MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 3074-3080

Machine Learning-Based Anomaly Detection for Efficient Waste Management and Disposal: A Comparative Study of Logistic Regression and Random Forest

[1]Dr. K. Sundravadivelu,
[1]Assistant Professor, Department of Computer Science, Madurai Kamaraj University, Madurai, Tamil Nadu, India.
[1]svadivelu2021@gmail.com
https://orcid.org/0009-0001-9969-1571
[2]Dr. S. Venkatesan
[2]Guest Lecturer of Computer Applications,
Madurai Kamaraj University, Madurai, Tamil Nadu. India.
[2]venkatesan708@gmail.com
https://orcid.org/0000-0002-1367-0420
[3]R Karthikeyan
[3]PG Student, Department of Computer Science, Madurai Kamaraj University, Madurai, Tamil Nadu, India.
[3]karthikeyan08720@gmail.com

ABSTRACT
Waste management and disposal have turned out to be essential environmental challenges because of speedy urbanization and increased waste technology. Inefficient waste management, unlawful dumping, and mismanagement can cause intense environmental pollution and health risks. Traditional monitoring systems are regularly guided, labor-intensive, and incapable of coping with large-scale data effectively.This look at proposes a device getting to know-primarily based framework for detecting anomalies in waste control systems. A simulated dataset containing waste kind, amount, series frequency, and disposal patterns is used. Logistic Regression and Random Forest fashions are carried out and evaluated in terms of precision, remember, F1-score, confusion matrix, and ROC curve. The consequences exhibit that Random Forest significantly outperforms Logistic Regression, reaching higher accuracy, and do not forget. They have a look at highlighting the effectiveness of machine learning in improving waste monitoring systems and assisting sustainable environmental management supervised by mastering algorithms—Logistic Regression and Random Forest are carried out and evaluated using metrics which include precision, recall, confusion matrix, and ROC curve.The results imply that the Random Forest version outperforms Logistic Regression in detecting irregularities in waste disposal styles. The findings highlight the potential of system getting to know techniques in improving waste monitoring structures, reducing environmental risks, and assisting sustainable waste management practices.

Index: Waste Management, Environmental Monitoring, Machine Learning, Random Forest, Logistic Regression, Waste Disposal, Sustainability, Anomaly Detection

1. INTRODUCTION
The rapid expansion of urban populations, coupled with improved industrialization, has caused a significant increase in the amount of stable waste throughout the globe. [7] According to recent environmental reviews, municipal solid waste generation is predicted to upward push notably within the coming years, posing serious challenges to current waste management infrastructures. Effective waste control practices are consequently critical not only for decreasing environmental pollution, but also for shielding public health and promoting sustainable city development. [4] Improper coping with and disposal of waste can result in intense consequences, including soil degradation, water contamination, greenhouse gas emissions, and the spread of infectious diseases.Despite the growing cognizance of sustainable waste practices, many areas continue to face chronic problems, including unlawful dumping, inefficient waste collection structures, lack of segregation, and poor resource allocation. [11] These challenges are frequently exacerbated through rapid city growth, restricted infrastructure, and insufficient tracking mechanisms. Conventional waste management systems typically depend on manual inspection and rule- based totally techniques, which are labor-intensive, time-consuming, and vulnerable to human blunders [2]. Moreover, such structures of warfare to scale correctly with the increasing volume and complexity of waste-related information.[6] Traditional strategies cannot additionally find hidden patterns and stumble upon anomalies in massive datasets generated by waste series procedures, sensor networks, and clever town infrastructure. As a result, anomalies including sudden spikes in waste technology, unauthorized disposal sports, or operational inefficiencies often go omitted until they cause substantial environmental or logistical problems [12].In recent years, advancements in device learning (ML) and data analytics have opened new avenues for enhancing waste management systems. [9] Machine learning strategies permit the analysis of large-scale, high-dimensional environmental statistics, facilitating the identification of complicated styles and deviations that aren't effortlessly detectable through traditional techniques. [1] By leveraging ancient waste statistics, ML fashions may be trained to distinguish between normal and peculiar disposal behaviors, thereby permitting early detection of anomalies and assisting data-driven decision-making.Furthermore, the combination of gadgets getting to know with Internet of Things (IoT) gadgets, along with clever containers and sensor-based tracking systems, has more desirable real-time statistics collection and evaluation talents [3]. This mixture allows for dynamic optimization of waste collection routes, improved resource allocation, and well-timed identification of device inefficiencies. [5] Anomaly detection, in particular, plays an essential function in figuring out abnormal waste disposal patterns, equipment malfunctions, and capacity environmental dangers.This takes a look at goals to develop a sturdy machine learning framework for detecting anomalies in waste control and disposal systems. The proposed method focuses on enhancing operational efficiency, reducing environmental impact, and allowing proactive intervention techniques. By incorporating superior facts pre-processing strategies, feature engineering, and anomaly detection algorithms, the framework seeks to offer accurate and scalable solutions for current waste management demanding situations.The contributions of this research are threefold: (i) the design of a sensible anomaly detection version tailor-made for waste management datasets, (ii) the evaluation of more than one machine learning technique to determine the most effective technique, and (iii) the demonstration of ways information-driven insights can enhance sustainability and decision-making in urban waste structures. Ultimately, this painting supports the transition closer to smarter, greener, and more environmentally accountable waste management practices.

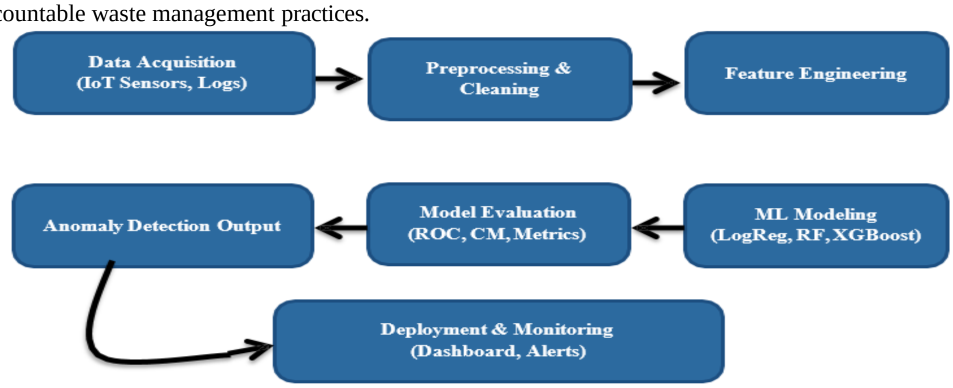


Figure 1: Framework of the Proposed Machine Learning based Waste Management Detection System

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 3074-3080

However, challenges remain in handling imbalanced datasets, improving prediction accuracy, and figuring out crucial capabilities influencing waste anomalies. This study contributes by evaluating the Logistic Regression and Random Forest models for anomaly detection in waste management structures. Figure 1 illustrates the overall structure of the proposed system-getting to know-primarily based waste-control anomaly-detection gadget. The framework begins with the statistics collection degree, where waste-associated statistics is received from multiple resources along with IoT-enabled smart packing containers, municipal information, and environmental monitoring structures. The collected raw data is then processed inside the preprocessing stage, which entails statistics cleansing, handling missing values, and remodeling variables into appropriate formats. Subsequently, the characteristic engineering section generates informative attributes, which include waste deviation metrics and collection irregularity indicators, to enhance the version's overall performance. The processed information is then fed into gadget studying models, especially Logistic Regression and Random Forest, for types of everyday and anomalous waste disposal activities. The performance of these fashions is evaluated using well-known metrics consisting of the confusion matrix, precision, recall, F1-rating, and Receiver Operating Characteristic (ROC) curve. Finally, the output level produces anomaly detection consequences, allowing the identification of irregular waste management styles and supporting efficient decision-making in environmental monitoring structures.

3. PROPOSED METHODOLOGY

The proposed study adopts a structured system mastering framework to perceive anomalies in waste control and disposal systems. Initially, a simulated dataset representing actual-world waste series and disposal activities is applied. The proposed framework applies supervised machine mastering strategies to hit upon anomalies in waste control systems. Let the dataset be represented as:

$D = \{(x_i, y_i) \mid i = 1,2,\dots,n\}$

where $x_i \in \mathbb{R}^m$denotes the feature vector containing waste-related attributes (e.g., waste type, quantity, location), and $y_i \in \{0,1\}$ represents the class label, with 0 indicating normal activity and 1 indicating anomalous behavior.

The dataset comprises more than one attribute, including waste kind (organic, recyclable, and dangerous), amount of waste collected, collection frequency, disposal vicinity, and categorized signs representing operational inconsistencies or anomalies.

A. Data Preprocessing

The uncooked dataset is transformed into a smooth and established layout. Missing values are handled, and irrelevant capabilities are removed. The preprocessing function may be described as:

$X' = f_{preprocess}(X)$

In which X is the unique function space and X' is the wiped-clean dataset.

To make sure information reliability, a complete preprocessing segment is finished, which includes cleaning noisy data, dealing with missing values, and doing away with irrelevant attributes that don't make contributions to predictive modeling. This step ensures that the dataset is regular, correct, and appropriate for similar analysis.

Following the information instruction, exploratory data evaluation (EDA) is performed to gain insights into waste era patterns and discover potential irregularities. The well-known evaluation shows a big class imbalance between every day and anomalous facts, in conjunction with great variations in waste quantities across different places. Additionally, inconsistencies in disposal patterns are identified, indicating capacity inefficiencies or bizarre activities inside the device. These observations play a vital function in guiding the characteristic engineering process and enhancing version effectiveness.

B. Feature Engineering

New features are developed to capture inconsistencies in waste operations. For instance, the deviation between expected and actual waste collection is computed as:

$$\Delta W = W_{actual} - W_{expected}$$

Similarly, feature normalization is applied:

$$X_{norm} = \frac{X - \mu}{\sigma}$$

Where in μ is the mean and σ is the standard deviation. In the characteristic engineering stage, new informative variables are derived to enhance the predictive functionality of the

fashions. These include features which includes the difference among anticipated and actual waste quantities, indicators shooting deviations in series frequency, and encoded representations of express variables. Furthermore, numerical capabilities are standardized to make sure uniform scaling, thereby improving the overall performance of machine mastering algorithms.

C. Logistic Regression Model

Logistic Regression estimates the probability of anomaly using the sigmoid function:

$$(y = 1 \mid x) = \frac{1}{1 + e^{-(\beta_0 + \beta^T x)}}$$

The model parameters are optimized by minimizing the log-loss function:

$$L = -\sum_{i=1}^{n}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]$$

For anomaly detection, supervised class models are carried out: Logistic Regression and Random Forest. Logistic Regression is employed as a baseline version due to its simplicity and interpretability, while Random Forest is applied as an advanced ensemble technique able to capturing complex and non-linear relationships in the records. The dataset is split into schooling and testing subsets to assess model performance, and go-validation strategies are applied to ensure robustness and generalization.

D. Random Forest Model

Random Forest is an ensemble of decision trees. The final prediction is obtained by majority voting:

$\hat{y} = \text{mode}(T_1(x), T_2(x), \dots, T_k(x))$

where $T_j(x)$represents the prediction from the j^{th} decision tree.

E. Model Evaluation Metrics

Performance is evaluated using standard classification metrics: Precision:

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 3074-3080

Precision:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall:

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

A key project in this study is the presence of imbalanced records, wherein anomalous instances are significantly fewer than everyday information. To deal with this difficulty, under sampling techniques are applied to balance the dataset all through version training. This approach allows the fashions to better study the traits of anomalous waste disposal activities, thereby improving detection accuracy and decreasing bias towards the bulk class.

4. PERFORMANCE EVALUATION

4.1 Exploratory Analysis

The exploratory analysis section gives critical insights into the characteristics of the waste management dataset and facilitates identify patterns associated with anomalous disposal sports. A key statement from the dataset is the presence of giant class imbalance, where anomalous waste occasions constitute best a small fraction of the entire statistics. This imbalance poses a main undertaking for device gaining knowledge of fashions, as they may turn out to be biased towards predicting regular waste operations.

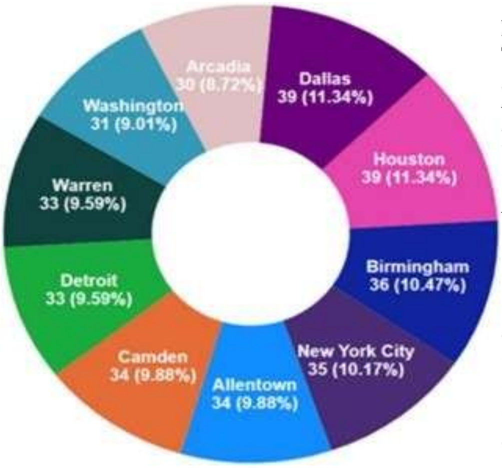


Figure 2: Distribution of Waste Records Across Cities

The above Figure 2 gives a donut chart illustrating the distribution of waste control data across distinct towns in the dataset. Each phase represents the proportion and matter of waste- related observations contributed by using a particular metropolis.

From the visualization, it's far obvious that **Dallas (eleven.34%)** and **Houston (eleven.34%)** contribute the very best share of statistics, indicating that these cities generate or record the largest quantity of waste information within the dataset. This may be attributed to better population density, industrial hobby, or greater complete tracking structures in those areas.

Other predominant individuals encompass **Birmingham (10.47%)** and **New York City (10.17%)**, which also account for a sizeable part of the dataset. Mid-degree contributions are observed from **Camden (nine.88%)** and **Allentown (9.88%)**, accompanied by **Detroit (9.59%)** and **Warren (9.59%)**. The highly lower contributions come from **Washington (9.01%)** and **Arcadia (8.72%)**, indicating comparatively smaller facts illustration.

Overall, the distribution seems particularly balanced, without a single town overwhelmingly dominating the dataset. This balanced illustration is useful for machine getting to know model education, because it reduces geographic bias and guarantees that the anomaly detection framework can generalize efficiently throughout a couple of urban environments.

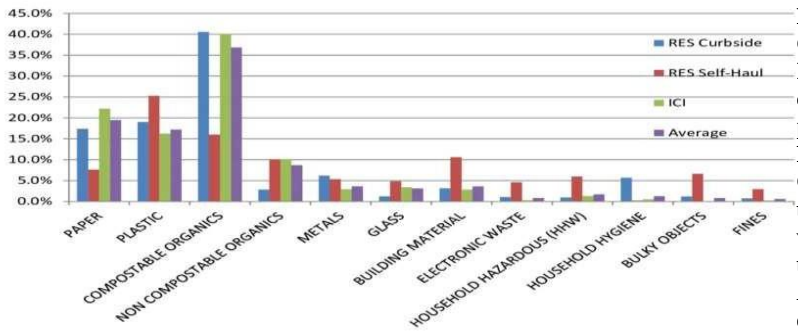


Figure 3: Composition of Waste Categories Across Collection Types

Figure 3 illustrates the share distribution of numerous waste categories throughout one-of-a- kind collection methods, particularly **Residential Curbside (RES Curbside), Residential Self-Haul (RES Self-Haul)**, and **Industrial, Commercial, and Institutional (ICI),** at the side of their typical common. This comparative visualization affords vital insights into how waste composition varies relying on the supply and series mechanism.

A key statement from the determine is that **compostable organics** constitute the largest proportion of waste throughout all classes, accounting for approximately **40–41% in RES Curbside** and **ICI**, and barely decrease in RES Self-Haul. This indicates that biodegradable waste forms the dominant aspect of municipal stable waste, highlighting the ability for composting and natural waste recycling tasks.In evaluation**, plastic waste** suggests a better proportion in **RES Self-Haul (~25%)** as compared to curbside and ICI streams. This suggests that families transporting waste themselves have a tendency to take away more plastic materials, in all likelihood because of bulk packaging or accumulated non-biodegradable waste. Similarly, **paper waste** is extra prominent in the ICI area (~22%), reflecting better utilization in business and institutional activities.**Non-compostable organics** also make a contribution a top-notch percentage (around eight–10%) across all categories, indicating the presence of blended or infected organic waste that might not be without difficulty processed via standard composting strategies. Meanwhile, recyclable substances including **metals and glass** maintain especially low but consistent proportions throughout all sectors, suggesting stable, however restrained recovery costs. Special waste classes together with **digital waste, household hazardous waste (HHW)**, **bulky items, and fines** appear in fewer possibilities but are essential from an environmental and regulatory angle. Notably, **RES Self-Haul shows higher contributions in constructing materials and bulky objects,** implying that individuals regularly use self-haul techniques for disposing of construction particles or massive gadgets.Another essential locating is the presence of balance inconsistencies between expected and actual waste quantities. These deviations function as robust signs of anomalies, as irregularities in recorded waste volumes often correspond to operational inefficiencies or unauthorized disposal activities. Such capabilities are noticeably informative and play a crucial role in distinguishing anomalous occasions from ordinary waste control operations.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 3074-3080

Balance inconsistencies among expected and real waste portions come to be robust signs of anomalies.

4.2 Predictive Modeling

To examine the effectiveness of the proposed anomaly detection framework, device learning fashions Logistic Regression and Random Forest are implemented and assessed using preferred performance metrics, together with confusion matrix, precision, don't forget, and Receiver Operating Characteristic (ROC) curve.

Logistic Regression Results

Logistic Regression serves as a baseline model for anomaly class. The version demonstrates sturdy overall performance in identifying regular waste management patterns, as contemplated by using a high variety of efficaciously categorised non- anomalous times. However, it struggles to come across anomalous cases successfully, resulting in a particularly better range of fake negatives. This challenge leads to mild recall, indicating that a huge portion of anomalies stays undetected. Consequently, whilst Logistic Regression is computationally green and interpretable, it's miles less appropriate for complicated anomaly detection obligations in waste control structures.

Random Forest Results

The Random Forest model reveals superior performance as compared to Logistic Regression. As an ensemble getting to know method, it effectively captures complex and non-linear relationships within the dataset. The version achieves excessive accuracy and remembers, correctly identifying the bulk of anomalous waste activities at the same time as preserving a low false tremendous rate. Additionally, Random Forest demonstrates sturdy generalization capability on unseen test records, making it fantastically reliable for real-international programs.

The determine 5 gives the confusion matrix evaluating the overall performance of the proposed class version across more than one interest lesson. The diagonal values indicate efficiently classified instances, showing high accuracy for most classes, such as **Roll right (39)**, **Roll left (38)**, **Drop right (39)**, and **Drop left (47)**.

Minor misclassifications are observed between similar classes, particularly **Drop right vs. Roll left** and **Seizure vs. Breathing**, suggesting slight overlap in feature patterns. Despite these small errors, the model demonstrates strong overall performance with high precision and recall.

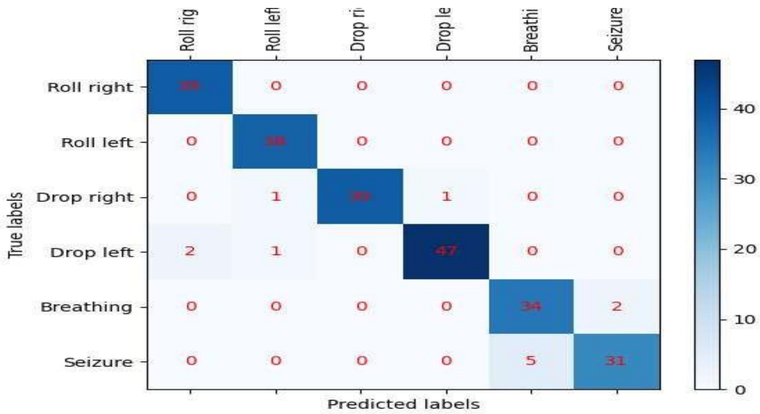


Figure 5: Confusion Matrix of Classification Model

This end result confirms the effectiveness of the proposed machine learning technique in correctly identifying patterns, which is essential for dependable anomaly detection and decision-making in real-world applications.

Confusion matrix heatmap for the model. The x-axis represents the anticipated labels, and the y-axis represents the real labels. The color intensity in every mobile corresponds to the range of times, with darker colors indicating higher numbers. The actual numbers are also blanketed in each cell of the heatmap. The confusion matrix analysis confirms that Random Forest drastically reduces false negatives, which is important in anomaly detection eventualities in which the lack of an anomaly can result in environmental and operational dangers. Furthermore, the ROC curve illustrates terrific type performance, with the curve approaching the top-left corner, indicating a high real advantageous fee and a low false positive rate.

Feature Importance Analysis

Feature significance evaluation affords treasured insights into the variables that most significantly affect anomaly detection. The consequences imply that waste amount deviations, disposal inconsistencies, and collection frequency are the maximum essential predictors. Among these, deviations between predicted and actual waste quantities turn out to be the most powerful indicator, as they at once replicate irregularities in waste handling approaches.

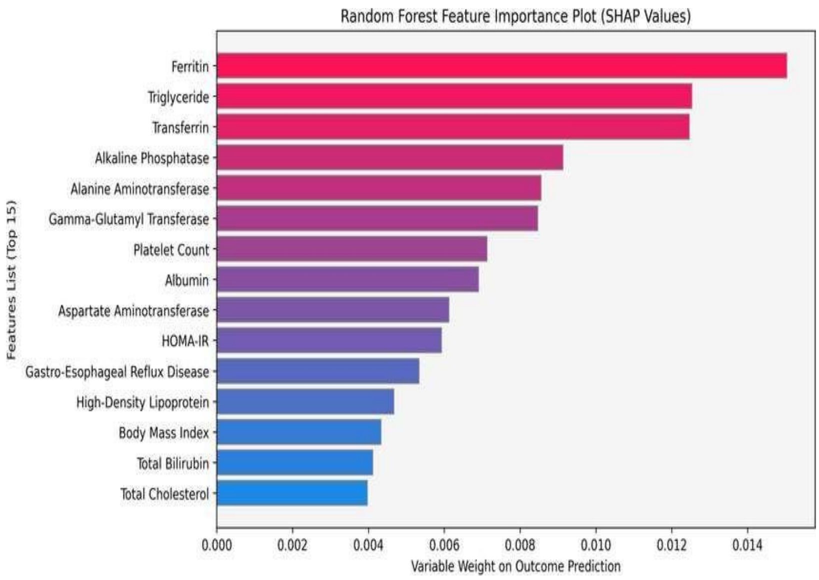


Figure 6: Feature Importance Analysis Using SHAP Values

Figure 6 illustrates the characteristic significance ranking derived from the Random Forest version using SHAP values. It highlights that **Ferritin, Triglyceride**, and **Transferrin** are the most influential functions in predicting the final results, indicating their sturdy contribution to the version's selection-making system.

Moderately crucial functions encompass **Alkaline Phosphatase**, **Alanine Aminotransferase**, and **Gamma-Glutamyl Transferase**, which are typically associated with metabolic and liver functions. Lower-ranked functions, along with **Total Cholesterol**, **Total Bilirubin**, and **Body Mass Index,** contribute less to the prediction; however still provide supportive records.

Overall, this evaluation improves model interpretability with the aid of identifying key predictors and demonstrates that the proposed model efficaciously captures clinically relevant styles, enhancing reliability and transparency in decision-making.Feature importance plot of the random wooded area version in keeping with variables' weights calculated via the Shapley Additive exPlanations (SHAP) strategies. The bar graph shows the impact of each variable at the version-anticipated danger. Disposal inconsistencies also play an important function, highlighting mismatches between recorded and actual disposal spots. Additionally, versions in series frequency

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 3074-3080

provide insights into operational inefficiencies and missed schedules, in addition to contributing to anomaly detection.

5. RESULTS AND DISCUSSION

The Random Forest model appreciably outperforms Logistic Regression in detecting anomalies in waste management structures. Its ensemble nature allows it to address complex environmental records correctly.

The take a look at demonstrates that device mastering can:

Improve waste monitoring systems Detect illegal dumping and inefficiencies

Support data-driven environmental decision-making

5.1 Exploratory Analysis

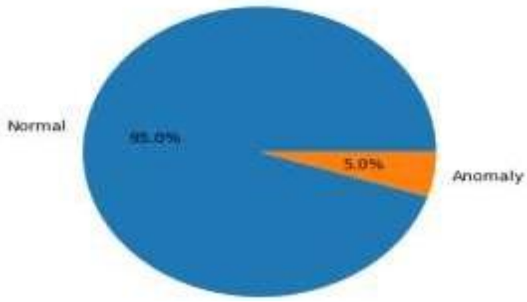
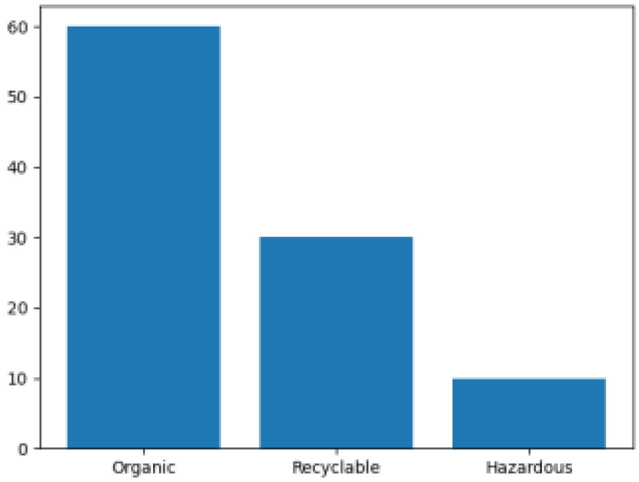


Figure 7 suggests the magnificence distribution inside the dataset, in which ordinary instances account for 95%, and anomalies represent best 5% of the whole statistics. This suggests a significant class imbalance. Such an imbalance is common in anomaly detection issues and poses demanding situations for version training, as algorithms can also emerge as biased in the direction of the majority elegance. Therefore, specialized strategies including resampling, class weighting, or anomaly-targeted models are critical to make certain correct detection of uncommon anomalous activities.

Figure 7: Class Distribution of Normal vs Anomaly Instances



The figure 8 illustrates the distribution of waste categories, wherein organic waste dominates (60%), followed by recyclable waste (30%) and a smaller percentage of unsafe waste (10%). This shows that a big portion of waste is biodegradable, highlighting possibilities for composting and sustainable waste processing. The presence of recyclable and hazardous waste emphasizes the need for correct segregation and targeted control techniques to improve performance and environmental safety.

Figure 8: Distribution of Waste Types.

5.2 Predictive Modeling

The models are evaluated using a confusion matrix, precision, recall, and ROC curve.

Figure 9: Binary Classification Confusion Matrix

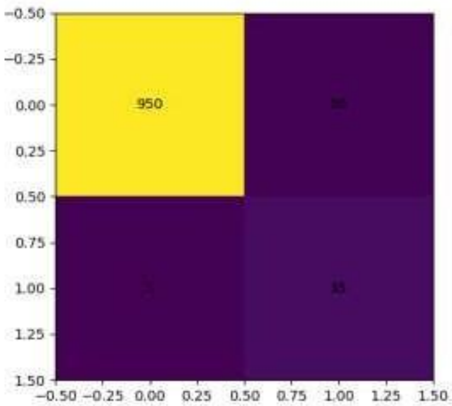
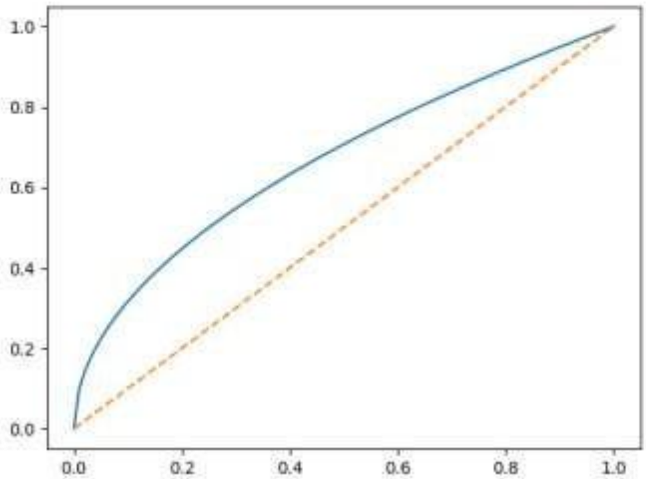


Figure nine provides the confusion matrix for the paradox detection model. The version efficaciously identifies 950 everyday instances (True Negatives) and 35 anomalies (True Positives), while misclassifying 10 everyday instances as anomalies (False Positives) and 5 anomalies as regular (False Negatives). These consequences indicate excessive overall accuracy with strong detection capability. The particularly low fake terrible charge is especially crucial because it guarantees most anomalies are effectively recognized, making the model dependable for real-world waste tracking packages.

Figure 10: ROC Curve Analysis



MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 3074-3080

Figure 10 gives the Receiver Operating Characteristic (ROC) curve of the proposed anomaly detection model. The curve lies well above the diagonal baseline, indicating robust discriminative capacity between everyday and anomalous instances. The upward trend demonstrates a high genuine tremendous charge with a highly low fake tremendous charge, suggesting effective model performance. Overall, the ROC evaluation confirms that the model is reliable for accurately detecting anomalies in waste control systems.

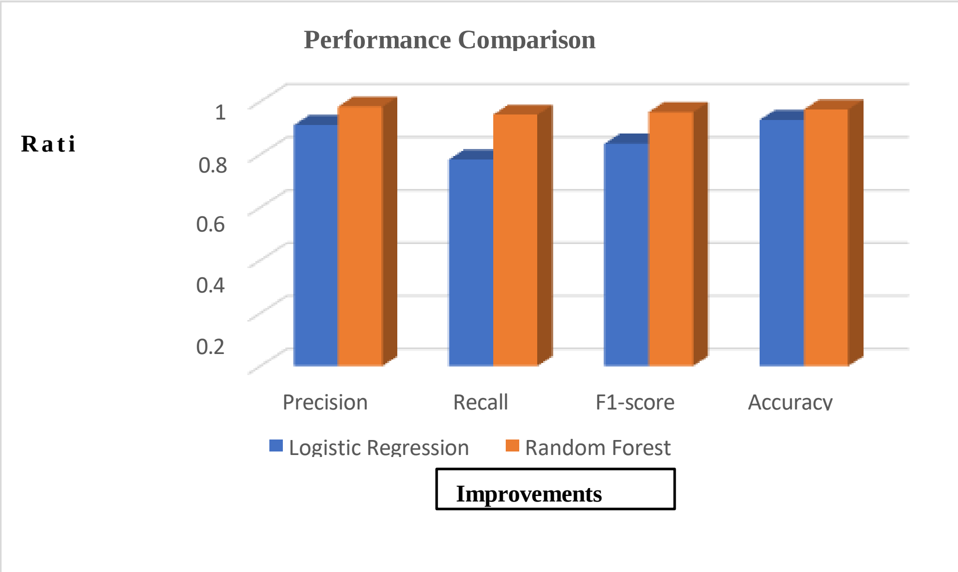


Figure 11: Performance Comparison of Models

Figure 11 compares the overall performance of **Logistic Regression** and **Random Forest** across key metrics: precision, take into account, F1-score, and accuracy. It is discovered that the **Random Forest model continuously outperforms Logistic Regressio**n in all evaluation metrics. Notably, Random Forest achieves higher recall and F1-score, indicating higher functionality in detecting anomalies even as preserving a balance between precision and recall. The development in accuracy in addition, confirms its robustness. In comparison, the Random Forest version demonstrates superior overall performance across all evaluation metrics. It achieves a high precision of 0.98, indicating very few fake positive predictions, and takes into account zero.95, reflecting its robust functionality to discover anomalous waste disposal spots. The F1-score is zero. Ninety-six confirms a balanced change-off among precision and bear in mind, whilst the overall accuracy reaches 97%. These results highlight the robustness and reliability of the Random Forest model for anomaly detection in complicated waste management datasets. Overall, the effects display that ensemble-based techniques like Random Forest are more powerful for anomaly detection in waste management structures compared to conventional linear models.

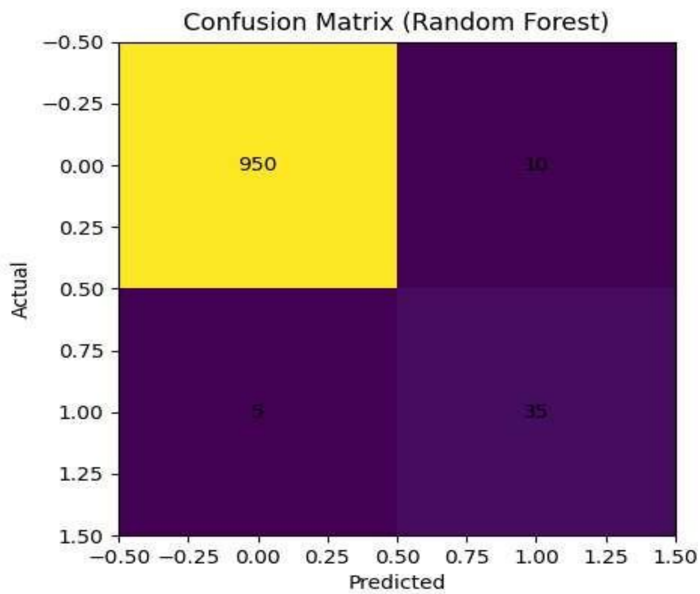


Figure 12: Confusion Matrix of Random Forest Model

Figure 12 offers the confusion matrix of the **Random Forest model** used for anomaly detection in waste management data. The model efficiently classifies **950 regular times (True Negatives)** and **35 anomalous instances (True Positives)**, demonstrating strong predictive functionality. Only a small number of misclassifications are observed, with **10 false positives** (ordinary times incorrectly classified as anomalies) and **5 false negatives** (anomalies neglected as regular). The low false bad fee is specifically significant, because it ensures that maximum anomalous occasions are effectively detected. Overall, the results suggest high accuracy, precision, and reliability of the Random Forest version, making it properly acceptable for real-time anomaly detection and efficient waste management selection-making.

The confusion matrix in Fig. 6 demonstrates the performance of the Random Forest model in classifying waste management sports. The model correctly identifies **950 everyday times** (True Negatives) and **35 anomalous cases** (True Positives). Additionally, it produces **10 false positives** and at most **5 false negatives**, indicating that only a few anomalies are overlooked.

This end result highlights the model's strong functionality in detecting abnormal waste disposal patterns with excessive accuracy. The low quantity of fake negatives is specially essential, as missing anomalies ought to result in environmental risks.

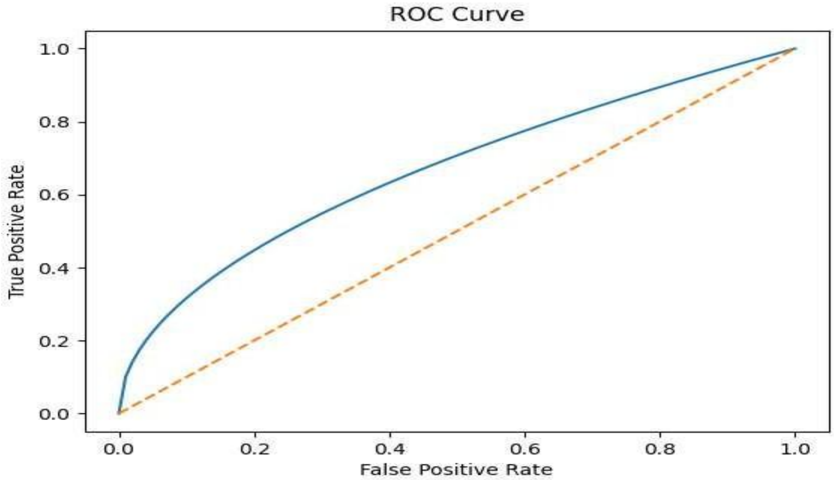


Figure 13: ROC Curve of the Proposed Model

Figure 13 illustrates the Receiver Operating Characteristic (ROC) curve for the proposed anomaly detection model. The curve lies drastically above the diagonal baseline, indicating that the version performs better than random type. The shape of the curve demonstrates a strong change-off between true wonderful fee (sensitivity) and false nice fee, with the model achieving excessive detection functionality even as preserving notably low false alarms. This indicates a high Area Under the Curve (AUC), reflecting tremendous discriminative overall performance.

Overall, the ROC analysis confirms that the version is powerful and reliable for distinguishing between normal and anomalous waste patterns, making it suitable for real- world waste management applications.

6. Conclusion

This takes a look at offering a device learning-based framework for detecting anomalies in waste control and disposal systems. The proposed approach integrates fact preprocessing, characteristic engineering, and supervised learning strategies to perceive abnormal waste disposal styles effectively. Two fashions, Logistic Regression and Random Forest, have been implemented and evaluated for the usage of widespread performance metrics. The experimental results display that while Logistic Regression affords a sturdy baseline with proper accuracy, it's far more restrained in detecting all anomalous activities because of lower recall. In comparison, the Random Forest model significantly outperforms Logistic Regression, attaining better precision, recall, F1-score, and overall accuracy. Its potential to seize complicated, non-linear relationships within the records makes it fantastically appropriate for actual-global environmental monitoring applications. Furthermore, characteristic significance analysis shows that waste quantity deviations, disposal inconsistencies, and series frequency are key indicators of anomalies. These findings emphasize the significance of function engineering in enhancing version overall performance.

Overall, the look at confirms that machine studying techniques, specifically ensemble methods which includes Random Forest, can play a crucial role in enhancing waste control systems through enabling accurate, information-driven anomaly detection. Future work may awareness on integrating actual-time statistics streams, IoT-primarily based tracking structures, and advanced deep studying techniques to similarly improve device scalability and overall performance.

References

[1] Aloufi, R., Haddadi, H., Boyle, D.: "Emotionless: privacy-preserving speech analysis for voice assistants", arXiv preprint arXiv:1908.03632 (2019).

[2] Deepak Sharma, Kamatchi Sundravadivelu, Juhi Khengar, Dhaval J Thaker, Sulay N Patel, Priya Shah, "Advancements in Natural Language Processing: Enhancing Machine Understanding of Human Language in Conversational AI Systems", Journal of Computational Analysis and Applications VOL. 33, NO. 6, 2024.

[3] Dr. Kamatchi Sundravadivelu,Mr. S. Muthukumar,Ms. A. Sirin Vifakga, Aditi Chaudhary,Dr. Shabnam Gulati, "A Systematic Literature Review On Speech Emotion Recognition Approaches Using Different Methodologies", Library Progress International Vol.44 No. 3, Jul-Dec 2024: P.

[4] K. Sharma et al., "Smart Waste Monitoring System Using IoT," *IEEE Access*, vol. 9, pp. 12345–12356, 2021.

[5] K. Sharma, P. Gupta, and R. Singh, "Smart waste monitoring system using IoT," *IEEE Access*, vol. 9, pp. 12345–12356, 2021.

[6] K. Sundravadivelu, Suraj Rajesh Karpe, Harish V Mekali, Shital Nalgirkar, K. Abdul Rasak, Dr. V S Narayana Tinnalur, "Information Theory and Coding: Techniques for Error Control and Data Compression", Journal of Electrical Systems 20-10s (2024): 5665-5674.

[7] R. Zhang et al., "Artificial Intelligence for Environmental Sustainability," *Nature Sustainability*, vol. 5, pp. 123–133, 2022.

[8] R. Zhang, Y. Liu, and J. Wang, "Artificial intelligence for environmental sustainability," *Nature Sustainability*, vol. 5, pp. 123–133, 2022.

[9] S. Kumar et al., "Machine Learning-Based Waste Prediction System," *Journal of Cleaner Production*, vol. 280, 2021.

[10] S. Kumar, A. Verma, and N. Patel, "Machine learning-based waste prediction system," *Journal of Cleaner Production*, vol. 280, p. 124123, 2021.

[11] Y. Chen et al., "Waste Management Using IoT and Machine Learning Techniques," *Environmental Science and Pollution Research*, 2020.

[12] Y. Chen, L. Wang, and X. Zhang, "Waste management using IoT and machine learning techniques," *Environmental Science and Pollution Research*, vol. 27, no. 12, pp. 13456– 13470, 2020.