FMDB ***Transactions on Sustainable Intelligent Networks***

PredictiWaste: An ML-Powered Framework for Sustainable Food Inventory Optimization in Restaurants

H. Rakesh[1,*], R. Harshad[2], K. Sanjay Kumar[3], R. Regin[4], A. Anithalakshmi[5], G. Padmapriya[6]

[1,2,3,4]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[5]Department of Science and Humanity, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
[6]Department of Chemistry, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India.
rh3510@srmist.edu.in[1], hr6127@srmist.edu.in[2], sk2520@srmist.edu.in[3], regin12006@yahoo.co.in[4], anithalakshmi@dhaanishcollege.in[5], ppstminex@gmail.com[6]

Abstract: Food wastage is a major concern in the food industry, affecting economic and environmental sustainability. This project introduces a predictive analysis model to reduce food waste in restaurant inventory and menu planning. The system uses machine learning techniques to forecast food wastage based on factors like food type, guest count, quantity, event type, storage conditions, purchase history, preparation methods, seasonality, location, and pricing. A dataset containing food wastage patterns is preprocessed to manage missing values and categorical variables. Label encoding is applied to categorical features, while numerical features are standardized. The data is split into training and testing sets, and a Random Forest Regressor is trained for accurate predictions. The model achieved 92.79% accuracy, with a Mean Absolute Error (MAE) of 1.63 and an R^2 score of 0.9279, proving its reliability. A Tkinter-based GUI lets restaurant managers input parameters and receive real-time wastage predictions. The system categorizes wastage as minimal, moderate, or high and suggests inventory control measures. Future improvements include integrating Long Short-Term Memory(LSTM) for time-series forecasting, IoT sensors for real-time storage monitoring, and blockchain for supply chain transparency. By minimizing food wastage at the inventory stage, this project helps businesses optimize resources while promoting sustainability.

Keywords: Predictive Analysis; Machine Learning; Random Forest Regressor; Mean Absolute Error (MAE); R^2 Score; Long Short-Term Memory (LSTM); IoT Sensors; Inventory Management.

Received on: 22/05/2024, **Revised on:** 19/07/2024, **Accepted on:** 30/09/2024, **Published on:** 14/12/2024

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSIN

DOI: https://doi.org/10.69888/FTSIN.2024.000289

Cite as: H. Rakesh, R. Harshad, K. S. Kumar, R. Regin, A. Anithalakshmi, and G. Padmapriya, "PredictiWaste: An ML-Powered Framework for Sustainable Food Inventory Optimization in Restaurants," *FMDB Transactions on Sustainable Intelligent Networks*., vol.1, no.4, pp. 201–219, 2024.

1. Introduction

Food waste is a critical global challenge that affects economic stability, social well-being, and environmental sustainability. Studies indicate that nearly one-third of all food produced is wasted annually, leading to significant financial losses and increased carbon emissions [1]. As the demand for food continues to rise, particularly in the hospitality and catering industries,

*Corresponding author.

addressing food waste has become essential. Researchers have explored predictive analytics and machine learning techniques to address this issue to enhance food demand forecasting, optimize inventory management, and minimize wastage. Several studies have demonstrated the effectiveness of machine learning models in tackling food waste. Rodrigues et al. [1] highlighted how predictive models like Random Forest, LightGBM, and LSTM reduced wasted meals by 14% to 52%, thereby improving efficiency in catering services. Similarly, Ifham et al. [3] emphasized the role of machine learning in forecasting ingredient demand and managing near-expiry food items to prevent unnecessary waste. Additionally, Nilashi et al. [6] investigated the impact of Big Data and Predictive Analytics (BDPA) on waste management, demonstrating how these technologies optimize supply chains while lowering operational costs. These findings emphasize the importance of predictive modelling in reducing food wastage across different sectors.

This project focuses on developing a machine learning-based predictive model for food wastage management, specifically in restaurant inventory and menu planning. By employing algorithms such as Random Forest, XGBoost, and Gradient Boosting, the system analyzes key factors influencing food wastage, including event type, storage conditions, pricing, seasonality, and guest count. The best-performing model, Random Forest Regressor, achieved an accuracy of 92.79%, outperforming other models in predicting wastage trends. Furthermore, a Graphical User Interface (GUI) built using Tkinter enables real-time predictions, allowing restaurants to make data-driven decisions to reduce food waste effectively. This study aims to support sustainable food management practices by integrating predictive analytics and machine learning. Future enhancements could incorporate IoT-based monitoring for real-time storage tracking and blockchain technology for improved transparency in supply chains. This research provides a practical, data-driven approach to minimizing food waste, aligning with global sustainability goals, and promoting responsible consumption.

1.1. Background Research

Food wastage has become a critical global issue, impacting economic stability, social well-being, and environmental sustainability. Studies show that nearly one-third of all food produced globally goes to waste yearly, resulting in financial losses, food insecurity, and increased pollution [1]. In particular, the hospitality and restaurant industries contribute much of this wastage due to inaccurate demand forecasting, inefficient inventory management, and improper storage conditions [2]. Predictive analytics and machine learning have emerged as powerful tools to address this problem, allowing businesses to make data-driven decisions that optimize food usage and minimize unnecessary waste [3].

Recent research highlights the role of artificial intelligence (AI) in improving food demand forecasting. Machine learning models such as Random Forest, XGBoost, and LSTM have been successfully applied to predict food consumption trends and reduce wastage in catering and restaurant industries [4]. For instance, Rodrigues et al. [1] demonstrated how machine learning models could reduce wasted meals by 14% to 52%, significantly improving operational efficiency. Similarly, Nilashi et al. [6] explored how Big Data and Predictive Analytics (BDPA) enhance food supply chain management, leading to better inventory control and lower waste. These studies highlight the importance of predictive modelling in transforming food waste management practices. Given the widespread impact of food wastage, there is a growing need for innovative solutions that integrate AI and data analytics. By developing an intelligent prediction model, businesses can proactively manage food inventory, reduce overproduction, and implement better portion control strategies. This project focuses on applying machine learning techniques to predict food wastage accurately, helping restaurants and food service providers make informed decisions to minimize waste and improve sustainability.

1.2. Problem Statement

Managing food supply efficiently remains a significant challenge in the food industry. Restaurants, catering businesses, and food suppliers often struggle to accurately estimate the required food quantity, leading to excessive waste or shortages. When demand is overestimated, large amounts of food go unused, resulting in financial losses and increased environmental impact due to disposal. On the other hand, underestimating demand leads to food shortages, affecting customer satisfaction and overall business operations. Wasted food increases costs and contributes to the growing issue of food insecurity and carbon emissions, making it essential to find a solution that balances supply and demand.

Many businesses still rely on outdated inventory management methods, using past experiences, manual tracking, or simple estimation techniques that often lack precision. These traditional approaches fail to consider external factors such as seasonality, event types, storage conditions, and pricing variations, significantly influencing food consumption patterns. Without accurate forecasting tools, businesses risk making costly errors in inventory planning. Adopting machine learning and predictive analytics presents an opportunity to address this problem by leveraging historical consumption data and external influences to generate reliable wastage forecasts. Implementing a data-driven predictive system can help food businesses reduce waste, optimize resources, and operate more efficiently while supporting sustainable food management practices.

1.3. Objectives

The primary objective of this project is to develop a predictive model that helps restaurants and food service providers minimize food wastage through machine learning techniques. The specific goals of this study include:

- **Developing an Accurate Predictive Model:** Implement machine learning algorithms such as Random Forest, XGBoost, and Gradient Boosting to forecast food wastage based on historical consumption data and key influencing factors.
- **Enhancing Demand Forecasting:** Utilize data-driven insights to predict food demand more accurately, preventing overproduction and reducing inventory losses.
- **Building a User-Friendly Application:** Develop a Tkinter-based Graphical User Interface (GUI) that allows restaurant managers to input relevant parameters and receive real-time predictions on food wastage levels.
- **Optimizing Inventory Management:** Identify key factors contributing to food wastage, such as seasonality, storage conditions, and pricing, to help businesses adjust their procurement and storage strategies effectively.
- **Improving Sustainability Practices:** Support environmental sustainability by reducing food waste, lowering greenhouse gas emissions, and promoting responsible consumption within the food industry.
- **Exploring Future Enhancements:** Investigate the potential integration of Long Short-Term Memory (LSTM) models for time-series forecasting and IoT-based sensors for real-time inventory monitoring to improve prediction accuracy and waste management.

By achieving these objectives, this project aims to create a reliable and efficient system that empowers food businesses with actionable insights, enabling them to make smarter decisions in food inventory planning and waste reduction.

2. Review of Literature

Rodrigues et al. [1] proposed an innovative approach utilizing machine learning models to boost demand forecasting within food catering services. Their findings illustrate a noteworthy capacity to diminish both food waste and instances of unmet demand markedly. The research indicates that by employing advanced predictive models such as Random Forest, LightGBM, and LSTM, wasted meals could be reduced by an impressive 14% to 52%. As forecasting accuracy improves, catering operations become more efficient and take on a more environmentally conscientious profile. The results show a considerable promise that sophisticated predictive methodologies hold good for food waste management in the food catering industry.

Arvindaraj et al. [2] talk about the importance of food in everyday life, underlining its place in our culture and the need for working ingredient verification in food recipes. The high demand in the food and catering sector, plus the competition of online recipe websites, shows the importance of technology. The study shows that the AdaBost Gradient Boosting Algorithm (AGBA) model can classify the food ingredients of a given recipe with an accuracy of 99.8%. Reduced Human Errors: This advancement minimizes human errors, allows for better resource utilization, and makes catering operations more efficient overall. The results emphasize how AI-enabled techniques can transform the planning, categorizing of food, and sustainability contributions of ingredients to the catering industry.

Ifham et al. [3] discuss the global challenge of food waste and its economic, social, and environmental consequences, emphasising the challenges faced by a developing country — Sri Lanka. Abstract — As food waste is a key problem in food security, the study emphasizes that approximately one-third of the world's food supply is wasted annually. This study indicates the importance of applying machine learning methods to successfully address food waste prediction and forecast ingredient demand and near-expiry food prices. We also show that the proposed models effectively minimize food waste, maximize sustainability, and improve efficiency in food management. Additionally, recycling optimization with AI is a viable solution for expired processed food. The results reflect an increased commitment to more sustainable and intelligent food waste management in catering and food service industries.

Thay and Chinda's [4] study on food waste management in Phnom Penh, Cambodia, highlights the significance of this practice. The use of biodigesters, family size, packaging size, population growth, sales promotion, shelf-life, frequency of shopping, and tax payment are among the eight major factors identified in the study as impacting food waste management. Better methods for minimizing food waste can be developed by better understanding these variables. As possible remedies, they also recommend using bio-digesters and changing packaging dimensions. According to these results, tackling these issues can significantly enhance food waste management procedures, promoting sustainability and preservation. Future studies might create a dynamic model based on these variables to monitor and control food waste over an extended period.

Anggraeni et al. [5] detail how the usage of Bayesian Networks (BN) and Agent-Based Modeling (ABM) have been instrumental in analyzing waste patterns and guiding policy interventions. At the household level, BN-ABM models reveal that structured economic incentives and awareness campaigns could reduce waste by up to 20%. Meanwhile, in the retail sector, ABM can highlight that with strong retailer networks and consumer awareness increasing adoption rates—leading to waste reductions of 2–3.6% in the Netherlands and 0.8–2.1% in Italy. Supply chain simulations further illustrate a trade-off between efficiency and resilience, where highly efficient networks are more vulnerable to shocks, but less efficient structures demonstrate better recovery.

Nilashi et al. [6] investigate how Big Data and Predictive Analytics (BDPA) have transformed waste management, particularly in the food recycling industry, improving efficiency and sustainability. With landfill-related methane emissions accelerating climate change, BDPA adoption offers a solution by enhancing environmental and economic performance. Applying Partial Least Squares Structural Equation Modeling (PLS-SEM) to a survey of 130 respondents in Saudi Arabia showed that employee knowledge and competitive pressure are significant factors driving the adoption of BDPA. However, nearly 80% of BDPA projects failed due to inadequate strategic planning, thus highlighting the importance of a structured implementation approach. The Resource-Based View (RBV), Technology-Organization-Environment (TOE), and Human-Organization-Technology (HOT-fit) theories indicating organizations that effectively integrated BDPA could optimize their supply chains and lower operational costs.

Dagiliūtė and Musteikytė [7] discuss in detail the significant issue in Lithuania regarding Food wastage, with 76% of Europeans recognizing individual responsibility for waste prevention. In a survey of 174 respondents, 91% had heard about food waste issues, yet 66% lacked awareness of its environmental impact. 52% of respondents acknowledged throwing away edible food, 12% converted it to animal feed, and 88% disposed of food waste with regular municipal trash, with a mere 9% composting. 14% of respondents wasted food daily, while 28% threw out food two to three times each week. Expired food (44%), loss of freshness (18%), and over-purchasing (59%) were the leading causes of food wastage. Restaurants wasted 14,744 kg being discarded over six months, where 40% of grain products and 5% was meat. Diners frequently avoided taking leftovers home due to shame (43%), despite 73% finishing meals.

Pereira et al. [8] researched Industrial Kitchens (IKs) operations, specifically on waste generation, electricity, and water consumption. The project was implemented in Portuguese luxury hotels and the food preparation sector to improve operational efficiency, staff experience, and sustainability. The main objective was understanding the interactions between waste production and resource consumption. The study collected organic (undifferentiated) and inorganic (glass, paper, plastic) waste data using ultrasonic sensors attached to waste bins. The data collected provided valuable insights into how waste production correlates with electricity and water usage in industrial kitchens. The study collected organic (undifferentiated) and inorganic (glass, paper, plastic) waste data using ultrasonic sensors attached to waste bins.

Turker [9] explores food waste reduction in campus dining through data-driven demand prediction. This study applied six machine learning models—Linear Regression, Extra Tree Regressor, Lasso, Decision Tree Regressor, XGBoost Regressor, and Gradient Boosting Regressor—to forecast food preferences and daily meal consumption. The lasso model achieved the highest accuracy ($R^2 = 0.999$), and the XGBoost Regressor with accuracy ($R^2 = 0.882$). Using attributes such as meal variety, revenue, campus mobility, temperature, and food wastage was reduced to 28%, ensuring better demand-supply alignment. Data from Suleyman Demirel University included 215,504 campus movements and 55,006 meal counts over a semester. Menu composition had a stronger impact on dining hall attendance than overall campus density.

Nijloveanu et al. [10] analyze the issue of food loss and waste (FLW) in consumer segments of the agrifood chain, emphasizing its economic and social impacts. The study introduces a predictive model using machine learning methods. Specifically, Decision Tree and Neural Network approaches are used to determine the effects of food loss and waste based on collected statistical data from Romania. The findings indicate that the Neural Network approach achieved relatively high-accuracy predictions(Accuracy=78%), outperforming the Decision Tree model(Accuracy=31.5%). Statistical analysis of the dataset, which includes responses from 365 participants, demonstrated significant correlations between consumer behaviour, purchasing habits, and food waste tendencies. Targeted interventions in consumer awareness and supply chain optimization can enhance sustainability and economic efficiency.

Fatorachian and Pawar [11] explore the AI-driven predictive models for optimizing waste management in cold chain logistics using ARIMA and Multiple Linear Regression (MLR). ARIMA effectively captured seasonal demand trends, whereas MLR used multivariable attributes such as temperature, product type, and promotional activity. The models were trained on 350,000 sales records across 174 products and 33 customers and helped reduce food loss by 15%. ARIMA achieved a Mean Absolute Error (MAE) of 4.5%, and MLR had an MAE of 5.1%, which was useful for mid-sized cold chains. The Resource-Based View (RBV) and Supply Chain Resilience Theory highlighted AI’s effectiveness in reducing overproduction spoilage. The research suggests IoT sensor integration could further enhance forecasting precision, supporting food and pharmaceutical logistics sustainability.

Jansson-Boyd and Mul [12] examine the most important drivers contributing to household food waste minimization in the UK from semi-structured interviews with participants between 26 and 65 years old. The results showed that 61% of the food waste in the world comes from households alone. The thematic analysis confirmed that extensive thinking, flexible behaviours, and emotional involvement were the most important drivers. Volunteers who actively cut down on waste indicated that they thought about wasting food 3-4 times a week, which was related primarily to cooking. Guilt and regret were significant factors, particularly among those with strong anti-waste sentiments. Unlike past studies, social influence was not very significant, with most participants not considering their behaviour based on other's opinions.

Leal Filho et al. [13] analyze university food waste reduction strategies, emphasizing the environmental, social, and economic consequences of excessive food disposal in higher education institutions. The global waste generation varies from 0.12 to 50 kg per person daily. Gender, age, season, and consumer habits play a role in this issue. However, initiatives like composting, trayless dining, and educational programs have cut food waste by 13% to 50%. A qualitative analysis of case studies from North America, Asia, and Europe found that many universities do not have standardized methods for measuring waste, with over 60% not accurately tracking food waste. A study from China revealed that 46.8% of the waste was vegetables, 36.2% grains, and 13.9% meat. Approaches such as pre-booked meals, better portion control, and food recovery programs have proven effective in minimizing waste.

Onyeaka et al. [14] explain how artificial intelligence improves the circular economy and reduces food wastage by maximizing resource efficiency and reducing the negative environmental impact. It is to be noted that every year, almost 1.3 billion tonnes of food are wasted and generate a whopping 3.3 gigatons of carbon emissions. Food insecurity and environmental negative effects are exacerbated in high-income nations, where each individual produces an average of 79 kilograms of food waste annually. AI applications like blockchain, IoT, and machine learning in waste management, food redistribution, and predictive analytics could be investigated for more accurate resource allocation.

Chowdary et al. [15] employed five machine-learning models for food waste prediction: XGBoost, AdaBoost, CatBoost, SVM, and Gaussian Naïve Bayes. The best performance was shown by XGBoost, with the values MSE, RMSE, and MAE being 31.05, 5.57, and 4.13, respectively, and an R^2 score of 0.73. Feature importance analysis showed that guest count, food type, and seasonality significantly contributed to food waste. Integrated diverse datasets to improve prediction accuracy. Data quality, model interpretability, and computational efficiency were still hard to tackle.

3. Proposed Methodology

The methodology to forecast food wastage in restaurant stocks is a formal process. It starts with data collection and cleansing, feature selection of the most important features, and training, testing, and model evaluation. The overall aim is to establish a sound system that can be used to forecast food waste in restaurants. Ensemble learning, conventional regression models, and deep learning are among the methods employed to improve the accuracy of predictions. The data used in this project comes from various food-related documents, which collect variables influencing food wastage. These include event type, storage conditions, location, seasonality, buying history, and price. By knowing these, one can detect patterns that contribute to food wastage. To ensure the dataset is clean and ready, missing values are handled using forward fill for time series, mean or median imputation for numerical attributes, and mode imputation for categorical attributes. The data uses both numeric and categorical values, and thus, encoding schemes such as Label Encoding and One-Hot Encoding are used wherever necessary. Numeric attributes are also normalized to ensure consistency during the development of model training.

The models used in the project are regression models, ensemble learning models, and deep learning models. Regression models such as Linear Regression, Ridge Regression, and Lasso Regression specify the relationship between input features and food wastage with the inclusion of techniques to prevent overfitting. Ensemble learning algorithms such as Random Forest, Gradient Boosting, XGBoost, and Extra Trees Regressor use numerous decision trees and boosting techniques to enhance accuracy and stability. Deep learning techniques like Long Short-Term Memory (LSTM) networks are used to analyze time trends in the wastage of foods. The models are trained in an 80-20 train-test split supported by an additional 10% validation set to fine-tune parameters for deep learning models. Hyperparameter tuning techniques such as grid search, random search, and Bayesian optimization are employed to enhance the model's performance by attempting various combinations of parameters. These observations assist us in validating that predictions are correct and reliable. After identifying the highest-performing model, then the model is implemented as a web application using Flask. The trained model is saved in an easily accessible and combined format. A Flask API is created to process user inputs, perform calculations, and predict food wastage.

To make this model user-friendly, an interface is created so that users can easily provide data and see results without any complexity. This not only makes the model precise but also feasible for practical use. With various modelling methods and efficient data preprocessing, the system offers valuable information that can be utilized to enhance food management policy. The forecasts assist organizations and businesses in making informed decisions, reducing food waste, optimizing stock planning, and establishing a sustainable food culture of consumption. The performance of each machine learning model has been measured by utilizing primary performance metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), R-squared (R^2) Score, and Root Mean Squared Error (RMSE). These performance metrics help us to ensure predictions are reliable and accurate. Once the best-performing model is determined, the model will be deployed as a web application through Flask. The model is saved in a convenient, integrated format. A Flask API handles user input, calculates, and generates food waste predictions.

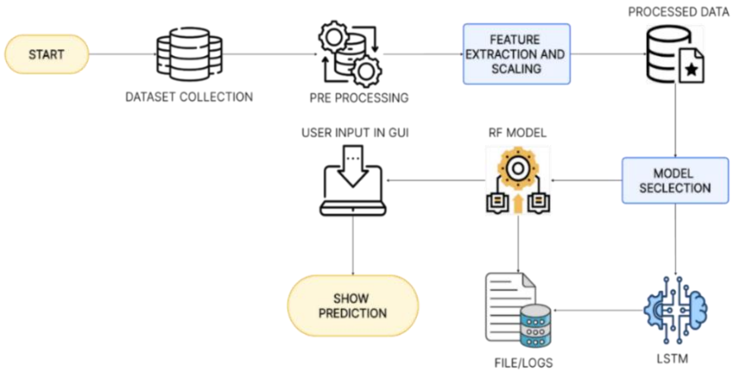


Figure 1: PredictiWaste Framework

Figure 1 PredictiWaste Framework diagram describes a machine learning-based prediction system that elaborates on the processes that run from data collection for prediction outputs and followed by cleaning and preparation. Feature extraction and scaling are performed to refine the data and make it suitable for model training. The preprocessed data is sent for model selection, after which a model is selected based on the prediction needs. Once selected, the model acts on the data and logs files for record-keeping and analysis. The user interacts with the model through a GUI to input real-time predictions. To ensure that the system is user-friendly, an easy-to-use and user-friendly interface is designed so that users can easily provide input and interpret results without any complication. This renders the model not only accurate but also viable for real-world usage. The system provides valuable information that can promote food management policies through different modelling techniques and effective data preprocessing. These predictions benefit companies and organizations by helping them make informed decisions, reducing food wastage, optimizing stock planning, and maximizing a more sustainable food consumption culture.

3.1. Data Preprocessing and Feature Engineering

The model's data preprocessing and feature engineering steps are determined to ensure the dataset is clean, suitable to algorithms, and structured. The data is loaded from the CSV file using the Pandas library. This dataset contains various food wastage features, such as preparation methods and target variables. The main feature is the geographical location and seasonality. These features provide intuition about contributing food wastage and are used to train the predictive model. Since the machine learning model requires numerical inputs, these entities are transformed using label encoding. The Label encoding assigns a unique numerical value to each entity in the dataset. For example, an entity like "Type of food" with values such as destroyable or undestroyable is converted into numerical values like zero denotes destroyable and 1 denotes the undestroyable. A dictionary is managed to store these mappings, which ensures consistency when encoding new data. This makes the model split the data more efficiently without getting into unnecessary problems.

After categorizing entities is encoded, the dataset is split into features and target variables. This is applied to all independent entities that exclude wastage of food amount, which is a dependent variable. This separation ensures that the machine learning model understands the patterns for all the inputs to be evaluated and predicts the wastage amount accurately. By structuring the dataset, the model can find the associations and dependencies among the various features affecting food wastage. Feature scaling is an important step in preprocessing, as a dataset consists of numerical values with different ranges. To standardize the dataset, the Standarscaler from the sklearn library is used. This transforms numerical features into a mean of 0 and a standard deviation 1. It also ensures that all entities contribute equally to the model learning process. Also, this prevents the features with larger numerical value ranges from controlling the model's learning process, thereby improving the model's performance. The changes are applied separately to training and test datasets to prevent data outflow and ensure unbiased model evaluation.

This model's feature engineering process primarily focuses on encoding categorical data, scaling numerical features, and structuring the dataset for effective learning. The preprocessing transformations, including label encoders and scalers, are saved using the joblib library. This library allows for reusing the encoders and scalers when making new predictions to ensure precision in data transformation. This model retains all features to maximize predictive accuracy. Even though more advanced techniques, such as correlation analysis or mutual information selection, are possible, the current method supports the datasets so that no information is lost or omitted.

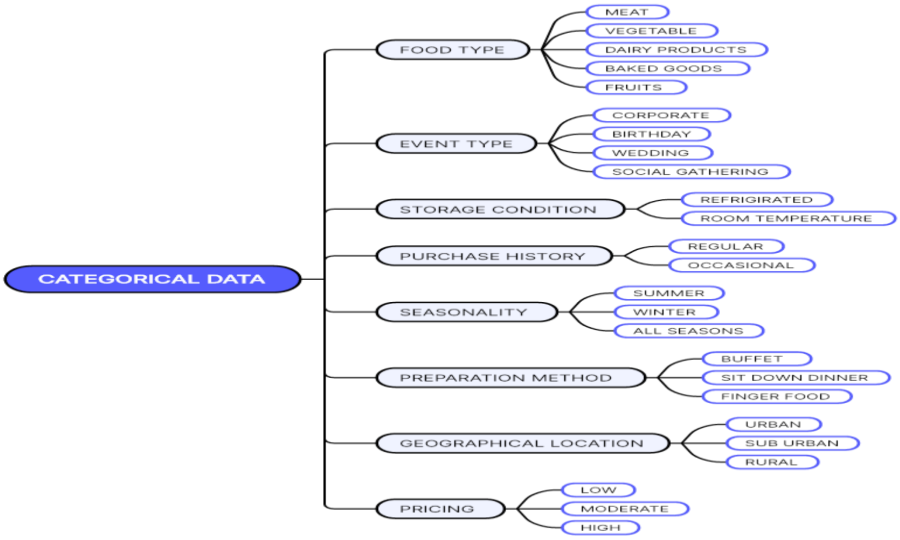


Figure 2: Categorical Data Mapping of the food wastage dataset

Figure 2 is a mind map of the categorical attributes of the dataset, which are significant in predicting and understanding food wastage trends. They are food types that classify various items, including perishable and non-perishable, and affect their propensity for wastage. Event Type influences food consumption since various events, such as weddings, corporate conferences, or festivities, have varying food requirements. Storage Condition influences wastage since proper refrigeration, freezing, or room temperature storage influences the shelf life of food products. Purchase History helps understand consumers' purchase patterns and thereby identifies excess stock that may contribute to increased levels of wastage. Seasonality of food availability impinges on demand and supply and, by association, wastage levels. Method of Preparation is also a contributor, with variations in cooking style like frying, boiling, or baking influencing food spoilage and resulting wastage levels. Geographical location is an influence due to varying climatic conditions, Available supplies, and storage facilities. Finally, pricing influences consumer consumption behavior, such that increased food prices can result in more careful consumption and less wastage. These categorical data attributes offer useful information regarding food waste trends, enabling businesses and organizations to adopt improved planning and resource management practices to reduce waste.

Table 1: Food Wastage Dataset Description

Attribute Name	**Type**	**Description**
Type of Food	Categorical	Type of food being served (e.g., Meat, Vegetables, etc.).
Number of Guests	Numerical (int)	The number of guests attending the event. Ranges from 207 to 491.
Event Type	Categorical	The type of event (e.g., Corporate, Birthday, etc.).
Quantity of Food	Numerical (int)	The total quantity of food prepared for the event. Ranges from 280 to 500.
Storage Conditions	Categorical	How the food is stored (e.g., Refrigerated, Room Temperature, etc.).
Purchase History	Categorical	The frequency of food purchases (e.g., Regular, Occasional, etc.).
Seasonality	Categorical	The season during which the event occurs (e.g., Summer, Winter, All Seasons).
Preparation Method	Categorical	The food preparation method (e.g., Buffet, Finger Food, etc.).
Geographical Location	Categorical	The location where the event is held (e.g., Urban, Suburban, Rural).
Pricing	Categorical	The cost of food (e.g., Low, Moderate, High).
Wastage Food Amount	Numerical (int)	The amount of food wasted after the event. Ranges from 10 to 63.

Table 1 provides a detailed description of the dataset used for food wastage prediction. It includes various attributes such as the type of food served, number of guests, event type, and food quantity, which influence food wastage. Factors like storage conditions, purchase history, seasonality, and geographical location are considered in analyzing their impact. The dataset also incorporates pricing and preparation methods, helping to understand patterns in food wastage and optimize inventory planning.

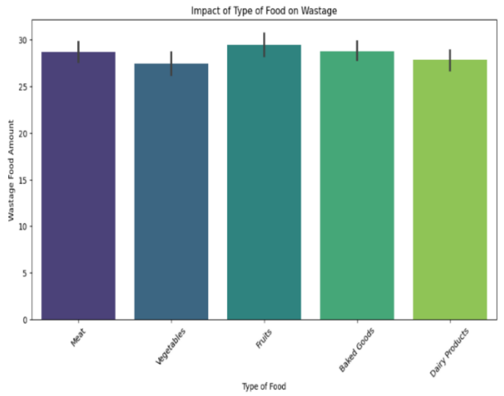


Figure 3: Bar Plot - Type of Food vs. Food Wastage

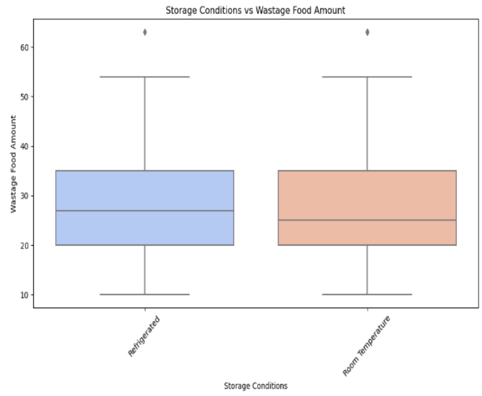


Figure 4: Box Plot - Storage Conditions vs. Food Wastage

In Figure 3, the bar chart visually represents the relationship between different food types and the amount of food wastage; among the five categories—Meat, Vegetables, Fruits, Baked Goods, and Dairy Products—Meat and Fruits exhibit the highest levels of wastage, while Dairy Products and Baked Goods show slightly lower amounts. The presence of error bars suggests variations in data distribution, indicating that some food types experience more inconsistent wastage patterns than others. In Figure 4, the box plot highlights the influence of storage conditions on food wastage. It compares refrigerated storage with room temperature storage, showing that food stored at room temperature tends to have higher wastage. The wider range and presence of outliers in the room temperature category suggest that perishable food deteriorates faster when not stored properly. This reinforces the importance of optimal storage conditions in minimizing food wastage.

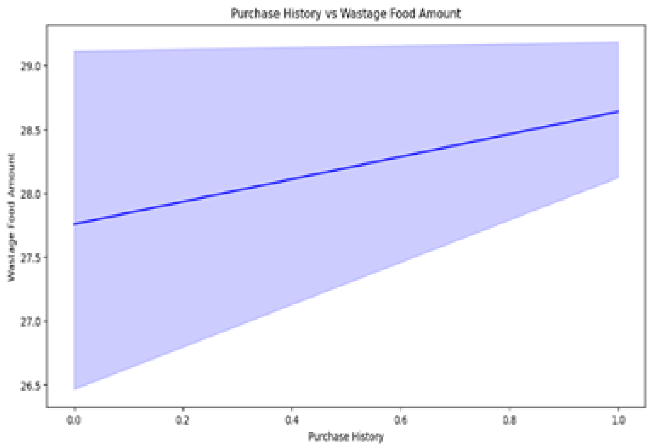


Figure 5: Line Plot - Purchase History Trend

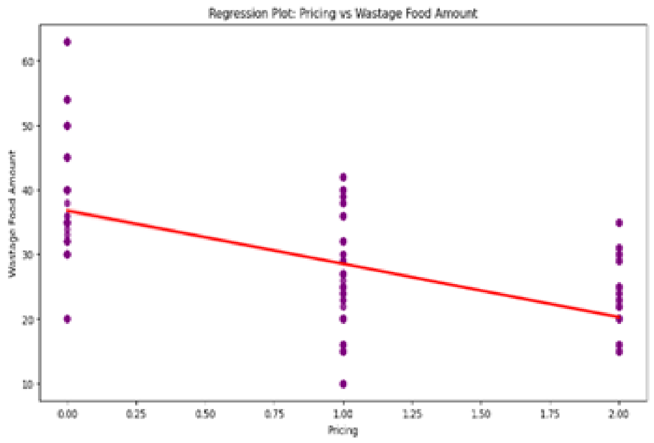
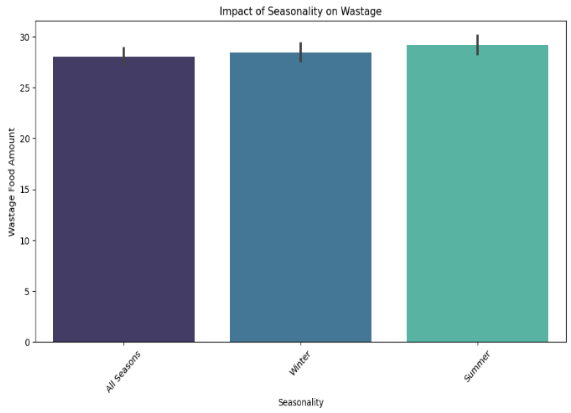


Figure 6: Regression Plot - Pricing vs. Food Wastage

In Figure 5, the diagram illustrates the relationship between purchase history and food wastage. The positive trend in the regression line indicates that food wastage also tends to rise as purchase frequency increases. This suggests that customers who buy more frequently may waste more food due to over-purchasing or lack of proper storage. The shaded region represents the confidence interval, showing variability in wastage patterns. In Figure 6, the negative slope of the regression line suggests that higher pricing is associated with lower wastage. This implies that when food items are priced higher, people tend to be more cautious and waste less. The scattered data points indicate some variability, meaning that while pricing influences wastage, other factors may also play a role. These insights can help design better pricing strategies and awareness programs to minimize food waste in restaurants or households.



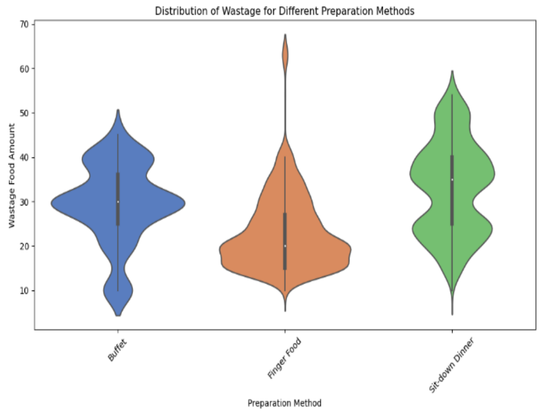
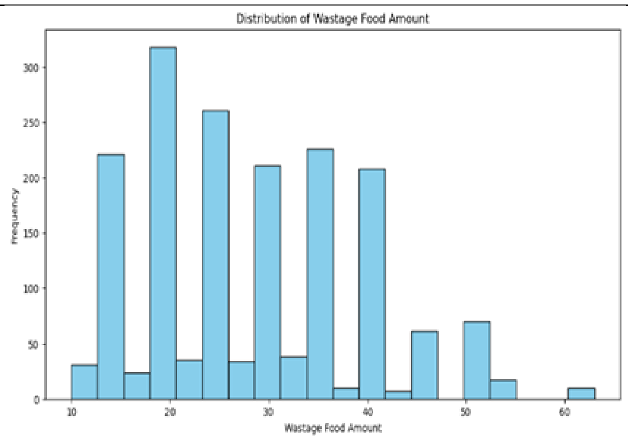


Figure 7: KDE Plot - Seasonality vs. Food Wastage

Figure 8: Violin Plot - Preparation Method vs. Food Wastage

In Figure 7, the diagram illustrates the impact of seasonality on food wastage. It compares wastage levels across different seasons, namely all seasons combined, winter and summer. The bar heights suggest that food wastage remains fairly consistent across seasons, with only minor variations. This implies that seasonal changes do not significantly affect food wastage, which may be due to factors like proper storage and preservation methods. In Figure 8, the diagram represents the distribution of food wastage across different preparation methods using a violin plot. The three preparation methods—buffet, pre-prepared food, and on-demand preparation—show varying levels of wastage. Buffet-style meals exhibit a wider distribution, indicating a more varied wastage pattern. Pre-prepared food also has significant wastage, whereas on-demand preparation shows a relatively controlled distribution, suggesting it leads to less waste. This analysis highlights that food preparation methods are crucial in determining wastage, with on-demand methods potentially more efficient in reducing excess food.



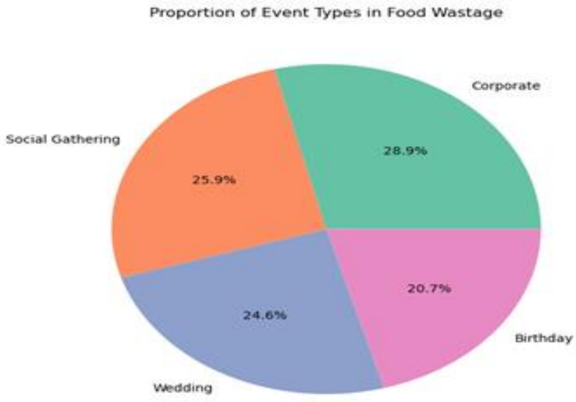


Figure 9: Histogram- Distribution of Food Wastage

Figure 10: Pie Chart - Proportion of Event Types

In Figure 9, the histogram shows the distribution of food wastage amounts. The x-axis represents different amounts of wasted food, while the y-axis shows their frequency. The chart indicates that most wastage occurs between 10 and 40 units, with peaks around 20 and 30. Higher wattage values beyond 40 are less frequent. The uneven distribution shows that some food waste is more common than others.

In Figure 10, the pie chart illustrates the proportion of different event types that contribute to food wastage. Corporate events have the highest share (28.9%), followed by social gatherings (25.9%) and weddings (24.6%). Birthday events contribute the least (20.7%). This suggests that large formal events generate more waste than smaller gatherings.

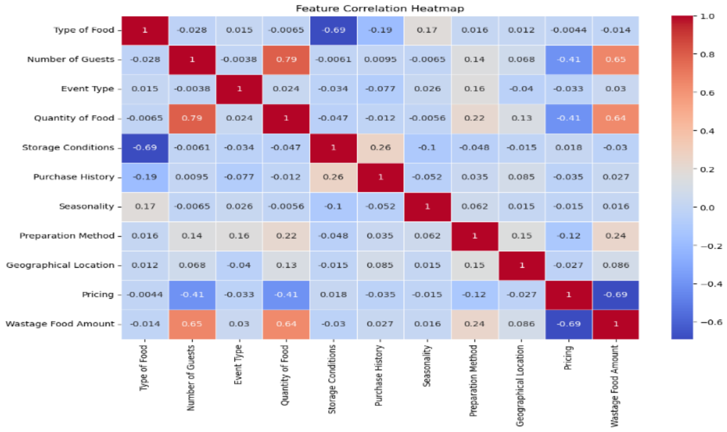


Figure 11: Heatmap - Feature Correlation in Food Wastage

Figure 11 heatmap displays correlations between different factors affecting food wastage. Darker red indicates a strong positive correlation, while blue shades show a negative correlation. Key insights include a strong correlation between the number of guests, event type, and food wastage and a negative correlation between storage conditions and wastage, suggesting better storage reduces wastage.

3.2. PredictiWaste RF-LSTM Algorithm

The food wastage prediction system follows a structured process that involves data preparation, model training, comparison, and GUI implementation using Tkinter. This algorithm integrates traditional machine learning models, an advanced LSTM-based deep learning approach, and an interactive user interface for real-time predictions.

Step 1: Import Required Libraries

- Load necessary libraries such as pandas, numpy, matplotlib, seaborn, sci-kit-learn, TensorFlow.keras, and Tkinter. Use joblib for model saving and its themes for UI enhancements.

Step 2: Load and Preprocess the Dataset

- Read the food_wastage_data.csv file. Handle missing values using forward fill, where missing data points are replaced with the previous valid value in time-series data.
- Identify categorical and numerical features. Convert categorical columns into numerical values using LabelEncoder and store the mappings for reference. Normalize numerical features using StandardScaler to ensure uniformity.

Step 3: Data Splitting and Scaling

- Define X (features) and y (target variable). Split the dataset into 80% training and 20% testing for model training. Use StandardScaler to normalize feature values to improve the performance of models.

Step 4: Train and Evaluate Machine Learning Models

- Train multiple regression models, including:
  - Linear Regression, Ridge Regression, and Lasso Regression (traditional models).
  - Random Forest Regressor, XGBoost, Extra Trees Regressor, Gradient Boosting Regressor, AdaBoost Regressor (ensemble learning models).
- Evaluate each model using the following metrics:

In the following equation, we use,

$$y_i = \text{Actual value of food wastage for the } i^{th} \text{ data point.}$$
$$\hat{y}_i = \text{Predicted value of food wastage for the } i^{th} \text{ data point.}$$
$$\bar{y} = \text{Mean (average) of all actual values.}$$

$n =$ Total number of observations (data points).

3.2.1. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \quad (1)$$

In Equation 1, MAE calculates the average absolute difference between the actual and predicted values. It shows how much the predictions deviate from the real values without considering direction.

Interpretation:

- Lower MAE means the predictions are closer to actual values.
- It is easy to interpret since it maintains the original unit of measurement.

3.2.2. Mean Squared Error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \quad (2)$$

In Equation 2, MSE measures the average squared differences between actual and predicted values. Squaring the errors gives more weight to larger errors, making them sensitive to outliers.

Interpretation:

- A smaller MSE means better accuracy.
- The squared term penalizes larger errors, so models with more extreme mispredictions will have a higher MSE.

3.2.3. Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \quad (3)$$

In Equation 3, RMSE is simply the square root of MSE, making it more interpretable by converting the error back to the original unit of measurement.

Interpretation:

- RMSE provides a clear understanding of the average magnitude of error.
- It is useful when larger errors must be penalized more than smaller ones.

3.2.4. R^2 Score (Coefficient of Determination)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (4)$$

In Equation 4, the R^2, or the coefficient of determination, refers to measures of how well the independent variables explain the variance in the dependent variable. It ranges from 0 to 1, where 1 indicates perfect predictions.

Interpretation:

- $R^2 = 1 \rightarrow$ The model perfectly predicts the values.
- $R^2 = 0 \rightarrow$ The model performs as poorly as a random guess.
- A higher R^2 score means the model explains more variation in food wastage.

Step 5: Select the Best Model

- Compare models based on their R^2 score and error metrics and select the model with the highest accuracy and lowest error values. Save the best-performing model using joblib.dump() for future use.

Step 6: Implement Deep Learning with LSTM

- Reshape input data into a 3D format (samples, time steps, features) to fit LSTM requirements. Construct an LSTM-based neural network with:
  - LSTM layers to capture sequential dependencies. Dropout layers to prevent overfitting. Activation functions such as ReLU, Tanh, or Softmax.
- Compile the model using the Adam optimizer and Mean Squared Error (MSE) loss function. Train the LSTM model using a defined number of epochs and batch size.
- Evaluate performance using MAE, MSE, RMSE, and R^2 Score. Save the trained LSTM model in .h5 format for later predictions.

Step 7: Develop a Tkinter GUI for Predictions

- Create an interactive GUI using Tkinter with ThemedTk() for an improved user experience.
- Add input fields for user-provided features such as event type, storage conditions, location, and purchase history.
- Implement a dropdown menu for categorical variables with pre-defined options.
- Provide a "Predict" button to generate real-time predictions.

Step 8: Handle User Input and Predictions

- On button click, fetch user inputs and encode categorical values using the stored mappings. Scale numerical inputs using StandardScaler to match the trained model format.
- Load the saved machine learning model or LSTM model for prediction. Predict food wastage levels and display the result using a message box.showinfo().

Step 9: Visualize Insights and Trends

- Generate graphical representations to compare model performance and provide meaningful insights to help restaurants optimize food management and minimize wastage.

Step 10: Deploy and Run the Application

- Start the Tkinter event loop (root. main loop ()) to keep the GUI running. Ensure smooth interaction between user inputs and model predictions.
- Optimize for real-time usability, allowing restaurant owners to make informed decisions on food inventory.

4. Result

The results of this project provide valuable insights into food wastage prediction using machine learning and deep learning models. Various algorithms were evaluated based on accuracy, mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). The findings highlight the effectiveness of advanced models like Random Forest and LSTM in predicting wastage patterns. These results can help optimize inventory management and reduce unnecessary food loss in restaurants. The study revealed that machine learning models can significantly improve demand forecasting, enabling restaurants to make more data-driven decisions. The Random Forest Regressor achieved the highest accuracy of 92.79%, with an MAE of 1.63, demonstrating its effectiveness in predicting food wastage. The LSTM model also showed promising results by capturing time-dependent patterns, making it suitable for long-term forecasting. By implementing these predictive models, businesses can reduce operational costs and contribute to sustainability by minimizing food waste. Integrating this technology into restaurant management systems can lead to more efficient inventory planning, ultimately improving overall efficiency in the food industry.

4.1. Performance Metrics Evaluation

Table 2 presents the performance metrics of various machine learning models for food wastage prediction. Random Forest Regressor achieved the highest accuracy (92.79%) with the lowest MAE and MSE, making it the most effective model. Simpler models like Linear and Lasso Regression performed poorly, highlighting the need for more complex algorithms for accurate predictions.

Table 2: Model Performance Metrics

Model	Accuracy (%)	MAE	MSE	RMSE	R^2 Score
Random Forest Regressor	92.791006	1.631482	7.472684	2.733621	0.927910
XGBoost Regressor	92.034686	1.925293	8.256666	2.873442	0.920347
Extra Trees Regressor	91.892286	1.575434	8.404277	2.899013	0.918923
Gradient Boosting Regressor	90.766746	2.341782	9.570987	3.093701	0.907667
AdaBoost Regressor	77.826506	3.774461	22.984555	4.794221	0.778265
Ridge Regression	64.407032	4.744169	36.894885	6.074116	0.644070
Lasso Regression	64.405123	4.739213	36.896864	6.074279	0.644051
Linear Regression	64.403628	4.744874	36.898414	6.074406	0.644036

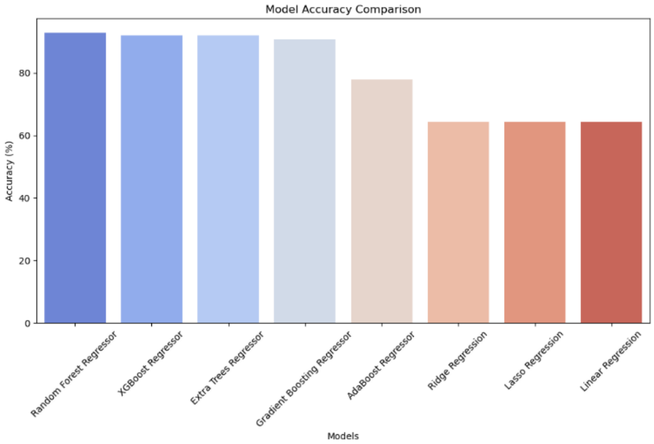


Figure 12: Model Accuracy Comparison

In Figure 12, the bar chart compares the accuracy of different regression models. Random Forest, XGBoost, and Extra Trees Regressors exhibit the highest accuracy, making them the best-performing models. Gradient Boosting follows closely, while AdaBoost shows a noticeable drop in accuracy. Ridge, Lasso, and Linear Regression models perform the worst, indicating that simpler regression techniques are less effective in this scenario. The chart highlights the advantage of ensemble learning methods in achieving higher predictive accuracy.

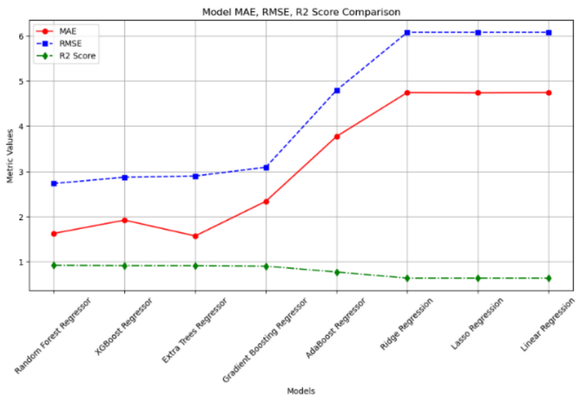


Figure 13: Model MAE, RMSE, R2 Score Comparison

In Figure 13, the graph compares different regression models based on three key performance measures: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 Score. Among all models, the Random Forest Regressor performs the best, showing the lowest error values and making more accurate predictions. As we move towards simpler models like Ridge and

Linear Regression, the errors increase, indicating that these models are less precise. The R² Score, which tells us how well a model explains the data, is highest for Random Forest and gradually decreases for other models, confirming its reliability.

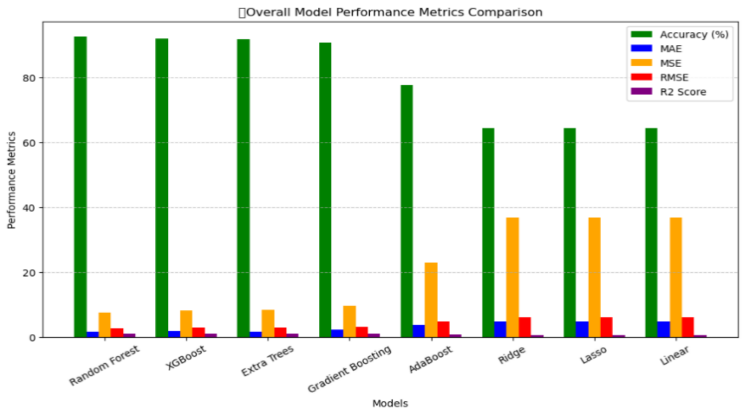


Figure 14: Overall Model Performance Metrics Comparison

In Figure 14, the graph compares different regression models based on multiple performance metrics, including accuracy, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R² Score. The best-performing models are Random Forest, XGBoost, and Extra Trees, which show the highest accuracy (green bars) and lowest error values (MAE, MSE, and RMSE), indicating their strong predictive capabilities. Moderate performance is shown by Gradient Boosting, which also performs well but has slightly higher errors compared to the top models. Weaker Models are Simpler models like Ridge, Lasso, and Linear Regression, which show lower accuracy and higher errors, meaning they struggle more in making precise predictions. The graph highlights that ensemble models (Random Forest, XGBoost, and Extra Trees) are the most effective, while traditional regression models have comparatively lower performance.

4.2. Food Wastage Prediction Application Across Different Scenarios

The Food Wastage Prediction App is designed to help restaurants and event planners minimize food waste by analyzing key factors such as food type, event type, storage conditions, preparation method, and guest count. By leveraging a Random Forest Regressor model, the app predicts potential food wastage and provides actionable suggestions for optimization. Users input details about an event, including the number of guests and food quantity, while categorical variables like seasonality, pricing, and geographical location are automatically encoded. The app then processes this information, scales the features, and predicts the expected food wastage. Based on the prediction, the system provides colour-coded recommendations, from maintaining current practices for minimal wastage to revising storage and portioning strategies for high wastage.

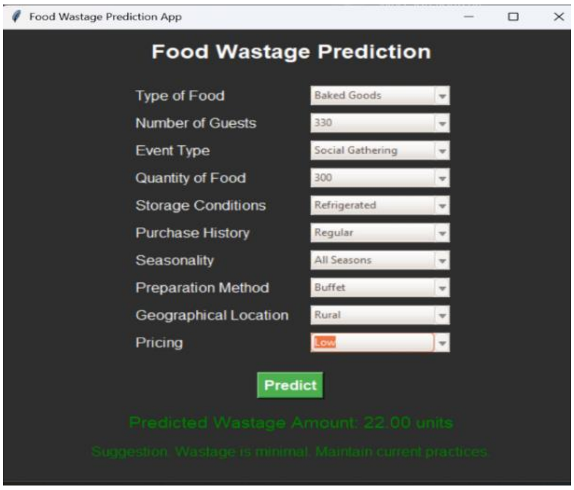


Figure 15: Case 1 - Minimal Wastage (Lesser than 30.00 units)

Figure 15 is a Scenario where a social gathering in a rural area with baked goods served as a buffet, with low pricing. Key takeaways are it’s a buffet-style event with a moderate guest count and proper storage conditions (refrigerated), and people likely take only what they need. Low pricing (22.00 units) means careful food management. Suggestions are no major changes needed, and current practices are effective, so they are indicated in green. Figure 16 is a Scenario where the same event type and food as Case 1, but with high pricing, a different season (summer), and food served as finger food. Key takeaways from case 2 are that higher pricing might lead to over-ordering or less concern for wastage. Finger foods may also contribute to waste since they are often left unfinished or discarded more frequently. Suggestions are to optimize portion sizes to reduce waste, as the food wastage amount is 30.48 units, which is indicated in orange.

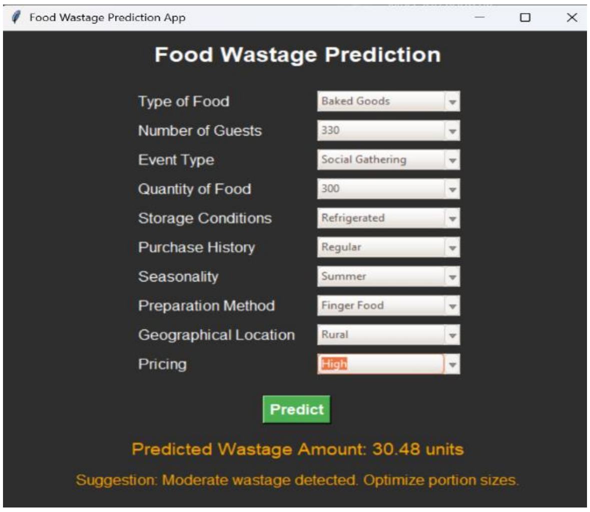


Figure 16: Case 2 - Moderate Wastage (Lesser than 40.00 units)

Figure 17 is a Scenario where a birthday event in an urban area serves fruits stored at room temperature, with a buffet-style serving, occasional purchasing, and high pricing. A key takeaway from this case is that fruits are highly perishable and storing them at room temperature increases spoilage risk. The urban setting and high pricing might lead to over-purchasing, while occasional purchasing could mean less experience in estimating food requirements. Suggestions are to improve storage and inventory practices to prevent spoilage as the food wastage amount is 40.77 units, which is indicated in red.

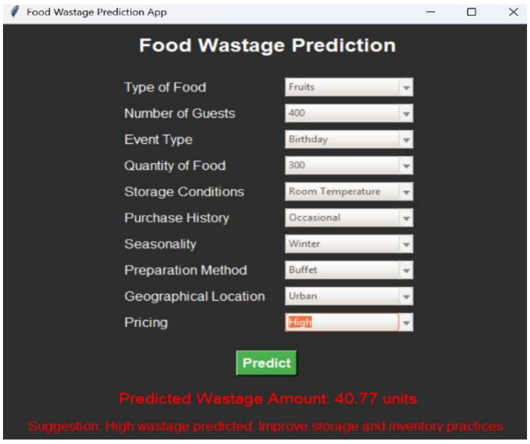


Figure 17: Case 3 - High Wastage (Greater than 40.00 units)

4.3. Comparison of Training and Testing Loss and MAE for LSTM Models with Different Architectures

We trained two LSTM models, one with more units (LSTM 1) and another with fewer units (LSTM 2), and tracked their loss (MSE) over 300 epochs. Initially, both models had high training and testing losses, but they gradually decreased as training progressed. The model with more units (LSTM 1) performed slightly better, achieving lower loss values than the smaller model (LSTM 2) across all epochs. However, after 250 epochs, the difference in performance between the two models became less significant. This suggests that while a larger model learns faster and performs better early on, a smaller model can still achieve comparable performance with sufficient training. Increasing the number of LSTM units helps reduce loss, but the improvement may not always be substantial beyond a certain point.

Table 3: MAE for LSTM Models

Epochs	Train MAE (LSTM 1)	Test MAE (LSTM 1)	Train MAE (LSTM 2)	Test MAE (LSTM 2)
50	3.753	3.8835	3.7445	3.8608
100	3.3728	3.5035	3.4996	3.6571
150	3.2558	3.3744	3.2409	3.3913
200	3.0419	3.2024	3.0898	3.2311
250	2.9044	3.0823	2.9772	3.1017
300	2.7599	2.9541	2.9654	3.1209

Table 3 provides numerical training and testing of MAE values for both LSTM models at different epochs (50 to 300). It shows that as the number of training epochs increases, the MAE decreases for both models, with LSTM 1 consistently having lower values than LSTM 2. The difference between training and test MAE indicates the models' generalization ability. While both models improve over time, LSTM 1 achieves a better final performance, reinforcing the importance of using more LSTM units.

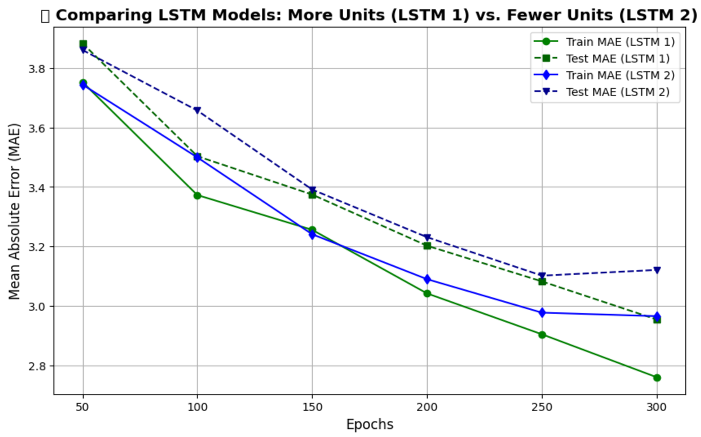


Figure 18: Comparison of LSTM Models Based on Mean Absolute Error (MAE)

Figure 18 graph compares the Mean Absolute Error (MAE) of two LSTM models, one with more units (LSTM 1) and another with fewer units (LSTM 2) over 300 epochs. Initially, both models have high MAE, but the error decreases steadily as training progresses. LSTM 1 (with more units) achieves lower MAE values than LSTM 2, indicating that increasing the number of LSTM units helps the model learn better and generalize well. The test MAE follows a similar pattern but remains slightly higher than the train MAE, showing some overfitting. Overall, the model with more units performs better in reducing error.

Table 4: Loss for Different LSTM Architectures

Epochs	Train Loss (LSTM 1)	Test Loss (LSTM 1)	Train Loss (LSTM 2)	Test Loss (LSTM 2)
50	23.7822	24.7190	23.7986	24.4710
100	19.5805	20.3940	20.7001	21.7837
150	18.8113	19.5662	18.8567	19.5264
200	16.0999	17.3316	16.5461	17.4956
250	15.4551	16.9537	16.0206	16.9915
300	14.2006	15.7602	15.4487	16.8340

Table 4 presents the numerical train and test loss (MSE) values for both LSTM architectures across different epochs. It shows a steady decrease in loss over time, with the larger model (128→64) achieving a slightly lower loss than the smaller model (64→32). However, the test loss remains higher than the train loss, indicating some overfitting. The decreasing trend suggests that both models benefit from longer training, but their performance difference becomes marginal after 250 epochs.

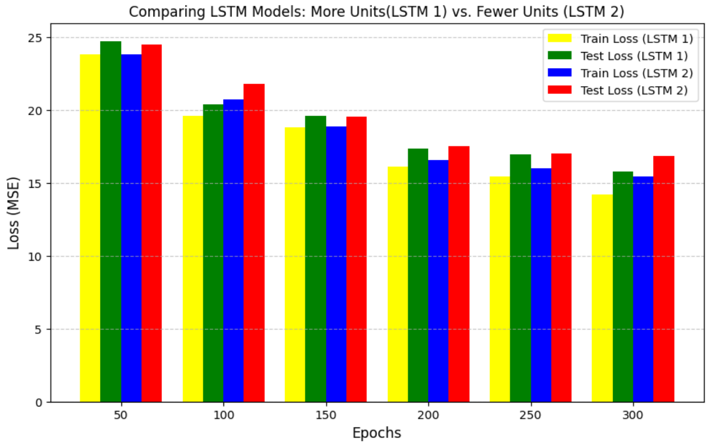


Figure 19: Comparison of LSTM Models Based on Losses

In Figure 19, the bar chart compares the training and testing loss (MSE) for two LSTM architectures, one with more units (LSTM 1) and another with fewer units (LSTM 2). Initially, both models have high loss values, which decrease as the epochs progress. The LSTM model with more units (LSTM 1) shows slightly lower loss values than the smaller model (LSTM 2), indicating better learning efficiency. However, after 250 epochs, the gap between the two models reduces, suggesting that while a larger model learns faster, a smaller model can still achieve similar performance with sufficient training.

5. Discussion and Finding

Food loss is a prominent worldwide issue, and improving restaurant inventory and menu planning with the help of predictive analysis is a positive solution. This project concerns historical food intake pattern analysis and creating a machine learning model for assessing the success and offering insight into their performance. This presentation provides primary findings, issues, and results from model performance and real-world applicability. Significance of Data Preprocessing: Data Preprocessing is crucial to ensuring our predictive model is accurate and authentic. The treatment of missing values, data normalization, and feature engineering contributed highly towards improved model performance. The data employed a range of variables affecting food wastage, such as price, patterns of demand, and seasonality. Carefully structuring and preparing the dataset ensured that the machine-learning model could bring meaningful insights.

- **Model Performance Analysis:** Multiple regression models were tried out, and their performance was noted based on key metrics such as accuracy, Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared scores. The results showed that ensemble learning models like Random Forest Regressor and XGBoost exceeded traditional linear regression models. The findings show that complex models with better feature-handling capabilities are more effective in predicting food wastage trends.
- **Best-Performing Model:** Random Forest Regressor: Among all the tested models, the Random Forest Regressor displayed the highest accuracy of 92.79%. It effectively handled non-linear relationships and provided strong predictions. The model's ability to reduce overfitting through multiple decision trees made it suitable for restaurant food wastage prediction. The proportional low MAE (1.63) and MSE (7.47) display that the model can accurately forecast food demand and allow restaurants to improve inventory levels.
- **Comparison with Other Models:** XGBoost Regressor trailed the Random Forest model by a percentage point with an accuracy of 92.03%. Extra Tree Regressor also had an accuracy of 91.89%, evidence of the reliability of ensemble models. Gradient Boosting and AdaBoost had slightly lower accuracy percentages, at 90.76% and 77.82%, respectively, making it clear that boosting techniques will not be a good choice for this dataset. Linear regression-based models, including Ridge, Lasso, and standard Linear Regression, achieved an accuracy of around 64%, which was unsuccessful. The high MSE values convey that simple regression models fail to represent the complex entities of the food consumption trends.
- **Feature Importance and Impact:** Feature analysis disclosed that the pricing, demand fluctuations, and past consumption patterns were the most important factors affecting food wastage. Regarding the factors, the pricing had a robust correlation with food wastage, so providing planned pricing adjustments could reduce the extra food loss. Seasonal demand disparity also played a major role in food loss, which shows restaurants should be considered when planning inventory.
- **Challenges Faced:** Firstly, the data quality issues in some records required careful data cleaning techniques to maintain data integrity. The computational complexity encountered while training the ensemble models, Random Forest and

XGBoost, required substantial computational resources. Then comes the feature selection. Finding the most suitable features from a large dataset was challenging, but it helped improve model accuracy and efficiency. Generalization about the restaurant types and locations required broad cross-validation.

5.1. Findings

The study builds up the value of data-driven decision-making in the food industry. By using machine learning models, restaurants can predict food wastage and provide improved purchasing decisions. Such a predictive system can save money and reduce food wastage. The findings suggest that adding the predictive analysis into the restaurant's inventory management can drastically reduce over-purchasing and food corruption. For example, by using the Random Forest model, restaurants can demand variations, make adjustments according to their acquisition, and ensure prime stock levels. Menu optimization is another primary and key takeaway from the study. This is too important to understand which items contribute most to food wastage and allow the restaurant managers to alter the menu offerings. An item that continuously generates high wastage can be re-evaluated, and recipe modifications and pricing adjustments can be applied to tune the menu to good. Reducing food wastage will significantly help the environment, including lower carbon footprints and reduced demands on natural resources. Restaurants and other food services adopting this predictive analysis will contribute to sustainability by minimizing the discarding of consumable food. The economic benefits of minimizing food waste are high. By saving on holding extra amounts and maximizing supply chain processes, restaurants can save a lot. This enhances profit margins alongside customer satisfaction by providing fresh and quality food.

6. Conclusion

This project uses machine learning and deep learning techniques to predict restaurant food wastage. The main goal is identifying the most effective model to help restaurants manage their inventory more efficiently and reduce food waste. From the trained models, Random Forest Regressor was the best, with 92.79% accuracy, an MAE of 1.63, and an RMSE of 2.73. XGBoost Regressor was next at 92.03% accuracy, even though its MAE was a bit higher at 1.92. Extra Trees Regressor 91.89% and Gradient Boosting Regressor 90.76% also gave good results. The LSTM model proved useful in recognizing the long-term patterns in food wastage, making it a strong choice for time-series forecasting. The models such as AdaBoost, Ridge Regression, Lasso Regression, and Linear Regression did not perform well, with accuracy levels falling below 80%, making them less reliable for this task. The poorest accuracy was that of the Linear Regression model, at 64.40%, with a high MAE of 4.74, making it unable to predict food wastage accurately. For more convenience, a Tkinter GUI interface was designed and implemented, through which users can provide inputs to obtain real-time predictions regarding food wastage. This interface makes the system easy to use for technical and non-technical users. The results show that ensemble learning models like Random Forest and XGBoost are appropriate for this task, and LSTM provides valuable insights while working with sequential data. The system can be further improved using real-time data updating, adjusting model parameters, and deeper or more advanced deep learning techniques. Through such enhancements, the model can be transformed into an efficient tool for restaurants and dining establishments to save wastage, maximize stocks, and generally optimize efficiency.

6.1. Future Scope

Future models integrated with IOT-enabled and smart sensors to improve prediction accuracy. These sensors can be integrated into refrigerators and storage units. Exploring neural networks, such as RNNs or CNNs, to dive deep into time series forecasting of food wastage. Expanding the dataset by collecting data from multiple restaurants that are under different circumstances to improve the model generalization. Perform a consumer behaviour analysis by including customer preferences and feedback to further understand the demand and predictions. Developing an AI-based automation model that directly adjusts the performance of Inventory Management and scheduling based on the model recommendation. This project highlights the potential of machine learning in tagging food wastage problems in the restaurant and food service sectors. Using sophisticated predictive models, restaurants can optimize inventory, significantly reduce food wastage, and enhance bottom-line performance. The results demonstrate ensemble learning methods, particularly Random Forest. This method provides the most valid predictions, which makes them the preferred choice for real-world implementation. Future Advancements in AI, IoT Integrations, and deep learning grab great potential for further research and expanding this approach. This covers the way for a more sustainable and efficient food industry.

Acknowledgment: I am deeply grateful to my co-authors for their expertise and dedication, which greatly enriches this work.

Data Availability Statement: The data for this study can be made available upon request to the corresponding author.

Funding Statement: This manuscript and research paper were prepared without any financial support or funding.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

Ethics and Consent Statement: This research adheres to ethical guidelines, obtaining informed consent from all participants.

References

1. M. Rodrigues, V. Miguéis, S. Freitas, and T. Machado, "Machine learning models for short-term demand forecasting in food catering services: A solution to reduce food waste," Journal of Cleaner Production, vol. 435, no.1, p.1-16, 2024.
2. A. G. Arvindaraj, B. Manikandan, R. Rokesh, and M. Abinaya, "AI-Automated System for Ingredient Planner using Machine Learning Algorithms," in 2023 IEEE International Conference on Smart Systems and Advanced Computing (I-SMAC), Kirtipur, Nepal, pp. 511-515, 2023.
3. N. M. Ifham, M. I. M. Abdulla, M. R. M. Rifadh, M. N. M. Hussain Ajward, N. H. P. R. S. Swarnakantha, and K. Rajendran, "Machine Learning Based Approach to Reduce Food Wastage," in 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, pp. 119-124, 2023.
4. S. Thay and T. Chinda, "Factors Influencing Food Waste Management in Phnom Penh, Cambodia: Data Collection," in 3rd Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), Bangkok, Thailand, pp. 1-4, 2018.
5. M. C. Anggraeni, C. A. Silaban, M. S. Anggreainy, and E. Cahyadi, "Role of Artificial Intelligence in the Management of Food Waste," in 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS), IPOH, Malaysia, pp. 1-6, 2021.
6. M. Nilashi, A. M. Baabdullah, R. A. Abumalloh, K.-B. Ooi, G. W.-H. Tan, M. Giannakis, and Y. K. Dwivedi, "How can big data and predictive analytics impact the performance and competitive advantage of the food waste and recycling industry?" Annals of Operations Research, pp. 1–42, 2023, Press.
7. R. Dagiliūtė and A. Musteikytė, "Food Waste Generation: Restaurant Data and Consumer Attitudes," Environmental Research, Engineering and Management, vol. 75, no. 2, pp.7-14, 2019.
8. L. Pereira, V. Aguiar, and F. Vasconcelos, "FIKWaste: A Waste Generation Dataset from Three Restaurant Kitchens in Portugal," Data, vol. 6, no. 3, p. 25, 2021.
9. G. F. Turker, "Reducing Food Waste in Campus Dining: A Data-Driven Approach to Demand Prediction and Sustainability," Sustainability, vol. 17, no. 2, p.379, 2025.
10. D. Nijloveanu, V. Tița, N. Bold, D. A. Popescu, D. Smedescu, C. Smedescu, and G. Fîntîneru, "The Development of a Prediction Model Related to Food Loss and Waste in Consumer Segments of Agrifood Chain Using Machine Learning Methods," Agriculture, vol. 14, no.7, p.1837, 2024.
11. H. Fatorachian and K. Pawar, "Sustainable Cold Chain Management: An Evaluation of Predictive Waste Management Models," Applied Sciences, vol. 15, no. 2, p.770, 2025.
12. C. V. Jansson-Boyd and C.-l. Mul, "Drivers That Affect Households to Reduce Food Waste: A UK Qualitative Study," sustainability, vol. 16, no. 5, p. 2188, 2024.
13. W. Leal Filho, P. C. C. Ribeiro, A. F. F. Setti, F. M. S. Azam, I. R. Abubakar, J. Castillo-Apraiz, U. Tamayo, P. G. Özuyar, K. Frizzo, and B. Borsari, "Toward Food Waste Reduction at Universities," Environment, Development and Sustainability, vol. 26, no.7, pp. 1-22, 2023.
14. H. Onyeaka, P. Tamasiga, U. M. Nwauzoma, T. Miri, U. C. Juliet, O. Nwaiwu, and A. A. Akinsemolu, "Using Artificial Intelligence to Tackle Food Waste and Enhance the Circular Economy: Maximising Resource Efficiency and Minimising Environmental Impact: A Review," sustainability, vol. 15, no.13, pp. 1-20, 2023.
15. K. S. Chowdary, L. K. Praneetha, S. Holika, D. B. Priya, S. Venkatrama Phani Kumar, K. Venkata Krishna Kishore, "Prediction of Food Wastage using XG Boost," in 2024 8th International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, pp. 307-312, 2024.