analyzing the data. Using the above Machine Learning techniques like Logical Regression cum Genetic Algorithm to compute the paths of wastes collection with different time schedules, it is cumbersome to get efficient route optimization; hence the aim of this paper was to present an IoT cloud solution combining device connection, data processing, control and ensuring route optimization. Genetic Algorithm ensures perfect route optimization.*

**KEYWORDS:** IoT, Machine Learning, Deep Learning, Waste Management, Genetic Algorithm, Logical Regression.



## INTRODUCTION

Waste and the risks associated with it are becoming a problem that is increasingly serious environmental protection. There is an expanding interest in waste management in the world, in both the development of technologies to minimize their quantity and those related to their disposal and economic use. The main reason for extreme waste generation is absurd materials management. The assembling of garbage in landfills may be used as secondary materials that are raw the value of which is estimated at a duo hundred million dollars (Batanic et al. 2016). Limiting the quantity of generated waste to a level that guarantees balance between raw material, ecological, and waste that is sanitary not possible without extensive synchronization of technologies and the manner people live with the formation and working of an ecological structure in the area. Actions aimed at reducing the amount of waste produced and placed in the surroundings should include recycling raw materials, minimizing waste production from end to end, the use of modern low-waste or non-waste technologies, and replacing traditionally used raw materials (Roy et al. 2022). The target system for solving the problem of production waste polluting the environment that is natural low and waste-free technologies. Non waste technology (NWT) is based on preventing waste and full comprehensive use of the raw material. It involves a number of technological processes that lead to total management and, consequently, the elimination of pollution without harmful effects on the environment. The condition here is that waste ought not to be deposited. The implementation of NWT has its economic justification, because the full use of materials and, consequently, the reduction of the amount of waste, allows for increased production and allows for the reduction of imports of raw materials. In some cases, it is also possible to reduce the consumption of electricity, heat, or technology by reducing energy-consuming waste treatment processes. The benefits of using non waste technology also include reducing material consumption, environmental losses, and costs that are operating.

Another method to reduce waste is recycling. Its basic job is to maximize the recycle of the same materials, including reduction of expenditure on their processing. The recycling process takes place in two areas: the production of goods and the generation that is subsequent of from them. Its assumptions assume the imposition of appropriate attitudes among manufacturers, conducive to the production of the most recoverable materials, and the creation of appropriate behavior among recipients. Recycling of waste from used post-consumer products can take place, among others through the secondary use of raw material combined with a change in its condition and composition. For this, it is necessary to sort waste not only into fractions such as metal, bio, plastic paper, or glass.

The internet of things, or IoT, is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers (UIDs) and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction.

A thing in the internet of things can be a person with a heart monitor implant, a farm animal with a bio chip transponder, an automobile that has built-in sensors to alert the driver when tire pressure is low or any other natural or man-made object that can be assigned an Internet Protocol (IP) address and is able to transfer data over a network.



Increasingly, organizations in a variety of industries are using IoT to operate more efficiently, better understand customers to deliver enhanced customer service, improve decision-making and increase the value of the business.

### **How does IoT work?**

An IoT ecosystem consists of web-enabled devices that are smart use embedded systems, such as processors, sensors and communication hardware, to collect, send and act on data they acquire from their environments. IoT devices share the sensor data they collect by connecting to an IoT gateway or other edge device where data is either sent to the cloud to be analyzed or analyzed locally. Sometimes, these devices communicate with other devices that are related act on the information they get from one another. The devices do most of the work without human intervention, although people can interact with the devices -- for instance, to set them up, give them instructions or access the data (Munir et al. 2023) .

Machine learning is behind chat bots and predictive text, language translation apps, they show Netflix suggested to you, and how your social media feeds are presented. It triggers autonomous vehicles and machines that can diagnose medical conditions based on images.

When companies today deploy artificial intelligence programs, they are most likely machine that is using, so much so that the terms are often used interchangeably, and sometimes unclearly. Machine learning is a sub field of artificial intelligence that gives computers the ability to explicitly learn without being programmed.

“In just the last five or 10 years, machine learning has become a way that is precarious arguably the most crucial way, most parts of AI are done,” said Abbas et al. (2019). “So that's why some people use the terms AI and machine learning almost as synonymous ... most of the current advances in AI have involved machine learning.”

Machine learning (ML) proffers effective solutions, such as regression, classification, clustering, a correlation that is perception for IoT-based waste management (Munir et al. 2023), There are three reasons for this: firstly, in IoT applications, all of the devices are connected, and greater amount of data is collected every day. Furthermore, they may be programmed to elicit some events, based either on some predefined conditions or exciting feedback from the collected data. Secondly, computer systems can learn to accomplish certain tasks, such as classification, clustering, predictions, and pattern recognition. Additionally, these systems are trained using algorithms that are numerous statistical models to analyze sample data. Thirdly, reckon-able characteristics (called features) usually illustrate the sample data; some ML algorithms attempt to find correlations between the features and some output values (called labels). The information obtained through training is then used to identify patterns or make decisions based on new data.

However, human intervention is typically required to analyze the collected data, extract meaningful information, and build smart applications. IoT devices must not only collect data and transfer it to other devices but be self-sufficient also. They must be able to make context-based decisions and learn from their collected data . Hence, waste management systems using machine that is IoT-based should consider more variables for task prediction and forecasting. On the other hand, regression is a model that is mathematical can represent or recognize the connection between two or more variables. The dependent or response variable is the system's output. The relationships between the dependent variables and the independent ones can be



detected by applying a regression model to a system. Logistic regression (LR) is a kind of regression analysis used to explain the connection between a dependent variable and one or more independent variables. LR is appropriate when the dependent variable takes values that are binary. As such, it is suitable for prediction tasks in ML and IoT application (Abbas, 2020)

IoT-based waste management models perform a vital function in improving the standard of living and human well-being by increasing energy-efficiency, enhancing governance, and cost that is reducing. The application of IoT technology to waste management is one aspect of this model with the successful model and dream of a top pollution free environment.

This article presents an innovative podium for smart trash control at Owerri Municipal, which is able to keep the city clean at low cost, requiring low labour resources. A novel micro controller system is designed with a sensor module for measuring the height that is filling of using ultrasound and geolocation of collected data based on LoRa technology. Furthermore, the paper presents a new method for predicting the probability of the filling level of each trash bin by applying Logical Regression in ML; furthermore, a graph theory-based optimization solution is proposed to compute the paths of waste collection with different time schedules in order to minimize the environmental and socio-economy micro impacts, as well as supporting the inhabitants of the city.

The contributions of our work are: (i) Previous articles have mostly evaluated results on a test board; our work introduces the design of a microcontroller that is single, which is of low cost and straightforward, with an ultrasound sensor which can measure the filling height of a garbage can and send information using LoRa technology. (ii) present a novel IoT-based machine learning method, which is employed to predict and forecast the probability of collecting waste in the real environment based on the historical input data. (iii) Our article is the first to propose the use of the function that is sigmoid predicting the probability of waste collection and to apply Genetic algorithm to optimize the path for waste collection from trash bins Auto ML (Goddard et al. 1995). In a research that is recent and there were no extensive work applying it in industrial applications. Therefore, this section concentrates on the state-of-the-art that is current of Smart Waste Management both in research and practical domains. The Smart Waste Management is very broad and includes many aspects. For the survey that is comprehensive of area as an integral part of the Smart City concept readers are kindly referred to (Pardini, 2019). A systems that are commercial the Smart Waste Management combines elements of Clean technology and Internet of Things as well as partially addresses the environmental issues, it is not surprising that there are several commercial systems competing in similar business niches. The list of commercial systems can be divided into two types: retrofitted sensors and containers that are smart. Smart containers are specially made containers, e.g., urban dustbins or containers for cardboard measuring the level that is filing. On the other hand, some approaches have developed waste management strategies based on the optimization to achieve an efficient system. In Bueno Delgado et al 2019, the authors presented a waste control and management stage to be applied in rural areas LoRaWAN that is using technology route optimization. Additionally, an implementation based on IoT was set up, but the system did not provide clarity about communication and optimization for all trash bins in the system. Based on logistic regression (LR) and genetic algorithm (GA) methods, the authors in Kolekar et al. (2016) presented a new method to check the status of smart trash bins and select a collection path in Philadelphia, USA. Moreover, they did not offer any technologies for the transmission of data from the trash bin to the other devices in the system. In particular, optimization algorithms have been clearly defined for IoT-based waste management, such as



the neighbour search that is nearest, colony optimization, genetic algorithm, and particle swarm optimization methods (Folianto & Yeow 2015). In Dinh & Yeow (2016), the authors proposed a solution to manage a garbage system integrated with IoT technology, which was a line-following that is autonomous with a robotic hand for garbage gathering, in which they did not apply any algorithms to optimize the waste collection. In , an IoT platform for an automated waste collection system provided by the project allowed real-time monitoring and interface with the system. However, the aim of this paper was to present an IoT cloud solution combining device connection, data processing, and control, rather than the design and optimization of waste collection. A food waste collection method was presented in Kolekar et al. (2016), where the information was gathered using radio frequency identification (RFID) technology and transmitted using a wireless mesh network. However, the disadvantages of this technology were severe in the long range, especially considering the aim of the smart city is management with an area that is large. Finally, the results of the optimization algorithm became too vague and could not be applied to a system that is real such as a city.

For practical waste management, an impressive architecture was proposed for a sensor node in Kolekar et al. (2016), which used a micro controller (ATMega328P), an ultrasonic sensor (SRF05), and a LORA E32 TTL—100 433 MHz module. Nonetheless, they only tested the test board as a platform to provide sensor nodes and did not apply any methods for waste management in a city that is smart such as optimized waste collection. The micro controller board was very problematical; moreover, certain routines and functionalities are needed for each particular application while the target of this article was to propose an IoT application, with their design. Our article applies machine learning and graph theory to optimize the waste collection processes, avoid bins that are overfilled and reduce the work load.

In this paper, we propose an IoT system that will optimize the segregation and collection of waste products by predicting and forecasting the probability of the waste level in trash bins. By using genetic algorithm cum machine learning, the system can optimize the collection of waste with the shortest(optimal) path.. The primary purpose is to reduce the total cost of transport, transfer, save labour, and reduce the dependency on used vehicles, while maximizing service quality, as well as improving general quality of life. The other algorithm optimizes waste collection, instead of considering low cost and energy savings, and supports the city's waste management network to perform with high efficiency by applying ML and graph theory methods.

## METHODOLOGY

The system under consideration consists of smart trash bins with a monitoring that is real-time which integrates multi choices, such as ultrasound distance, along with a LoRa E32 TTL-100 433 MHz transmission module. Low energy use was considered throughout the design process. Each node is consequently supposed to be powered by multiple sources; for instance, solar energy or batteries. For flexibility, we have designed hardware that can use either energy source, as trash bins are often put in places where direct sunlight is not available.

LoRa is designed to work in the appropriate band for each country or region because of license-free bands in each different country. The LoRa module depending on the physical layer used can be classified as in the 433, 868, or 915 MHz frequency bands (Batinic et al. 2016).



The topology of LoRa WAN network architecture is star, where the end devices can only communicate with LoRa WAN gateways and not directly with each other. The LoRa WAN gateways are used to respond to forwarding raw data packets from end nodes towards the network server. In our network architecture, the LoRa WAN is used by us MAC layer, class A, that provides the medium access control mechanism that enables the communication between multiple devices and the LoRa WAN gateway.

The IoT gateway node is responsible for establishing communications with all the nodes in 24 hours. Its hardware requirements are actually simple, since it only has to be able to host and send the data for the server and offer a communication interface with the node. Besides the hardware previously mentioned, the most important is the power system. For flexibility, the system can be chosen to power the nodes with Li-ion 18650 batteries or energy that is solar as shown in Table 3, such that the input voltage is within an acceptable range for operating the LoRa E32 TTL-100 433 MHz module.

### Optimal Path Planning Algorithm for Waste Collection

In this section, we discuss the waste collection data, along with their states, positions, and system design, and test data that are real verify the output result. First, the obtained information is transmitted through the communication link to the server, where data are processed, saved, and forwarded to the cloud, as studied in the next section. Second, we apply Algorithm 1 and 2 to predict the status and route of each trash bin daily. Note that we consider the location of the aggregate waste collection location of each building, instead of each small trash bin on each floor of a big compound; furthermore, the system will enable a node that does not work to be turned off. We have started by trying the algorithm that is genetic all the problems. The VNS (Variable Neighbor Search) consists of applying a local search method repeatedly in the neighborhood of the actual solution. When an optimal that is local reached, the algorithm changes the system of neighborhood with the aim of escaping from local optima and reaching a better one. For this reason, it can be said that the VNS performs well both when searching local and optima that is global. In Algorithm 2, the process followed is explained. We have implemented two systems of neighborhoods: the 2-opt neighborhood and a neighborhood that is swap-based. The neighborhood structure 2-opt consists of changing a pair of edges between cities the neighborhood that is swap-based we have created swaps the first element of the permutation with all the rest.

Algorithm 1 Algorithm to determine solutions using Genetic Algorithm

Procedure genetic

    Create all the permutations

for  $i=1:7!$  do

Algorithm 2: Algorithm to determine solutions using Simple Genetic Algorithm procedure Simple algorithm (that is genetic matrix, generations, population size, probabilities)

    Create a population that is initial of different permutations randomly

    Compute the evaluation function of each permutation



```

for i=1:generations do
for j=1:m/2 do
    Select two parents randomly from the population
    Cross with a certain probability to produce two descendants
    Correct the descendants to be permutations
    Mutate each individual with a certain probability
    Compute the evaluation function of each descendant if evaluation function of
the descendants smaller than the evaluation that is largest then
if the descendants are not repeated in the population then
    Replace the descendants by the permutations with evaluation that is largest
end if
end if
end for
    Output the shortest distance and the corresponding route of each generation
end for
    Output the shortest distance and the corresponding route of the process that is overall
End Procedure

```

Algorithm 1 and 2 : Optimal cum shortest path Algorithm that is planning waste collection.

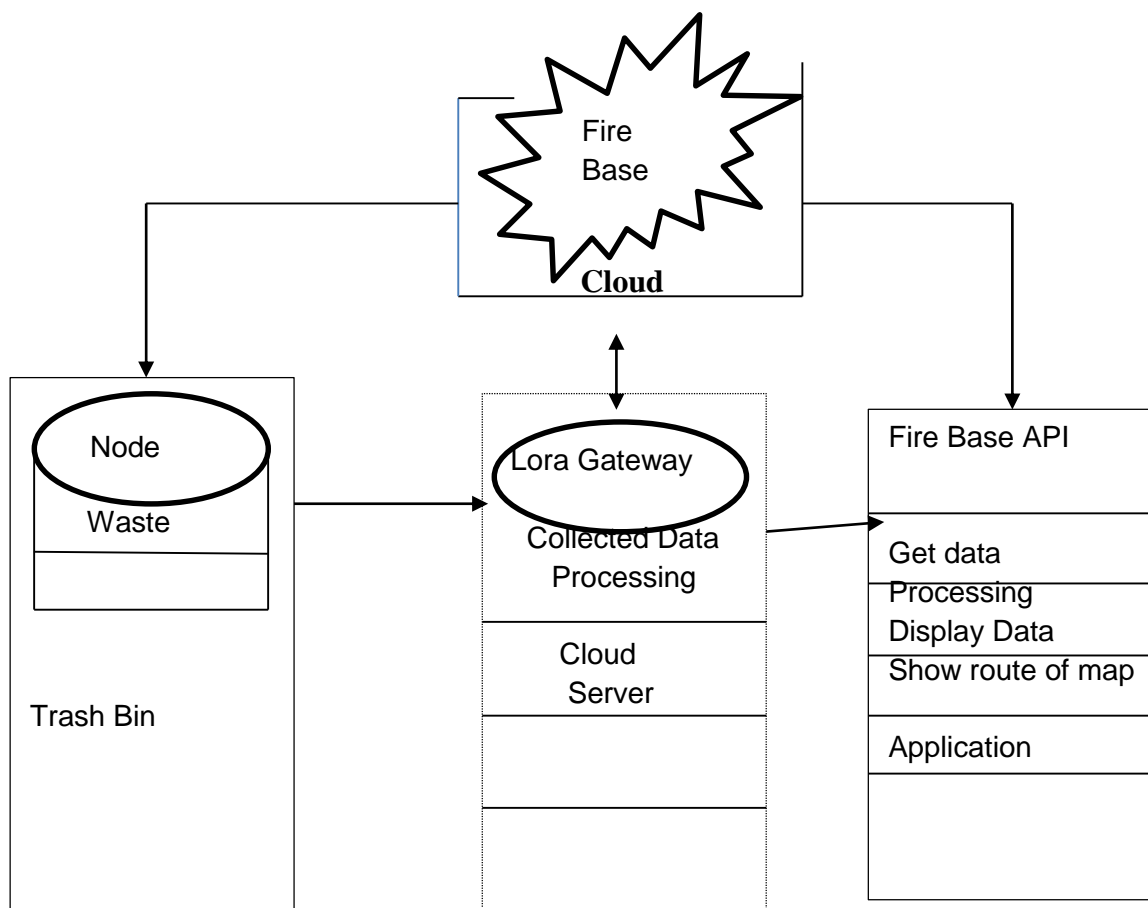
### Data Collection

To collect the data set of waste, we observed several classes for all buildings within working hours (from 7 A.M. to 21:00 P.M . from Monday to Sunday) for 4 months. The data set is obtained. A(node A Additionally, the weight of waste is increased in case of more occupants for instance, on Monday, at building. We conclude occupants in the building that the data set of waste is depended by the true number of occupants in a building. Therefore, we put one waste that is large on the ground floor to handle 13 buildings (nodes) in a given city area.

Python was used to carry out the data analysis and implement the forecasting models. Notebooks technology was employed to visualize and display the data, which codes interactive digital files that envelop both the models' code and description, while visualizing the results obtained. This facilitates the analysis and manipulation of the given information contained therein. Information in these types of tools may be analyzed separately and correlated among the same. This comparative analysis that is behavioral of led to the determination that, for this type of forecasting with a limited amount of data, support vector machines are more efficient given the type and quantity of available information.

With respect to data structuring, improvements were made to the data to achieve an automated reading of the same, and to be able to demonstrate its quality in terms of its missing elements. A characteristic of solid waste management databases is the amount that is limited of, as well as their continuity issues. Estimates were made of the missing data based on observations with the same characteristics. For example, values from earlier years or neighboring values in the data column were used. Once the data was organized and structured, a visual analysis tool was then developed. A file was developed in Notebook that has the reading, description and visualization of the data.

The information obtained is displayed in a manner that is vibrant which facilitates its analysis and handling. The information can be separately analyzed by locality and zone or correlated within the same. Initially, data exploration was performed to determine relationships and patterns of the stored figures. The best method to visualize and understand data is to convert stored data into a time series. Similarly, peaks and small periodic ranges can be identified. For example, the zones that generate the waste that is most are crowded zones.



**Fig.1 Illustration of the overall smart waste Collection System**



The sub systems that form the connection of the IoT application include 1. Lora Mac. 2. FireBase API Host for ESP, 3. Fire Base API/ Host for Application. The IoT nodes will have three in number. This connects to the Gateways which are two in number, then to the server that is hosting. This is illustrated in Figure 2.

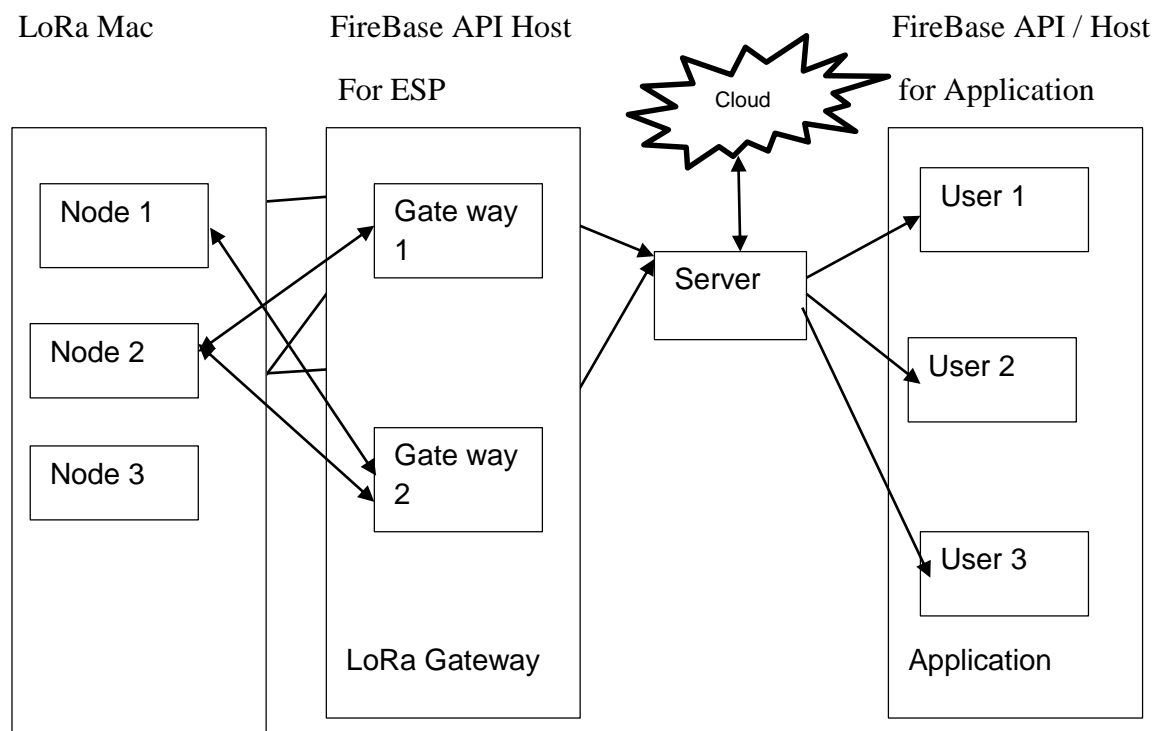


Fig. 2: Sub systems that form connections of IoT Application with the Application

## RESULTS AND DISCUSSION

The modeling results using the aforementioned data is shown above in figure 2 . First, the behavior of the results of each model was analyzed. Then, taking into consideration this behavior, an analysis that is predictive of waste generation in the city was carried out.

### Forecasting with decision trees

Decision trees model data separation limitations based on learning decision rules made for the model's input characteristics. The depth of the learning tree is determined based on the sensibility that is required of model, which contains a set of decision thresholds that separate the data contained in each characteristic. The input characteristic is waste production per year for example, in this particular problem. The learning algorithm will attempt to separate the waste into two subgroups and there would only be one separation value if one decision tree is defined from two trees. As the determined depth increases, the separation becomes narrower and the data can be separated in a more manner that is optimal. However, a very large tree depth value makes separation lines on the adjustment, which limits the generality of the model for new values.



This means that the model learned noisy input data and will not be able to forecast coherent values for values with a determined variance with respect to the training data. Decision trees perform recursive partitioning of the characteristics space  $j_1 \dots j_m$ , which in this case, corresponds to the waste percentage. Each candidate partition is defined as  $\theta = (j, \tau_m)$ , where  $j$  is the value in terms of waste and  $\tau_m$  is the prediction threshold. Each tree node  $Q$ , is therefore defined as the partition  $Q_{left}(\theta)$  and  $Q_{right}(\theta)$ . An optimized version of the CART algorithm (classification and regression trees) was used in the implementation.

## CONCLUSIONS AND FUTURE WORKS

A genetic algorithm has been described, with the possibility of assessing the probability of a trash bin being fully based on the number of buildings in a compound in this paper. . Vector support machines and genetic algorithm were also implemented to calculate regression models based on local radial functions, using points (support vectors) in a specific neighborhood as a secondary alternative given the possibility of obtaining suitable results with this tool due to the limited amount of data. Lastly, methods based on neural networks were implemented to estimate the trajectories of the points. This algorithm presents many advantages, as compared with the old waste collection methods. In addition, this scholarly study also provides improvements over classical algorithms. The algorithm is integrated into the system with a low-cost design circuit and LoRa WAN technology, enabling its application in practical use-cases, in which changing the sensor components can be done quickly. This study also presented three experiments: first, a test was conducted by simulating the location of arbitrary trash bins and finding the path that is shortest between the trash bins. Second, the logistic regression equation will be applied to estimate the probability of collecting the waste. Finally, a practical-use case of the waste collection process will be tested. The algorithm used the database of buildings used in e compound that is ach the city of Owerri and data received from smart buckets (e.g., occupancy rate) as input data. In summary, this system provides better operations for optimizing occupants use, saving operating costs, and collecting data on time. The system can be built cheaply, simply, and effectively and will be extensively applied in any planned city, to unify the filling level of the trash bins by using an ultrasonic sensor. Again, machine learning will be used, in particular, using multi classification methods.

On the basis of this quantitative evidence, it can be concluded that SVM ( Support Vector Machine) is the method with the best behavior in terms of the local trend of the points and assuming a high reliability in the recorded values. Likewise, this tool will enable the calculation that is correct analysis considering the limited amount of data which is often the case in comprehensive solid waste management processes in which records are limited or non-continuous due to the characteristics of the manner in which these types of practices are carried out in cities. These types of tools enable the calculation of trends for future values of waste that a city may thus generate to be able to design more precise and different types of strategies for the transport, collection, use and disposal of solid waste generated by a city.



## REFERENCES

- Abbas A.K., Al-haideri N.A., Bashikh A.A.(2020). Implementing artificial networks that are neural support vector machines to predict lost circulation. Egypt.
- Batinic B., Vukmirovic S., Vujic G., Stanisavljevic N., Ubavin D., Vukmirovic G.(2016) Using ANN model to determine future waste characteristics in order to achieve specific waste management targets - case study of Serbia. *J. Sci. Ind. Res. (India)* ;70(7):513–518.
- Bueno-Delgado, J.-L. Romero-Gázquez, P. Jiménez, and P. Pavón-Mariño, (2019)“Optimal path planning for selective waste collection in smart cities,” *Sensors*, vol. 19, no. 9, p.112-128
- Goddard J.C., Cornejo J.M., Martínez F.M., Martínez A.E., Rufiner H.L., Acevedo R.C.(1995). Proceedings of the International Computer Symposium Organized by the National Polytechnic Institute. Neural networks and decision trees: a hybrid approach; pp. 1–7.
- Kolekar K.A., Hazra T., Chakrabarty S.N.(2016). Prediction of municipal waste that is solid models. *Procedia Environ. Sci.* 2016;35:238–244.
- Munir, I. S. Bajwa, M. A. Naeem, and B. Ramzan, G.M.(2023). Implementation and“Design of an IoT system for smart energy consumption and smart irrigation in tunnel farming,” *Energies*, vol. 11, no. 12, p. 3427.
- Minir, I.S., Bajwa, M.A., Naeem, B. & Ramzam, G.M.(2020). Implementation and Design of an IoT system for smart energy consumption and smart irrigation bin tunnel farming.*International Journal of Energies*, vol. 11, no. 12, p.27 -34.
- Roy S., Rafizul I., Didarul M., Asma U., Shohel M., Hasibul M. (2022)..Prediction of municipal solid waste generation of khulna city using artificial neural network: a case study. *Int. J. Eng Res-Online*.
- Semtech (2021), LoRa (Phy). <https://www.semtech.com/lora/what-is-lora>. Accessed on 18 Dec, 2021