

Project Name

Project ID: R26-IT-126

Project Proposal Report

Wijesiri A.M – IT22088482

B.Sc. (Hons) Degree in Information Technology Specialized in
Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

March 2026

DECLARATION

I declare that this is my own work, and this proposal does not knowingly incorporate any material previously submitted for a degree or diploma at any other university or higher education institution, nor does it contain any material that has been previously published or written by another person apart from where proper acknowledgment is given in the text.

Name	Student ID	Signature
Wijesiri A.M.	IT22088482	

The above candidate is conducting research for their undergraduate Dissertation under my supervision.

Name	Date	Signature
Uchidha D	20/03/20	

ABSTRACT

Rapid urbanization and population growth have significantly increased the volume of municipal waste generated in modern cities. In many developing countries such as Sri Lanka, conventional waste management systems rely on fixed waste collection schedules that do not consider the actual fill levels of waste bins. As a result, waste bins may overflow before collection or remain partially empty when collection occurs, leading to inefficient resource utilization and increased operational costs. Recent smart waste management solutions have introduced IoT (IoT) technologies to monitor waste levels in bins in real time. However, most of these systems focus on detecting bin status after waste accumulation has occurred rather than predicting future waste levels.

A significant research gap exists in the ability to forecast waste bin fill levels in advance. Existing studies mainly emphasize real-time monitoring using sensors or optimization of collection routes, while limited research focuses on predictive analytics for anticipating when waste bins will become full. Addressing this limitation can enable waste management authorities to adopt proactive collection strategies and prevent overflow situations.

This research proposes a ML-based Waste Level Prediction System that forecasts the expected fill level of smart waste bins for upcoming time periods. The proposed system utilizes IoT sensor data collected from smart bins together with historical waste generation patterns to train predictive models. ML algorithms such as RF and Gradient Boosting will be used to analyze the collected data and estimate when bins are likely to reach critical fill levels.

The expected outcome is an intelligent decision-support system that enables waste management authorities to schedule waste collection more efficiently and prevent bin overflow. The proposed approach aims to improve operational efficiency, reduce unnecessary collection trips, and support the development of more sustainable smart city waste management systems.

Keywords: Smart Waste Management, IoT Sensors, Waste Level Prediction, ML, Predictive Analytics.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my family, especially my parents, for their continuous support, encouragement, and understanding throughout my academic journey. Their patience, motivation, and belief in my abilities have been a constant source of strength that helped me overcome many challenges during my studies.

I would also like to extend my heartfelt appreciation to my supervisors and lecturers for their valuable guidance, constructive feedback, and professional advice during the preparation of this research proposal. Their expertise and encouragement played a significant role in shaping the direction and quality of this research.

My sincere thanks also go to my colleagues and friends who provided support, shared ideas, and engaged in meaningful discussions during the development of this study. Their cooperation and positive contributions made this research process both productive and enjoyable.

Finally, I would like to acknowledge all individuals and organizations who generously shared their knowledge and insights related to this research area. Their support and willingness to assist greatly contributed to enhancing my understanding of the subject.

TABLE OF CONTENT

Table of Contents

DECLARATION	Error! Bookmark not defined.
ABSTRACT	3
ACKNOWLEDGEMENT	4
TABLE OF CONTENT	5
LIST OF TABLES	7
LIST OF FIGURES	8
LIST OF ABBREVIATIONS	9
LIST OF APPENDICES	10
1. INTRODUCTION	10
1.1 Background and Context	10
1.2 Literature Review	11
1.2 Research Gap	13
2. OBJECTIVE	15
2.1 Main Objective	15
2.2 Specific Objectives	15
3. METHODOLOGY	16
3.1 Research Approach	16
3.2 Technical Methods and Algorithms	17
3.3 Tools and Platforms	17
3.4 Data Collection and Ethical Considerations	18
3.5 Validation Strategy and Performance Metrics	19
4. HIGH-LEVEL SYSTEM ARCHITECTURE	20
4.1 Overview of the Proposed System	20
4.2 System Components	22
4.3 Data Flow Explanation	23
4.5 Scalability and Security Considerations	24
5. USER REQUIREMENTS	25
5.1 Stakeholder Identification	25
5.2 Functional Requirements	26
5.3 Non-Functional Requirements	27
6. COMMERCIALIZATION PLAN	28
6.1 Target Market and Customer Persona	28

6.2 Value Proposition	29
6.3 Revenue Model.....	30
6.4 Cost Estimation	30
6.5 Pricing Strategy	31
6.6 Competitive Advantage.....	32
6.7 Intellectual Property Considerations	32
7. BUDGET AND JUSTIFICATION	33
8. WORK BREAKDOWN STRUCTURE (WBS).....	36
9. GANTT CHART	38
10. REFERENCES LIST.....	39
11. APPENDICES.....	40

LIST OF TABLES

Table 1:Research Gap Summary table	14
Table 2:Technologies and Tools that will be used	18
Table 3:Target Market Segments	28
Table 4:Customer Persona	29
Table 5:Potential Revenue Streams.....	30
Table 6:Comparison with Existing Solutions	32
Table 7:Estimated Hardware Costs	33
Table 8:Software Licenses Costs	34
Table 9:Cloud Services Costs	34
Table 10:Estimated Human Effort hours.....	35
Table 11:Work Breakdown Structure of the component	37
Table 12:Workload Distribution	37

LIST OF FIGURES

Figure 1:High-Level Architecture Diagram of the System	21
Figure 2:Component Architecture Diagram	22
Figure 3:Work Breakdown Structure of the overall project	36
Figure 4:Gantt chart	38

LIST OF ABBREVIATIONS

Abbreviation	Meaning
AI	Artificial Intelligence
IoT	Internet of Things
ML	Machine Learning
API	Application Programming Interface
DB	Database
UI	User Interface
FR	Functional Requirement
NFR	Non-Functional Requirement
RF	Random Forest
XGBoost	Extreme Gradient Boosting
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
DSR	Design Science Research
HTTPS	Hypertext Transfer Protocol Secure
SaaS	Software as a Service
SDG	Sustainable Development Goal
RAM	Random Access Memory
AWS	Amazon Web Services

LIST OF APPENDICES

Appendix 1:Detailed System Architecture	40
Appendix 2:Dashboard UI Mockups	40
Appendix 3:IoT Hardware Setup	41
Appendix 4:Risk Analysis Diagram.....	41
Appendix 5:Data Collection Process	42

1. INTRODUCTION

1.1 Background and Context

Municipal solid waste management has become a critical environmental and operational challenge in rapidly urbanizing regions around the world. Increasing population growth, urban development, and changing consumption patterns have significantly increased the volume of waste generated in cities. Developing countries in particular face considerable challenges in managing waste effectively due to limited infrastructure, inefficient collection strategies, and the lack of intelligent technologies to support data-driven decision making [2],[11]. In many regions, waste collection operations still rely on fixed schedules and manual monitoring rather than intelligent systems that can anticipate waste accumulation.

Sri Lanka faces similar challenges in municipal waste management. According to national environmental reports, the country generates approximately 7,000–8,000 metric tons of municipal solid waste per day, with a large proportion originating from urban regions such as Colombo and other metropolitan areas . A considerable share of this waste consists of organic materials, plastics, paper, and other recyclable resources. However, waste collection in many areas still follows predetermined schedules without considering the actual fill levels of waste bins. This often results in inefficient collection operations where some bins overflow before collection while others remain partially empty when waste collection vehicles arrive.

These inefficiencies affect multiple stakeholders including municipal councils, waste management authorities, waste collection service providers, and urban communities. Poorly planned waste collection increases fuel consumption, operational costs, and environmental impact while also contributing to overflowing bins and public health risks [3],[11]. In addition, inefficient waste collection may increase landfill dependency and reduce opportunities for sustainable waste management practices.

Recent advancements in smart waste management have introduced IoT technologies that enable real-time monitoring of waste bins using sensors. These sensors can measure parameters such as bin fill levels, weight, temperature, and gas emissions to provide real-time data about waste conditions [1],[4]. Such technologies allow waste management authorities to detect when bins are nearing capacity and improve operational visibility. However, most

existing smart waste management systems remain reactive because they primarily detect waste conditions only after waste accumulation has occurred.

To address this limitation, predictive analytics techniques can be used to anticipate future waste accumulation patterns. By analyzing historical waste data collected from IoT-enabled smart bins, ML models can learn patterns in waste generation and estimate when waste bins are likely to reach critical fill levels. Algorithms such as RF and Gradient Boosting can be applied to analyze historical bin data and predict future waste levels based on time patterns and environmental factors.

This research therefore proposes a ML-based Waste Level Prediction System that forecasts the expected fill levels of smart waste bins in advance. By predicting when bins will reach critical capacity, the system can support proactive waste collection planning, reduce bin overflow incidents, and improve operational efficiency. The proposed system aims to transform waste collection operations from a reactive process into a proactive and intelligent waste management approach.

This research also contributes to global sustainability initiatives such as the United Nations SDGs. The proposed system supports SDG 11: Sustainable Cities and Communities by improving urban waste management efficiency and SDG 12: Responsible Consumption and Production by promoting more efficient resource utilization and waste management practices. From a technological perspective, this study falls within the Technology Integration and Management cluster in Information Technology, where predictive analytics and IoT technologies are applied to address environmental sustainability challenges.

Overall, the proposed system aims to improve waste collection efficiency by integrating IoT-based monitoring with ML-based prediction. By enabling proactive waste management planning, the system has the potential to reduce operational costs, prevent bin overflow, and support more sustainable urban waste management practices.

1.2 Literature Review

Recent literature shows that smart waste management research has gradually shifted from traditional waste collection practices toward intelligent and data-driven waste management systems. Early studies mainly focused on estimating the total amount of municipal solid waste generated rather than predicting waste accumulation at the bin level. Abbasi and El Hanandeh developed AI-based models for municipal solid waste forecasting and demonstrated that ML approaches can outperform traditional statistical methods when modeling complex and nonlinear waste generation patterns [6]. Similarly, Kannangara et al. applied ML techniques to predict regional municipal solid waste generation and diversion using demographic and socioeconomic variables, showing that predictive analytics can significantly support waste management planning and policy decisions [7].

More recent research has extended waste prediction by incorporating additional contextual and spatial factors. Lu et al. proposed a multi-city ML model for municipal solid waste generation forecasting, demonstrating that predictive accuracy improves when models are

trained using datasets from multiple urban environments rather than a single locality [6]. Wan et al. further showed that district-level forecasting models can provide better decision support for municipal authorities by capturing local variations in waste generation patterns [7]. These studies indicate that ML models can effectively analyze complex patterns in waste generation when multiple contextual factors are considered.

Another line of research has investigated the application of advanced ML models such as artificial neural networks and deep learning architectures for waste prediction. Puntarić et al. developed an artificial neural network model to estimate separated municipal solid waste fractions and demonstrated that ANN-based approaches can produce reliable predictions when socio-demographic and economic variables are available [6]. Zhang et al. introduced a deep learning model based on Bidirectional Long Short-Term Memory (BiLSTM) for municipal solid waste generation forecasting and reported that temporal models can better capture dynamic relationships among influencing variables over time [7]. Similarly, Liu et al. applied an LSTM-based forecasting model using multiple input indicators and found that deep learning techniques can improve predictive accuracy in complex urban waste management environments [8]. These findings highlight the potential of ML and deep learning approaches in predicting future waste accumulation patterns.

In parallel with forecasting research, many studies have explored the use of IoT technologies for smart waste management. IoT-enabled smart bins equipped with sensors can measure parameters such as fill level, weight, temperature, and gas emissions, enabling real-time monitoring of waste conditions. Cheema et al. reviewed smart waste management and classification systems and found that many existing solutions focus on bin-level sensing, route optimization, and automated waste classification using ML techniques [2],[11]. Fang et al. also reported that AI applications in smart-city waste management include smart bins, waste monitoring platforms, robotic sorting systems, and waste-to-energy optimization models, although the level of technological maturity varies across these applications [2]. While these systems significantly improve operational visibility, they mainly focus on detecting waste conditions after waste accumulation occurs.

Recent review studies further emphasize the growing importance of predictive analytics in waste management systems. Pillai and Singh observed that ML applications in municipal waste management have primarily been used for waste generation forecasting, automated sorting, and route optimization, while integrated predictive waste monitoring systems remain limited [7]. Dawar et al. similarly reported that ML and deep learning models are increasingly applied in waste generation forecasting and logistics optimization, but many existing implementations do not fully integrate predictive capabilities with real-time monitoring systems [8]. This observation is particularly relevant to the present research because it suggests that predictive models for anticipating future bin fill levels are still underexplored.

Commercial smart waste management platforms demonstrate a similar trend. Solutions such as Bigbelly, Sensoneo, Ecube Labs, and Compology provide smart waste management technologies including real-time bin monitoring, sensor-based waste detection, route optimization, and analytics dashboards. These systems help waste management authorities reduce bin overflow, optimize collection schedules, and improve operational efficiency. However, most commercial platforms primarily focus on monitoring current bin status rather than predicting future waste accumulation. As a result, waste collection planning remains largely reactive instead of proactive.

Overall, the literature shows substantial progress in three main areas: municipal waste generation forecasting, IoT-based waste monitoring, and intelligent waste management platforms. However, most studies either predict total waste generation at the city level, monitor bin conditions in real time, or optimize collection routes after waste accumulation has occurred [1]–[12]. Likewise, commercial smart waste platforms mainly focus on monitoring and logistics rather than predicting future bin fill levels. Therefore, there is a clear need for a ML-based approach that integrates IoT monitoring data with predictive analytics to forecast waste accumulation at the bin level. Such a system could enable proactive waste collection planning, reduce overflow incidents, and improve the overall efficiency of urban waste management operations.

1.2 Research Gap

Although significant progress has been made in the development of smart waste management systems, several limitations remain in current research and practical implementations. Many existing studies focus on improving waste monitoring, collection logistics, and municipal waste quantity forecasting rather than predicting waste accumulation at the bin level. For example, several ML-based studies have been developed to estimate the total amount of municipal solid waste generated within cities using demographic, socioeconomic, or environmental variables [6],[7],[8]. While such models are useful for long-term policy planning and strategic waste management decisions, they do not provide operational insights regarding when individual waste bins are likely to reach capacity.

Furthermore, recent studies have explored advanced prediction techniques such as artificial neural networks and deep learning architectures for municipal waste forecasting [6],[7]. These approaches demonstrate improved predictive accuracy by capturing complex relationships between influencing variables and waste generation patterns. However, the majority of these models focus on predicting overall waste quantities or regional waste trends rather than forecasting the fill levels of individual waste bins. Consequently, their ability to support real-time operational decision making for waste collection scheduling remains limited.

Another important limitation can be observed in currently deployed smart waste management systems. Many commercial platforms, including Bigbelly, Sensoneo, Ecube Labs, and Compology, utilize IoT sensors to monitor bin fill levels, detect overflow conditions, and optimize waste collection routes. Although these systems improve operational visibility and collection efficiency, they are fundamentally reactive systems because they respond only after waste has already accumulated inside the bins. Waste collection operations therefore still depend largely on real-time monitoring rather than predictive planning.

The lack of predictive models capable of forecasting future waste bin fill levels represents a significant research gap in smart waste management. Without the ability to anticipate when waste bins will reach critical capacity, waste collection authorities must

rely on fixed schedules or reactive monitoring, which may lead to inefficient routing, unnecessary collection trips, or bin overflow incidents.

Therefore, this research proposes a ML–based Waste Level Prediction System designed to forecast the expected fill levels of smart waste bins in advance. By analyzing historical IoT sensor data together with temporal and contextual factors, the proposed system aims to predict when waste bins are likely to reach critical capacity. This predictive capability enables waste management authorities to plan collection schedules proactively rather than reacting to overflow situations.

The novelty of this research lies in integrating IoT-based waste monitoring with ML–based predictive analytics to forecast future bin fill levels. By addressing the identified limitations in existing monitoring and forecasting systems, the proposed approach introduces a predictive intelligence layer that can support proactive waste collection planning and improve the efficiency and sustainability of urban waste management operations.

Table 1: Research Gap Summary table

Research Area	Existing Approaches	Identified Limitations	Proposed Contribution
Waste Generation Forecasting	Many studies use ML models to forecast the total volume of municipal solid waste using demographic, socioeconomic, and environmental variables [6],[7],[8].	These models mainly estimate overall waste generation at city or regional levels and provide limited insight into waste accumulation at the bin level.	The proposed research utilizes IoT sensor data together with historical waste generation patterns to predict waste accumulation in individual bins.
Deep Learning-Based Waste Prediction	Advanced techniques such as artificial neural networks and deep learning models are applied to improve prediction accuracy of waste generation trends [6],[7],[8].	Most models focus on predicting overall waste quantity rather than forecasting the fill level of individual waste bins in real-time waste management systems.	This research applies ML algorithms to predict future bin fill levels using historical IoT sensor data.
Smart Waste Monitoring Systems	Commercial systems such as Bigbelly, Sensoneo, Ecube Labs, and Compology provide IoT-based monitoring of bin fill levels and collection logistics [1],[2],[4],[11],[12].	These systems provide real-time monitoring but operate reactively after waste has already accumulated in bins.	The proposed system introduces predictive analytics to forecast when bins are likely to reach critical capacity.
Recycling Operation Planning	Waste collection schedules are often fixed or adjusted based on real-time monitoring data[9],[10].	Reactive collection strategies may lead to inefficient routing, unnecessary trips, or bin overflow situations.	The proposed system enables proactive waste collection scheduling based on predicted bin fill levels.
Intelligent Waste Management Research	Most intelligent waste management studies focus on IoT monitoring, route optimization, and waste quantity estimation[2],[11].	Limited research integrates IoT monitoring with predictive ML models for forecasting bin fill levels..	The proposed research integrates IoT-based monitoring data with ML models to support proactive waste collection planning..

2. OBJECTIVE

2.1 Main Objective

The main objective of this research is to design, develop, and evaluate a ML-based waste level prediction system capable of forecasting the future fill levels of smart waste bins using historical IoT sensor data and contextual factors such as time patterns, environmental conditions, and location characteristics.

The proposed system aims to generate accurate predictions of bin fill levels in order to support proactive waste collection planning and improve the operational efficiency of waste management systems. By predicting when waste bins are likely to reach critical capacity, the system can assist waste management authorities in scheduling collection activities more effectively and preventing bin overflow situations.

The performance of the predictive model will be evaluated using established ML evaluation metrics such as MAE and RMSE to measure prediction accuracy and assess the reliability of the forecasting system.

2.2 Specific Objectives

- To collect and preprocess historical waste bin data obtained from IoT-enabled smart bins, including sensor measurements such as bin fill levels and timestamps, together with relevant contextual information such as weather conditions and location characteristics.
- To analyze historical waste accumulation patterns and construct meaningful predictive features such as time of day, day type (weekday, weekend, and holiday), environmental conditions, and past waste generation trends that influence bin fill levels.
- To design and train ML models, such as RF and Gradient Boosting algorithms, capable of predicting future waste bin fill levels based on historical IoT sensor data.
- To evaluate the performance and reliability of the developed prediction models using appropriate ML evaluation metrics such as MAE, RMSE, and cross-validation techniques.
- To develop a visualization interface or dashboard that displays predicted waste bin fill levels to support waste collection planning and operational decision making.
- To analyze how predictive waste level insights can improve waste collection efficiency by enabling proactive scheduling and reducing bin overflow incidents in smart waste management systems.

3. METHODOLOGY

This research adopts a data-driven approach to design and evaluate a ML-based system capable of predicting future waste bin fill levels in smart waste management systems. The proposed methodology integrates IoT-based data collection, data preprocessing, ML model development, and experimental evaluation in order to generate reliable predictive insights that support proactive waste collection planning.

IoT-enabled smart bins equipped with sensors provide continuous measurements of waste accumulation levels. These sensor readings, together with historical waste data and contextual factors such as time patterns and environmental conditions, are used to construct a structured dataset suitable for ML analysis. The collected data is then processed and analyzed to identify patterns in waste accumulation behavior.

ML algorithms are applied to learn relationships between historical sensor readings and future bin fill levels. The developed prediction model aims to estimate when waste bins are likely to reach critical capacity, allowing waste management authorities to schedule collection activities more efficiently and prevent bin overflow situations.

The overall methodology therefore combines IoT-based monitoring with predictive analytics techniques to transform traditional reactive waste management practices into a proactive and intelligent decision-support system. The performance of the proposed predictive model will be evaluated using standard ML evaluation metrics to assess its accuracy and reliability in forecasting waste bin fill levels.

3.1 Research Approach

The study follows a DSR approach combined with experimental ML evaluation. DSR focuses on developing and evaluating technological artifacts that address real-world problems. In this research, the artifact is a ML-based predictive system designed to forecast future waste bin fill levels using IoT sensor data, which supports proactive decision-making in smart waste management operations.

The research process consists of several phases including problem identification, data collection and preparation, predictive model development, experimental evaluation, and performance validation. IoT-enabled smart bins generate sensor data that captures waste accumulation levels over time. This data is collected and organized to form a structured dataset suitable for ML analysis.

ML techniques are applied to identify patterns within historical waste bin sensor readings together with contextual variables such as time patterns and environmental conditions. The trained models are then used to forecast future bin fill levels and estimate when bins are likely to reach critical capacity.

Experimental evaluation is conducted to measure the prediction accuracy and reliability of the developed models. Standard ML evaluation metrics such as MAE and RMSE are used to

assess model performance. The results of the evaluation provide insights into the effectiveness of predictive analytics in improving proactive waste collection planning within smart waste management systems.

3.2 Technical Methods and Algorithms

The proposed system utilizes supervised ML techniques to predict future waste bin fill levels using historical IoT sensor data. The prediction task is treated as a regression problem, where the model estimates the expected fill level of a smart waste bin for a future time period. By forecasting bin capacity levels in advance, the system can help waste management authorities schedule collection activities more efficiently and prevent bin overflow situations.

Two ML algorithms are selected for this study due to their strong performance in predictive analytics tasks.

RF Regression

RF is an ensemble learning algorithm that constructs multiple decision trees and combines their predictions to improve accuracy and reduce overfitting. It is particularly effective in handling nonlinear relationships and complex datasets with multiple influencing variables. In this research, RF is used to analyze historical IoT sensor data such as bin fill levels and time patterns in order to predict future waste accumulation levels [6],[7].

Extreme Gradient Boosting

XGBoost is an advanced gradient boosting algorithm that builds a sequence of decision trees and optimizes model performance through iterative learning. XGBoost is widely used in predictive analytics due to its high accuracy, computational efficiency, and ability to handle large datasets. In this study, XGBoost is applied to improve prediction accuracy by learning complex relationships between historical bin data and contextual variables [6],[7].

Both algorithms analyze historical waste bin sensor data together with contextual factors such as time patterns, environmental conditions, and location characteristics to identify patterns influencing waste accumulation. By learning these patterns, the models can estimate future bin fill levels and predict when waste bins are likely to reach critical capacity.

3.3 Tools and Platforms

The implementation of the proposed system will utilize widely adopted data science tools, ML frameworks, and web technologies to support the development of a predictive waste bin monitoring system. These tools enable efficient data processing, ML model development, visualization of prediction results, and deployment of the system for real-world use.

Table 2: Technologies and Tools that will be used

System Component	Tools / Technologies
Programming Language	Python
Data Processing Libraries	Pandas, NumPy
ML Frameworks	Scikit-learn, XGBoost, TensorFlow
Data Visualization	Matplotlib, Seaborn, Plotly
Dashboard / Web Interface	Streamlit, MERN Stack (MongoDB, Express, React, Node.js), Flask
Data Storage	MongoDB
External Data Sources	Weather APIs (OpenWeatherMap, WeatherStack)
Development Environment	Jupyter Notebook, Google Colab, VS Code
Deployment Platform	Cloud platforms such as AWS, Microsoft Azure, or Google Cloud

These technologies are selected due to their reliability, scalability, and strong community support. Python-based ML frameworks enable efficient model development and experimentation, while modern web frameworks allow the integration of predictive results into interactive dashboards. The use of cloud deployment platforms further enables the system to scale and support real-time data processing for smart waste management applications.

3.4 Data Collection and Ethical Considerations

The predictive model requires multiple sources of data to identify patterns in waste accumulation within smart waste bins. The dataset will include several types of data collected from IoT-enabled monitoring systems and contextual sources.

The dataset will include:

- Historical waste bin fill-level data collected from IoT-enabled smart bins equipped with sensors such as ultrasonic sensors that measure the level of waste inside the bin over time.
- Timestamp information associated with each sensor reading, including date and time of measurement, which enables the analysis of temporal patterns in waste accumulation.

- Contextual environmental data such as temperature, rainfall, and humidity obtained from weather APIs (e.g., OpenWeatherMap or WeatherStack), which may influence waste generation patterns.
- Calendar information including weekdays, weekends, public holidays, and special events that may affect waste disposal behavior.
- Location characteristics such as area type (e.g., residential areas, university campuses, or commercial zones) that influence the rate of waste generation.

In situations where real-world sensor data is limited, simulated datasets may be generated based on realistic waste accumulation patterns and statistics reported in previous research studies [6],[7]. These datasets will be structured to resemble real IoT sensor data in order to support ML model development and experimental evaluation.

From an ethical perspective, the research primarily utilizes environmental and operational data rather than personal or sensitive information, which minimizes privacy risks. Nevertheless, appropriate data management practices will be followed to ensure secure storage, responsible data usage, and compliance with institutional research guidelines and ethical standards.

3.5 Validation Strategy and Performance Metrics

The predictive models will be evaluated using standard ML validation techniques in order to assess the accuracy, reliability, and generalization capability of the developed waste level prediction system.

Cross-Validation

K-fold cross-validation will be used to ensure that the model generalizes well to unseen data and to reduce the risk of overfitting. In this method, the dataset is divided into multiple subsets (folds), where the model is trained on a portion of the data and tested on the remaining portion. This process is repeated several times so that each subset is used as a testing set at least once.

Performance Metrics

The prediction performance of the developed models will be evaluated using several standard regression evaluation metrics:

- **MAE:** Measures the average absolute difference between predicted and actual bin fill levels. Lower MAE values indicate better prediction accuracy.
- **RMSE:** Evaluates prediction accuracy by penalizing larger prediction errors more heavily. RMSE is useful for identifying models that produce large deviations from actual values.

• **Coefficient of Determination** : Measures how well the model explains the variance in the dataset. Higher R^2 values indicate that the model better captures patterns in waste accumulation data.

A **comparative analysis** will be conducted between the RF and XGBoost models to determine which algorithm provides the best predictive performance for **waste bin fill level forecasting**.

Through this methodology, the research aims to develop a reliable ML-based prediction system capable of supporting proactive waste collection planning and improving operational efficiency in smart waste management systems.

4. HIGH-LEVEL SYSTEM ARCHITECTURE

4.1 Overview of the Proposed System

The proposed Waste Bin Fill-Level Prediction System is designed as a data-driven predictive analytics module within a smart waste management framework. The main purpose of this component is to predict future waste bin fill levels using historical IoT sensor data together with contextual factors such as weather conditions, time patterns, and location characteristics.

IoT-enabled smart bins continuously collect waste level measurements using sensors. These measurements are transmitted to a cloud-based data platform where the collected data is processed and prepared for ML analysis. The predictive models analyze historical patterns in waste accumulation and estimate when bins are likely to reach critical capacity.

By predicting waste bin fill levels in advance, waste management authorities can schedule waste collection proactively, prevent bin overflow situations, and optimize waste collection operations.

The architecture of the proposed system follows a layered structure consisting of five major components:

1. Data Sources Layer – collects IoT sensor data and contextual information.
2. Data Processing Layer – prepares and cleans the collected data for analysis.
3. ML Prediction Engine – trains predictive models to forecast future bin fill levels.
4. Prediction Output Module – generates predicted waste bin levels and alerts.
5. Visualization and Monitoring Interface – displays prediction results through a dashboard to support decision making.

These components collectively enable the system to collect relevant data, analyze waste accumulation patterns using ML techniques, generate predictive insights, and present the results in a user-friendly format for proactive waste collection planning.

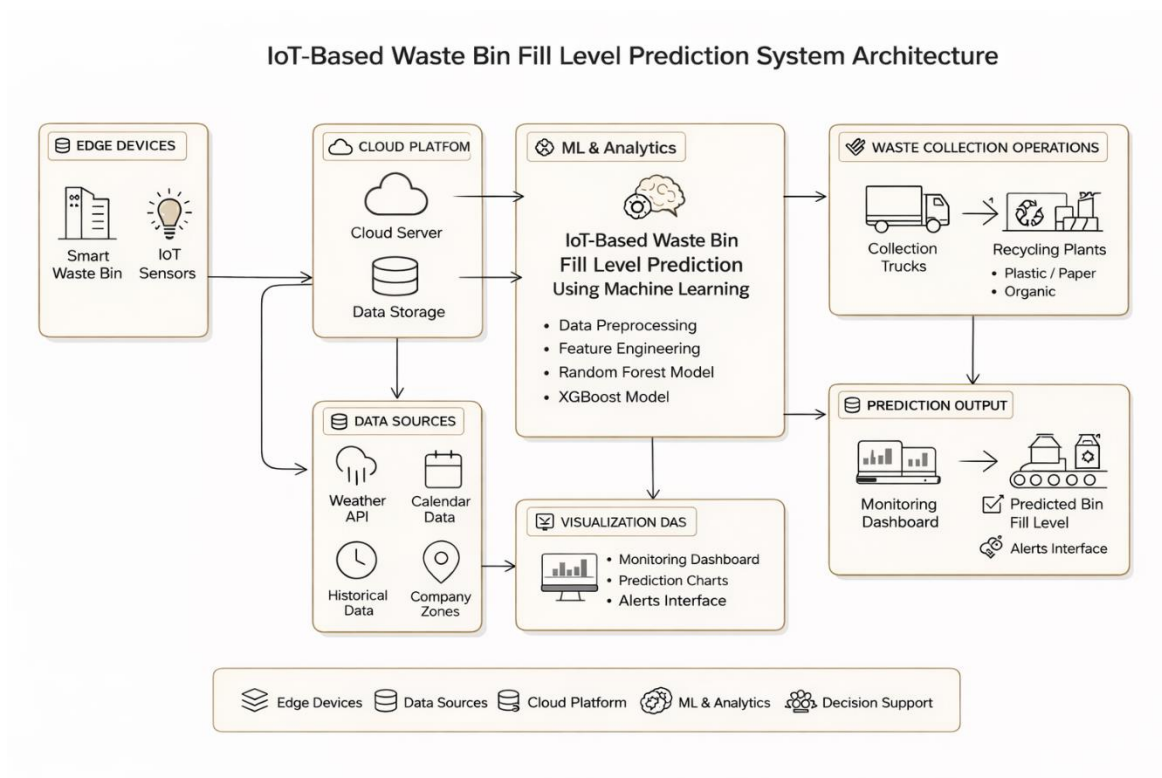


Figure 1: High-Level Architecture Diagram of the System

4.2 System Components

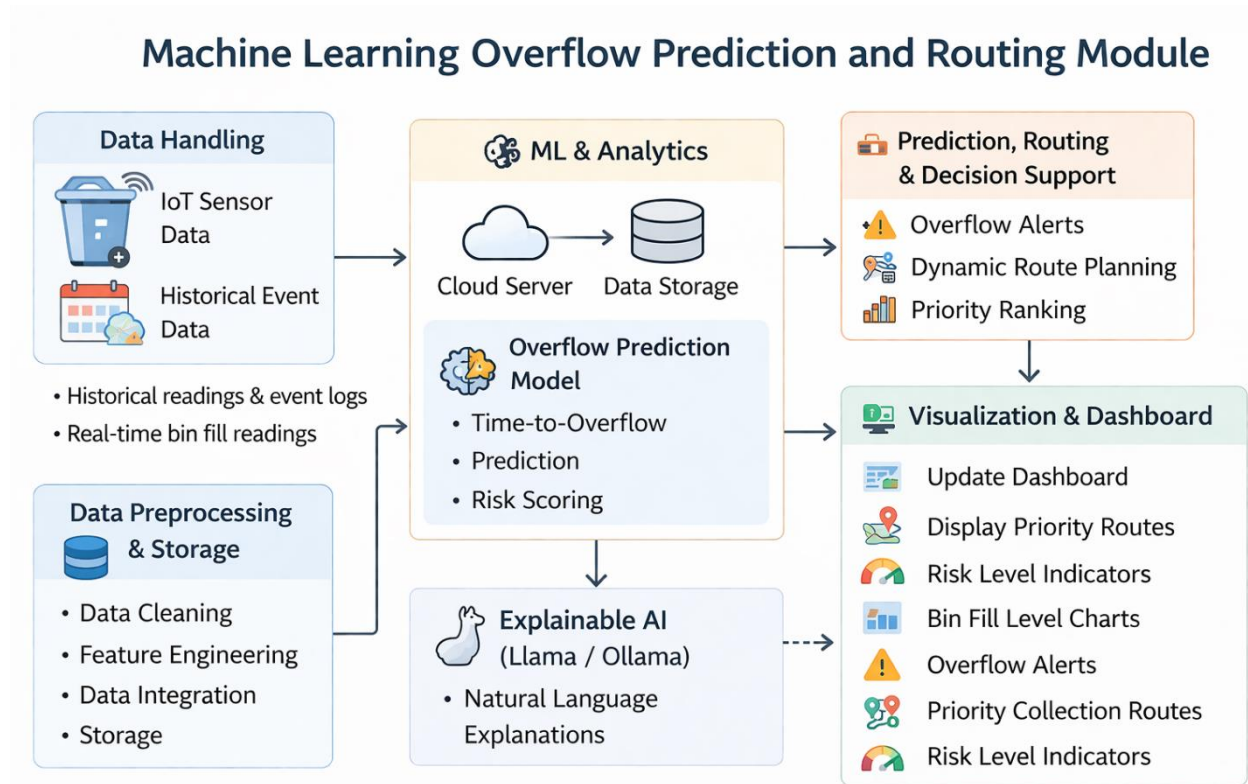


Figure 2: Component Architecture Diagram

The architecture of the proposed Waste Composition Forecasting System consists of several interconnected components that work together to collect data, process information, generate predictions, and present results for decision-making. The main components of the system are described below.

Data Sources Layer

The Data Sources layer provides the input data required for the overflow prediction process. The system receives data from IoT-enabled smart bins and historical operational records. These inputs include real-time bin fill level readings collected from IoT sensors, historical bin usage patterns, timestamps, and contextual operational data such as location identifiers and collection history. The collected data provides valuable information about waste accumulation patterns and enables the ML models to learn how quickly different bins reach critical capacity levels.

Data Processing and Feature Engineering Module

The Data Processing and Feature Engineering module prepares the collected data for ML analysis. This component performs several preprocessing operations including data cleaning,

handling missing values, normalization, and transformation of raw sensor readings. Feature engineering techniques are used to derive predictive variables such as fill rate, time since last collection, historical bin usage trends, and time-based patterns such as daily and weekly waste accumulation cycles. These engineered features help the ML model better understand waste accumulation behavior and improve prediction accuracy.

ML Prediction Engine

The ML Prediction Engine is the core component of the proposed system. This module applies supervised ML algorithms to analyze historical bin data and predict future waste accumulation levels. In this research, algorithms such as Random Forest Regression and XGBoost are used to identify complex relationships between bin fill levels, time patterns, and contextual variables. The trained models generate predictions indicating when a waste bin is likely to reach critical capacity or overflow conditions.

Forecasting Output Module

After the ML model generates predictions, the Prediction and Decision Support module converts these predictions into actionable operational insights. The predicted results include estimated time-to-overflow and overflow risk levels for each bin. Based on these predictions, bins can be categorized into priority levels such as low risk, medium risk, and high risk. This information supports intelligent waste collection planning by identifying bins that require urgent collection and those that can be scheduled later.

Visualization and Decision Support Dashboard

The Visualization and Decision Support Dashboard presents the prediction results through an interactive web-based interface. The dashboard displays bin fill level charts, overflow alerts, and priority collection routes generated from the prediction module. Risk level indicators help waste management operators quickly identify critical bins that require immediate attention. The dashboard also provides route planning recommendations to help waste collection teams optimize their collection schedules and reduce operational inefficiencies.

4.3 Data Flow Explanation

The data flow within the proposed Waste Composition Forecasting System follows a structured sequence of processing stages, beginning with data collection and ending with prediction visualization for decision support.

1. Data Collection

Raw data is collected from multiple data sources including historical waste collection records, weather information obtained through external weather APIs, calendar data such as weekdays and holidays, and location characteristics such as residential or commercial zones. These datasets provide contextual information that influences waste generation patterns.

2. **Data Storage and Preprocessing**

The collected data is stored in the system's data storage environment. Preprocessing techniques are applied to clean the data, handle missing values, remove inconsistencies, and transform the raw datasets into structured formats suitable for ML analysis.

3. **Feature Engineering**

After preprocessing, feature engineering techniques are applied to generate meaningful predictive variables. These features may include weather attributes, date-based indicators, and historical waste generation trends that help the ML model understand patterns influencing waste composition.

4. **ML Prediction**

The processed dataset is then provided to the ML prediction engine. Algorithms such as RF and XGBoost analyze the relationships between contextual variables and waste generation patterns. Based on these patterns, the model generates predictions for the expected composition of waste categories in the next collection cycle.

5. **Prediction Visualization and Decision Support**

The predicted results are transmitted to the visualization dashboard where they are displayed through charts and graphical interfaces. Waste management authorities can use this information to understand expected waste composition patterns and prepare recycling operations accordingly.

Through this structured data flow, the system transforms raw environmental and operational data into actionable predictive insights that support efficient waste management planning.

4.5 Scalability and Security Considerations

The proposed Waste Composition Forecasting System is designed with scalability and security in mind in order to support future expansion and reliable system operation. As waste management systems evolve and data volumes increase, the architecture must be capable of handling larger datasets and more complex operational environments.

Scalability

The architecture is designed to support scalable data processing and model deployment. Since the system utilizes modern data science tools and ML frameworks, it can be extended to

process larger datasets and support multiple geographical regions. Cloud-based storage and processing platforms can be integrated to manage increasing volumes of waste collection data, weather information, and contextual variables.

In addition, the ML models used in the system, such as RF and XGBoost, can be retrained periodically as new data becomes available. This enables the system to continuously improve prediction accuracy and adapt to changing waste generation patterns. The modular architecture of the system also allows additional data sources, sensors, or predictive models to be integrated without significantly modifying the overall system structure.

Security and Data Integrity

Although the system mainly processes environmental and operational data rather than personal information, maintaining data security and integrity is still essential. Secure communication protocols should be used when accessing external data sources such as weather APIs. Access control mechanisms should also be implemented within the system dashboard to ensure that only authorized users can access or modify system data.

Furthermore, data protection measures such as secure data storage, periodic system monitoring, and regular data backups should be implemented to prevent data loss and ensure system reliability. These security practices help maintain the integrity and availability of the system while supporting safe and reliable operation.

Summary

Overall, the proposed architecture provides a scalable and secure framework for implementing predictive waste composition forecasting. By supporting flexible system expansion and implementing appropriate data protection mechanisms, the architecture is capable of supporting modern smart waste management systems that rely on data-driven decision-making.

5. USER REQUIREMENTS

5.1 Stakeholder Identification

The proposed IoT-Based Waste Bin Fill Level Prediction System supports several stakeholders involved in municipal waste management operations. These stakeholders use the system to obtain predictive insights that improve waste collection planning, operational efficiency, and resource utilization within smart waste management environments.

Municipal Waste Management Authorities

Municipal councils and government waste management departments are responsible for managing public waste collection services and maintaining urban waste infrastructure. These authorities can use the proposed system to monitor predicted waste bin fill levels across different locations. By identifying bins that are expected to reach critical capacity, municipal authorities can plan waste collection schedules more effectively and prevent overflow

incidents. This enables more efficient resource allocation and improves the overall effectiveness of waste management operations.

Waste Collection Service Providers

Waste collection companies, whether public or private, are responsible for transporting waste from bins to waste processing or disposal facilities. The prediction system provides insights into future bin fill levels and potential overflow risks. These predictions allow collection service providers to optimize collection routes, prioritize bins that require immediate attention, and reduce unnecessary collection trips. As a result, operational costs, fuel consumption, and time inefficiencies can be significantly reduced.

Recycling and Waste Processing Facilities

Recycling plants and waste processing facilities depend on the timely arrival of collected waste for processing and material recovery. Predictive insights generated by the system help these facilities anticipate incoming waste volumes based on predicted bin fill levels. This information enables better preparation of processing equipment, workforce allocation, and operational scheduling.

Smart City and Urban Planning Authorities

Urban planners and smart city initiatives aim to develop sustainable and efficient urban infrastructure. The predictive waste management insights generated by the system can help planners identify areas with high waste generation patterns and evaluate the performance of existing waste collection strategies. These insights support the development of long-term waste management policies and contribute to building smarter and more sustainable cities.

5.2 Functional Requirements

FR1: The system shall collect waste bin fill-level data from IoT sensors installed in smart waste bins.

FR2: The system shall collect and integrate historical waste bin fill-level records from the waste management DB.

FR3: The system shall retrieve contextual environmental data such as temperature, rainfall, and humidity using external weather APIs.

FR4: The system shall incorporate calendar-based contextual information including weekdays, weekends, holidays, and special events that may influence waste generation patterns.

FR5: The system shall preprocess and transform the collected data through data cleaning, normalization, and feature engineering to create structured datasets suitable for ML analysis.

FR6: The system shall train ML models such as RF Regression and XGBoost Regression using historical waste bin data and contextual variables.

FR7: The system shall generate predictions of future waste bin fill levels and estimate the likelihood of bin overflow.

FR8: The system shall present predicted bin fill levels and overflow alerts through an interactive visualization dashboard.

FR9: The system shall allow users to monitor predicted bin status, view historical trends, and obtain insights that support efficient waste collection planning.

5.3 Non-Functional Requirements

Security Requirements

The system shall ensure secure access to the monitoring and prediction dashboard through authentication mechanisms. Access to system data and prediction results shall be restricted to authorized users only. Data retrieved from external services such as weather APIs shall be transmitted using secure communication protocols (e.g., HTTPS) to protect data integrity and prevent unauthorized access.

Performance Requirements

The system shall generate waste bin fill-level predictions within an acceptable response time to support efficient waste collection planning. ML models shall be optimized to ensure that prediction results are generated quickly without excessive computational delays, even when processing larger datasets.

Scalability Requirements

The system architecture shall support scalability to accommodate increasing volumes of waste bin data and additional geographic regions. The architecture should allow integration with cloud-based infrastructure to support larger datasets, additional IoT devices, and expanded waste management operations as the system grows.

Usability Requirements

The system shall provide a user-friendly dashboard interface that presents predicted bin fill levels and overflow alerts in a clear and understandable format. Visualizations such as charts, graphs, and status indicators shall be used so that users without advanced technical knowledge can easily interpret prediction results and take appropriate action.

Reliability Requirements

The system shall maintain consistent prediction performance and reliable data processing across different operational conditions. Mechanisms such as regular data backups, system monitoring, and error-handling procedures shall be implemented to ensure system stability and prevent data loss.

6. COMMERCIALIZATION PLAN

6.1 Target Market and Customer Persona

The proposed IoT-Based Waste Bin Fill Level Prediction System has strong commercial potential for organizations involved in waste collection, public sanitation, and smart city management. The system is especially valuable for institutions that require proactive waste collection planning, overflow prevention, and efficient resource utilization in waste management operations.

Table 3: Target Market Segments

Target Market	Description	Potential Benefits
Municipal Councils	Local government authorities responsible for managing public waste collection services	Improved waste collection planning, overflow prevention, and efficient resource allocation
Private Waste Management Companies	Organizations providing waste collection and transportation services	Optimized collection schedules, reduced fuel costs, and improved operational efficiency
Recycling and Waste Processing Facilities	Facilities responsible for receiving and processing collected waste	Better anticipation of incoming waste volumes and improved coordination with collection schedules
Universities & Campuses	Large institutions managing internal waste systems	Smart waste monitoring, sustainability reporting, and efficient campus waste collection planning
Smart City Projects	Urban digital infrastructure initiatives that integrate intelligent services	Integration with smart city waste monitoring platforms and improved urban waste management efficiency

Table 4: Customer Persona

Persona	Role	Responsibilities	System Benefits
Waste Management Operations Manager	Municipal authority or waste company	Oversees waste collection schedules and operational planning	Uses predicted bin fill levels to schedule timely waste collection and prevent overflow
Environmental Sustainability Officer	Universities, campuses, or large organizations	Manages sustainability and waste reduction initiatives	Uses prediction insights to improve waste management efficiency and support sustainability goals
Waste Collection Supervisor	Waste collection service provider	Coordinates collection teams, routes, and vehicle allocation	Uses overflow alerts and predicted fill levels to optimize collection routes and reduce unnecessary trips
Smart City Program Manager	Smart city project or urban innovation office	Oversees digital environmental services and infrastructure	Integrates predictive waste monitoring into larger smart city systems for improved decision making

6.2 Value Proposition

The proposed IoT-Based Waste Bin Fill Level Prediction System provides a data-driven predictive solution that enables proactive waste collection planning. Unlike traditional waste management systems that only monitor bin status after waste has already accumulated, the proposed system predicts future waste bin fill levels in advance using IoT sensor data and ML models.

Key value propositions include:

- Proactive waste collection planning through advance prediction of bin fill levels and overflow risk
- Reduced operational costs by minimizing unnecessary collection trips and improving vehicle utilization
- Overflow prevention by identifying bins that are likely to reach critical capacity before waste accumulation becomes a problem
- Improved collection efficiency through better scheduling and prioritization of waste collection activities

- Data-driven decision support for municipal waste authorities, waste collection providers, and smart city operators
- Contribution to sustainable waste management by optimizing resource use, reducing fuel consumption, and improving service reliability

These benefits provide both operational value and environmental value to organizations responsible for managing waste collection services. By introducing predictive intelligence into smart waste management operations, the proposed system supports a shift from reactive waste collection toward a more efficient and sustainable planning approach.

6.3 Revenue Model

The system can be commercialized using a SaaS business model where customers subscribe to the forecasting platform.

Table 5: Potential Revenue Streams

Revenue Source	Description
Subscription Service	Monthly or annual subscription for access to the prediction dashboard, smart bin monitoring interface, and analytics platform
Data Analytics Services	Customized waste bin fill-level analytics and reporting for municipalities, institutions, and private waste operators
System Integration	Integration of the prediction system with municipal waste management systems, IoT smart bin platforms, or smart city infrastructure
Consulting Services	Consulting support for waste collection planning, operational optimization, and predictive waste management strategies
Maintenance and Support Services	Technical support, dashboard maintenance, model updates, and system performance monitoring for subscribed customers

6.4 Cost Estimation

The development and operational costs of the proposed IoT-Based Waste Bin Fill Level Prediction System mainly involve software development, cloud-based infrastructure, IoT data integration, and ongoing system maintenance.

The key cost components include:

- **Cloud storage and data processing infrastructure** for storing historical waste bin data and running predictive ML models
- **API subscription costs** for external contextual data sources such as weather and calendar information
- **Software development and system integration** for implementing the ML prediction engine, dashboard, and data pipelines
- **IoT device and sensor integration costs** related to collecting waste bin fill-level data from smart bins
- **Maintenance and technical support** for model updates, dashboard management, and system reliability

Although the system requires integration with IoT-based monitoring infrastructure, the overall operational cost remains relatively lower than large-scale hardware-intensive waste management systems because the main value of the proposed solution is delivered through software-based predictive analytics and decision support. This makes the system commercially attractive for municipalities, private waste management operators, campuses, and smart city projects seeking cost-effective waste collection optimization.

6.5 Pricing Strategy

A tiered subscription pricing model can be adopted to accommodate different customer segments based on the scale of deployment and the level of predictive functionality required.

For example:

- **Basic Plan** – Designed for small institutions, campuses, or private facilities with a limited number of smart waste bins and basic monitoring requirements. This plan may include access to the prediction dashboard, basic fill-level forecasting, and overflow alerts.
- **Standard Plan** – Intended for municipal councils, medium-sized waste management organizations, or private waste operators managing multiple collection areas. This plan may include expanded dashboard access, predictive analytics for multiple zones, and operational reporting features.
- **Enterprise Plan** – Suitable for large cities, smart city projects, and integrated waste management systems requiring large-scale deployment, advanced analytics, multi-zone prediction, system integration, and enhanced decision-support capabilities.

Pricing can vary depending on:

- the **number of monitored bins or locations**
- the **amount of historical and real-time data processed**
- the **level of predictive analytics and reporting provided**

- the need for system integration, customization, and technical support

This pricing approach allows the system to remain flexible and commercially attractive to a wide range of customers, from small institutions to large-scale urban waste management environments.

6.6 Competitive Advantage

Most existing smart waste management systems focus on real-time bin monitoring and, in some cases, collection route optimization. These systems provide operational visibility by showing the current status of waste bins after waste has already accumulated. However, they rarely offer predictive insights about future bin fill levels before overflow occurs.

The proposed IoT-Based Waste Bin Fill Level Prediction System provides a significant competitive advantage by introducing ML-based forecasting into smart waste management operations. Instead of relying only on real-time monitoring, the proposed system predicts future waste bin fill levels using historical IoT sensor data and contextual variables such as weather, calendar information, and location characteristics.

Table 6: Comparison with Existing Solutions

Feature	Existing Smart Waste Systems	Proposed System
Real-Time Bin Monitoring	✓	✓
Overflow Detection	✓	✓
Future Bin Fill-Level Prediction	✗	✓
Proactive Collection Planning	Limited	✓
Data-Driven Forecasting	Limited	Advanced ML-based forecasting
Decision Support Dashboard	Basic Monitoring	Predictive Monitoring + Alerts

The predictive fill-level forecasting capability provides a unique advantage by enabling proactive waste collection planning, overflow prevention, and better operational decision-making. This differentiates the proposed system from conventional smart waste monitoring platforms that mainly operate in a reactive manner.

6.7 Intellectual Property Considerations

The proposed IoT-based waste bin fill-level prediction system may involve intellectual property related to its ML models, data processing pipeline, dashboard design, and overall software architecture. The developed code, prediction logic, and system implementation can be protected through software copyright.

If the system shows strong commercial value, the predictive framework and analytics workflow may also be treated as proprietary technology within a commercial platform. In addition, licensing agreements can be used to control how the software is distributed, integrated, or reused by third parties.

Proper documentation, ownership definition, and licensing strategies will help protect the intellectual property of the system while supporting future commercialization.

7. BUDGET AND JUSTIFICATION

The development of the proposed Machine Learning–Based Overflow Prediction and Dynamic Routing Module requires resources such as computing hardware, software tools, data infrastructure, cloud services, and development effort. Since this component is mainly focused on predictive analytics, routing optimization, and software-based decision support, the overall implementation cost is expected to remain relatively affordable Hardware Costs

The system requires a development computer for data processing, ML model training, and dashboard implementation. Additional storage and internet connectivity may also be needed for dataset backup and cloud access.

Table 7: Estimated Hardware Costs

Hardware Resource	Description	Estimated Cost (Rs)
Existing Laptop	Used for programming, ML model training, and data analysis	Already available
RAM Upgrade (Optional)	Additional memory to support ML model execution	15,000
External Storage	Backup storage for datasets and project files	10,000
Estimated Hardware Upgrade Cost: Rs.25,000		

A reliable computing device is necessary to train machine learning models, process historical event data, and test routing algorithms efficiently. External storage supports secure backup of datasets, logs, and development resources. Internet connectivity is required for accessing cloud services, external APIs if used, repositories, and development tools.

Software Licenses

The proposed system will primarily use open-source software tools, which helps minimize development costs while providing strong support for ML, data processing, and dashboard development.

Table 8: Software Licenses Costs

Software Tool	Purpose	License Cost
Java / Python	Core development for prediction and backend services	Free
PostgreSQL	Storage of historical readings, event logs, and prediction data	Free
Kafka	Event streaming and ingestion of bin/app data	Free
gRPC	Service-to-service communication	Free
XGBoost / Scikit-learn	Overflow prediction model implementation	Free
Google OR-Tools	Route optimization and dynamic routing	Free
React	Dashboard and visualization interface	Free
Docker	Containerization and deployment	Free
Llama / Ollama	Natural language explanations for alerts and route decisions	Free / Optional
Diagram Tools (Draw.io / Lucidchart)	System architecture and workflow diagrams	Free / Optional
Total Software Cost: Rs.0		

The selected technologies are widely used open-source tools for predictive analytics, backend service development, route optimization, and dashboard implementation. PostgreSQL and Kafka support reliable data storage and event processing, while XGBoost and Scikit-learn enable accurate overflow prediction. Google OR-Tools is used for dynamic routing, and React supports interactive visualization of prediction outputs. Llama can be optionally integrated to provide human-readable explanations for overflow alerts and routing decisions. The use of open-source software significantly reduces development cost while providing the required functionality for the proposed component.

Cloud Services

Cloud infrastructure may be used to store historical and real-time waste bin data received from IoT-enabled smart bins, execute the overflow prediction model, support routing-related computations, and host the visualization dashboard. Although the IoT data collection layer is handled by other components of the overall system, this component relies on those incoming sensor readings for predictive analytics and decision support.

Table 9: Cloud Services Costs

Cloud Service	Purpose	Estimated Monthly Cost (Rs)
Cloud Database / Storage (PostgreSQL, Firebase, AWS, Azure, or Supabase)	Store historical readings, real-time sensor data, event logs, and prediction outputs	3,000

Cloud Computing Services	Run overflow prediction models, backend services, and dynamic routing logic	4,000
Dashboard Hosting Services	Host the monitoring dashboard and API services	2,000
External API Services (if applicable)	Access contextual data such as weather information	Free / minimal
Estimated Monthly Cloud Cost: 7,000–9,000		

Cloud services provide scalable storage and computing resources, which are useful for processing continuous IoT sensor data, running prediction services, and hosting the dashboard for real-time monitoring. However, many cloud providers offer free-tier plans that may be sufficient for prototype-level research systems. For initial development and testing, parts of the component can also be executed in a local environment before moving to cloud deployment.

Human Effort Estimation

Table 10: Estimated Human Effort hours

Development Activity	Estimated Hours
Historical and real-time data preparation	110 hours
Feature engineering and dataset preparation	80 hours
Overflow prediction model development	180 hours
Dynamic routing / decision-support logic	90 hours
Model evaluation and validation	70 hours
Dashboard integration and visualization	70 hours
Total Estimated Development Effort: 600 hours	

The estimated effort represents the time required to design, develop, and evaluate the proposed Machine Learning–Based Overflow Prediction and Dynamic Routing Module. This includes preparing historical and real-time input data received from other system components, performing feature engineering, developing the overflow prediction model, implementing decision-support and routing logic, validating model performance, and integrating outputs into the monitoring dashboard.

8. WORK BREAKDOWN STRUCTURE (WBS)

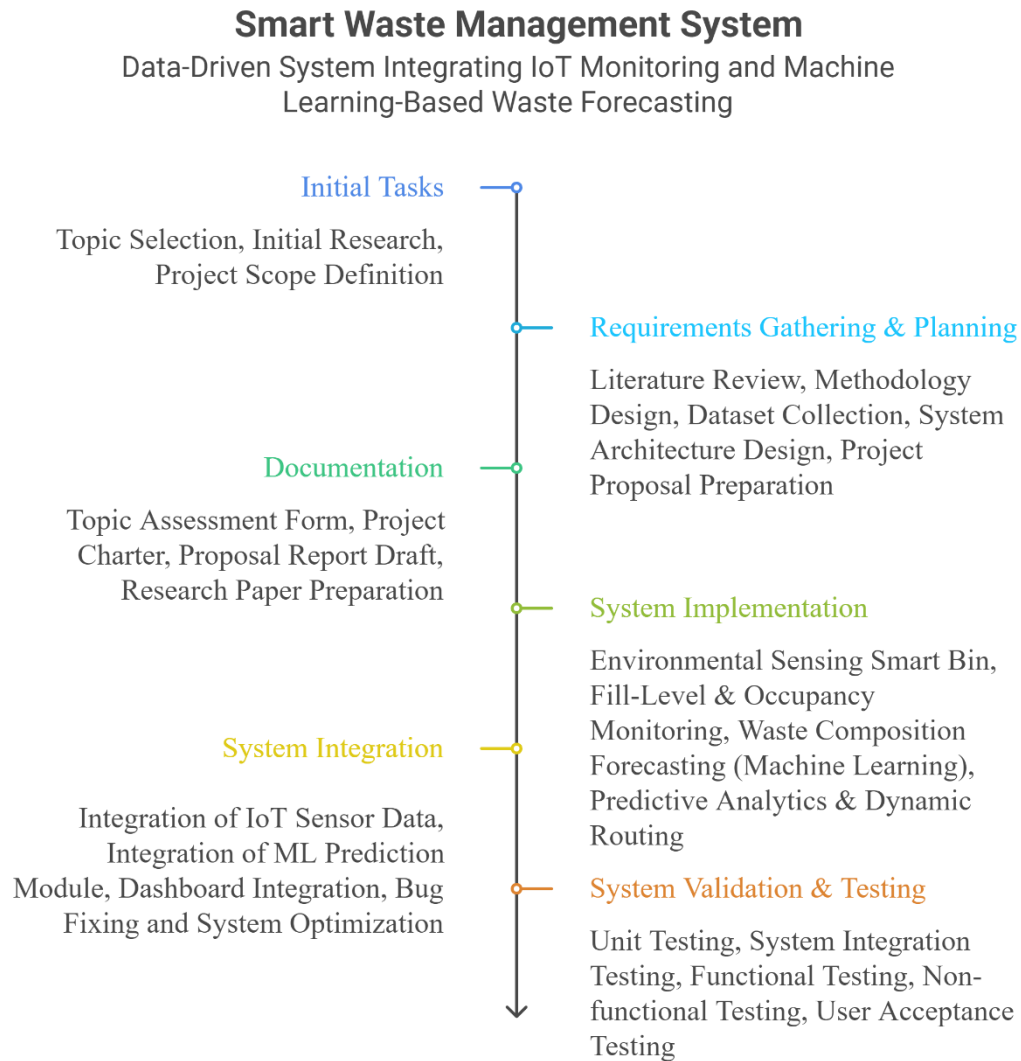


Figure 3: Work Breakdown Structure of the overall project

The Work Breakdown Structure (WBS) outlines the distribution of tasks among the project team members for the development of the Smart Waste Management System. Each member is responsible for a specific component of the system, ensuring balanced workload distribution and efficient project execution.

Table 11: Work Breakdown Structure of the component

Task ID	Task Description	Timeline	Deliverable
T1	Collection of historical bin readings, event logs, and contextual data	March – April	Initial dataset
T2	Data preprocessing and cleaning	April	Cleaned dataset
T3	Feature engineering and dataset preparation	April – May	Training dataset
T4	Development of overflow prediction models (RF / XGBoost)	May – June	Initial prediction models
T5	Model training, tuning, and risk score generation	June – July	Optimized prediction model
T6	Development of routing / decision-support logic	July – August	Initial routing module
T7	Model evaluation using MAE, RMSE, and cross-validation	August – September	Evaluation results
T8	Dashboard integration and visualization	September – October	Dashboard prototype
T9	System testing and performance analysis	October	Testing results

Table 12: Workload Distribution

Development Phase	Tasks	Estimated Effort
Data Preparation	T1 – T3	30%
Prediction Model Development	T4 – T5	30%
Routing / Decision Support	T6	15%
Evaluation & Testing	T7, T9	15%
Dashboard & Visualization	T8	10%

9. GANTT CHART

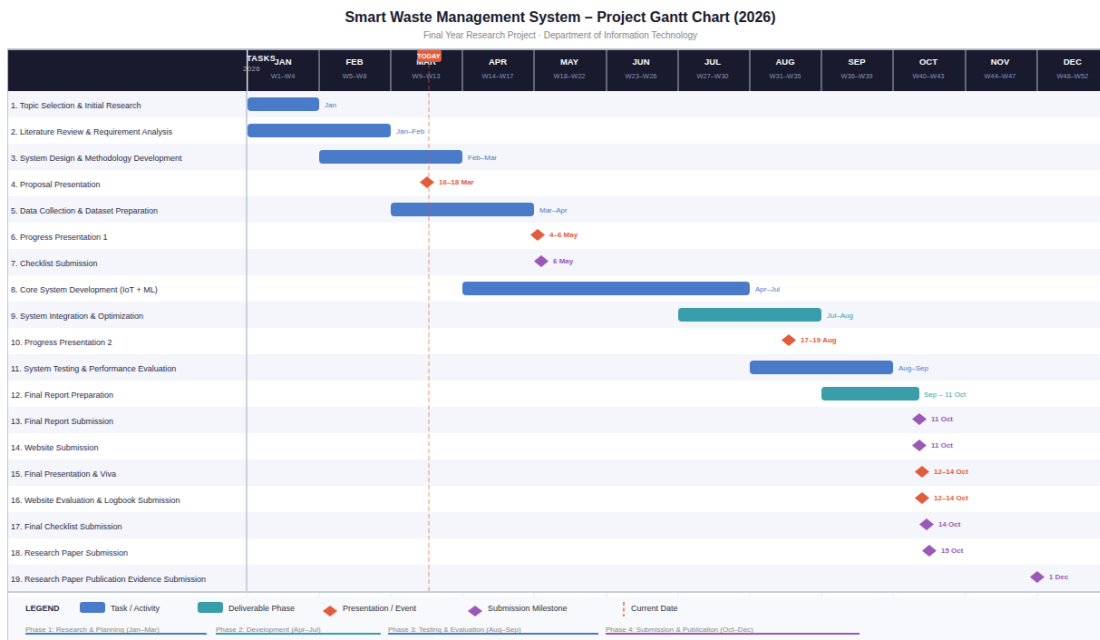


Figure 4: Gantt chart

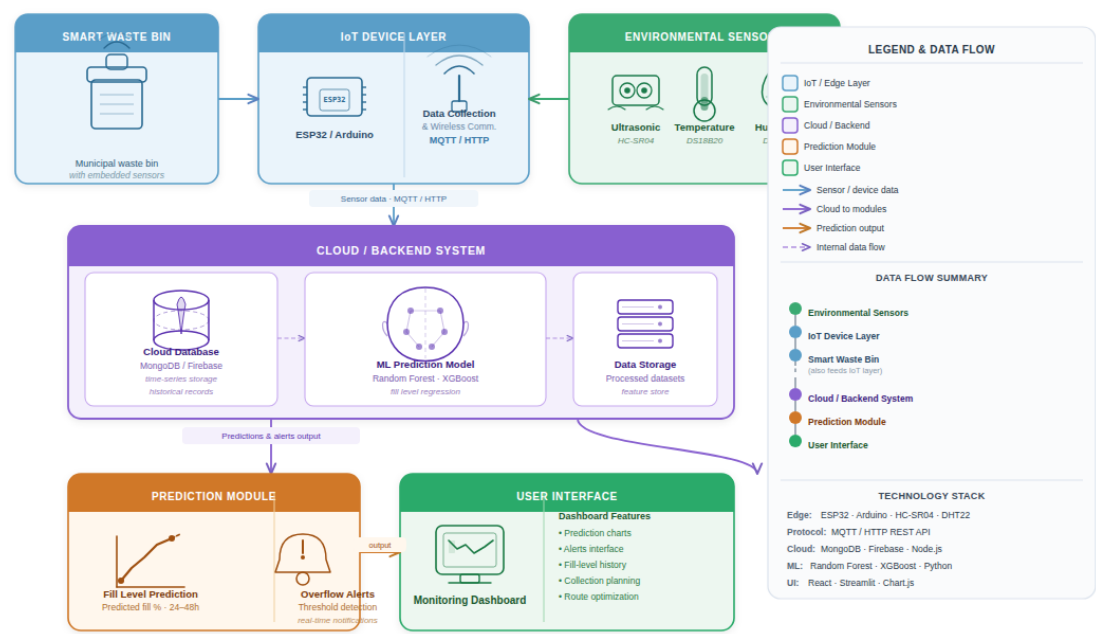
10. REFERENCES LIST

- [1] J. Guo, "IoT Technology in Urban Waste Management: Smart Waste Bin Systems," *Highlights in Science, Engineering and Technology*, vol. 111, pp. 452–458, 2024.
- [2] S. Adeoye, "Smart Waste Management Systems: An IoT and AI-Driven Approach for Urban Sanitation," *Computer and Information Science*, vol. 18, no. 2, pp. 30–45, 2025.
- [3] A. Arul Anitha and L. Arockiam, "Promoting a Clean and Hygienic Environment using IoT," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 5, pp. 4722–4726, Jan. 2020.
- [4] R. Rajasree, P. Suganya, and J. Ebens Nikshya, "Smart Garbage Monitoring System Using IoT," *Journal of Information and Computational Science*, vol. 13, no. 4, pp. 363–368, 2023.
- [5] S. M. Usha and H. B. Mahesh, "Accurate and High Speed Garbage Detection and Collection Technique using Neural Network and Machine Learning," *IOP Conference Series: Materials Science and Engineering*, vol. 1258, no. 1, Art. no. 012055, 2022.
- [6] E. B. Priyanka, S. Vijayshanthi, S. Thangavel, R. Anand, G. B. Bhavana, B. Khan, K. Jeyanthi, and A. Ambikapathy, "Prediction of waste generation forecast and emission potential on the Erode City solid waste dump yards based on machine learning approach," *Scientific Reports*, vol. 15, Art. no. 37021, 2025.
- [7] C. G. Samal, D. R. Biswal, G. Udgata, and S. K. Pradhan, "Estimation, Classification, and Prediction of Construction and Demolition Waste Using Machine Learning for Sustainable Waste Management: A Critical Review," *Construction Materials*, vol. 5, no. 1, Art. no. 10, 2025.
- [8] S. S. Chaudhari, K. Kale, M. Raghuvanshi, T. Kamble, R. Thakur, S. A. Dalvi, P. A. Patil, S. Y. Waware, and A. S. Kurhade, "Machine learning for waste-to-energy processes: Resource evaluation, conversion efficiency, and environmental effects," *Applied Chemical Engineering*, vol. 9, no. 1, 2026.
- [9] I. B. K. P. Arimbawa, I. G. A. Novitasari, and P. N. A. Permana, "Application of Ant Colony Optimization on CVRP for Waste Collection Route Optimization in Marga Village," *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 2, pp. 940–951, Nov. 2025.
- [10] A. M. Restrepo-Franco, O. Valencia-Rodríguez, and E. M. Toro-Ocampo, "The vehicle routing problem as applied to residential solid waste collection operations: Systematic literature review," *International Journal of Industrial Engineering Computations*, vol. 16, 2025.
- [11] S. R. Laha, K. Al Smadi, A. K. Habboush, B. K. Pattanayak, S. Pattnaik, and B. Mohanty, "Smart Waste Management Framework for Green Cities: Integrating IoT, LoRa, and Deep Learning for Efficient Waste Classification and Management," *Tikrit Journal of Engineering Sciences*, vol. 32, no. SP1, Art. no. 2162, 2025.

[12] O. Mugisho Mugaruka, J. R. Kala, J. Ndikumagenge, R. van Schadewijk, and R. Nibitanga, “Design of an IoT sensor-based network architecture for waste collection in Bukavu city (DRC),” *Research Square*, preprint, Feb. 2026, doi: 10.21203/rs.3.rs-8603843/v1.

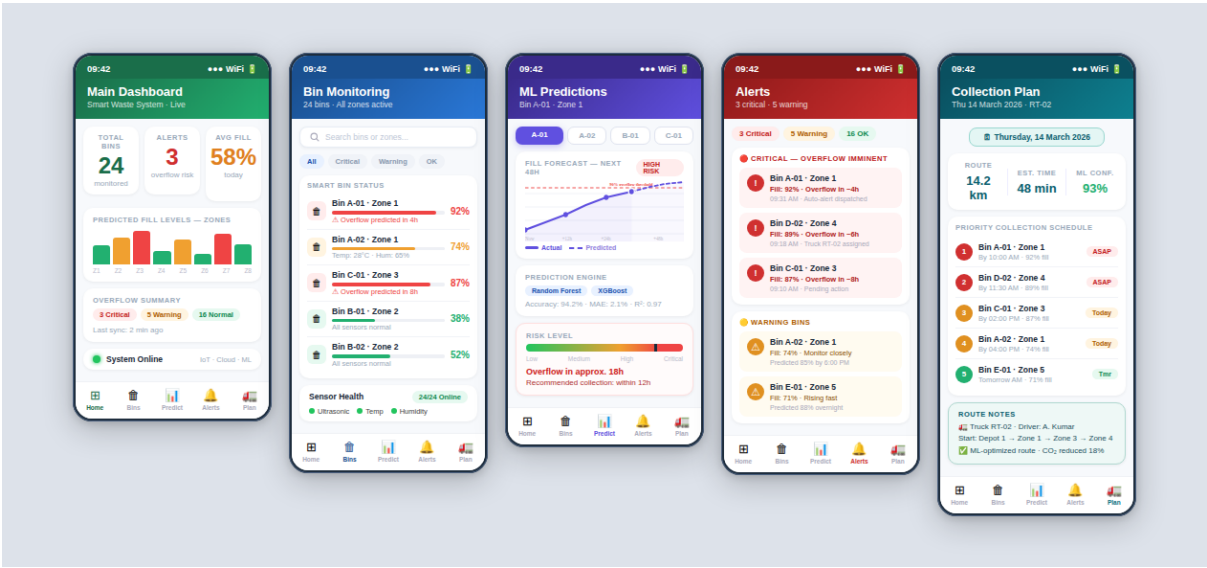
11.APPENDICES

□ Detailed System Architecture



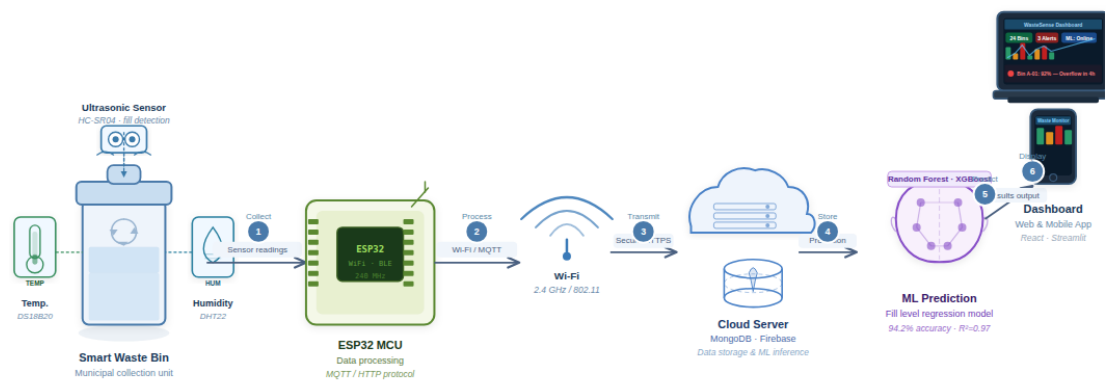
Appendix 1: Detailed System Architecture

• Dashboard UI Mockups



Appendix 2: Dashboard UI Mockups

- IoT Hardware Setup



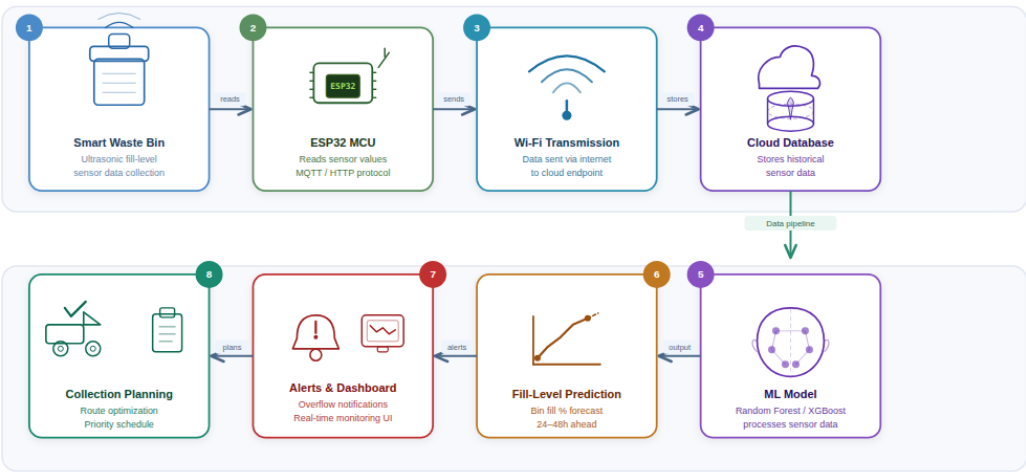
Appendix 3:IoT Hardware Setup

- Risk Analysis Diagram

 RISK FACTOR	 MITIGATION STRATEGY
 Sensor Failure Ultrasonic / DHT22 malfunction HIGH	 Sensor Calibration & Maintenance Regular testing, redundant sensors, auto-alerts
 Network Connectivity Issues Wi-Fi drops / MQTT failures MEDIUM	 Offline Storage & Retry Mechanism Local buffer on ESP32, auto-retry with exponential backoff
 Inaccurate ML Prediction Low model accuracy / concept drift HIGH	 Model Validation & Retraining Cross-validation, scheduled retraining on new data
 Data Loss Risk Server crash / corruption / deletion MEDIUM	 Regular Backup & Secure Cloud Storage Automated daily backups, redundant cloud replication
 Dashboard Access Issues Unauthorized access / downtime LOW	 Authentication & Monitoring Role-based access control, uptime monitoring alerts
 Overflow Alert Failure Missed notifications / false alerts HIGH	 Threshold Testing & Alert Verification End-to-end alert pipeline tests, SMS/email fallback

Appendix 4:Risk Analysis Diagram

- Data Collection Process



Appendix 5: Data Collection Process