Doctoral Dissertation

Study on smart garbage bin system for understanding household garbage disposal and contents estimation

Eunice Likotiko

March 3, 2023

Graduate School of Science and Technology
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Eunice Likotiko

Thesis Committee:

Professor Keiichi Yasumoto	(Supervisor)
Professor Kiyoshi Kiyokawa	(Co-supervisor)
Assistant Professor Yuki Matsuda	(Co-supervisor)

Study on smart garbage bin system for understanding household garbage disposal and contents estimation*

Eunice Likotiko

Abstract

Much garbage is produced daily in homes due to living activities, including cooking and eating. The garbage must be adequately managed for human well-being and environmental protection. Although the existing IoT-based smart garbage systems have gained high classification accuracy, they still have problems: (1) they can not learn the amount of garbage disposed of each time (2) They can not understand households' routine behaviour of garbage disposal; and (3) they provide a small number of garbage categories, not enough for reasonable practices of household garbage separation. Therefore, a new IoT-based garbage management system and classification tool must improve existing systems. In this dissertation, we present a new smart garbage bin system, SGBS in short, embedded with multiple sensors to realize its benefits on three challenges: (1) How to learn the amount of garbage disposed of each time and predict garbage growth behaviour for a single house?; (2) How to understand household garbage disposal behaviour and identify the type of garbage contents?; and (3) How to substantially improve the automation of garbage classification? To tackle challenge (1), we chose distance and weight sensors to track the garbage disposed of each time. For evaluation, we experimented with the SGBS in a student laboratory for over one month. An autoregressive integrated moving average (ARIMA) model was applied, providing MAE of 5.17 cm and an SD of 0.33 cm, thus considered satisfactory accuracy on the garbage growth prediction. For challenge

*Doctoral Dissertation, Graduate School of Science and Technology,
Nara Institute of Science and Technology, March 3, 2023.

(2), we deployed distance and weight sensors to learn garbage growth during disposal. Later we identified necessary garbage categories and contents in each category through user voice, and designed and implemented a smartphone annotations application comprised of 8 garbage categories and 25 garbage content identities to allow households user to annotate their daily garbage content. Afterwards, we conducted an initial experiment in three households to evaluate our approach. Our findings show that households' garbage disposal behaviour depends on the amount and contents of garbage and the routine of disposing of such garbage content. For challenge (3), we introduce a new garbage content estimation method by training a machine learning model using daily collected fuse sensor readings combined with detailed household garbage contents annotations to perform garbage classification tasks. For evaluation, we deployed the designed SGBS in five households over one month. We confirmed that the leave-one-house cross-validation results showed an accuracy of 91% in 5 kitchen waste contents, 89% in 5 paper/softbox contents, and 85% in 8 garbage categories for the classification tasks.

Keywords:

Classification, disposal behaviour, garbage category, garbage content identity, household, IoT, machine learning algorithms, SGBS.

Contents

List of Figures

1 Introduction

1.1 Background and Motivation

Much garbage is produced daily in homes due to living activities, including cooking and eating. Therefore, garbage must be adequately managed for human well-being and environmental protection. In the standard municipal garbage management system, households are responsible for sorting and managing garbage produced in their home. However, it is hard to depend solely on public awareness to provide the correct garbage management at the source. Therefore, an automation tool that can reflect the home's daily life and understand households' routine behaviour of garbage disposal would be necessary to influence behaviour change on garbage disposal and increase home monitoring for the case of elderly anomaly detection and healthy living.

Furthermore, it would improve garbage management services through proper garbage separation practices for the well-being of people and the environment. It is reported that the world generates 2.01 billion tonnes of municipal solid waste annually, with at least 33% of that not managed environmentally safely [1]. In fact, daily waste generated per person ranges widely, from 0.11 to 4.54 kilograms [2]. Furthermore, only 17% of electronic garbage is collected and recycled [3]. Moreover, 32% of plastic packages still need to be managed, which leads to severe implications for ecological balance and human well-being. But, again, garbage separation by the person who disposes of garbage has been widely accepted as ethical behaviour and best practice for reducing, reusing, and recycling [4].

Several existing IoT-based smart garbage systems and the classification using computer vision and artificial intelligence have been conducted to improve household garbage management [5–7] and [8]. However, the existing systems have the

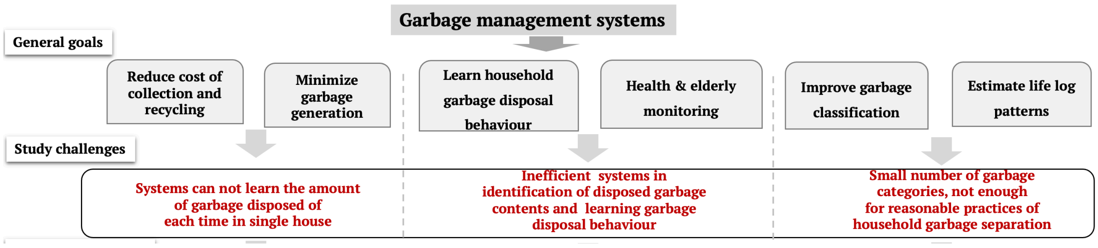


Figure 1.1: Garbage management systems general goals and challenges

following problems: first, they can not learn the amount of garbage disposed of each time; second, they provide a small number of garbage categories, not enough for reasonable practices of household garbage separation; and third, they can not understand the routine behaviour of garbage disposal by households. Fig. 1.1 depicts the general goals of garbage management systems and the challenges to be solved in this study.

To solve the three challenges, this study aims to develop a newly designed and developed smart garbage bin system (SGBS) embedded with multiple sensors to learn the amount of garbage disposed of each time, identify the garbage contents disposed of by households and understand the routine behaviour of garbage disposal by households. Before designing and developing the SGBS, challenges must be conveyed, and solutions must be evaluated for effectiveness.

1.2 Problem statement

To implement the SGBS, we must first describe the requirements for designing the SGBS architecture. Then, the proposed approach must be evaluated to prove its effectiveness in learning the amount of garbage disposed of each time, identifying the garbage contents disposed of by households and understanding the routine behaviour of garbage disposal by households. We determine these challenges to be solved in this dissertation, and we organize them as follows:

Challenge 1: How to learn the amount of garbage disposed of each time and predict garbage growth behaviour for a single house?
To tackle this challenge, we designed and developed a smart garbage bin sys-

tem (SGBS) embedded with distance and weight sensors to track the amount of garbage during disposal. For evaluation, we experimented with the SGBS in a student laboratory for over one month. As a result, an autoregressive integrated moving average (ARIMA) model was applied, providing MAE of 5.17 cm and an SD of 0.33 cm, thus considered satisfactory accuracy on the garbage growth prediction.

Challenge 2: How to understand household garbage disposal behaviour and identify the type of garbage contents?

In order to learn behaviour on garbage disposal and influence behaviour change, it is necessary to consider families of different living styles and sizes to track the amount of garbage produced and garbage content disposal patterns. To solve this challenge, first, we designed and developed SGBS embedded with distance and weight sensors to detect the amount of garbage disposed of each time. Afterwards, we designed and developed a garbage annotations application to allow households user to annotate their daily garbage content. The annotation application comprised 4 garbage categories and 10 garbage content identities. Afterwards, we experimented with the SGBS in three households to evaluate our approach. However, through the user's voice as feedback from the three households, some important garbage content we missing in the garbage annotations application. Therefore to improve the garbage annotation tasks, we redesigned garbage annotation applications to have 8 garbage categories and 25 garbage content identities, and conducted a new experiment with five houses to learn the annotation frequency of deposing different garbage contents. Our findings show that households' garbage disposal behaviour depends on the amount and contents of garbage and the routine of disposing of such garbage content.

Challenge 3: How to substantially improve the automation of garbage classification?

To tackle the challenge, we designed and developed a smart garbage bin system (SGBS) embedded with multiple sensors to identify the garbage contents disposed of. We chose moisture and air quality sensors to identify and distinguish disposed garbage content. Afterwards, we introduce a new garbage content estimation method by training a machine learning model using daily collected fuse sensor readings combined with detailed household garbage contents annotations to perform garbage classification tasks. For evaluation, we deployed the designed SGBS in five households over one month. We confirmed that the leave-one-house cross-validation results showed an accuracy of 91% in 5 kitchen waste contents, 89% in 5 paper/softbox contents, and 85% in 8 garbage categories for the classification tasks.

1.3 Dissertation Organization

This dissertation is organized as follows: we present a review of related literature in Chapter 2. Then, in Chapter 3, we present the smart garbage bin system for learning and predicting garbage growth behaviour for a single house. We then describe the systems requirements, methods and tools needed. Then, we discuss the implementation, the evaluation experiment and the results. In Chapter 4, we extend our SGBS with garbage annotation to guide users during garbage disposal. Then we describe the evaluation experiment and its results to show the effectiveness in understanding household garbage disposal behaviour and identifying the type of garbage contents. Chapter 5 presents a new garbage content estimation method by training a machine learning model using daily collected fuse sensor readings combined with detailed household garbage contents annotations to perform the garbage classification task. Finally, in Chapter 6, we present our conclusions and future work.

2 Related work

In this chapter, we present a review of studies and discuss the concepts related to our study.

2.1 Study 1: Garbage growth prediction

IoT based smart garbage solutions have been implemented at the heart of major cities in the world such as Seoul-Republic of Korea, Varese- Italy, Hong Kong, Barcelona- Spain, Singapore, and Stockholm- Sweden [9–11]. In these cities, smart bins are equipped with sensors that provide users with ability to know the fill-level (volume) of each waste container in real time. These bins are equipped with a live monitoring platform which helps the waste collection staff to plan ahead on how collections should be implemented, targeting only the locations of full garbage bins [11].

Mostly, there are different technological approaches for implementing such application solutions. For instance, studies by Thakker *et al.* [12] and Kumar *et al.* [13] developed a smart and wireless waste management system using a load cell, ultrasonic sensors, and GSM module, which used to notify either the bin is full or emptied. Besides, the work by Talha *et al.* [14] developed a cloud-integrated and wireless waste management system for smart cities involving a combination of infrared (IR), ultrasonic sensor, temperature sensor, (MQ2) gas sensors and load cell in monitoring and storing the information about waste status in a bin.

On the other hand, some approaches, including Chowdhury *et al.* [15], Kumar *et al.* [16], and Papalambrou *et al.* [17] focused on RFID technology where the smart bin embedded with RFID tags, and the collection vehicle is installed with the RFID reader to detect smart bins during waste collection in the city. Likewise, Reis *et al.* [18] introduced iBags using RFID to implement waste reduction and recovery measures. Also, Hong *et al.* [11] used RFID technology to build a smart garbage management system evaluated in Seoul, Korea. Similarly, the stud by Lee *et al.* [10] applied RFID technology in producing an intelligent garbage management system at the Hong Kong Polytechnic University, where users used an NFC card to open the bin. In this system, a verification process is done by pairing up with the user's card, and the waste weight measured by Loadcell is used as a factor for charging the user during waste disposal.

On the other hand, more efforts in improving garbage management with IoT systems have been devoted at the municipal and city levels to predict garbage generation for future planning. For instance, Kannangara *et al.* [19] presented modelling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches. Similarly, Ali *et al.* [20] investigated waste generation in the metropolitan city using an artificial neural network time series model, while Sun *et al.* [21] studied the development of an appropriate model for forecasting municipal solid waste generation in Bangkok. These approaches benefit the city authorities in estimating and allocating essential resources needed in the future for garbage management and formulating alternative strategies to influence the attainability of sustainable goals [22].

Existing approaches have been helping users to conduct real-time monitoring, take data-driven action ahead of time, send a notification for full waste bins, and predict and plan for the best garbage collection route. Yet, they are inefficient in learning the amount of garbage disposed of each time. Therefore chapter 3 of this dissertation proposes to develop a new smart garbage bin system to learn the amount of garbage disposed of and predict growth behaviour.

2.2 Study 2: Understanding household garbage disposal behaviour

Recently, several IoT-based smart garbage management systems have been developed. Most of the existing work invests effort in using the amount of garbage to estimate its future growth and provide dynamic garbage collection. We discussed our related work by considering both aspects found in the existing work.

Table 2.1: A summary of studies used different technologies in the development of IoT-based smart garbage systems

Reference	Wi-Fi	GSM	RFID	ZigBee	Ultrasonic sensor	Infrared sensor	Load cell	DHT11/22	MQ2/135	Camera	ToF
[23]	O	X	X	X	X	O	O	X	X	X	X
[24]	X	O	X	X	X	X	O	X	X	O	X
[16]	X	O	O	X	O	X	X	X	X	X	X
[18]	X	X	O	O	X	X	X	X	O	O	X
[25]	X	O	X	O	O	X	O	X	X	X	X
[11]	X	X	O	X	X	X	X	X	X	X	X
[12]	X	O	X	X	O	X	O	X	X	X	X
[10]	X	X	O	X	O	X	O	X	X	X	X
[14]	X	O	X	X	O	X	X	O	O	X	X
[26]	O	X	X	X	O	X	X	X	X	X	X
[17]	X	X	O	X	O	X	X	X	X	X	X
[27]	X	O	X	X	O	X	X	O	X	X	X
[13]	X	O	X	X	O	X	O	X	X	X	X
[28]	O	X	X	X	O	X	X	X	X	X	X
our system *	O	X	X	X	X	X	O	O	X	X	O

* *O = YES, X = NO*

Kristanto *et al.* [29] proposed a dynamic polling algorithm for low-energy garbage level measurement to eliminate the high cost and inefficiency of the existing static garbage collection systems. The designed smart trash bin prototypes were embedded with an ultrasonic range finder to measure the garbage level. Then a polling algorithm estimated the maximum height of garbage based on historical information on garbage height gathered previously. Moreover, the dynamic polling algorithm was used to reduce the device's power consumption.

Furthermore, Faye *et al.* [30] suggested a novel smart waste management approach for business IoT "SWAM" the system was elaborated in the city of Luxembourg, targeting businesses and large entities. The system used ultrasonic sensors to measure garbage levels in smart bins. Driver mobile data and customer profile were combined, and advise the driver on the best times to visit a customer and collect garbage. Also, the study proposed a multi-objective optimization layer, which compiles the collection routes that minimize the impact on the environment and maximize the service quality. Likewise, Hossain *et al.* [31] demonstrated an optimal route planning model based on Dijkstra's algorithm as one of the city's most important factors in the smart waste management system. Both authors [30], [31] considered the status of the amount of garbage level in a bin as one of the real-life parameters in calculating optimal distance link cost and other parameters such as road congestion status and distance travelled by the driver.

Idwan *et al.* [32] also advance the use of IoT technology to determine the schedule and pathways of waste collection trucks. First, the study simulated multiple route trucks using a heuristic algorithm. Later, developed a smart dumpster equipped with an ultrasonic sensor and GSM module to measure the level of waste and send updates to the central management system using a wireless network. The author asserted that data regarding the garbage status in the developed dumpsters are used to determine the most effective route for the truck, reducing the cost and time taken.

Additionally, Memon*et al.* [33] demonstrated the dynamic features of IoT on human innovation with the remedy of an ever-increasing amount of garbage. In this study, the developed smart garbage bin system was able to record the status of the amount of garbage level periodically in a cloud server and send a report to waste management authorities, thus automatic instruct drivers for the garbage collection on full path bins only.

Ferrer*et al.* [34] presented a software system for predicting fill-level containers, namely BIN-CT (BIN for the city). The study pivoted on paper waste containers due to their variability in collection frequency rather than organic waste. The system combined two main algorithms first for next-day prediction of the containers' fill level based on each container's historical fill level data using machine learning algorithms. Second, it computes the best routes to visit them; BIN-CT prioritized containers with fill levels greater than 80 % for the collection schedule. The predictive system was designed to improve municipal waste collection planning. Similarly, Faye et al. [30] also realized the future strategy to include the predictive models to estimate the fill level at least 48 hours in advance and plan for collection.

Previous studies on IoT-based smart garbage management systems have evidently used garbage amounts to predict garbage growth and provide a dynamic and optimal route for garbage collection to reduce cost and improve garbage operational services in a city. However, they can not understand household garbage disposal behaviour and identify the type of garbage contents disposed of by households, as explored in chapter 4 of this dissertation, which has yet to be considered.

2.3 Study 3: Garbage content estimation model

This Section gives an overview of related work from two different perspectives. First, we provide an overview of the separation and disposal of garbage with an emphasis on municipals in Japan, where this study was conducted. Secondly, we discuss recent work on garbage classification from images using machine learning and deep learning to recall existing approaches to assess it.

2.3.1 Separation and disposal of garbage in Japan

Garbage separation has been a major challenge across developing countries than in developed countries where there are various collection systems for house-separated garbage, such as in Sweden and German [35], China [36], and Japan [37]. While in other developed countries, garbage separation is often classified into three categories: recyclable, household and vegetation garbage. In Japan, the garbage separation and disposal system is different and complex. The rules for separating and disposing garbage depend on the particular local municipality, whereby each city in Japan provides a well-documented pamphlet explaining the garbage disposal rules. In general, garbage is divided into four categories: Burnable garbage (Kitchen waste, paper scraps, clothing, etc.), non-burnable garbage (Metal, glass, ceramics and pottery, etc.), recyclable (Plastic bottles, container jars, cans, newspapers, etc.), and oversized (Large furniture, etc.) [37]. Therefore, each municipality uses such a general garbage division to classify garbage for their residents. (see burbalegarbage) provides an overview of the division of burnable garbage content in four cities in Japan; Kashihara [38], Ikoma [39], Nara [40] and Kyoto [41]. Apart from garbage descriptions from the municipal pamphlets, residents use designated plastic garbage bags of up to 45 litres to dispose of garbage. Moreover, garbage collection for each category of garbage is set by the municipal for instance, Mondays and Thursdays in Ikoma city [39] are used for the collection of burnable garbage only. The above facts show that families in Japan play a hand role in their municipal rules for garbage separation and disposal systems. However, The failure of households to sort the garbage renders the whole system useless [7]. Therefore, automation tools are necessary to monitor daily family garbage disposal and improve garbage separation.

Table 2.2: Overview of burnable garbage separation in Japan

City Name	Burnable garbage separation
Kashihara city [38]	1:Kitchen scraps
	2:Small plastic e.g DVDs/CDs,toys
	3:Waste paper (tissue, mixed papers)
	4:Weeds, twigs, and leave (30 cm)
	5:Containers e.g for mayonnaise, oil
	6:Polystyrene
	7:Shoes,Bag, Cloth
Ikoma city [39]	1:Kitchen waste
	2:Plastic/unclean products e.g storage containers
	3:Paper/textile e.g Tissues,Milk box, cloth
	4:Others
Nara city [40]	1:Kitchen garbage
	2:Styrofoam e.g noodle cup
	3:Cassette tape/videotape (less than five)
	4:Waste Woodchips e.g chopsticks, pencils
	5:Waste Paper
Kyoto city [41]	1:Raw garbage
	2:Non-recyclable paper
	3:Broken glasses
	4:Small plastics

2.3.2 Garbage classification from images with deep learning models

A possible solution to overcome the existing challenges in household garbage separation and management is to adopt sustainable automation tools to improve garbage separation. Presently, several works have been devoted to the automation and detection of garbage from images, which has now become a popular choice to replace manual garbage separation while taking advantage of the rapid advances in computer vision and artificial intelligence. Various standard CNN architectures have been recently proposed to perform image classification tasks with high accuracies, such as VGGNet [42], AlexNet [43], ResNet [44] and DenseNet [45].

Nnamoko*et al.* [5] investigated the problem of manual household garbage separation into two categories, namely, organic and recyclable. Experiments presented in this paper were conducted with Sekar's waste classification image dataset available in the Kaggle library [46]. Later, a bespoke 5-layer CNN architecture was used to perform image classification tasks. In this work, the training was conducted on two datasets, smaller model (80x45 pixels) and a larger model (225x264 pixels), for performance comparison, thus obtaining similar cross-validation accuracy of 79%. Likewise, Mookkaiah*et al.* [47] proposed a model to identify and classify two types of garbage, biodegradable and non-biodegradable. First, the images were collected in the respective garbage bin by Raspberry Pi Camera Module v2. Then garbage classification task was done by CNN architecture. However, separating garbage into two categories is insufficient for logical household garbage separation. Besides, there is still a shortage of publicly available garbage image datasets and an information gap in their experimental procedures.

Furthermore, Wang*et al.* [7] revealed garbage sorting and classification at the source, the beginning of garbage collection while utilizing the combined method of IoT and CNN. The study used experimental data available in the Trashnet [48] dataset, merged with other datasets thus, resulted in nine categories of garbage (Kitchen waste, other waste, hazardous waste, plastic, glass, paper or cardboard, metal, fabric and other recyclable waste). In addition, the study developed an intelligent bin embedded with ultrasonic sensors, MQ9, and MQ135 gas sensors to monitor the garbage's running state in the bin. Finally, the CNN model was deployed in mobile phones and cloud computing servers for garbage classification. The system required citizens to take pictures of garbage using their mobile phones and send them to a cloud server to run the deep-learning algorithm to recognize categories. Despite the high-performance accuracies of 92.44% and 92.00% achieved by Xception and MobileNetV3 models on classifying nine types of garbages, the author presented more generalizable garbage categories that need to be improved for proper household garbage separation.

Besides, a distributed architecture for smart recycling using machine learning was realized by Ziouzios*et al.* [6] as a solution for garbage classification in collection facilities to solve the problem of non-segregated garbage, which exists more in developing and developed countries. The Trashnet [48] dataset was used for training the models by utilizing computation offloading to the cloud. The CNN architecture classified the garbage materials into five categories: Paper, glass, plastic, metal, carton, and trash. Similarly, Sami*et al.* [49] used the Trashnet [48] dataset to automate the garbage classification problem into six classes: glass, paper, metal, cardboard, and trash using a Support Vector Machine, Random Forest, Decision tree, and CNN to find the optimal algorithm that best fits garbage classification solution. However, the available public garbage image datasets need more classes of garbage categories for proper garbage classification. Therefore, the garbage categories presented in both studies [6, 49] are not practical for household garbage separation and for improving the garbage management systems.

Despite the high accuracies achieved by the existing solutions on garbage classification through the automation and detection of garbage from images by the deep learning models. Yet, they provide a small number of garbage categories, not enough for reasonable practices of household garbage separation. Therefore, to our knowledge, an automation tool that can learn and identify the daily garbage content disposed of in homes and perform classification tasks, as investigated throughout this work, has yet to be considered.

2.3.3 Key issues addressed in existing IoT garbage management system

This sub-section summarises the issues addressed in the existing IoT garbage management system as described in the literature and specific questions in each study challenge to investigate in this dissertation.

Study 1

- In the first study of garbage growth predictions, the existing IoT In the first study of garbage growth predictions, the existing IoT garbage management systems have been exclusively helping users conduct real-time monitoring, taking data-driven action ahead of time. Also, send a notification for full garbage bins, and predict and plan the best garbage collection route. However, such existing systems are city-based and unfriendly in learning garbage growth behaviour for a single house at different times of the day, a week, or a month.

  Study Challenge 1: How to learn the amount of garbage disposed of each time and predict garbage growth behaviour for a single house?

  Specific questions

  1. How to learn the amount of garbage disposed of each time?
  2. How to predict garbage growth behaviour?

  Study 2

- In the second study of understanding household garbage disposal behaviour. Literature has proven that IoT garbage management systems can use garbage amounts to predict garbage growth and provide dynamic and optimal routes for garbage collection to reduce costs and improve garbage operational services. However, IoT garbage management systems are inefficient in identifying the type of garbage contents disposed of daily by households and learning disposal patterns.

  Study Challenge 2: How to understand household garbage disposal behaviour and identify the type of garbage contents?

  Specific questions

1. How to understand household garbage disposal behaviour?
2. How to identify the type of garbage contents?
3. What are the methods to save energy on IoT systems?

Study 3

- In the third study of the garbage content estimation model. Studies have been dedicated to the automation and classification of garbage from images by utilizing computer vision and artificial intelligence. However, they provide a small number of garbage categories, not enough for reasonable practices of household garbage.

  Study Challenge 3: How to substantially improve the automation of garbage classification?
  Specific questions

  1. What sensor values are relevant to the identification of garbage content?
  2. How to perform garbage content estimation from daily disposed of garbage content?

3 Smart garbage bin: Garbage growth prediction

This chapter lays the groundwork for the first challenge of garbage growth predictions for a single house. First, we describe the requirements needed to answer the main study questions. Then we design the proposed smart garbage bin architecture and discuss the tools and methods required to develop a smart garbage bin prototype based on the requirements. Afterwards, we build a model for garbage growth prediction and evaluate its effectiveness.

3.1 Methods and tools

3.1.1 System requirements

This section describes the system requirements for the proposed customized smart garbage bin system. Based on the discussions in Chapters 1 and 2, there are the following two main requirements for a smart bin system:

Req 1: It should be able to detect specific garbage values.
Req 2: It should upload the detected garbage value into the specified cloud storage at a defined programming time interval via a wireless gateway and ensure real-time data visualization.
Req 3: It should be able to analyze data and predict future garbage growth.

To achieve Req 1 and Req 2, we designed and developed a smart garbage bin system (SGBS) embedded with sensors for distance, weight, temperature and humidity sensors to monitor garbage growth. Also, we selected the Wi-Fi network as a gateway to the cloud server for real-time data visualization. To address Req 3,

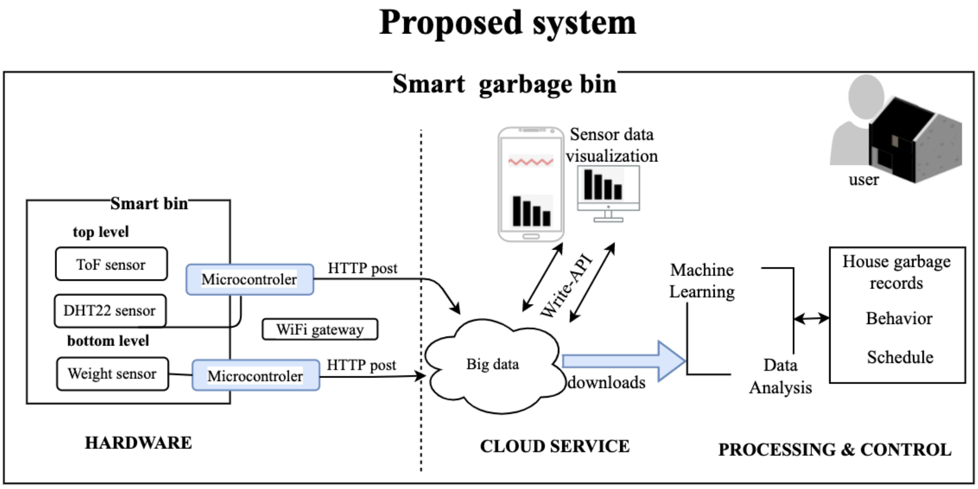


Figure 3.1: Smart garbage system architecture design

we apply a time-series machine learning algorithm for continuous data to predict future garbage growth.

3.2 Proposed smart garbage system

Fig. 3.1 shows the proposed SGBS architecture design comprises three layers: Hardware, Cloud service, and Processing and control. First, the SGBS collects its status data and sends it to the cloud platform via a gateway. Later, machine learning methods are applied for data analysis and prediction. Fig. 3.2 shows an overview of the proposed smart garbage bin system. The following subsections details each part of the proposed SGBS.

Hardware layer

The Hardware layer comprises the hardware used in developing the proposed SGBS Fig. 3.1. An Adafruit feather m0 Wi-Fi atsamd21 + atwinc1500 Micro-controller is placed in the heart of the system connected to the sensors. The Adafruit feather m0 has a built-in Wi-Fi module, hence providing the system's ubiquity and ease of setup. We considered exploring other related studies that used different technologies in developing IoT-based smart garbage systems, as

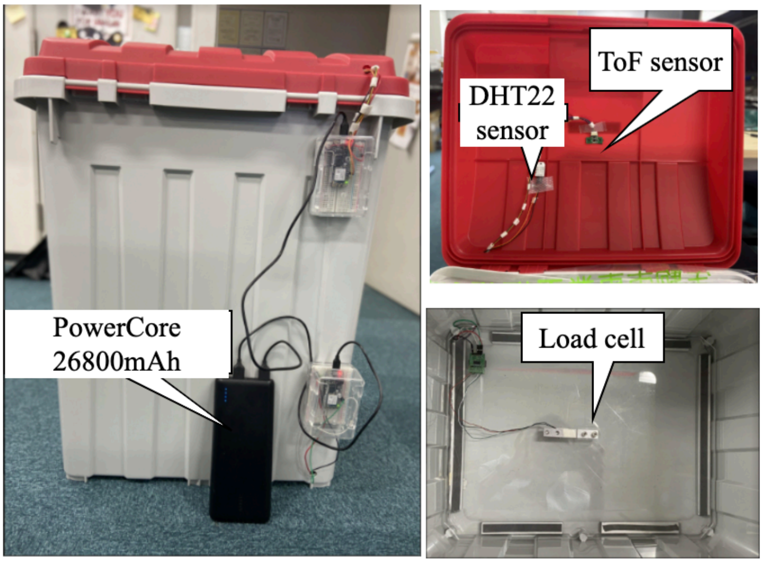


Figure 3.2: Smart garbage bin overview

shown in the Table 2.1. Communication technologies such as RFID, GSM, Wi-Fi, and ZigBee enabled garbage data transfer for real-time monitoring. Besides, sensor technologies such as ultrasonic, infrared, load cell, and DHT22, provided the measurements of the garbage status. The review from other related studies found that the ToF sensor has not yet been commonly considered in developing an IoT-based smart garbage system. In contrast, ultrasonic and infrared sensors are found to be popular. Thus, we ultimately chose the time of flight (ToF) sensor in this study. The ToF sensor is not affected by the colour of the target object compared to the infrared sensor. Also, compared to the ultrasonic sensor, the ToF sensor does not critically depend on the angle of incidence and is not disturbed by environmental noise; thus, it has greater readings and accuracy. Therefore, the smart bin cover in the proposed system is embedded with the ToF, DHT22 (temperature, and humidity) sensors (Fig. 3.2. The ToF sensors measure the increase of garbage fill levels in a smart garbage bin in the centimetre unit of measurements plus DHT22, which measures the inside temperature and humidity condition of garbage in the bin. The temperature and humidity need to be monitored because the garbage may decompose and produce a pungent smell. Further, the bottom part of the smart garbage bin Fig. 3.1 comprises the load cell to detect the increase of the garbage weight in the bin. Power Core 26800mAh Anker external battery supplies sufficient power to the system. Arduino IDE software was used as the programming environment for the sensors. Table 5.1 shows the type of sensor used and its purpose.

Table 3.1: Sensor used in the developemnt of smart garbage bin

Sensor	Purpose
1. AE-VL53L0X(ToF)	Measure fill level of garbage in a bin
2. DHT22	Measure temperature and humidity in a bin
3. Load cell	Measure increase of weight of garbage in a bin

Cloud service layer

To achieve the real-time monitoring and data visualization as defined in Req2 of the proposed SGBS, we utilized a ThingSpeak cloud platform. ThingSpeak is an open-source cloud platform that provides cloud space for IoT projects. Therefore, using the Wi-Fi gateway, the detected garbage data sensor data found in the smart garbage bin were continually uploaded, stored, and visualized into the Thingspeak cloud space with the Write-Application Programming Interface *(W-API)*. In addition, a ThingView mobile application linked to Thingspeak via *(W-API)* allowed easy data visualization through a smartphone in real-time garbage monitoring.

Processing and control layer

The processing and control layer supports the achievement of the second requirement of the smart garbage bin system, where a time-series machine learning algorithm for continuous data named an autoregressive integrated moving average (ARIMA) was applied. Fig. 3.3 is a flow chart of the predictive model for a single house's future garbage growth behaviour. The flow chart includes data preparation, machine learning model building, performance measurement, and the model's deployment.

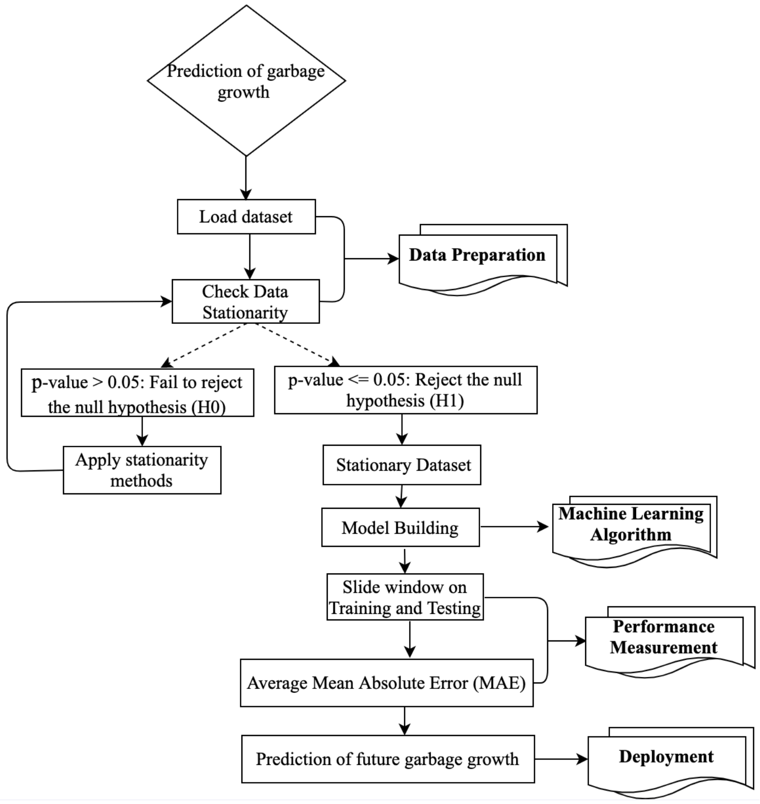


Figure 3.3: The flow chart of modelling steps used in this study

3.2.1 Deployment and data collection experiment one

In evaluating the performance of the developed smart garbage bin see Fig. 3.2, we explored a way of successfully utilizing the SGBS in terms of the number of days *(times)*, type of user, and garbage growth. Therefore for SGBS validation and feasibility, key questions were studied (1) How is a big data collection of garbage growth of a single house being conducted? (2) How to learn the patterns of garbage growth? (3) How to predict garbage growth for the following schedule and changing garbage bags? To respond to the questions above, we conducted experiments number 1 as a preliminary deployment of the smart garbage bin system, (SGBS) in a university laboratory consisting of 42 research students who use the laboratory daily. As a routine, students visit the laboratory from Mondays to Fridays, on or after morning to night hours, while few visit the lab on Saturdays and Sundays. At these times, they do different activities, including eating, drinking, and cleaning, thus producing garbage. The lab is placed with three types of garbage bins: burnable, cans, and plastics. However, for this study, the focus stood on burnable garbage only. Burnable garbage includes food waste, paper waste, fruit, vegetable peel, eggshells, old clothes, and other food items that may leave an unpleasant odour if left in the bin for too long.

3.3 Model building

To model continuous time series data using an autoregressive integrated moving average (ARIMA), first, we need to check the stationarity of the observation, which can be used in the feature selection process on the time series problem. Therefore, we applied an augmented statistical dickey-fuller (ADF) test to test the stationarity on our dataset with a 2-day rolling window. Then, using the p-value of the ADF as shown in Fig. 3.4, we interpreted the results. The time series dataset is considered stationary if the p-value is $p \leq 0.05$, and the critical values at 1%, 5% and 10% confidence intervals are as close as possible to the ADF statistics. In our test, the ADF test gave the p-value of 0.006, and the test statistic was less than the 1% critical value. Thus, it suggests we can reject the null hypothesis with a significance level of less than 1%; therefore, the dataset is stationary. Therefore, we used the ARIMA model to predict future garbage growth using a fixed-sized slide forecast window on the train and test the model. ARIMA is a popular algorithm widely used in the statistical method for continuous-time series forecasting [50]. The ARIMA model consists of three components: autoregression (AR), integrated (I), and moving average (MA), which is explicitly specified in the model as a parameter like ARIMA*(p, d, q)*. An autocorrelation function (ACF) provides the MA value, and a partial auto-correlation function (PACF) provides the AR. The Akaike information criterion (AIC) value allows us to compare how well the model fits the data. The lower the value, the better the model. Therefore, we built our model with the ARIMA *(2, 1, 0)* with the AIC value of *205*.

```
Results of Dickey-Fuller Test:
Test Statistic                  -3.548993
p-value                          0.006814
#Lags Used                       7.000000
Number of Observations Used     21.000000
Critical Value (1%)             -3.788386
Critical Value (5%)             -3.013098
Critical Value (10%)            -2.646397
dtype: float64
```

Figure 3.4: Dickey-Fuller test statistic results

3.4 Results and discussion

3.4.1 Cloud data visualization

In our preliminary deployment of a smart garbage bin in a university laboratory, data from three sensors, ToF, DTH22, and load cell, were continuously collected and stored in the ThingSpeak cloud over 30 days at a 1-minute interval. For smooth live streaming and garbage data monitoring, a ThingView mobile application shown in Fig. 3.5 linked to the ThingSpeak cloud storage via *W-API* provided good data visualization.

3.4.2 Daily garbage growth

During the deployment, we studied the garbage's growth using the fill-level values sensed by the ToF sensor from the top level of the smart garbage bin in a time series interval of each day. As a reference, Fig. 3.6 reveals both slow and high variations of the garbage growth behavior, which depended on the use of smart garbage bin on a particular day by the students in the laboratory. Further, we determined the trended frequency of changing the garbage bags on different peak values as Fig. 3.6 illustrates, whereas, the small green boxes indicate the lowest peak value *(5.6 cm)* from the bottom when the smart bin was empty. The dark red small boxes indicate the highest peak values of garbage where the change of garbage bag occurred, such trend frequency of changing the garbage bag learned as irregular behavior in the smart bin system. The change of the garbage bag can also be due to bad smell resulting from decomposed garbage in the smart garbage bin.

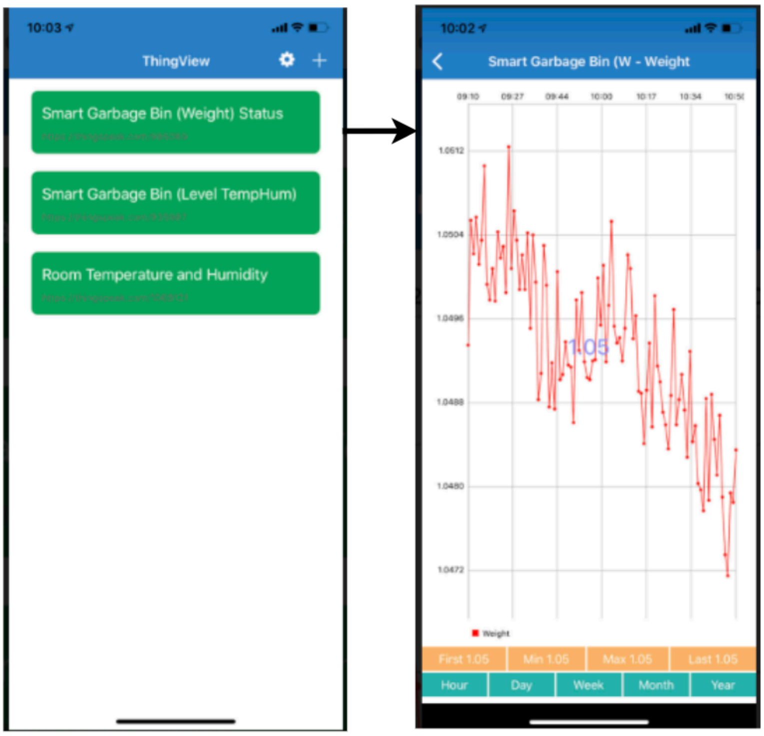


Figure 3.5: Daily garbage growth data visualization using ThingView app

3.4.3 Garbage growth prediction

To predict the future garbage growth for the following collection schedule and the change of garbage bags. We applied the ARIMA model using a sliding window forecast method to train the model on the fill-level dataset, consisting of 30 days of observations as provided by the ToF-level sensor. Initially, we started by splitting the whole number of observations into different training sizes. Each train size takes a given forecast window size as an input for testing and predicting **(Forecast)** future garbage growth behaviour. As shown in Table 5.4, the number of observations (N) was split into ten, fifteen, and twenty days as training size. Therefore, a fixed forecast slide window size of 2, 4 or 7 days was applied to each N-number of observations to forecast the garbage growth. Fig. 3.7 illustrate the prediction of garbage growth on training the ARIMA model with the ten, fifteen, and twenty N-number of observations. The result shows that the predicted garbage values follow the actual values of the initial observations. Additionally,

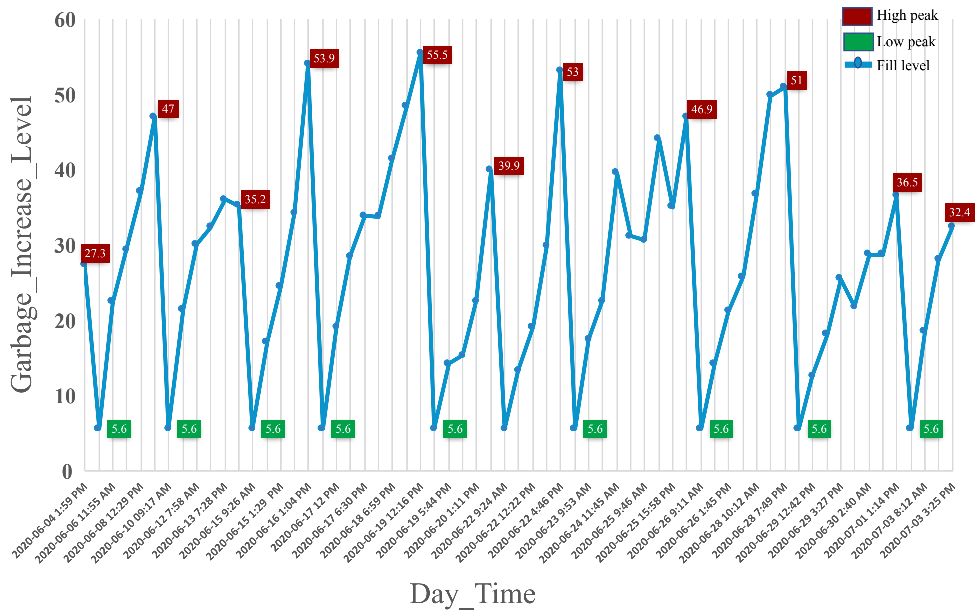


Figure 3.6: Garbage growth and frequency of change of a bag in the smart bin

Fig. 3.8 demonstrates prediction outcomes of garbage growth on each training size; N=10, N=15, N=20 with the given forecast window size. We have observed that the ARIMA model is suitable for predicting future garbage growth behaviour in the single house because the prediction follows the actual observations and can provide predictions with few amounts of data; also, the prediction model was capable of gathering the fluctuations on observed data, and the averaged accuracy error decreases with respect to the simultaneous increasing forecast window size and training size. Moreover, the ARIMA model is a flexible method which uses past data to predict the future where its application does not require much data. Thus, in this study, the ARIMA model using the slide forecast window provided flexibility and functional result.

3.4.4 Performance measurements

Given the slide forecast window, we calculated the model's error on each training size (N-number of observations). There are different ways to evaluate the performance accuracy of the ARIMA model. In that regard, we conducted a performance measurement of our model using mean absolute error (MAE) and standard deviation (SD) in the centimetre unit of measurement (cm). As shown in Table 5.4, the forecast window started with a two-days size, then slid to four-day and, finally, seven-day window sizes on the same training size; N=10, N=15, N=20. Therefore, the performance was observed and compared from a few to higher numbers of observations (N). We obtained the model performance accuracy using three training iterations; thus, MAE was recorded and averaged. Fig. 3.9 is an error bar graph achieved during model performance measurements. Training the model with ten observations (N=10) using the two-days window size indicates satisfactory performance, but the number of observations was low. Thus, training the model with twenty observations (N=20) given with the four-days window size, as shown in Table 5.4, provided the best accuracy on the garbage growth prediction, i.e. N=20, Average MAE=5.17 cm, SD=0.33 cm. The results show that the simultaneous increase of both the training and forecast window sizes provides fewer errors. The fewer prediction errors during the ARIMA model's performance measurement indicates the predicted values are closer to the actual observation, which offers high efficiency in predicting the future amount of garbage growth behaviour on daily practical use. In contrast, the higher error value above 10% of the smart bin's maximum fill level, which the model best achieved, can impact the timing of garbage bag change during garbage disposal and garbage collection schedule.

3.5 Chapter summary

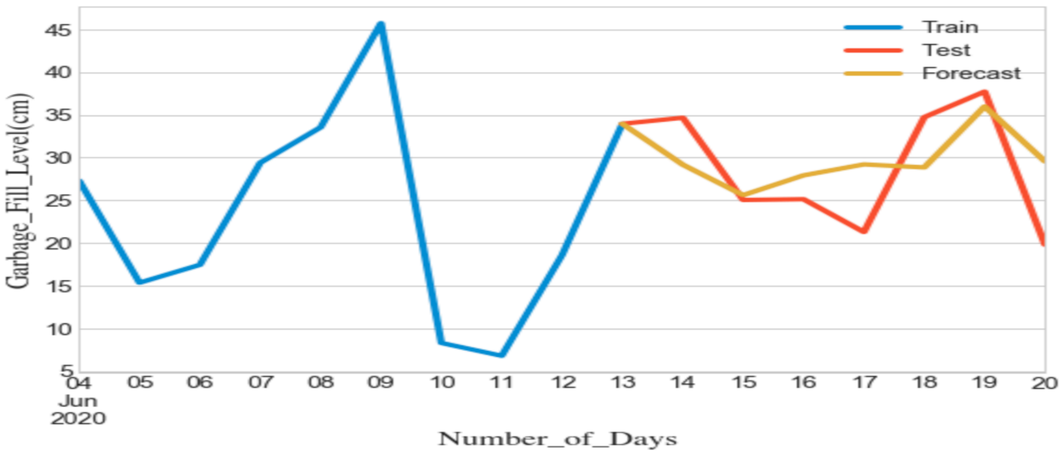
In this chapter, we have presented the smart garbage bin system (SBGS) to learn garbage growth behaviour in a single house and predict its growth. First, the SGBS embedded with distance, loadcell, temperature and humidity sensors via Wi-Fi gateway were collected and stored in the cloud platform. Then we conducted a preliminary deployment of the SGBS in a student laboratory over one month. Later, we applied an ARIMA model with a fixed slide forecast window size on each N-number of observations to predict the garbage growth behaviour for the collection schedule and change the garbage bag during the day of garbage disposal. The result found that the ARIMA model was suitable for predicting

Table 3.2: ARIMA model performance measurement

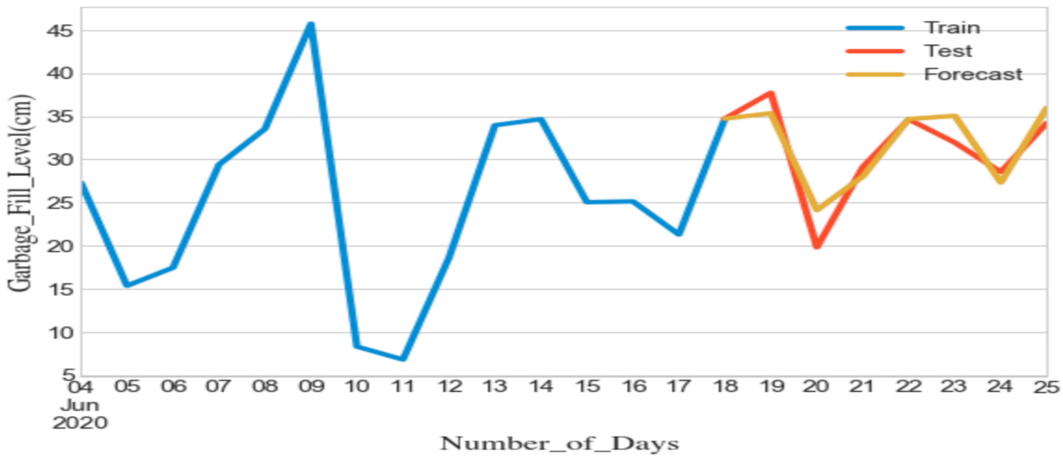
Training size (N) (day)	**Forecast window size** (day)	**Average MAE** (cm)	**SD** (cm)
10	2	3.67	0.75
10	4	4.63	0.76
10	7	5.34	0.62
15	2	7.54	0.18
15	4	6.11	1.42
15	7	5.56	1.19
20	2	5.08	3.24
20∗	4	5.17	0.33
20	7	5.34	0.55

∗ *is considered as the best performance accuracy*

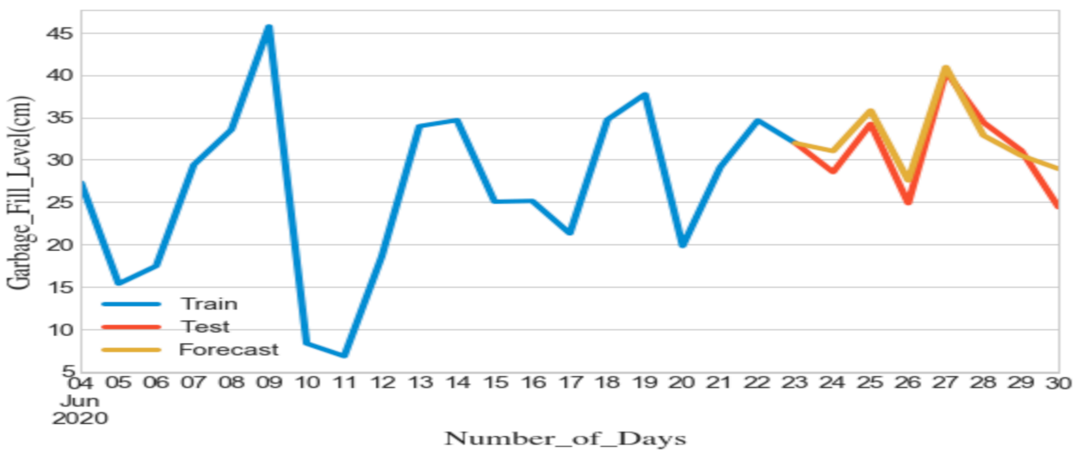
garbage growth behaviour in a single house. Moreover, the model was capable of gathering the fluctuations in observed data, and its application did not require much data. However, to build a generalizable (versatile) model first we need to investigate if the ARIMA can be used as the versatile model. Nonetheless, in this Chapter we have confirmed that the ARIMA model can be used for the garbage growth behaviour prediction in a single house. In the next chapter of this dissertation, we will investigate the routine of garbage disposal behaviour in houses and identify the disposed of garbage contents.



(a)



(b)



(c)

Figure 3.7: Garbage growth prediction with ARIMA model using:(a) 10 days training (a) 15 days training (a) 20 days training

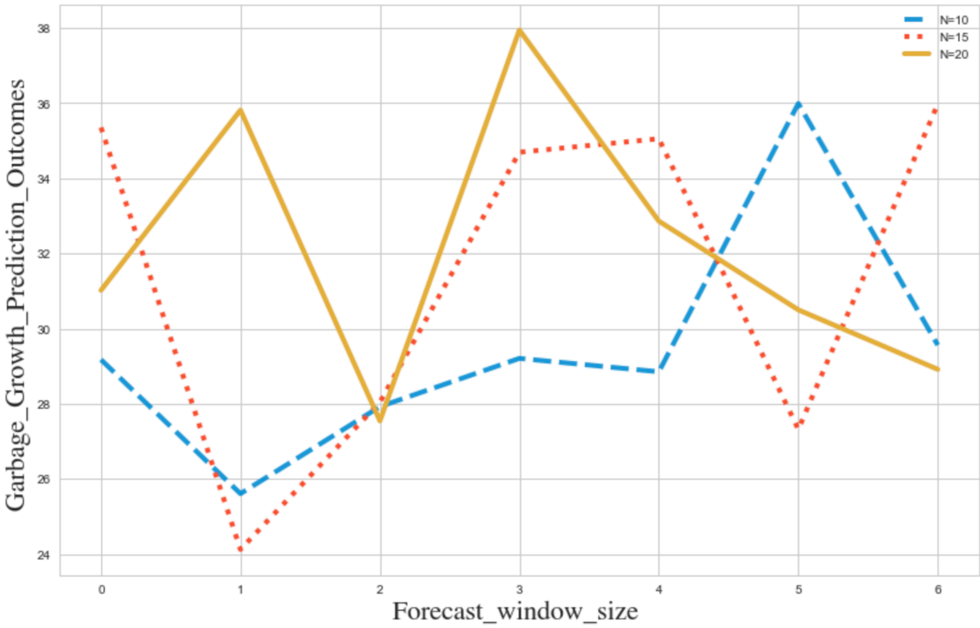


Figure 3.8: Prediction outcomes of garbage growth on different N-number of observation

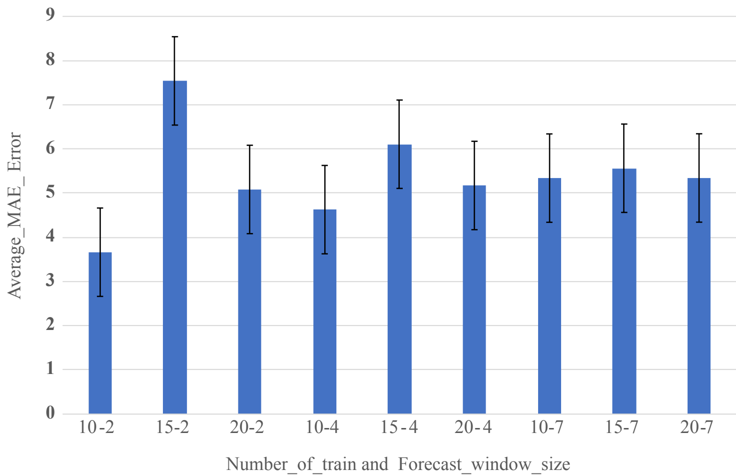


Figure 3.9: An error bar graph during model performance measurement

4 Smart garbage bin: Understanding household garbage disposal behaviour and content identification

In this following chapter of the study, we respond to the second challenge of household behaviour on garbage disposal and identification of important garbage content disposed of daily in households. Therefore, we begin by describing the requirements needed for designing and developing a smart garbage bin system prototype and constructing an architectural design for the systems while considering methods and tools that will save the purpose of low power and wide range data transfer of the developed system. Eventually, we realize garbage disposal behaviour and identification of garbage contents daily disposed of in families.

4.1 Methods and tools

4.1.1 System requirements

In this section, we identified and described the requirements for the proposed Smart garbage bin system "SGBS" as follows;

1. It should detect the amount of garbage, moisture condition and air quality in the smart bin.
2. It should be able to provide desirable low energy usage on devices

3. It should be capable of transfer the detected data into a cloud server at a long-range.

4. It should identify the type of garbage contents and guide users on how to dispose of garbage.

5. It should allow safe installation and replacement of smartphone.

6. It should allow household users to clean their smart bins for healthier living.

Table 4.1: Sensor properties and purposes

	Sensor device	**Max current (mA)**	**Voltage (V)**	**Purpose**
1.	DHT22	2.5	$3 \sim 5$	Temperature and humidity
2.	ToF	10	$2.8 \sim 5$	Filling level
3.	Load cell and HX711	1.6	$2.6 \sim 5.5$	Weight
4.	CCS811 Air quality sensor	26	$3 \sim 5$	TVOC and CO2

In order to achieve requirements 1, 2, and 3 for the SGBS we selected hardware devices, where in the heart of the proposed SGBS, an ATmega328-Arduino Pro Mini microcontroller was used. The ATmega328 uses a 3.3V, 8 MHz and consumes 16mA before the sleep mode state; this is suitable for implementing a low-energy smart garbage bin. Also, we selected the Low Power Wide Area Network sigfox antenna module programmed with the ATmega328 to transfer small chunks of measured garbage data into a specified cloud data service. In addition, we selected small, low power and lightweight sensors for determining the state of garbage amount (fill level and weight), moisture and air quality (Total Volatile Organic Compounds (TVOCs) and Equivalent CO2 (eCO2)). Further, to achieve requirements 4 and 5, we designed and developed a garbage annotation application to guide users during garbage disposal. The annotation application was installed in Google Pixel 3a 64 GB simple smartphone with Wi-Fi capability; thus, the annotations of daily garbage contents were stored in a cloud server. Finally, for requirement 6, we designed our system to be easily installed in houses, allowing household users to clean their smart garbage bins healthily. Table 5.1 provides the purpose and power consumption properties of the chosen sensors used in developing the smart garbage bin.

4.1.2 Architecture design

Fig. 4.1 demonstrates the high-level architecture of the SGBS consists of three primary services. First is the smart garbage bin operation service, embedded with ToF (time of flight), DHT22 (temperature and humidity), HX711-load cell, CCS811 air quality sensor and a solar panel battery. We Used a sigfox antenna module as a gateway to send garbage data to the cloud server through API calls. In addition, the smart garbage bin operation service consists of the garbage annotation application installed on smartphones that appropriately guides users in disposing of their garbage in the smart garbage bin. Secondly, the Cloud data service collects, stores and processes all sensor data from the smart garbage bin. Thirdly, the garbage log service "GLS" comprises the garbage amount information and disposed contents from households to learn the garbage disposal behaviour of households which is the focus of this study.

4.1.3 System design

This subsection describes the design of the SGBS. Fig. 4.2 illustrates the design of the Smart garbage bin system. The SGBS design consists of four components:

- ***Smart garbage bin:*** Where sensors, solar panel battery, sigfox antenna, and smartphones are attached for garbage data collection.
- ***Sigfox antenna module:*** That uses low-power ultra narrowband and sends data to the backend of the sigfox cloud.
- ***Sigfox cloud:*** That uses a custom callback service type to integrate with google cloud API through the created google sheet URL to receive the bytes of data via the sigfox antenna module.

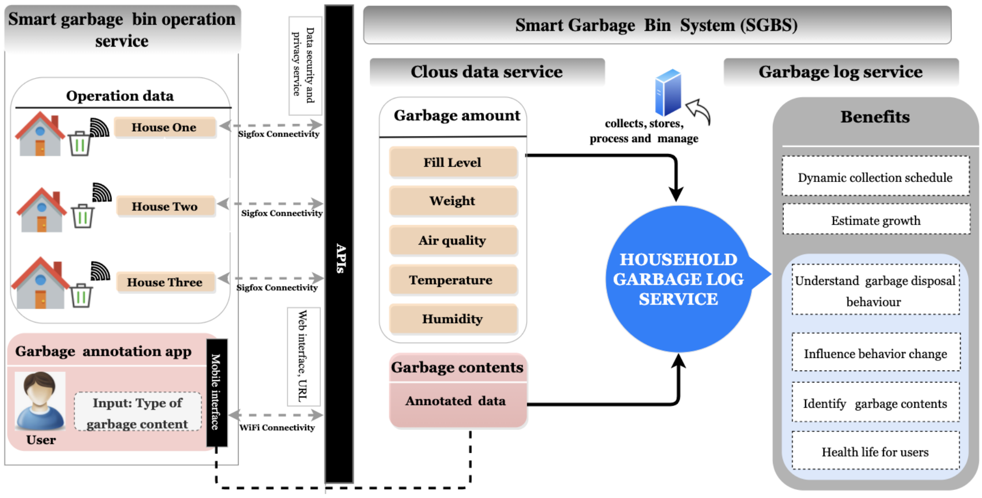


Figure 4.1: High level architecture of smart garbage bin system

- ***Google sheet cloud service:*** That receives data from the sigfox cloud using the *doGet* parameter function, storing and processing all sensor data from the smart garbage bin that are sent in a defined system time interval.

4.1.4 Sigfox as the enabling data communication infrastructure for SGBS

Sigfox network is part of the Low Power Wide Area Network(LPWAN) with an ultra-narrowband technology that uses a standard radio transmission method called binary phase-shift keying (BPSK). Sigfox operates in unlicensed bands worldwide, with radio frequencies of 868 to 869 MHz and 902 to 928 MHz using a data rate of 100bps to 600 bps [51] depending on the region. Sigfox network is employed mainly for developing IoT more reliable than Wi-Fi because it can handle a data transmission without object obstruction to 6km/h from the installed location. Conversely, Wi-Fi technology standards 802.11abgn involve short-range, high cost and high power consumption. As mentioned in subsection 4.1.1 of the requirements, we wanted to achieve low energy usage on the proposed SGBS. Therefore, we installed a Kit breakout board sigfox BRKWS01 (RC3+915Mhz) antenna in this study. The registered antenna contains a low-cost one-year contract for the network provision and subscription. With the sigfox module, the

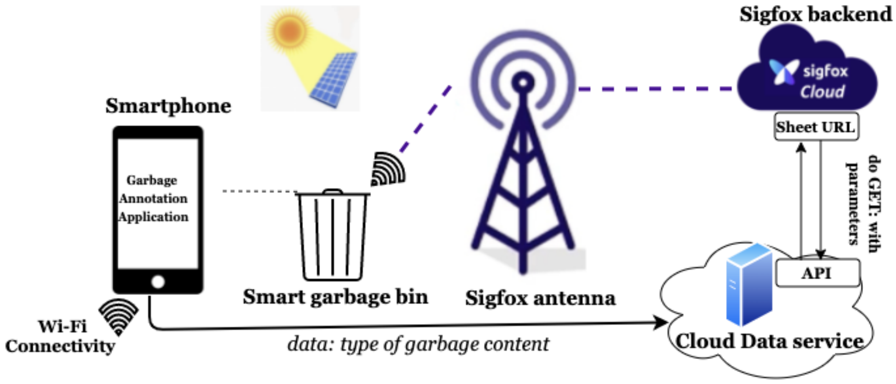


Figure 4.2: Smart garbage bin system design

devices can send a maximum of 6 messages per hour (36/6) in a 10-minute duty cycle, which means 144 messages per day. So we used this scenario to define the SGBS time interval for sleep and active mode. This way, the SGBS sends new garbage data into the cloud service using the designated 10-minute break continuously in a day. To receive data, we used the custom call back service at the back end of sigfox cloud to integrate with the google sheet cloud server through google API, whereas data are received with the help of the *doGet* function query.

4.2 Energy saving algorithm

Herein, an energy-saving algorithm is proposed and described to reduce power consumption on the smart garbage bin. We assume the contents of the garbage differ from one household to another in a day. Thus, the proposed SGBS has to record such patterns without missing the user's important garbage disposal behaviour in specified time intervals. Therefore, as detailed below, the energy-saving algorithm operates in active and sleep modes to collect information about garbage contents.

Active mode

During the active mode, garbage contents are measured using the embedded sensors. First, the smart garbage bin starts by waking up for two seconds, and all the sensors measure the available garbage values. Because the frequency of measurement directly affects power consumption, it is important to reduce the system's active time as much as possible. Next, the measured garbage values are uploaded as bytes of a message into the cloud server via the sigfox antenna module. This process requires a delay time of about 1.5 seconds before the system enters sleep mode.

Sleep mode

In sleep mode, we achieved a low energy usage of the smart garbage bin. During this state, the ATmega328-Arduino Pro Mini microcontroller enters sleep for 8 minutes. We used a Low Power Mode (LPM) from Arduino's lower power library to set inactive all functions that consume power to run in a microcontroller, including; Timer 0, 1, 2, SPI and UART Communication, and External Oscillator. This method results in low usage of power in smart bin devices. To avoid confusion on the solar battery and consistently power the system devices using this algorithm, we introduced a 1-minute wake delay in the energy-saving duty cycle. Thus, the energy-saving algorithm efficiently reduces power consumption on the proposed SGBS.

4.2.1 Evaluation of power consumption

In this subsection, we introduce a method for measuring the smart garbage bin's power consumption to confirm its low energy attributes, as discussed in the previous section. The proposed SGBS aims to track garbage amounts, including; garbage level, weight, moisture, and air quality. Thus, the energy-saving algorithm was used to increase the lifetime of sensor devices. Using Cen-Tech digital multimeters, we measured the total current consumption by the smart garbage bin during the active mode, when it performs measurements and communicates the data to the cloud server. Also, we measured the power when the smart garbage bin was in sleep mode.

Table 4.2: Power consumption characteristics

Mode		Current consuption
Active mode	Wakeup current	50.1 mA
	Wakeup time	3.5 sec
Sleep mode	Sleep current	38 mA
	Sleep time	9 min

Power Measurement method

To calculate the power measurement by the SGB, we measured the total current consumption by the devices over time (3,600 seconds in one hour). As discussed above, we divided the measurement into two ways; during measurements (Active mode) and in sleep mode. Table 4.2 recorded the power consumption characteristics of the SGBS during active mode and sleep mode.

Results of power consumption

By referring to Table 4.2 in active mode, the SGB consumes 50.1 mA as a wakeup current in a total of 3.5 seconds, six times every hour. If it is expressed into milliamp-hours (mAh) units, the power consumption of the smart garbage bin can be calculated as follows;

$$50.1\,\mathrm{mA} \times (3.5 \text{ sec} \times 6 \,/3600 \text{ sec}) = 2.9 \times 10^{-1}\,\mathrm{mAh} \tag{4.1}$$

Again from Table 4.2 during the system sleep mode, the SGB consumes 38 mA as a sleep current in a total of 9 minutes, equal to 540 seconds. If it is expressed into milliamp-hours (mAh) units, the power consumption of the SGB can be calculated as follows;

$$38\,\mathrm{mA} \times (540 \text{ sec} \times 6 \,/3600 \text{ sec}) = 34.2\,\mathrm{mAh} \tag{4.2}$$

Estimate Battery life

To confirm the lifetime of smart bin devices using our energy-saving algorithm, we estimated the battery life of the solar panel battery powering the SGB. We used a 26,800 mAh solar panel battery to power the SGB. Therefore, we used power consumption characteristics as demonstrated in Table 4.2 and assumed the percentage wasted capacity of the solar battery is less than 5%. Thus, the battery life of the SGB was calculated as follows;

$$\begin{aligned}
\text{Average current } &= \frac{(38\,\text{mA} \times 540 \text{ sec}) + (50.1\,\text{mA} \times 3 \text{ sec})}{(540 \text{ sec} + 3.5 \text{ sec})} \\
&= 38.0779\,\text{mA} \\
\text{Battery life time } &= \frac{26,800\,\text{mAh} \times 0.95}{38.0779\,\text{mA}} \\
&= \frac{668.63\,\text{h}}{24\,\text{h}} \\
&\approx 29\,\text{days}
\end{aligned} \tag{4.3}$$

4.3 Deployment and data collection experiment two

We conducted experiment number 2 to verify the feasibility of the SGBS for 11-21 days in three households to understand households' behaviours on the garbage disposal and identify the type of garbage contents disposed of. We used age group, family size, family type and cooking and eating habits as criteria to select the participants for the experiment. Table 4.3 outlines the information of the study participants. We ensured safe data collection and storage for users by protecting users' anonymity and confidentiality, and the collected data was used only for the intended purposes of this study. Therefore, we did not sample for specific experience with smart home technology. However, we assumed all participants have smartphone experience and general knowledge of sensors Fig. 4.3 shows the deployed SGBS in the household. The following section illustrates how households use a mobile application that guides them during garbage disposal.

Table 4.3: Information of the participants

Household one	**Particulars** - Age group:(23) - Type of family: Single - Size of family: 1 - Life style: 1 absence on day on weekdays - Cooking & Eating habits: often evening
Household two	**Particulars** - Age group:(29) - Type of family: Young married couple - Size of family: 2 adults, 1 child - Life style: 1 absence on day on weekdays - Cooking & Eating habits: often all the days
Household three	**Particulars** - Age group:(55-50) - Type of family: Married Couple - Size of family: 2 adults - Life style: 1 absence on day on weekdays - Cooking & Eating habits: often all the days

Figure 4.3: Smart garbage bin system deployed in a household

4.3.1 Garbage annotation mobile application

Fig. 5.2 depicts the designed and developed garbage annotation mobile application to guide households during garbage disposal and collect information about the type of garbage contents disposed of by households. The annotation application consists of four categories with ten different types of burnable garbage contents; kitchen garbage; (*all food garbage*), plastic/unclean products; (*storage containers, toys, unclean packages, unclean containers*), paper/textile; (*tissues, mixed papers, milk/juice box, unclean cloth*) and other related types of garbage contents. The classification of burnable garbage is based on the catalogue as instructed by Ikoma city in Japan. The garbage annotation application allows individual households to select the type of garbage content each time they dispose of garbage in the SGB from a handy smartphone fixed outside on top of the SGB cover. Then data about the garbage content are sent to the cloud data server using a Wi-Fi network.

4.3.2 Result and discussion

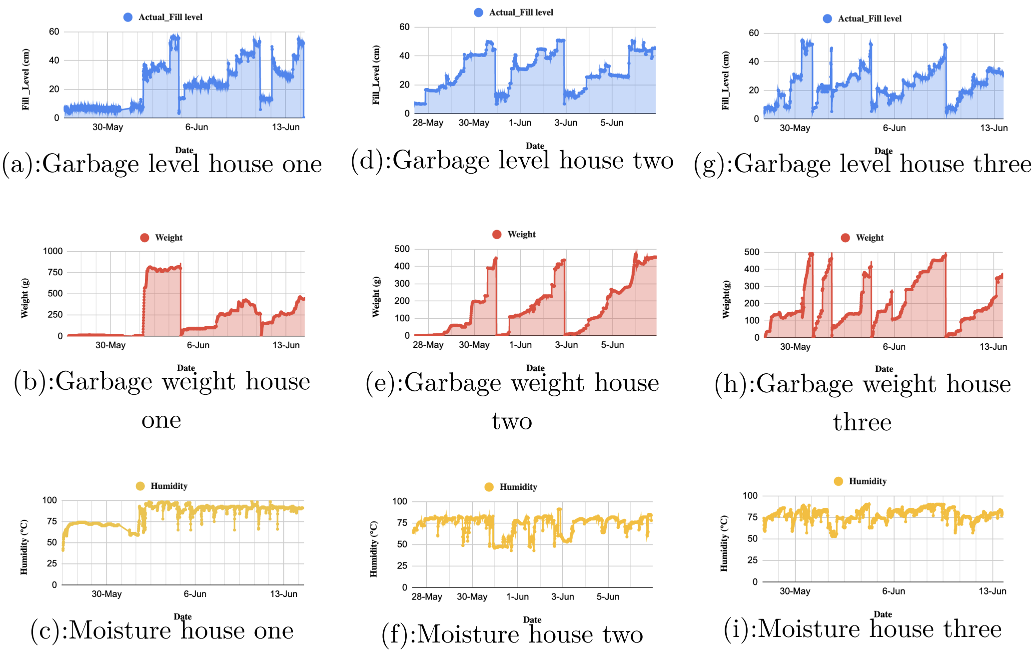
The discussion below is on the general aspects of understanding the behaviour of household users on garbage disposal from the established garbage log service "GLS" of each family. First, we discuss tracking garbage growth amounts from the users of the three households. Afterwards, we discuss the identified important

Figure 4.4: Garbage annotation mobile application

types of garbage content for each household, and lastly, we describe the routine behaviour of disposing of such garbage content. This discussion is more on general households' adoption, experience and limitation in using the SGBS.

Tracking garbage growth amount

The objective of this initial experiment was to verify the feasibility of the SGBS in understanding the behaviour of households on garbage disposal. Fig. 4.5(a) to Fig. 4.5(i) depicts the results of tracking the garbage amount for all three household users when using the SGBS. The size of the smart garbage bin used in the experiment was Width 31 $\times$ Depth 39 $\times$ Height 57.5 cm and 45 L capacity. From Fig. 4.5(a) to Fig. 4.5(i), the Y-axis of the garbage level shows the fill level in cm unit, and the Y-axis of garbage weight graphs shows the weight of the



(a):Garbage level house one (d):Garbage level house two (g):Garbage level house three

(b):Garbage weight house one (e):Garbage weight house two (h):Garbage weight house three

(c):Moisture house one (f):Moisture house two (i):Moisture house three

Figure 4.5: Garbage filling level, weight and moisture condition for household one, two and three

garbage in gram units respectively. The X-axis of both shows the date of the collected garbage data. The actual fill level was measured between the lid of the smart garbage bin and the garbage disposal bag inside the smart garbage bin. So a narrow distance shows that the garbage is reaching the maximum fill level, and a long-distance shows that the smart garbage bin is almost empty. The actual behaviour of the garbage fill level amount is directly proportional to the growing weight of the garbage; however, the fill level was much affected by compression behaviour causing huge fluctuations of the fill level from the steady state. Therefore, the behaviour of changing garbage disposal bags inside the smart garbage bin differs from household to household. As proposed in our previous work [52], the garbage growth amount from the households helps predict growth behaviour. Furthermore, predicting future garbage growth behaviour can be scaled from the households to a larger scale of the city in a different season of operation, thus providing more efficient garbage management. The subsequent section provides a brief understanding of household behaviour on garbage disposal using the tracked amount of garbage.

- **Household one:** It was a single-family occupied by one individual. During the first seven days (27-May to 2-June) of the experiment in this household, the garbage amount, as illustrated in Fig. 4.5(a) and Fig. 4.5(c), remained at the lowest steady-state with slight sensor fluctuations. This behaviour validated the assumption that either the participant was absent from home or the participant did cook or eat at home at all. Thus, there was no garbage disposal behaviour observed. However, in the following days of the experiment, the garbage growth started to be observed using the collected sensor data and annotation app. Surprisingly, the garbage amount grew faster than expected and remained at high peaks for four days before changing the garbage disposal bag in the smart garbage bin, whereas the sensor values began from their initial value on around 4-June. Further on the experiment, we also observed that the garbage amount continued to grow for six days, around 6-June to 11-June, with some fluctuation due to garbage compression behaviour, thus verifying that the household did not change the garbage disposal bag for all six days. This was revealed as the sensor values did not begin from their initial values but continued. Such behaviour was found strange and interesting compared to the observed behaviours of households two and three, as described below.

- **Household two:** While household one had an unpredictable garbage growth behaviour, household two consisted of a family size of two adults with one child who had constant behaviour on the garbage disposal. As shown in Fig. 4.5(d) and Fig. 4.5(f), the participants used the smart garbage bin every day and changed the garbage disposal bag mostly after every two days. This occurrence was also realized by using the collected sensor values as they changed and began from their initial values every time the disposal bag was changed. However, the garbage compression behaviour was also highly perceived, which explains that probably participants always wanted to keep the garbage bin from reaching its maximum thresholds or heavier garbage came after lighter garbage.

- **Household three:** The household consisted of two married couples. The behaviour of garbage growth, as shown in Fig. 4.5(g) and Fig. 4.5(i), looked

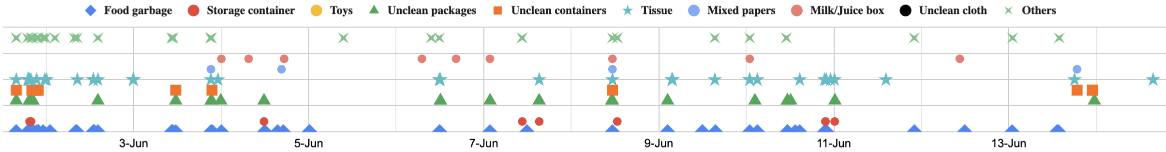


Figure 4.6: Identified garbage contents in household one

similar to household two, although the peak, trends and patterns of garbage growth were different. In this household, the behaviour of changing garbage disposal bags varied from one day (31-May to 1-June) to a maximum of four days (4-June to 9-June) as depicted in both Fig. 4.5(g) and Fig. 4.5(i). This trend was also realized using the collected sensor values as they changed and began from their initial values.

Identification of garbage content

Furthermore, we identified the type of garbage contents disposed of by the households. The households used the garbage annotation app see (Fig. 5.2) to input data about their daily garbage disposal. The garbage annotation app had four categories of burnable garbage with ten different types of garbage contents, as described in subsection 4.3.1. Therefore, households used the garbage annotation app to select and input the type of garbage contents each time they disposed of garbage in the smart garbage bin. First, we identified the type of garbage content through users' input data from the garbage annotation app. Later, the garbage contents were realized through the moisture conditions and the air quality found in the smart garbage bin since the type of garbage contents affects the moisture and air quality in the smart garbage bin. Therefore, using data from the garbage annotation application and moisture inside the smart garbage bin, we identified and ranked the type of garbage contents that were more important to the households (highly produced and disposed of) than the others (low produced and disposed of).

- **Household one:** As shown in Fig. 4.6 we found that food garbage contents ranked as the most important garbage produced and disposed of by the participants. In this household, the participant often disposed of the food

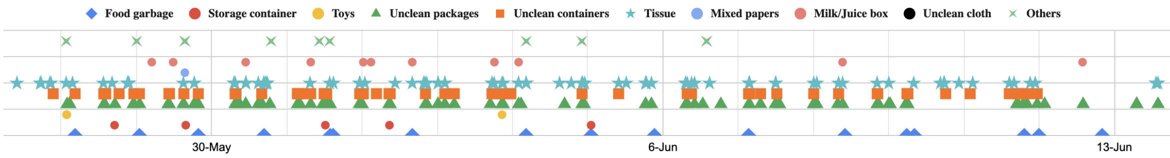


Figure 4.7: Identified garbage contents in household two

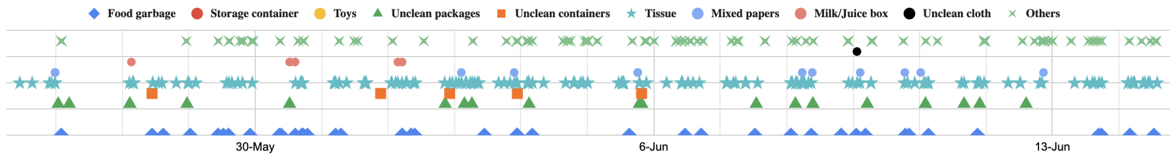


Figure 4.8: Identified garbage contents in household three

garbage content causing the moisture to rise in the smart garbage bin. Therefore, as depicted in Fig. 4.5(c), the moisture condition was consistently high. Furthermore, Tissue contents were second in rank, followed by other types of burnable garbage content. The unclean package was the fourth in high ranking. The mixed paper was the lowest in the rank of garbage content produced and disposed of only three times by the household. The result has shown that mixed paper was only disposed of three times between 3-June and 5-June and once on 13-June. An unclean container was disposed of six times, whereas storage containers appeared seven times during the experiment.

- **Household two:** Like in household one (Fig. 4.6), it also observed that tissue garbage content ranked second in household two as illustrated in Fig. 4.7. The unclean package garbage content ranked first, followed by the unclean container. Lastly, the food garbage content ranked fourth among the important garbage produced and disposed of by the participant. In comparison, only five times the storage containers were disposed of, making it the lowest in rank. Although toy garbage seemed unimportant in other households, it was exciting to find it a few times in household two. Moreover, in this household, the moisture inside the garbage bin was observed with a shifting tendency, as shown in Fig. 4.5(f), because the participant used to collect and park garbage in a small disposal bag before disposing of it in the smart garbage bin. Therefore, it justifies the observation that the unclean package was identified as the first garbage content produced.

- **Household three:** Contrary to the observation in households one and two, where tissue garbage content was the second in rank, in household three, as shown in Fig. 4.8, tissue was ranked first as the most important garbage content, followed by other types of burnable garbage content. Food garbage contents ranked third, whereas unclean packages ranked fourth. The unclean cloth was the lowest in rank as it was observed only once on 9-June throughout the experiment. Even though tissue contents are frequently disposed of, the moisture inside the smart garbage bin was high, as observed in Fig. 4.5(i); this proves the assumption that the tissues were slightly wet and also the food contents contributed to the rise of the moisture in a smart garbage bin.

Routine behaviour of garbage disposal

In addition, we have also learned the routine of garbage disposal by the households. In general, the study found that households can dispose of different types of garbage simultaneously and annotate all types of garbage content at the exact incidence. The study further observed that certain types of garbage content were frequently disposed of and annotated daily by households. For instance, food garbage contents in household one, Unclean packages garbage content in household two and tissues garbage content in household three see (Fig. 4.6 to Fig. 4.8) thus, were identified as the most important type of garbage contents disposed of by the household every day in the experiment.

4.4 Chapter summary

In this dissertation chapter, we focus on understanding households' garbage disposal behaviour and identification of the type of garbage contents disposed of. First, we designed and developed a smart garbage bin system, "SGBS", fastened with distance and weight sensors to detect the amount of garbage disposed each time. Then, we designed and developed a garbage annotations application to allow households user to annotate their daily garbage content. The annotation application comprised 4 garbage categories and 10 garbage content identities. To evaluate our approach, we conducted an initial experiment on the smart garbage bin system in three households. Later we identified necessary garbage categories and contents in each category through the user's voice and redesigned the annotations application to have 8 garbage categories and 25 garbage content identities to allow households user to annotate their daily garbage content. Therefore we conducted another experiment on the smart garbage bin system in five households. Our findings show that households' garbage disposal behaviour depends on the amount and contents of garbage and the routine of disposing of such garbage content. Finally, we discuss the potential of our system to be scaled in a smart city to influence behaviour change, provide healthier life, and improve garbage management operational efficiency. The next chapter introduces a new garbage content estimation model and improves the garbage classification task using machine learning algorithms.

5 Smart garbage bin: Garbage content estimation model

Throughout this chapter, we respond to the third challenge of improving the accuracy of garbage classification by training a machine learning model using daily collected fuse sensor readings combined with detailed household garbage contents annotations. First, we describe the requirements for designing and developing smart garbage bin as the primary tool for data collection experiments. Next, we discuss rules for garbage separation from municipals in Japan, where we lay our ground for identifying garbage content for households. Afterwards, we build a garbage content estimation model. Eventually, we realize the identification and classification of garbage using the built content estimation model.

5.1 Methods and tools

This section presents the details of the system requirements necessary for designing and developing a smart garbage bin system (SGBS), tools and the procedure for selecting important garbage categories for developing garbage annotation application design.

5.1.1 System requirements

In this subsection, we describe the system requirements for the proposed system. Based on the discussions in Chapter 1 and Chapter 2, we find the following four requirements for a smart garbage bin system:

1. The smart garbage bin system should automatically collect sensor data without any additional activities by users.

2. The smart garbage bin system should estimate detailed garbage categories and garbage content identities corresponding to each disposal behaviour.

To address requirement (1), we designed and developed a smart garbage bin system which is always connected to the internet, uploads all sensor data to the cloud to store them. To address requirement (2), we built a new machine learning model for estimating garbage categories and garbage content identities with high accuracy.

5.1.2 Architecture design

Fig. 5.1 demonstrates a designed and developed SGBS architecture to revolutionize the existing household garbage management system by tracking daily household garbage disposal information and identifying the type of garbage contents disposed of at the source. The smart garbage bin system architecture consists of two subsystems: the smart garbage bin (SGB), embedded with distance and weight sensors to detect the timestamp of newly disposed of garbage content during garbage disposal. On the other hand, the smart garbage bin (SGB) is embedded with temperature, humidity, and gas sensors to identify and distinguish disposed of garbage contents. Secondly, SGBS architecture comprises the garbage annotation mobile application (GAA) with a smooth interface that allows users to annotate their daily disposal of garbage content during garbage disposal. The two subsystems (SGB and GAA) later create a daily garbage log data for each house. Moreover, the designed architecture comprises the analysis part that uses machine learning algorithms to classify garbage contents found in the house logs. The outcome of the analysis produces a garbage content estimator for each home which helps identify and classify garbage content at the source.

5.1.3 Smart garbage bin

Fig. 5.3 shows the overview of a designed and developed smart garbage bin system (SGBS). Considering the significant roles of the proposed SGBS architecture described in Section 5.1.2, a set of lightweight, low-cost, high-precision IoT sensors were chosen and embedded in the smart garbage bin (SGB). The selected devices have different hardware configurations and purposes. In our SGB prototype, we

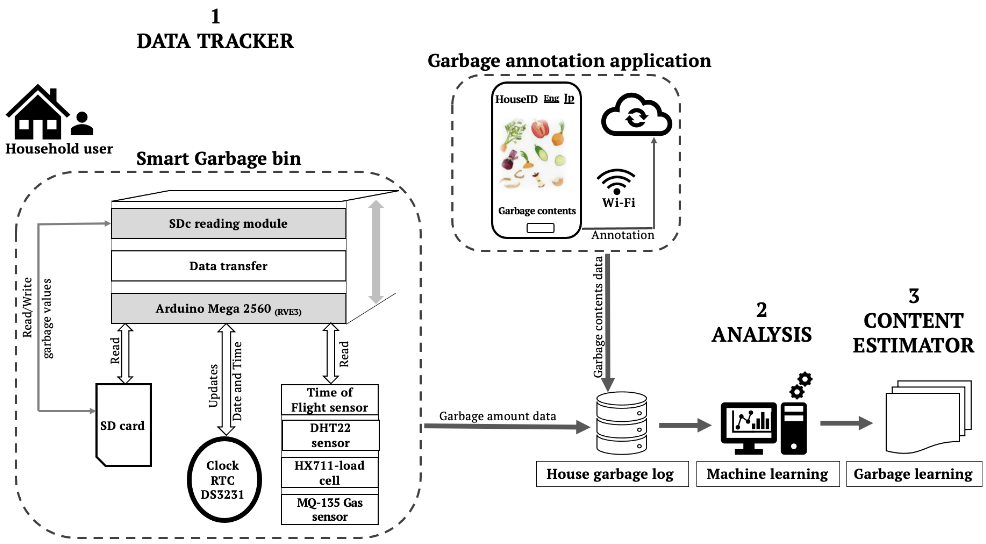


Figure 5.1: Smart garbage bin system architecture design

used a DHT22 (temperature and humidity) and MQ135 gas sensors to monitor the moisture and air quality of the disposed garbage content in the smart garbage bin. Furthermore, we used a ToF (time of flight) and HX711-load cell to track the garbage filling level and weight at each time of disposal. Using a Wi-Fi gateway, the smart garbage bin system is always connected to the internet, uploads all sensor data to the cloud, and stores them. In addition, the Secure Digital non-volatile flash memory card format (SD), connected to an I2C real-time clock with 32.768 kHz frequency (DS3231 RTC) module data are also collected and stored in the SD-created file in one-minute intervals daily. On the other hand, the SGB comprises the 2×16 character LCD Module with a blue backlight, which uses an I2C interface to communicate with the host Arduino Mega 2560 microcontroller Rev3. Therefore, the LCD module displays the garbage's current filling level and temperature data of the smart bin. The proposed smart garbage bin prototype allows easy tracking of garbage amount information at the source.(see Table 5.1) provides the purpose of the chosen sensors used to develop the smart garbage bin.

Table 5.1: Sensor used in development of smart garbage bin

Sensor	Purpose
1. AE-VL53L0X(ToF)	Measure fill level of garbage in a bin
2. DHT22	Measure temperature and humidity in a bin
3. HX711-load cell	Measure increase of weight of garbage in a bin
4. MQ135 gas sensor	Measure CO2,NH3,Smoke in a bin

5.1.4 User feedback from deployment and data collection experiment two

From Chapter 4 of the dissertation. Apart from the achieved garbage content identification results through household users annotation. Yet, as feedback from users in experiment number 2, some important garbage contents were missing in the initial design of the garbage annotation application. Through the user's voice, a short survey and studying rules for garbage disposal from municipal pamphlets, important categories were identified in the study. Therefore, we redesigned the garbage annotation application to have 8 garbage categories and 25 garbage content identities. The subsequent section details the design and development of a garbage annotation application interface for the deployment and data collection experiment number 3 of the SGBS.

5.1.5 Garbage annotation application

To provide a smooth and easy way for households to annotate garbage content they dispose of daily. We further present a garbage annotation mobile application (GAA). The GAA designed and installed in a handy smartphone made a significant value consideration to household users by allowing annotation in a more efficient and tailored way through a smooth interface. The selection of the garbage categories in our proposed study is based on the rules for separating and disposing of burnable garbage as provided in four random selected municipal's pamphlets in Japan that explain the garbage disposal rules described in Section 2.3.1, including the city of Kashihara [38], Ikoma [39], Nara [40], Kyoto [41]. Additionally, we conducted a short survey with fifteen (15) students living in the city of Ikoma and Nara for one week. The survey participants were asked to annotate their daily burnable garbage disposal on paper. The annotation included the name of the garbage contents and the frequency of disposing of such garbage. Thus, by analyzing the survey results and the rules for disposing of the garbage from municipal pamphlets, we established important categories of burnable garbage with specific content identities for the mobile annotations application. The garbage annotations application interface comprises the garbage categories and a menu with two languages, English and Japanese, giving users flexibility to switch between the languages. Also, the interface consists of house numbers as an identification for the experimental data collection.

Fig. 5.2 demonstrates the garbage annotation application interface whereby vertically depicts 8 garbage categories (i.e., Kitchen waste, Meal garbage, Paper/softbox, Fabric/textile, Plastic, Dust, Plant, and All others) and horizontally depicts 25 garbage contents identities (i.e., Food garbage, Edible food, Sink basin, Kitchen waste bag, Unclean cup, Unclean container, Unclean packages, Waste wood, Tissues, Mixed Papers, Milk/Juice box, Masks, Clothes, Shoe, bag, Rubber products, Disposable diapers, Plastic product, Toys, CD, Cigarette ashes/stick, Vacuum cleaner, Plant and Others) belonging to each category. The garbage annotation application provides a guide knowledge that allows individual households to smoothly select the type of garbage content each time they dispose of garbage in the SGB from a handy smartphone fixed outside on top of the SGB cover. Then, data about the garbage category and its specific identity content are sent to the cloud data server using a Wi-Fi network.

5.2 Deployment and data collection experiment three

Herein we present the experimental setup and data collection, including datasets, the data preprocessing steps undertaken to build the garbage contents estimation model, and the methods adopted to address the study aims.

Figure 5.2: Garbage annotation application interface

5.2.1 Experiment and Participant information

We conducted the evaluation experiment from June to August 2022 in five households of heterogeneous characteristics in the city of Nara, Ikoma, and Kyoto in Japan for 3-5 weeks. We considered family size, type of family, age group, number of children, and city as the criteria for selecting participants for the experiment. Table 4.3 outlines the participant's information. All participants were well informed about the experiment and provided their own consent to participate in the experiment. In addition, smart garbage bins were distributed and installed in each house. Fig. 5.3 shows the overview of the deployed SGBS.

Table 5.2: Information of participants

House ID	Family size	Family type	Child/infant	Age group	City
1	2	Couple	0	50-55	Kyoto
2	4	Couple	2	7-50	Nara
3	2	Couple	0	28-29	Ikoma
4	2	Single-shared	0	22-23	Ikoma
5	3	Couple	1	27-29	Ikoma

5.2.2 Datasets

The experiment resulted in five garbage logs data from the five households. The garbage log consists of data from the SGB (i.e., timestamp, filling level, weight, temperature, humidity, and air quality), collected every one-minute interval. Also, data from the GAA (i.e., timestamp, garbage categories, and content identities) collected only when a user disposes of and annotates the garbage in a smart garbage bin. The frequency of garbage disposal and annotation of garbage contents differ in each household due to household characteristics.Table 5.13 details the full annotations of garbage contents found in houses 1 to 5 by the household users during the experiment. Therefore, we define the following rules to merge the multiple sensor data from the smart garbage bin **(as features)** and garbage content annotations by the households **(as labels)** to create a single dataset of each house. We considered a time stamp of 10-minute intervals from the disposal time recorded by the annotation application to calculate features for the particular label. The features include maximum, minimum, and rate of change of the garbage filling level, weight, temperature, humidity, and air quality. At the same time, the label consists of 8 garbage categories and 25 garbage identities. Thus, we obtained the total original datasets of each house for both garbage categories and content identities. Below are the rules used to merge the collected data;

1. Every 10 minutes, if a new garbage label is input, and then calculate new features for the label.

2. If at the same time or in less than 10 minutes, another new label is input, then use the previously calculated features for the new label (Overlap features).

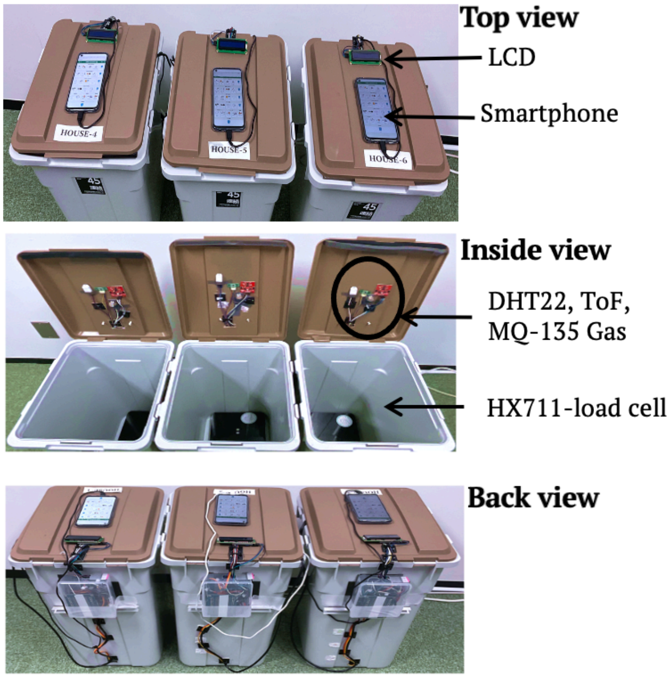


Figure 5.3: Smart garbage bin system overview

5.2.3 Class balance

A lower frequency of disposing of a particular type of garbage content than the others experienced in all houses leads to a minority of such garbage content. Therefore, the minority class labels affect the model-building process, i.e., a model that always chooses the majority class regardless of the corresponding feature. To solve this, we utilize the resampling technique to enhance the classifier model's size and quality and avoid biases class during training. There are two main approaches for random resampling: Oversampling, which duplicates the minority class, and Undersampling, which deletes the majority class. In our case, due to the low number of annotations in garbage category 4 (Fabric/textile), garbage category 5 (Plastic), garbage category 6 (Dust), and garbage category 7 (Plant) experience in all five houses (see Table 5.13), we applied the Oversampling technique to increase the minority class using the imbalanced-learn sci-kit-learn library. Table 5.3 and Table 5.4 show the total number of datasets of garbage categories and content identities before and after resampling.

Table 5.3: Re-sampling and cross-validation split for the 8 garbage categories

	Before re-sampling		After re-sampling	
House ID	Original dataset	25% splits	New dataset	25% splits
1	682	170-171	2618	654-655
2	360	90	1400	350
3	538	134-135	1281	320-321
4	121	30-31	305	76-77
5	449	112-113	1064	266

Table 5.4: Re-sampling and cross-validation split for the 25 garbage content identities

	Before re-sampling		After re-sampling	
House ID	Original dataset	25% splits	New dataset	25% splits
1	687	170-172	1930	482-483
2	364	91	570	142-143
3	541	134-136	1295	323-324
4	121	30-31	360	90
5	450	112-113	1165	291-292

5.3 Garbage content estimation model

This study aims to identify garbage contents disposed of and perform the garbage classification from garbage contents disposed of daily in the household by adopting IoT and data-efficient machine learning algorithms. Therefore we present a garbage content estimation model to classify 8 categories of garbage and a total of 25 garbage contents identities relating to a particular category, as demonstrated in Fig. 5.2 of the garbage annotation application. The subsequent section details the process of building classification models.

5.3.1 Model building

Fig. 5.4 demonstrates model building steps and order of operations. We performed the classification tasks from daily collected fuse sensor readings combined with detailed household garbage contents annotations intending to find the class

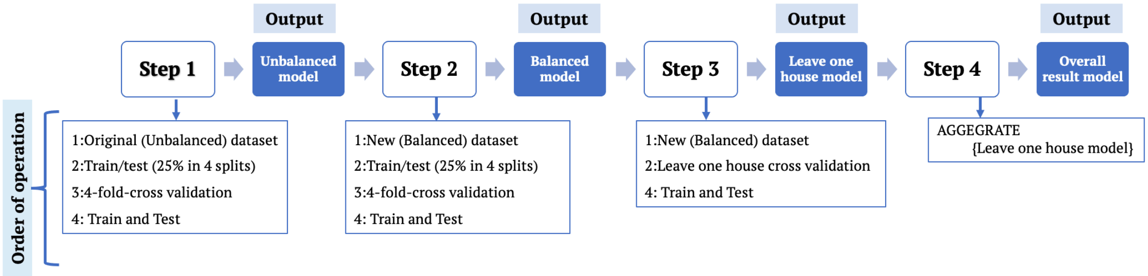


Figure 5.4: Model building steps and order of operations

(i.e., 8 garbage categories: *Kitchen waste, Meal garbage, Paper/softbox, Fabric/textile, Plastic, Dust, Plant, and All others*) and (i.e., 25 garbage content identities: *Food garbage, Edible food, Sink basin, Kitchen waste bag, Unclean cup, Unclean container, Unclean packages, Waste wood, Tissues, Mixed Papers, Milk/Juice box, Masks, Clothes, Shoe, bag, Rubber products, Disposable diapers, Plastic product, Toys, CD, Cigarette ashes/stick, Vacuum cleaner, Plant and Others*) to which a new unseen observation belongs. During the model-building steps in Fig. 5.4, we only consider utilizing data-efficient methods, namely: Random forest, Naive Bayes, Extreme Gradient Boosting (Xgboost), and Decision tree algorithms to build the garbage content estimation model, for the reasons such as the comparison of the machine learning classifiers, the small number of available datasets, the popularity of the classifier and data preprocessing to avoid minority class labels. We eventually defined the order of operations applied to the selected classifiers during the model-building steps.

More precisely, we train and test by spliting the dataset of each house into four (4) chucks of 25% equal size dataset as shown in the Table 5.3 and Table 5.4 for garbage categories and content identities. To avoid overfitting as much as possible, first, we utilize repeated k-fold cross-validation to evaluate the machine learning models in steps 1 and step 2 (see Fig. 5.4). Then, we averaged the results with 4-fold cross-validations to compute the final validation score for each investigated model configuration. Therefore, the model created in step 1 used the original (unbalanced) datasets, i.e., before resampling (see Table 5.3). While the model developed in step 2 used the balanced class dataset, i.e., after resampling (see Table 5.4), as discussed in Section 5.2.3 Thus, for performance comparison of balanced and unbalanced datasets, our model-building process output two models, an unbalanced model and a balanced model (see Fig. 5.4).

Afterwards, for better comparison reasons of the cross-validation methods applied to the classifiers, and, in order to increase the training set, in step 3 (see Fig. 5.4), we changed the cross-validation method to leave one house out cross-validation method where we repeatedly trained our models with total balanced datasets from the four houses and testing the model with the remaining one house. Thus, we obtained the Leave one house out model.

Furthermore, we built the overall result models in step 4 (see Fig. 5.4) of the classification tasks for both class garbage categories and content identities for each house to investigate the overall performance of the classifiers. We first made the overall result model on all 8 garbage categories, *i.e. Kitchen waste, Meal garbage, Paper/softbox, Fabric/textile, Plastic, Dust, Plant, and All others* found in House 1, House 2, House 3, House 4 and House 5. Nonetheless, because each garbage category comprises 5 to 2 specific garbage content identities (see Fig. 5.2), in total, there are 25 different garbage content identities belonging to the eight categories expected to be annotated by the users daily using the garbage annotations application. Therefore because of the majority number of garbage content identities and differences in frequency behaviour of garbage disposal and annotation exhibited from each house (see Table 5.13). In this study, we first selected the five garbage content identities from the Kitchen waste (category 1) as it has had a higher frequency of annotation in house 3, house 4 and house 5. Also, we chose the five garbage content identities from the paper/softbox (category 3) as it has had a higher frequency of annotation in house 1 and house 2 to learn the performance of the classifiers on garbage content identities. Therefore, to this point of the study, we created three overall result models for garbage content estimation, namely;

1. Overall result model for general garbage categories

2. Overall result model for kitchen waste contents identities

3. Overall result model for paper, softbox contents identities

5.3.2 Performance evaluation

Our model evaluation performance is based on accuracy, which is the percentage of correct comparison classifications. Moreover, we evaluate the performance of our models using other metrics, such as confusion matrices, Precision, Recall and F1-score. We will especially give the most informative metrics for the overall result models because they aggregated the garbage class label results from all houses belonging to the same classification and averaged the result into a single metric measurement. Furthermore, the model parameters tuning was applied on all classifiers, Random forest, Naive Bayes, Extreme Gradient Boosting (Xgboost), and Decision tree. As a result, the accuracy slightly increased by increasing the number of parameters such as estimators, criterion, and random state for each model separately. Therefore, we independently investigated the model performance on all experimental datasets found in House 1, House 2, House 3, House 4, and House 5 on garbage categories and garbage content identities classification tasks. The percentage performance accuracy results using 4-fold cross-validation and leave-one-house-out cross-validation as applied to the four machine learning classifiers for the 8 garbage categories and 25 garbage identities are summarized in Table 5.5, Table 5.6, Table 5.7 and Table 5.8.

Table 5.5: 4-fold cross-validation performance accuracy for the 8 garbage categories

	Accuracy (%) of unbalanced model				Accuracy (%) of balanced model			
House ID	Random forest	Naive bayes	Xgboost	Decision tree	Random forest	Naive bayes	Xgboost	Decision tree
1	**90**	89	85	67	65	71	68	67
2	85	85	74	71	73	72	73	63
3	**86**	72	85	82	86	79	79	72
4	80	79	73	77	75	76	78	63
5	87	86	83	81	80	78	62	69

5.3.3 Results

Throughout this subsection, we describe results obtained from the classification tasks as detailed in Section 5.3.2. Specifically, we look into and compare the

Table 5.6: Leave one house cross-validation performance accuracy for the 8 garbage categories

	Accuracy (%) of Leave one house model			
House ID	Random forest	Naive bayes	Xgboost	Decision tree
1	83	83	79	57
2	84	84	84	69
3	**88**	81	81	80
4	80	80	78	72
5	84	78	78	69

Table 5.7: 4-fold cross-validation performance accuracy for the 25 garbage content identities

	Accuracy (%) of unbalanced model				Accuracy (%) of balanced model			
House ID	Random forest	Naive bayes	Xgboost	Decision tree	Random forest	Naive bayes	Xgboost	Decision tree
1	**93**	89	86	86	**88**	79	63	72
2	89	87	84	85	**88**	81	68	76
3	87	87	87	82	83	85	68	76
4	83	84	80	80	79	73	71	72
5	86	85	82	83	75	78	65	62

performance accuracy from the unbalanced, balanced, leave one house, and overall result models using the four machine learning classifiers..

Unbalanced model

We see from the results of the unbalanced model (seeTable 5.5 and Table 5.7) using the 4-fold cross-validations that Random forest performs slightly better than other classifiers (Naive Bayes, Xgboost, and Decision tree), for classification tasks of both garbage categories and garbage content identities. For garbage categories, the highest accuracy was 90% obtained in house 1, and the 67% lowest accuracy resulted from the Decision tree in the same house. Also, 93% for garbage content identities was the highest accuracy found in house 1 by Random forest, and the lowest accuracy was 80% by the Decision tree found in house 4

Table 5.8: Leave one house cross-validation performance accuracy for the 25 garbage content identities

	Accuracy (%) of Leave one house model			
House ID	Random forest	Naive bayes	Xgboost	Decision tree
1	**91**	89	88	74
2	90	87	84	78
3	89	89	89	84
4	86	86	86	81
5	88	85	85	65

Balanced model

Afterwards, we compared the four classifiers with the same 4-fold cross-validations method in all five houses on a balanced dataset with the approaches discussed in Subsection 5.2.3 to deal with the unequal class balance. The results can be seen in Table 5.5 and Table 5.7. We observed that the performance accuracy slightly decreased compared with the unbalanced model performance. Yet, Random forest manifested the highest accuracy and thus outperformed the rest of the classifiers. For the garbage categories, the Random forest exhibited 86% in house 3, and 63% by the Decision tree in house 2 was the lowest accuracy. While for garbage content identities, the accuracy was 88% by Random forest from house 1 and house 2, and the most insufficient accuracy was 62% by a decision tree in house 5.

Leave one house model

In the next step, we compare the results of the repeated 4-fold cross-validation in step 2 to the Leave one house out (LoH) cross-validation approaches in step 3 (see Fig. 5.4). In order to investigate the classification performance in all five houses. Therefore, we applied the LoH on the balanced class datasets using the four classifiers in step 3. However, we maintained the same order of operation as in step 2. With this approach, the sum of four houses increases the size of the training set during repeated testing with only one house dataset. The results for Random forest, Naive Bayes, XGBoost, and Decision tree in the case of the garbage categories and garbage content identities for all four classifier sets are shown in Table 5.6 and Table 5.8. We see an apparent accuracy increase in each house compared to the balanced model of 4-fold cross-validation in Table 5.5 and Table 5.7. For the garbage categories, the Random forest revealed the highest accuracy of 88% in house 3, while the decision tree showed the lowest accuracy of 57% in house 1. In addition, garbage content identities in the leave one house model achieved the highest accuracy of 91% and 90% by random forest in house 1 and house 2, respectively. On the other hand, the decision tree exhibited unsatisfactory performance, 65% in house 5. Moreover, Random forest again steadily outperformed the rest of the classifiers.

Overall result model

To realize the performance of the three overall result models described in Section 5.3.1 above Overall result model of garbage categories, (2) Overall result model of kitchen waste contents identities and (3) Overall result model of Paper/softbox contents identities. The performance accuracy results for the three models are shown in Table 5.9. Moreover, we compared the Recall, Precision, and F1-score for the overall result models as they can better judge the performance by showing the metric measurements of each class label.

For the garbage categories overall result model (see Table 5.9), Random forest achieved the highest accuracy of 85%, followed by Naive Bayes at 82% and Xgboost at 80%, while the decision tree lags with the least accuracy of 64%. Table 5.10 summarises the metric accuracies of the 8 garbage categories overall result model with Recall, Precision, and F1-score using the Random forest classifier.

Further, for the overall result model of kitchen waste contents identities (see Table 5.9) *(i.e., food garbage, edible food, sink basin, kitchen waste bag, and others)*. The random forest has steadily revealed the best classification accuracy of 91%, while the accuracies of the rest of the models are; 88% Naive Bayes, 84% Xgboost and 76% Decision tree. Likewise, the overall result model of the paper/softbox contents identities (see Table 5.9) *(i.e., tissues, mixed papers, milk/juice box, masks, and others)* are 85% Naive Bayes, 83% Xgboost and 71% Decision tree were outperformed by the Random forest at 89%. The summary of the Recall, Precision, and F1-score for the overall result models of the 5 kitchen waste and the 5 paper/softbox content identities are shown in Table 5.11 and Table 5.12, using the Random forest as it has been portrayed as the best classifier.

The aggregated confusion matrix plots using the Random forest of each overall result model are shown in Fig. 5.5, where the columns represent the actual values (Truth) of the target class label. The rows represent the predicted values (Predicted) of the target variable class label. The number of validation samples that were correctly classified are demonstrated in the diagonal cells, and that were incorrectly classified are demonstrated in the off-diagonal cells.

In addition, to investigate the impact of the collected multiple sensor readings on the garbage content estimation model, we applied the features importance method using a random forest classifier as our chosen classifier for the garbage content estimation model. The results in Fig. 5.6 show that air quality, humidity, temperature, and fill level values are more relevant features for identifying garbage content in the smart bin. Therefore, the identified garbage content disposed of daily and annotation procedures contributes to the garbage classification tasks. Furthermore, the cross-validation approaches provided satisfactory results, especially for the leave-one-house cross-validation, which performed better than the 4-fold cross-validation.

Table 5.9: Accuracy performance of the three overall result models

	Accuracy(%) of Overall result models			
Overall result model of:	Random forest	Naive bayes	Xgboost	Decision tree
Kitchen waste	**91**	88	84	76
Paper/softbox	**89**	85	83	71
Garbage category	**85**	82	80	64

5.4 Discussion

Throughout this section, we discuss our findings and possible implications. Due to the sufficient classification outcomes, we chose the Random forest algorithm as the best classifier. We also decided on the overall result models as the final model for our garbage content estimation tasks. Generally, the highest accuracy

Table 5.10: Summary of 8 garbage categories overall result model

Name	Precision	Recall	F1-Score
Kitchen waste	0.84	0.85	0.85
Meal garbage	0.79	0.81	0.80
Paper/softbox	0.69	0.89	0.78
Fabric/textile	0.96	0.88	0.92
Plastic	0.99	0.88	0.93
Dust	0.90	0.88	0.89
Plant	0.95	0.78	0.86
All other	0.87	0.84	0.86

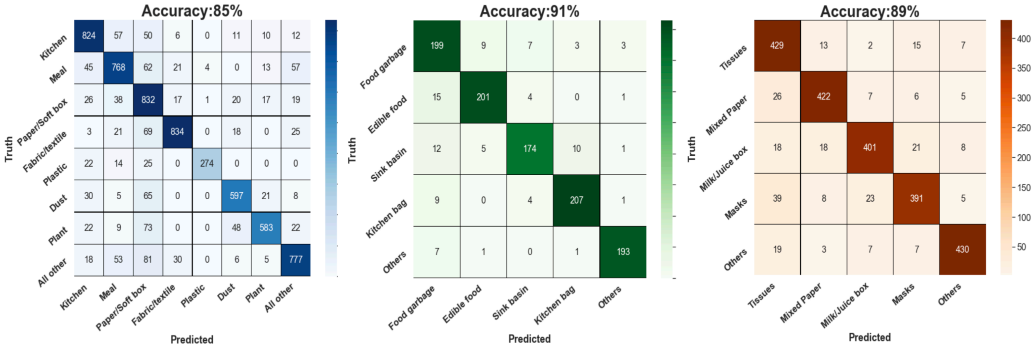
Table 5.11: Summary of 5 Kitchen waste contents identities overall result model

Name	Precision	Recall	F1-Score
Food garbage	0.82	0.90	0.86
Edible food	0.93	0.91	0.92
Sink basin	0.92	0.86	0.89
Kitchen bag	0.94	0.94	0.94
Others	0.97	0.96	0.96

Table 5.12: Summary of 5 Paper/softbox contents identities overall result model

Name	Precision	Recall	F1-Score
Tissues	0.81	0.92	0.86
Mixed paper	0.91	0.91	0.91
Milk juice box	0.91	0.86	0.89
Masks	0.89	0.84	0.86
Others	0.95	0.92	0.93

is between 85% and 91%, and the lowest is 64%, which is satisfactory for garbage content classification tasks. However, the lowest amount of annotation on certain class (imbalance) labels makes the classification task difficult. We start the detailed discussion by comparing garbage annotations from each house and then classification tasks by the machine learning algorithms, followed by the usefulness of the garbage content estimation model. Finally, we look at the comparison of our approach to the literature.



(a) Garbage category (b) Kitchen waste contents (c) Paper/softbox content

Figure 5.5: Confusion matrices of the three overall result models

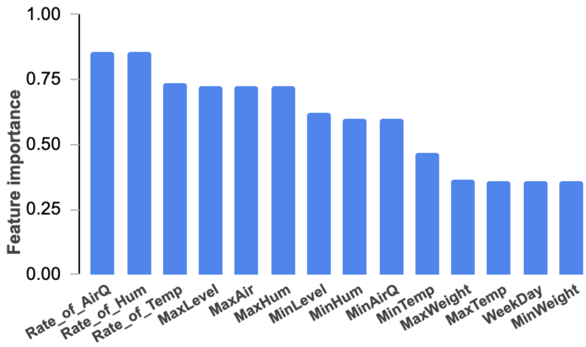


Figure 5.6: Features importance on sensor readings

5.4.1 Comparison of house garbage disposal annotation and classification

In general, we observed different behaviour of garbage disposal in all five houses, which is due to the heterogeneity behaviour in each family, such as living style, size of the family, type of the family, number of children/infants, age group, and city. In this case, the study observed differences in the routine frequency of garbage disposal and the type of garbage content disposed among the houses. Therefore, using the smooth garbage annotation interface (see Fig. 5.2) that allowed household users to annotate garbage contents during disposal, the study found that certain garbage contents were important in some houses, i.e., daily disposed and annotated, compared to others. Table 5.13 shows the annotation frequency of garbage category disposal among houses, as briefly detailed below.

- **House 1:** as shown in Table 4.3, this house consists of a married couple in Kyoto prefecture. In this house, garbage category 3 (Paper/softbox) was the most important category compared to other categories annotated 374 times during the experiment (see Table 5.13). In comparison, garbage category 5, which consisted of plastic contents, appeared as the least important annotated only 5 times. In addition, other categories had almost a similar frequency of annotation, such as Kitchen waste (78), Meal garbage (66), All others (74), and Dust (50). On the other hand, fabric/textile had 21 annotations, while the plant had 19 annotations.

- **House 2:** consists of a married couple with two children living in Nara city (see Table 4.3). Like in house 1 (see Table 5.13), garbage category 3 (Paper/softbox) was the most important category in this house, annotated 200 during the experiment, and Category 5 (Plastic) was the least annotated, only 4 times. Compared with other categories, Kitchen waste had 37 annotations, Meal garbage 63, All others 24, Fabric/textile had 16, dust 11, and Plant 9. House 2 had fewer annotations than house 1

- **House 3:** as shown in Table 4.3, this house comprises a young married couple in Ikoma city. Even though garbage category 3 (Paper/softbox) is steady as the most important and Plastic as the minor category observed in houses 1 and house 2, in this house, the study observed a slight difference in annotation frequency exhibited among Kitchen waste, Meal garbage, and Paper/softbox categories. The result in Table 5.13 shows that the annotations frequency kept, such as Paper/softbox (183), was the most important, followed by Meal garbage (125), and Kitchen waste (104) was the third in the garbage category importance ranking.

- **House 4:** While Houses 1, 2, 3, and 5 comprise married couples, house 4 consists of two singles living in a shared house in Ikoma city (see Table 4.3). The study observed less annotations frequency in this house than in other houses. However, similar to houses 1, 2, and 3, garbage category 3 (Paper/softbox) had the highest annotation frequency and ranked as the most important, while the plastic was minor. Therefore, the annotation frequency in Table 5.13 is as follows: Paper/softbox had 61 annotations, followed by Kitchen waste (23) and Meal garbage (11), which similarly ranks with house 3. In addition, not only Plastic was the minor but also dust which was annotated only once each. Moreover, category 7 (Dust) was not annotated in this house.

- **House 5:** This house comprises a young married couple with an infant in Ikoma city (see Table 4.3). Contrary to all other houses, the study observed a fewer annotation frequency of garbage category 3 (Paper/softbox), which prevailed in houses 1, 2, 3, and 4 as the most important garbage category (see Table 5.13). Instead, kitchen waste was the most important category in this house, with 152 annotations, followed by Meal garbage (135) and Fabric/textile (77) third in the ranks. The high annotation frequency of category 4 (Fabric/textile) was due to the disposal frequency of disposable diapers the fourth garbage content in the Fabric/textile category 4 (see Fig. 5.2) thus increasing the number of fabric/textile. On the other hand, Plant category 7 was annotated only once and therefore appeared as a minor category, similar to house 3. Plastic had 9 annotations, and dust had 6 annotations.

Eventually, daily disposed garbage contents and detailed garbage annotation frequency by households impacted the classification tasks in each house. For instance, in Random forests, the chosen classifier for this study (see Table 5.5) and (see Table 5.7), the accuracies for classification tasks of both garbage category and content identities in house 1 were higher than in house 4, which had fewer annotations frequencies. Moreover, the study found that the decision tree was the insufficient classifier model compared to Random forest Naive Bayes, Xgboost applied on the datasets in all five houses. Over and above that, the leave-one-house cross-validation method showed better performance compared to the 4-fold cross-validation approach despite its computational cost (see Table 5.6 and Table 5.8). Therefore, in the overall result models, we aggregated the classification result of the same class label into one metric performance using the leave-one-house approach, which has manifested better performance than 4-fold cross-validation on the balanced model. The following section compares our approaches with the literature

Table 5.13: Garbage annotation frequency found in house 1 to 5

Category ID	Category name	House 1	House 2	House 3	House 4	House 5
1	Kitchen waste	78	37	104	23	152
2	Meal garbage	66	63	125	11	135
3	Paper/softbox	374	200	183	61	53
4	Fabric/textile	21	16	21	8	77
5	Plastic	5	4	14	1	9
6	Dust	50	11	3	1	6
7	Plant	19	9	13	0	1
8	All other	74	24	78	18	17
	Total	687	364	541	121	449

5.4.2 Comparison with literature

As discussed in the Chapter 2, similar approaches in other domains/applications were investigated, and we compare our strategies and experimental setups and those more similar to ours, as detailed below.

- Suitable practice for house garbage separation
  Our study has considered the identification of daily disposed of garbage content and provided a satisfactory garbage category suitable for burnable garbage separation practice for most families in Japan. However, Nnamoko *et al.* [5] and Mookkaiah *et al.* [47] investigated only two kinds of garbage, i.e., Organic and recyclable, which is not enough for rational garbage separation in houses. Likewise, apart from increasing the number of classes as demonstrated by Ziouzios *et al.* [6] and Sami*et al.* [49], to find respective garbage categories such as *(kitchen waste, other waste, hazardous waste, plastic, glass, paper or cardboard, metal, fabric, and other recyclable waste)*. Yet these studies provided a small number and more generalizable garbage categories, which is not the best practice for proper house garbage separation and can not fully solve the problem of profound implications for

ecological balance and threat to global sustainability, development, and human well-being.

- Use of daily garbage contents and experiment transparency
  Our study proposed to perform garbage content estimations from the daily collected fuse sensor readings and household annotations with transparency on experiments and thus can be reproducible in the field. On the contrary, the studies by [6,47] and [49] used publicly available garbage image datasets to improve classification tasks with less transparency information on their experimental setup. However, the publicly available image datasets are associated with problems such as resizing, resolutions, and inappropriate colour presentation, thus lowering the quality of the classification task.

- Use of efficient data models
  Our study applied more data-efficient methods, namely Random forest, Naive Bayes, Xgboost, and Decision tree, for the classification tasks. On the contrary, most of the previous works applied the existing standard models for the classification tasks, such as VGGNet [42], AlexNet [43], ResNet [44], and DenseNet [45]. A common issue associated with image classification using the existing standard model is high computational cost which often results in high development time and prediction model size because they are often pre-trained for more than one purpose [5]. In addition, CNN-based models are difficult to run on embedded systems suitable for garbage bins, and their architecture requires large amounts of data for training which is yet to be available.

5.5 Chapter summary

In this dissertation chapter, we presented a new smart garbage bin system (SGBS) embedded with multiple sensors to identify the disposed garbage content categories by households. First, we designed and developed a smart garbage bin system (SGBS) architecture comprised of the smart garbage bin (SGB) equipped with temperature, humidity, gas, ToF, and load cell sensors and the garbage annotation mobile application (GAA) consisting of a smooth interface of 8 garbage categories and 25 content identities to allow users to annotate garbage contents during garbage disposal. Finally, we introduce a new garbage content estimation method by training a machine learning model using daily collected fuse sensor readings combined with detailed household garbage contents annotations to perform garbage classification tasks. We deployed the designed SGBS in five households over one month and applied the leave-one-house-out cross-validation to the model trained and tested with the collected data. As a result, our proposed method achieved an accuracy of 91% in 5 kitchen waste contents, 89% in 5 paper/softbox contents, and 85% in 8 garbage categories for the classification tasks. Moreover, our results show that air quality, humidity, temperature, and fill level values are more relevant features in the garbage content estimation model.

6 Conclusion

6.1 Summary

In this dissertation, we have presented a smart garbage bin system, SGBS, beneficial for learning and predicting garbage growth for a single house, understanding household garbage disposal behaviour and identifying the type of garbage content disposed of daily. Although the existing IoT-based smart garbage systems and automation and detecting garbage from images by artificial intelligence have high accuracies. Yet, most current systems still have three major issues: (1) they can not learn the amount of garbage disposed of each time; (2) they provide a small number of garbage categories, not enough for reasonable practices of household garbage separation. and (3) They can not understand households' routine behaviour of garbage disposal. Therefore, we need a new IoT-based garbage management system and a classification tool which improves existing systems.

To realize the benefits of SGBS, three challenges have been tackled in this dissertation:

1. How to learn the amount of garbage disposed of each time and predict garbage growth behaviour for a single house?

2. How to understand household garbage disposal behaviour and identify the type of garbage contents?

3. How to substantially improve the automation of garbage classification?

Since, to our knowledge, there was no such system before, it is necessary to investigate the impact and feasibility of learning the amount of garbage disposed of each time, identifying the garbage contents disposed of by households and understanding the routine behaviour of garbage disposal by households.

For the first challenge, we presented a Smart garbage bin system for garbage growth behaviour prediction in Chapter 3. We designed and developed the initial smart garbage bin prototype embedded with ToF (time of flight) and load cell sensors to track the amount of garbage during disposal. Using a Wi-Fi gateway, data were sent to a cloud platform. For evaluation, we deployed the smart garbage bin in a student laboratory over one month. An autoregressive integrated moving average (ARIMA) model was applied, providing an average mean absolute error (MAE) of 5.17 cm and a standard deviation (SD) of 0.33 cm, thus was considered satisfactory accuracy for the garbage growth prediction. Therefore, our prediction model was suitable for predicting future garbage growth behaviour, enhancing flexibility in the garbage collection schedule and the frequency of changing garbage bags in the smart bin.

To examine the second challenge, in Chapter 4, we extended our designed and developed a smart garbage bin system, "SGBS", to track garbage amounts and identify the disposed garbage contents. The smart garbage bin was fastened with a ToF (time of flight) and weight sensors to detect the amount of garbage disposed each time. Then, we designed and developed a garbage annotations application to allow households user to annotate their daily garbage content. The annotation application comprised 4 garbage categories and 10 garbage content identities. To evaluate our approach, we conducted an initial experiment on the smart garbage bin system in three households. Later we identified necessary garbage categories and contents in each category through the user's voice and redesigned the annotations application to have 8 garbage categories and 25 garbage content identities to allow households user to annotate their daily garbage content. Therefore we conducted another experiment on the smart garbage bin system in three households. Our findings show that households' garbage disposal behaviour depends on the amount of garbage, type of garbage contents and the routine of disposing of such garbage content. Finally, we discuss the potential of our system to be scaled in a smart city to influence behaviour change, provide healthier life, and improve garbage management operational efficiency.

For the third challenge, in Chapter 5, we presented a new smart garbage bin system, SGBS, embedded with multiple sensors to solve the problem. We deployed DHT22 (temperature and humidity) and MQ135 gas sensors to know the condition and identify the garbage content disposed of. Then, we introduce a new garbage content estimation method by training a machine learning model using daily collected fuse sensor readingsn combined with detailed household garbage contents annotations to perform garbage classification tasks. For evaluation, we deployed the designed SGBS in five households over one month. As a result, we confirmed that the leave-one-house cross-validation results showed an accuracy of 91% in 5 kitchen waste contents, also, 89% in 5 paper/softbox contents, and 85% in the 8 garbage categories for the classification tasks. Fig. 6.1 demonstrates the big picture of the study by including the general goals of garbage management systems, the identified challenges to be solved, the achieved goals in this study, and what has remained for future work.

Finally, the contributions of this dissertation to academic knowledge are summarized in the following three aspects:

1. As a scientific aspect, this is the first study that clarified the feasibility and appropriate design and development of the SGBS prototype that can support an understanding of household garbage disposal behaviour and identification of daily disposed contents, which help improve garbage classification tasks.

2. This study provides a new automation tool for understanding the lifestyle of families,influencing families' behaviour change in the garbage disposal, providing healthy living and increasing home monitoring.

3. The study offers a tool for policy and decision-making to guide municipal governments and improve smart city services solving social problems such as food security by learning the amount of edible food disposed of in households. Next, as a technical aspect, the three different SGBS prototypes were implemented to build a system operated in practice by households. Finally, as a practical aspect, we conducted 3-week and 5-week experiments using the SGBS prototypes to investigate our hypothesis and demonstrate the system's effectiveness with 40 students in the laboratory and also with 3 and 5 families in Japan.

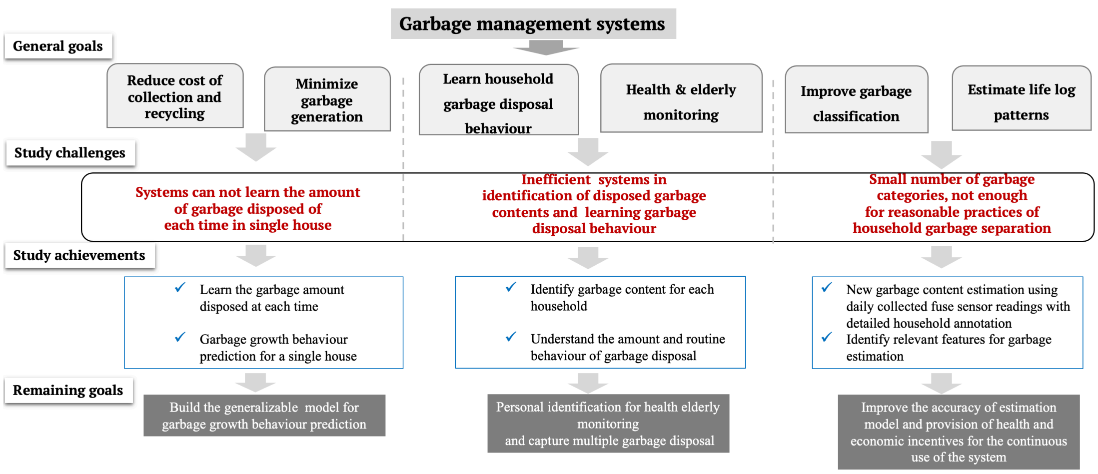


Figure 6.1: Garbage management systems goals, challenges and achievements

6.1.1 Study limitations

Apart from the achievements made in tackling the three challenges of the study, however, the study had some technical challenges and limitations in using the SGBS for garbage content estimation as follows;

6.1.2 SGBS technical challenges

- **Power problem:** During the experiment as detailed in the Chapter 4, we faced a power supply problem. The used solar panel battery could not achieve the power supply for 29 days as it was approximated by our energy-saving algorithm, unsteady it only supplied power to the smart garbage bin for half of the expected 13 days. We also experienced some short circuits caused by the design of the circuit on a breadboard.

- **Network connectivity:** Although the used sigfox antenna showed good area coverage, we still experienced a message delay problem with our server, which persisted for some hours and sometimes changed the transmission interval from 10 minutes as pre-defined in the SGBS to 20 time interval minutes. These technical challenges led to missing some household data points and disturbed sensor readings.

6.1.3 Limitations of using the SGBS for garbage content estimation

- **Generalizable garbage annotation interface:**
  The proposed garbage annotation application interface is designed for the purpose of garbage classification. Therefore it can be used as a generalizable interface by appropriately setting garbage categories, However, for other purposes other than classification tasks remain as the study limitations.

- **Personal identification and capture multiple garbage disposal:**
  The current SGBS captures single garbage disposal and deployed in single and couple households without identifying who disposes of what garbage. However, person identification and capturing of each event of garbage disposal is necessary to separate who disposes of the garbage contents and estimate an individual's life pattern on the garbage disposal to provide health care support or elder monitoring.

- **Few numbers of annotation:**
  The study provided sufficient burnable garbage identification to guide house users during garbage disposal through the mobile application interface. Yet, few annotations were recorded on some garbage categories because of the difference in disposal behaviour in each house. For instance, the low number of plastic, dust and plant categories in houses 1, 2, 3 and 4 (see Table 5.13), therefore, were removed during model building as they were affecting the performance accuracy. For that reason, more garbage annotation is required for additional training data to ensure a robust garbage estimation in application scenarios.

- **Learn correct annotation:**
  Even though the study determined the frequency of annotations for each category in every house, households need to learn and remember to correctly annotate garbage content for each category which is important act in the improvement of the garbage classification tasks.

- **Garbage compression behaviour:**
  The garbage compression in the smart garbage bin by the participant's hand

during the disposal as one way of keeping the garbage bin from reaching its maximum thresholds shifts the values from its steady state, thus disturbing both fill level and weight sensor readings.

- **Use several garbage bins:**
  Households often use several bins for different types of garbage; however, this initial study was based on burnable garbage only.

6.2 Future work

For challenge 1 in chapter 3, the existing garbage management systems can not learn the amount of garbage disposed of each time in a single house. Therefore, this study designed and developed an SGBS to detect the garbage amount disposed at each time and built an ARIMA model to predict garbage growth behaviour for a single house. In the future, we will consider building the generalizable (versatile) model by investigating if the ARIMA can be used as the versatile model or consider using other prediction model.

For challenge 2 in chapter 4, the existing garbage management systems are inefficient in identifying disposed garbage contents and learning garbage disposal behaviour. As a solution, this study designed and developed an SGBS capable of identifying important garbage content in households, understanding the routine behaviour of garbage disposal and the amount of disposed of garbage since the current SGBS design can not identify who disposes of the garbage. The first step in the future, we will consider introducing a name or identification tag embedded on SGBS or engaging other motion identification methods such as the use of pressure, accelerometer and gyroscope to learn the open and close behaviour of smart bin lid to identify and distinguish who disposed of the garbage contents for the estimation of individual's life pattern on the garbage disposal to provide health care support or elder monitoring.

In addition, for challenge 2 in chapter 4, since the current SGBS design captures a single garbage disposal, therefore, to capture multiple garbage disposal in the future system, one of the methods is to add tools such as depth cameras on the current design of the SGBS. Afterwards, combine the detection of the garbage disposal from all devices, such as the sensor readings, smartphone annotations and cameras at each time of garbage disposal.

For challenge 3 in chapter 5, The existing garbage management systems provide a small number of garbage categories, not enough for reasonable practices of household garbage separation and classification tasks. Therefore, in this study we introduced a new garbage content estimation method to solve this challenge by training a machine learning model using daily collected fuse sensor readings combined with detailed household garbage contents annotations to perform the garbage classification task. In the future, we will consider investigating if the machine learning performance can be improved for practical use when users assume a more active role in garbage annotation tasks. However, for purposes like knowing if the garbage content follows the guideline of the municipality, 85% to 91% accuracy, as achieved by the study in Chapter 5, might be satisfactory. Still, an ideal near 100% is needed for the anomaly detection situation.

In addition, for challenge 3, in Chapter 5, we suggest that the system estimate an individual's life log pattern on the garbage disposal to provide health and economic incentives to users for the continuous use of the SGBS.

In general, two ultimate goals can be drawn from this study, first, understand if the garbage content follows the municipality's guidelines and second, use garbage disposed garbage content for health or elderly monitoring. In both goals, our proposed interface design can be used for various groups and places by appropriately setting garbage categories. Therefore, the garbage categories data must be prepared for various groups and places since they are differently defined in different places.

Above all, we suggest a friendly industrial make-up of the SGBS for the ideal deployment of the system to include a customized interface design based on the group of users, installation of a waterproof case on the system, and a constant power supply.

Acknowledgements

Upon completing this dissertation, First and foremost, I would like to thank and Glorify the Almighty God for His countless graces and blessings throughout my studies. Second, I extend my deepest gratitude to all who have made this dissertation possible. I want to thank Professor Keiichi Yasumoto, my supervisor, for all his guidance and advice during my 4-year-and-a-half study in the Ubiquitous Computing Systems Laboratory (Ubi Lab) since 2019. After knowing I was seeking a supervisor in Japan through my Tanzanian friend “Jema -san,” who was studying in the Cyber Resilience laboratory, I appreciated that he kindly accepted me and let me study and work under his supervision. Moreover, I especially thank him for his invaluable advice, continuous support, and patience during my PhD study. His immense knowledge and plentiful experience have encouraged me throughout my academic research and daily life. Also, I thank him for his help with my writing and for carefully reading and commenting on countless revisions of all my manuscripts. Furthermore, his warmness to students and strict attitude to research makes him the most respectful Professor I have ever met. His advice will be beneficial in my future academic activity.

A BIG THANK YOU to Professor Yuki Matsuda for his close technical support of my study. From the beginning of the research, he has been showing me a way through the design and development of a Smart garbage bin system. When I learned about hardware, he provided me with sample sensors and teaching materials on IoT devices and helped me always make the best choices. I appreciate his patience with me from when I knew nothing until my first and last prototype development. When I was troubled by the development of the Garbage annotation application, he provided technical support and helped me complete the interface's development. Furthermore, he highly supported me in the experiment designs and deployments, even after the experiments and data collection. When I did not know how to analyze the data collected from the experiment, he provided teaching materials on statistical analysis and taught me how to use those analysis functions. With his help, I could accomplish my research goals and publish my publications. He is not only my Professor but also a friend. He contributed ideas not only for my research but also for my life here in Japan.

I also thank all Ubi Lab members, especially those who participated in my experiment in their homes and the laboratory. Ubi Lab was a family to me. I thank everyone for the joyous moments we have shared, study discussion, wisdom, encouragement, friendship and support for my life here in Japan. It is your help that made me today.

I want to express my deepest love, miss, and thanks to my family. To my Late beloved mother and father in heaven. Especially my mother, who always insisted I study until graduate school. Even though I was a child, I couldn't understand the meaning of it. So today, I dedicated this achievement to you. Thank you for always being by my side. Finally, My appreciation goes to all my other family members and friends for their encouragement and support throughout my studies.

Bibliography

[1] "What a waste 2.0," \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://datatopics.worldbank.org/what-a-waste/}, accessed: 2022-11-5.

[2] "The new plastics economy: Rethinking the future of plastics and catalysing action," \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://emf.thirdlight.com/link/cap0qk3wwwk0-l3727v/@/#id=1}, accessed: 2022-11-5.

[3] "The global e-waste monitor 2020: Quantities, flows and the circular economy potential," \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://https://collections.unu.edu/view/UNU:7737}, accessed: 2022-11-5.

[4] B. Wang, M. Farooque, R. Y. Zhong, A. Zhang, and Y. Liu, "Internet of things (iot)-enabled accountability in source separation of household waste for a circular economy in china," *Journal of Cleaner Production*, vol. 300, p. 126773, 2021.

[5] N. Nnamoko, J. Barrowclough, and J. Procter, "Solid waste image classification using deep convolutional neural network," *Infrastructures*, vol. 7, no. 4, p. 47, 2022.

[6] D. Ziouzios, D. Tsiktsiris, N. Baras, and M. Dasygenis, "A distributed architecture for smart recycling using machine learning," *Future Internet*, vol. 12, no. 9, p. 141, 2020.

[7] C. Wang, J. Qin, C. Qu, X. Ran, C. Liu, and B. Chen, "A smart municipal waste management system based on deep-learning and internet of things," *Waste Management*, vol. 135, pp. 20–29, 2021.

[8] M. Cubillos, “Multi-site household waste generation forecasting using a deep learning approach,” *Waste Management*, vol. 115, pp. 8–14, 2020.

[9] S. Longhi, D. Marzioni, E. Alidori, G. Di Buo, M. Prist, M. Grisostomi, and M. Pirro, “Solid waste management architecture using wireless sensor network technology,” in *2012 5th International Conference on New Technologies, Mobility and Security (NTMS).* IEEE, 2012, pp. 1–5.

[10] C. K. M. Lee and T. Wu, “Design and development waste management system in hong kong,” in *2014 IEEE International Conference on Industrial Engineering and Engineering Management.* IEEE, 2014, pp. 798–802.

[11] I. Hong, S. Park, B. Lee, J. Lee, D. Jeong, and S. Park, “Iot-based smart garbage system for efficient food waste management,” *The Scientific World Journal*, vol. 2014, 2014.

[12] S. Thakker and R. Narayanamoorthi, “Smart and wireless waste management,” in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS).* IEEE, 2015, pp. 1–4.

[13] S. V. Kumar, T. S. Kumaran, A. K. Kumar, and M. Mathapati, “Smart garbage monitoring and clearance system using internet of things,” in *2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM).* IEEE, 2017, pp. 184–189.

[14] M. Talha, A. Upadhyay, R. Shamim, and M. S. Beg, “A cloud integrated wireless garbage management system for smart cities,” in *2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT).* IEEE, 2017, pp. 175–179.

[15] B. Chowdhury and M. U. Chowdhury, “Rfid-based real-time smart waste management system,” in *2007 Australasian Telecommunication Networks and Applications Conference.* IEEE, 2007, pp. 175–180.

[16] N. S. Kumar, B. Vuayalakshmi, R. J. Prarthana, and A. Shankar, "Iot based smart garbage alert system using arduino uno," in *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2016, pp. 1028–1034.

[17] A. , D. Karadimas, J. Gialelis, and A. G. Voyiatzis, "A versatile scalable smart waste-bin system based on resource-limited embedded devices," in *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*. IEEE, 2015, pp. 1–8.

[18] P. Reis, R. Pitarma, C. Goncalves, and F. Caetano, "Intelligent system for valorizing solid urban waste," in *2014 9th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2014, pp. 1–4.

[19] M. Kannangara, R. Dua, L. Ahmadi, and F. Bensebaa, "Modeling and prediction of regional municipal solid waste generation and diversion in canada using machine learning approaches," *Waste Management*, vol. 74, pp. 3–15, 2018.

[20] S. A. Ali and A. Ahmad, "Forecasting msw generation using artificial neural network time series model: a study from metropolitan city," *SN Applied Sciences*, vol. 1, no. 11, p. 1338, 2019.

[21] N. Sun and S. Chungpaibulpatana, "Development of an appropriate model for forecasting municipal solid waste generation in bangkok," *Energy Procedia*, vol. 138, pp. 907–912, 2017.

[22] J. Ferrer and E. Alba, "Bin-ct: Urban waste collection based on predicting the container fill level," *Biosystems*, vol. 186, p. 103962, 2019.

[23] S. Navghane, M. Killedar, and V. Rohokale, "Iot based smart garbage and waste collection bin," *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, vol. 5, no. 5, pp. 1576–1578, 2016.

[24] G. Prajakta, J. Kalyani, and M. Snehal, "Smart garbage collection system in residential area," *IJRET: International Journal of Research in Engineering and Technology*, vol. 4, no. 3, pp. 122–124, 2015.

[25] M. A. Al Mamun, M. Hannan, A. Hussain, and H. Basri, "Wireless sensor network prototype for solid waste bin monitoring with energy efficient sensing algorithm," in *2013 IEEE 16th International Conference on Computational Science and Engineering*. IEEE, 2013, pp. 382–387.

[26] M. Adam, M. E. Okasha, O. M. Tawfeeq, M. A. Margan, and B. Nasreldeen, "Waste management system using iot," in *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*. IEEE, 2018, pp. 1–4.

[27] H. N. Saha, S. Gon, A. Nayak, S. Moitra *et al.*, "Iot based garbage monitoring and clearance alert system," in *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2018, pp. 204–208.

[28] P. J. Ayaskanta Mishra, Nisha Ghosh, "Internet of things based waste management system for smart cities: A real time route optimization for waste collection vehicles," *International Journal of Computer Sciences and Engineering*, vol. 7, pp. 496–503, 4 2019. [Online]. Available: https://www.ijcseonline.org/full_paper_view.php?paper_id=4064

[29] S. Kristanto, T. Yashiro, and Koshizuka, "Dynamic polling algorithm for low energy garbage level measurement in smart trash bin," in *Proceedings of the Second International Conference on IoT in Urban Space*, 2016, pp. 92–94.

[30] S. Faye, F. Melakessou, and W. Mtalaa, "Swam: A novel smart waste management approach for businesses using iot," in *Proceedings of the 1st ACM International Workshop on Technology Enablers and Innovative Applications for Smart Cities and Communities*, 2019, pp. 38–45.

[31] M. A. Hossain, I. Ahmedy, M. Z. M. Harith, and M. Y. Idris, "Route optimization by using dijkstra's algorithm for the waste management system," in *Proceedings of the 2020 The 3rd International Conference on Information Science and System*, 2020, pp. 110–114.

[32] S. Idwan, I. Mahmood, J. A. Zubairi, and I. Matar, “Optimal management of solid waste in smart cities using internet of things,” *Wireless Personal Communications*, vol. 110, no. 1, pp. 485–501, 2020.

[33] S. K. Memon, A. R. Memon, and A. A. Memon, “Smart garbage bin: An iot platform for smart waste management system in pakistan,” in *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE, 2020, pp. 1–5.

[34] J. Ferrer and E. Alba, “Bin-ct: Urban waste collection based on predicting the container fill level,” *Biosystems*, vol. 186, p. 103962, 2019.

[35] “Managing municipal solid waste 32 european countries eea report no 2/2013issn 1725-9177,” \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://www.eea.europa.eu/publications/managing-municipal-solid-waste}, accessed: 2022-11-5.

[36] W. Guo, B. Xi, C. Huang, J. Li, Z. Tang, W. Li, C. Ma, and W. Wu, “Solid waste management in china: Policy and driving factors in 2004–2019,” *Resources, Conservation and Recycling*, vol. 173, p. 105727, 2021.

[37] “Japanese waste management and recycling industry,” \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://www.env.go.jp/en/index.html}, accessed: 2022-11-5.

[38] “Separation and disposal of garbage in kashihara separation and disposal of garbage in kashiha,” \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://www.city.kashihara.nara.jp/documents/5c34c0f2f1a7f00f31b18cc1}, accessed: 2022-11-5.

[39] “Garbage collection schedule in ikoma city,” \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://www.city.ikoma.lg.jp/cmsfiles/contents/0000005/5895/gomical.pdf}, accessed: 2022-11-5.

[40] “How to divide and take out garbage (for household garbage),” \protect\relax$\@@underline{\hbox{https://www.city.nara.lg.jp/uploaded/attachment/34246.pdf}}\mathsurround\z@$\relax, accessed: 2022-11-5.

[41] “Combustible garbage,” \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{http://kyoto-kogomi.net/wp-content/uploads/2016/04/gg-eng.pdf}, accessed: 2022-11-5.

[42] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[43] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[44] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning,” *Image Recognition*, vol. 7, 2015.

[45] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[46] “Waste classification data,” \protect\relax$\@@underline{\hbox{}}\mathsurround\z@$\relax{https://www.kaggle.com/datasets/techsash/waste-classification-data}, accessed: 2022-11-5.

[47] S. S. Mookkaiah, G. Thangavelu, R. Hebbar, N. Haldar, and H. Singh, “Design and development of smart internet of things–based solid waste management system using computer vision,” *Environmental Science and Pollution Research*, pp. 1–15, 2022.

[48] “trashnet,” \protect\relax$\@@underline{\hbox{https://github.com/garythung/trashnet}}\mathsurround\z@$\relax, accessed: 2022-11-5.

[49] K. N. Sami, Z. M. A. Amin, and R. Hassan, “Waste management using machine learning and deep learning algorithms,” *International Journal on Perceptive and Cognitive Computing*, vol. 6, no. 2, pp. 97–106, 2020.

[50] “Time series forecasting using auto-arima,” https://towardsdatascience.com/time-series-forecasting-using-auto-arima-in-python-bb83e49210cd.

[51] A. Lavric, A. I. Petrariu, and V. Popa, “Long range sigfox communication protocol scalability analysis under large-scale, high-density conditions,” *IEEE Access*, vol. 7, pp. 35 816–35 825, 2019.

[52] E. Likotiko, Y. Matsuda, and K. Yasumoto, “Smart garbage bin: Garbage growth behavior prediction,” *The 28th Multimedia Communication and Distributed Processing System Workshop (DPSWS ’20)*, pp. 27–33, 2020.

Publication List

Journal paper

1. Likotiko, E., Matsuda, Y., Yasumoto, K. (2023). Garbage Content Estimation Using Internet of Things and Machine Learning. IEEE Access.
   Corresponds to Chapter 5 of this dissertation

International Conference

1. Likotiko, Eunice, Shinya Misaki, Yuki Matsuda, and Keiichi Yasumoto. "SGBS: A novel smart garbage bin system for understanding household garbage disposal behaviour." In 2021 Thirteenth International Conference on Mobile Computing and Ubiquitous Network (ICMU), pp. 1-8. IEEE, 2021.
   Corresponds to Chapter 4 of this dissertation

Japanese Conference

1. Likotiko, Eunice, Yuki Matsuda, and Keiichi Yasumoto. "Smart garbage bin: Garbage growth behavior prediction." In The 28th Multimedia Communication and Distributed Processing System Workshop (DPSWS'20), pp. 27-33. 2020.
   Corresponds to Chapter 3 of this dissertation