FMDB *Transactions on Sustainable Environmental Sciences*

Advanced Predictive Analytics for Global Food Waste, Environmental Conservation and Sustainable Resource Management

Varsha Ramkumar[1], S. Rubin Bose[2,*], J. Angelin Jeba[3], R. Regin[4], S. Suman Rajest[5]

[1]Department of Business Analytics, University of Galway, Galway, Ireland.
[2,4]School of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[3]Department of Electronics and Communication Engineering, S.A. Engineering College, Chennai, Tamil Nadu, India.
[5]Department of Research and Development, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
v.ramkumar1@universityofgalway.ie[1], rubinbos@srmist.edu.in[2], angelinjeba@saec.ac.in[3], reginr@srmist.edu.in[4], sumanrajest414@gmail.com[5]

Abstract: Food waste has become a serious problem and a loss of resources around the world, and hurts the balance of the environment, with an economic loss. This paper suggests a machine learning-based predictive model, which is hybrid with a quantum-inspired optimization approach, to estimate food waste and the associated economic loss more accurately. The system uses Linear Regression to assess the economic impact, while Random Forest is used to predict waste quantities, both enhanced by quantum optimisation for feature selection and hyperparameter tuning. The proposed architecture not only predicts food waste but also displays alert levels (Low, Moderate, High, and Critical) and provides intelligent suggestions to minimize food waste. Unlike classical models, this hybrid approach results in faster convergence, better generalisation, and more interpretable predictions. The key challenges encountered while developing the model were data imbalance, feature dependency, and fine-tuning the quantum layer with the classical ML pipeline to ensure stable model training and efficient computation. The quantum-enhanced learning approach achieved high predictive accuracy and demonstrated that these techniques can significantly improve decision-making in sustainable food management systems. Overall, the paper emphasizes QML's potential to solve real-world problems, such as food waste, and provides a scalable solution that supports global sustainability goals through data-driven insights, economic forecasting, and actionable alerts for effective food resource management.

Keywords: Resilient Food Chain; Economic Loss; Total Food Waste Prediction; Linear Regression; Random Forest; Ensemble Learning; Non-Linear Relationships; Prediction Accuracy; Quantum Machine Learning.

Received on: 10/05/2025, **Revised on:** 13/07/2025, **Accepted on:** 24/09/2025, **Published on:** 05/06/2026

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

DOI: https://doi.org/10.69888/FTSESS.2026.000718

Cite as: V. Ramkumar, S. R. Bose, J. A. Jeba, R. Regin, and S. S. Rajest, "Advanced Predictive Analytics for Global Food Waste, Environmental Conservation and Sustainable Resource Management," *FMDB Transactions on Sustainable Environmental Sciences*, vol. 3, no. 2, pp. 103–122, 2026.

Copyright © 2026 V. Ramkumar *et al.*, licensed to Fernando Martins De Bulhão (FMDB) Publishing Company. This is an open access article distributed under CC BY-NC-SA 4.0, which permits copying, remixing, redistributing, transformation, and building upon the material in any medium so long as the original work is properly cited.

1. Introduction

*Corresponding author.

Food waste is one of the most critical issues today, with economic, social, and environmental implications [1]. According to the Food and Agriculture Organisation (FAO), about one-third of the world's total food production, or almost 1.3 billion tons, is lost or wasted annually. This massive waste represents billions of dollars in economic losses and the unnecessary use of natural resources, including land, water, and energy. Furthermore, food waste accounts for approximately 8-10% of global GHG emissions, exacerbating climate change. Although there have been many awareness campaigns and worldwide initiatives to address food waste, the problem persists because most food waste reduction systems lack a reliable, data-driven method to predict and effectively address it before it occurs [2]. To address this need, the objective of this paper is to create a quantum-inspired machine learning model that can accurately forecast food waste (and economic losses) and generate alert levels and actionable recommendations that reduce food waste. This is a hybrid of traditional machine learning algorithms, linear regressions, and random forests, and uses quantum-inspired optimisation methods to boost performance. On the one hand, the economic impact is estimated based on the quantity of waste produced, using linear regression [3]. On the other hand, the quantity of expected food waste is predicted using a random forest model based on various food features, such as food type, storage duration, storage temperature, humidity, and consumption rate [4]. A layer of quantum optimization was thus added to the system for hyperparameter tuning and feature selection, further improving its predictive ability.

This integration makes the model operate more efficiently in high-dimensional, high-complexity data spaces, enabling greater accuracy and faster convergence than traditional optimisation approaches [5]. The proposed model is designed to be a complete prediction system that can analyse historical data and provide future waste forecasts. It's designed to measure waste, but also recognize it, and assign it to one of the alert levels: low, moderate, high, and critical, depending on the forecasted value. These alert levels provide users (households, restaurants, supermarkets) with information on the strength of waste production, enabling them to take preventive measures at the right time. In the case of a restaurant with a "high" alert for perishables, this can trigger a proactive adjustment to the purchasing plan. In the event of a supermarket that has a "critical" alert, it can trigger the implementation of a donation or redistribution mechanism [6]. It also has a suggestion engine that offers tailored tips for waste reduction, such as implementing First-In-First-Out (FIFO) inventory management, setting stock-level targets, tracking product expiration dates, and adopting IoT-based storage solutions [7]. This architecture combines classical and quantum computing paradigms in a hybrid configuration, a feature that distinguishes it from conventional architectures. While such machine learning models are powerful for pattern recognition and regression problems, they can be challenging in high-dimensional data, with correlated features, or in optimization problems where they may get trapped in local minima. Quantum machine learning overcomes these challenges by leveraging the principles of superposition and entanglement, enabling quantum bits (qubits) to store multiple states [8].

This parallelism enables quantum-enhanced optimizers to search a larger solution space more efficiently, resulting in higher accuracy and stability. In essence, it is the quantum parts that serve as intelligent optimisers, giving classical models an edge in prediction quality and speed. The value of this mixed architecture is that it offers an easily scalable and easily interpretable solution, that is, a combination of machine learning and computational power from quantum tech [9]. It aims to provide faster training times, better generalization, and low overfitting, while maintaining sufficient interpretability for practical use [10]. The system's interpretability enables users and decision-makers to understand the rationale behind predictions, which is essential for sustainability-oriented decision-making. Additionally, its modular design makes the architecture flexible and adaptable to a range of data sources, including IoT sensors, retail data management systems, and smart waste monitoring devices. The system provides real-time analytical and predictive insights that help to build a circular economy framework with responsible production and consumption [11]. There are many difficulties in developing such hybrid predictive models, however. The most prominent one is data imbalance. Most public food waste datasets are biased, with few examples of extreme food waste cases, making training and evaluation difficult. Another challenge of feature correlation is that some features are dependent on each other (e.g., temperature and humidity) and can affect model learning. Once the quantum layer is integrated into the classical pipeline, it is important to tune the parameters and ensure that the quantum layers (e.g., Qiskit or PennyLane) and machine learning libraries (e.g., scikit-learn) work together [12]. Some other challenges include hardware compatibility, cost, and the lack of quantum hardware.

When simulating quantum phenomena on a classical machine, it is challenging to find a balance between model complexity and dataset size that allows the simulation to run within a reasonable time and computational budget. However, these difficulties were overcome through iterative training, feature engineering, and quantum-assisted optimization, enabling the system to converge reliably and achieve enhanced predictive capabilities [13]. The motivation for this research was the recognition that, for effective food waste management, it is important to adopt a predictive rather than a reactive approach. Current systems use descriptive analytics (historical data to gain insight into waste patterns) but lack predictive and prescriptive features. To address this research gap, the proposed system will help to predict future waste and make recommendations for corrective actions. Beyond that, the fact that this socially relevant application is obtained through quantum optimisation demonstrates the applicability of quantum technology to solving problems beyond the realm of abstract computational problems. It proves to be a practical instrument for achieving the Sustainable Development Goals [14]. The model aligns well with the UN Sustainable Development Goals, specifically Goals 12 (Responsible Consumption and Production) and 13 (Climate Action), in promoting

waste reduction, resource conservation, and environmental protection. The system can be used in various real-world applications. The model can be used to predict food surpluses for a given day and to optimise purchasing plans in the restaurant and hospitality industries. It can help with inventory management in supermarkets and retail stores, preventing overstocking and product spoilage.

The system could also help households by tracking shopping habits and reminding or alerting them to the need to conserve food. Also, the model could have predictive capabilities for food distribution to NGOs or food banks, thereby optimizing food delivery and logistics before it spoils. Policymakers and researchers can also use the system's aggregated data to understand the waste situation in a particular area, which can inform policies and public campaigns. The contribution of this work is in two aspects. It proposes a scalable hybrid learning framework that incorporates these ideas into classical machine learning models to improve their predictive power [15]. At the societal level, it provides a pragmatic solution to a global issue: food waste. The model leverages advanced analytics and recommendations to minimise losses and advance environmental and humanitarian goals. This system has potent predictive and prescriptive capabilities, and the data, being a powerful tool for sustainable development, could transform the waste management industries. In this paper, researchers demonstrate the potential for combining AI and QC to make significant inroads on real-world sustainability problems. It extends theoretical aspects of quantum computing by adding practical application cases with measurable benefits. The combination of data science, environmental consciousness, and the latest computing technologies demonstrates the potential of quantum-improved machine learning models to create sustainable, efficient systems that can significantly impact the management of global resources. The project's success lies not only in its impressive predictive power but also in its role as a catalyst for further research in quantum computing, machine learning, and environmental sustainability, thereby paving the way for smarter technologies that promote responsible living and a greener planet.

2. Literature Review

Rodrigues et al. [1] addressed food waste in catering businesses by applying machine learning-based demand forecasting. They propose four predictive models (including random forests and LSTM neural networks) to predict short-term meal demand and compare them with two baseline models commonly used for meal demand estimation in catering services. Three case studies of true day-level data from 3–9 years in three cantinas (2 student and 1 company) are used in the experimental design. Their results show that the Random Forest and LSTM models are superior to the baseline models, potentially reducing the number of wasted meals by 14-52% and the number of meals that are not met by 3-16%. The study highlights the importance of more accurate demand forecasting to improve food service operations, with less waste and higher service levels. To predict and minimize food waste across sectors such as retail, food services, and domestic, Yang et al. [2] examined the use of ML, particularly Random Forest. They compiled predictors for food waste classification and regression, including retail demand, food service forecasts, household consumption patterns (per-capita and aggregate), and regional indicators. High precision, high accuracy (≈ 0.97), and recall (≈ 0.93) are good performance metrics for the Random Forest model, indicating that it can reliably distinguish between high- and low-waste instances. The research aims to reduce the environmental and economic impacts of food waste and to improve the effectiveness of a data-driven approach for estimating waste amounts. During the COVID-19 pandemic, Malefors et al. [3] explored the potential of machine learning models to reduce food waste and costs in Swedish public catering. The research indicated excess production in schools and preschools, where meals are often prepared for all students regardless of changing attendance. Using data on actual guest counts from 2010 to 2020, the researchers constructed forecasting models based on daily attendance to better align production with demand.

The research was carried out using three models: Random Forests, Artificial Neural Networks, and Poisson autoregression with a baseline "cook for all" approach. The results indicate that Random Forest consistently outperformed all other methods, achieving the highest accuracy and a mean absolute prediction error of less than 15%. Implementing these predictions would save food by considerable l amounts while generating healthy economic gains. This research demonstrates that ML-based forecasting for operational performance is an effective data-driven approach to boosting sustainability and resilience during crises. Suresh et al. [4] discuss the implementation of machine learning techniques to manage and forecast food waste, supporting improved planning and waste reduction. This research highlights that ~1.3 billion tons of food are wasted each year, and that per capita food waste is ~55 kg/year. It emphasizes identifying the causes of food waste, understanding existing waste management practices, and predicting future waste trends using a predictive model. A random forest prediction algorithm was used based on past and current data on food stock, sales patterns, supply, and storage. The model aims to predict waste so managers can modify purchasing, storage, and preparation practices. By examining patterns such as overstocking, spoilage, and demand variability, the model provides projections that can reduce waste by aligning supply chain decisions with future demand. Olawale et al. [5] explored how machine learning can contribute towards minimizing food loss in sustainable agricultural practices. The major factors causing loss in stored commodities are temperature, moisture, storage practices, pest/disease prevalence, delays in handling, and transport. Using historical storage conditions data and spoilage loss rates, they train machine learning models to predict spoilage risk and detect critical control points in perishable goods supply chains.

The research primarily focuses on feature importance analysis, identifying that storage temperature variations, storage duration, and transportation time have disproportionate impacts on waste prediction. Their suggested system enables farmers, storage managers, and supply chain participants to take proactive steps (e.g., modify storage conditions, accelerate transport, implement safeguarding measures) to reduce losses. The researchers suggest that such models could significantly reduce food waste and economic losses in farm-to-store supply chains, particularly in areas with weak cold-chain infrastructure. Thippa and Derek et al. [6] suggested a new combination of reinforcement learning with nanomaterials design at the quantum level to advance circular economy and waste management objectives. Their system referred to as RLQNWMCE, can be used for training RL agents to suggest and optimize quantum nanomaterial (QN) structures and compositions that maximize functionalities such as pollutant capture, sensing, and catalytic transformation The RL approach repetitively estimates prospective QNs through performance feedback (e.g., pollutant binding efficacy, stability, environmental footprint) optimizing design to material types that produces a healthy balance between effectiveness, cost and environmental responsibility. The article also mentions the possible combination of sensor networks and material feedback loops for tracking material degradation or ageing, facilitating adaptive replacement or regeneration. While predominantly theoretical, the authors highlight that RL-aided material discovery can minimise trial-and-error in laboratory work, accelerate the deployment of intelligent materials within waste streams, and enable more environmentally friendly material lifecycles. Senthil et al. [7] propose a quantum-boosted artificial intelligence framework for accurate, personalised nutritional prediction to aid food security and sustainability in healthcare bioengineering.

It integrates image analysis via a Convolutional Neural Network (CNN) with an accuracy of ~91.87% in classifying food types from images. It uses users' individual data (age, weight, and height, BMI) to customize calculations of nutritional needs. To improve predictive accuracy, the authors combine quantum machine learning algorithms, Quantum Support Vector Machines (QSVM), Quantum Neural Networks (QNN), and Quantum Reinforcement Learning (QRL). The quantum algorithms achieve high accuracy on nutritional prediction tasks across various foods (QSVM ~90%, QNN ~92%, and QRL ~ 93%). The method is suggested to improve scalability, model expressiveness, and flexibility in high-dimensional, complex environments. The article states that, with the development of quantum computing, these types of Quantum AI systems have a good chance of helping with personalized nutrition advice, preventing malnutrition, and enhancing food security. Wan et al. [8] addressed the challenge of managing reactive composting of food waste by developing a machine learning model to predict compost maturity indicators. They are based on two maturity metrics: the seed Germination Index (GI), indicative of phytotoxicity and the suitability of the compost for plant use, and the carbon-to-nitrogen (C/N) ratio, a common chemical maturity metric. Researchers compare the machine learning algorithms Random Forest (RF), XGBoost, LightGBM, and Multilayer Perceptron (MLP). The combination of these algorithms' predictions yielded a very high-accuracy model with $R^2 = \approx 0.977$ for GI and $\approx$ 0.986 for C/N. They also performed feature importance analysis to find the most influential inputs and interactions for maturity and validated the system in a composting reactor. Hatab et al. [9] investigated household food waste behavior in Addis Ababa, Ethiopia, by incorporating cultural, social, and economic variables into the Integrative Model of Behavioral Prediction (IMBP). The study tests the impact of attitudes, perceived norms, and self-efficacy on contextual factors such as income, education, and cultural food habits using household survey data.

The extended IMBP model indicates that household attitudes, social norms, and perceived control significantly influence food waste behavior. Moreover, socioeconomic status (household size and income level) strongly moderates behavioral outcomes. The study highlights the need for context-based interventions coupled with behavioral changes, as well as addressing individual food choices, for food waste reduction to succeed in urban Africa. An in-depth review of quantum machine learning (QML) has been presented by Zeguendry et al. [10], which has the potential to address scalability and complexity issues in classical machine learning for high-dimensional data processing. It discusses some fundamental quantum computing techniques, such as Quantum Neural Networks (QNN), Quantum Support Vector Machines (QSVM), and Quantum Reinforcement Learning (QRL), and how the principles of quantum computing, like superposition and entanglement, can help to optimise faster and provide a more expressive representation of features. Case studies include examples of QML in use for pattern recognition, medical data analysis, and performance optimisation, where it has outperformed traditional approaches in some scenarios. Salemdeeb et al. [11] conducted a comparative life-cycle assessment of different food waste management practices, with a specific focus on the environmental and health impacts of food waste as animal feed compared to anaerobic digestion, composting, and landfilling. They concluded that the reuse of food waste as animal feed was a major way to reduce GHG emissions and resource use, and to lower risks to public health from improper disposal. The study highlights the potential of applying circular economy principles to food systems to maximize resource recovery and minimize environmental impacts. It also refers to regulatory and safety issues, such as the need for strict monitoring and quality control. The research provides a solid foundation for implementing food waste valorization in sustainable waste management systems on a global scale.

In a life-cycle assessment (LCA) approach, Lundie and Peters [12] analysed trade-offs across various food waste management scenarios, including composting, anaerobic digestion, incineration, and landfilling. They showed that the environmental impacts of waste avoidance, recycling, and energy recovery are the lowest, especially in terms of greenhouse gas emissions and eutrophication potential. The study emphasized the necessity of evidence-based decision-making in waste management, requiring a holistic assessment of environmental impacts rather than relying on a single indicator. It further underscored the

significance of infrastructure, energy efficiency, and policy frameworks to maximize resource utilization. This work made important contributions to the design of sustainable waste systems and significantly influenced policy recommendations to minimize food waste footprints. Filimonau and Delysia [13] conducted a critical review of food waste management in the hospitality sector, noting the major challenges and behavioral issues contributing to food waste. The study compiled results from hotels, restaurants, and catering establishments to identify the operational, managerial, and consumer causes of waste generation. It highlighted the importance of staff training compliance, adequate inventory management, and cultural attitudes of abundance as significant factors in minimizing waste. They also highlighted the importance of developing policies and digital innovations for a sectoral approach to waste reduction. This review provided a detailed framework for sustainable hospitality operations and highlighted the economic, environmental, and ethical reasons to tackle food waste. Thi et al. [14] provided an overview of food waste management in developing countries, highlighting the lack of infrastructure, technology, and policy. Their work found that most food waste ends up in landfills or open dumping, resulting in environmental degradation, methane emissions, and the loss of valuable nutrients.

The study highlighted the potential of anaerobic digestion, composting, and waste-to-energy technologies as sustainable alternatives, but also noted barriers, including financial constraints, inadequate legislation, and limited awareness. They suggested a combination of policy changes, technology adaptation, and public involvement to enhance management efficiency. Food waste is a complex socio-technical problem that demands coordinated government and community action to achieve long-term sustainability, which this paper is important for advancing in developing countries. Food waste has been described as a "wicked problem" by Närvänen et al. [15], who highlight its multidimensionality and the social, economic, and environmental complexities it entails. It examined the dynamics among stakeholders (consumers, policymakers, and businesses) and their interactions in the food system, which can generate waste as feedback loops. The authors propose that holistic solutions can be developed by promoting cooperation across sectors and by adopting systems thinking that includes digital tracking, circular-economy models, and policy integration. They also discussed the cultural and ethical aspects of waste, emphasizing that, more important than technology, are changing behavior and educating people. The book combines theoretical and practical perspectives, presents concrete solutions to the global challenge of mitigating food waste, and puts sustainability at the center of the societal agenda rather than treating it as a technical issue. As revealed by the reviewed literature, the use of Machine Learning (ML) and emerging Quantum Machine Learning (QML) approaches for food waste prediction, management, and minimization to promote economic and environmental benefits is an evolutionary process. The most used algorithms are Random Forest (RF), Linear Regression, Artificial Neural Networks (ANNs), LSTM, XGBoost, Support Vector Machines (SVM), Quantum Support Vector Machines (QSVM), Quantum Neural Networks (QNN), and Quantum Random Forest (QRF).

Rodrigues et al. [1] and Yang et al. [2] developed predictive food waste models using Random Forest and LSTM models to forecast meal demand and minimize overproduction in the food service industry, particularly in canteens and hospitality settings. Among them, the best results were achieved by Rodrigues et al. [1], who reduced waste by up to 52% and unmet demand by 16%, and by Malefors et al., who achieved errors below 15% in real-world prediction, demonstrating the reliability of RF for real-life applications. Then, these techniques were applied across multiple industries, and high accuracy (~97%) and recall (~93%) were achieved across different datasets, demonstrating the remarkable classification ability of Random Forest. The same researchers, Suresh et al. [4], also used RF to predict future wastage trends from supply, stock, and sales data, highlighting a possible improvement to supply chain management policy through data. In agricultural contexts, Olawale et al. [5] used a comparable strategy to identify key spoilage factors on the farm using RF and feature importance analysis, with a special focus on the versatility of ML in addressing farm-level waste management. Wan et al. [8] demonstrated better prediction accuracy ($R^2 \approx 0.977–0.986$) than other algorithms and hybrid or fusion models. The application of ensemble learning (RF, XGBoost, LightGBM, and MLP) for compost maturity prediction yielded interesting results. These models achieved strong results, especially when combining multiple classifiers. To solve the scalability issues, Thippa and Derek et al. [6], Senthil et al. [7], and Zeguendry et al. [10] proposed Quantum ML (QML) techniques, such as Quantum Neural Networks (QNNs), Quantum Support Vector Machines (QSVMs), and Quantum Reinforcement Learning (QRL).

They use quantum superposition and entanglement to process high-dimensional data and achieve over 90% accuracy in nutritional and pattern recognition applications. Thippa and Derek's RLQNWMCE framework combines reinforcement learning and quantum nanomaterial design, demonstrating how AI systems can support sustainable business practices in the future. Finally, Senthil et al. [7] showed that QML is powerful for nutrition prediction and food security applications, paving the way for quantum models to outperform classical models in more complex, non-linear environments. Socio-behavioral research mainly focused on cultural, economic, and social factors. Hatab et al. [9] used behavioral prediction models (IMBP) to describe food waste at the household level in developing settings. It also fills the gap between technical models and the human factor in food waste, which algorithms have not yet studied. Thi et al. [14], Lundie and Peters [12], and Salemdeeb et al. [11] conducted classical environmental management studies and applied comparative analysis and techniques to assess the sustainability outcomes of various waste-handling practices. They may not have predictive ML models, but their conclusions (e.g., the environmental benefits of converting food waste into feed products for domestic animals and the promotion of circular economy models) provide the empirical basis for developing modern ML-based prediction systems. Filimonau and Delysia

[13] and Närvänen et al. [15] stress the importance of policy, management, and system-level integration, while also underscoring the critical role of digital monitoring, stakeholder coordination, and behavioral change as complementary to the AI-based forecasting.

The most common algorithm across the studies reviewed is Random Forest, as it is robust, interpretable, and achieves high accuracy across different applications and datasets. It is useful for predicting food waste, agricultural losses, and demand in both classification and regression problems. In the meantime, LSTM and hybrid ML-QML architectures have proven very promising for dynamic time-series data, enabling more accurate forecasting and optimization under uncertain, variable conditions. The accuracy and stability of ensemble learning methods can be improved by combining multiple algorithms to leverage their complementary strengths, as shown by Wan et al. [8]. In addition, there is a clear interdisciplinary move toward environmentally aware technological innovation with AI integration as observed in Salemdeeb et al. [11] and Närvänen et al. [15]. Lastly, Quantum Machine Learning (QML) is gaining momentum as a future frontier, promising significant computational benefits for handling large, complex, high-dimensional data sets in the context of a global food system, thereby contributing to the development of more efficient, scalable, and sustainable waste management. Comprehensive research has shown that machine learning models such as random forests and long short-term memory (LSTM) can provide immediate, practical, and highly accurate solutions for food waste prediction. Still, quantum machine learning models have the potential to deliver scalable, uncertainty- and high-dimensional optimization solutions in the future.

Nevertheless, there are some technological, infrastructural, and socio-economic constraints to their widespread use. The future lies in hybrid AI systems that combine data-driven prediction, quantum optimization, and behavioral strategies to enable sustainable, smart global food waste management. Based on analyses of previous studies, several gaps and challenges limit the application of smart food waste prediction systems at larger scales. One major problem is data heterogeneity and scarcity; many studies use limited, local data that do not capture the full range of food system variability across regions and seasons, thereby reducing the generalizability and robustness of models. A further constraint is the lack of cross-sector integration – most studies to date have focused on a single sector, such as agriculture, retail, or households, without creating an integrated framework that brings data together from production to consumption. Moreover, many prediction models remain batch-based, limiting their ability to adapt to changing demand, storage conditions, and waste patterns. Another implementation restriction is quantum machine learning (QML) due to the lack of quantum hardware, high computational costs, and scalability issues. Lastly, although there have been important advancements in technology, the integration of behavioral and policy aspects has been underdeveloped – especially, as noted by Hatab et al. [9] and Närvänen et al. [15], the social, cultural, and infrastructural factors that contribute to food waste behavior have not received sufficient attention in models.

2.1. Problem Statement

Although the fields of machine learning and new quantum technologies are making tremendous progress, food waste is a multifaceted and complicated issue. Previous studies suggest that current solutions have several limitations. First, data is heterogeneous and scarce; most data is local or partial, capturing only general variations in food types, storage conditions, consumer behavior, and seasonal demand. Second, limited integration across sectors – from agriculture to retail, and retail to households – makes it difficult to develop integrated forecasting systems that can help coordinate waste along the supply chain. Third, most forecasting models are batch-only and cannot adjust to changing patterns of demand, storage, and consumption in real time. Fourth, although quantum-based approaches are promising, they also face hardware challenges, high computational costs, and scaling issues, making their deployment impractical. Lastly, although very important for food waste outcomes, behavioural and policy factors (including cultural habits, policies and regulations, and infrastructure constraints) are not included. Considering these, the main issue is that existing food waste forecasting systems are either technologically limited or context-sensitive and/or lack social information, making them less suitable for practical use. There is a need for an integrated, scalable, and adaptive forecasting system that leverages classical machine learning, quantum optimisation, and social-behavioral factors to accurately forecast food waste, estimate economic losses, and offer actionable recommendations to reduce waste across the entire food supply chain.

2.2. Proposed System: An Overview

Designed to predict food waste and estimate economic losses, this system combines classical machine learning with quantum-enhanced optimisation to provide an accurate, scalable, and adaptive solution. The system is based on data collected from many sources, including agricultural production, retail, restaurants, and households, as well as environmental and storage conditions, among others, to ensure comprehensive coverage of the entire food supply chain. Data preprocessing and feature engineering clean and transform the inputs and select the most influential predictors to increase the model's robustness. The core predictive models are random forests (for predicting waste), linear regression (for predicting economic losses), and long short-term memory (LSTM) networks (for sequence or time-series data). For hyperparameter tuning and feature optimization, researchers employ quantum machine learning techniques, such as quantum neural networks (QNNs) and quantum support vector machines

(QSVMs), to improve performance and efficiently process high-dimensional datasets. The system predicts waste levels, classifies them into different alert levels, and provides actionable recommendations for adjusting procurement, production, and storage. In addition, the system integrates behavioral and policy considerations to ensure the solution is tailored to the specific context. The proposed framework, addressing issues such as data heterogeneity, sector integration, immediate adaptability, and scalability, offers a holistic, practical approach to sustainable food waste management that integrates classical machine learning, quantum optimization, and social-behavioral insights.

3. Methodology

The approach to food waste prediction in this paper is intended to systematically cover the entire lifecycle of machine learning models from data acquisition to deployment, while also exploring quantum machine learning (QML) as a novel dimension in predictive modeling. The pipeline includes both classic machine learning algorithms (e.g., linear regression, random forests) and experimental QML techniques to provide benchmarks for baseline and advanced performance. This hybrid approach ensures that the paper is technically sound and research-oriented, adding to practical applications of machine learning and emerging technologies. The paper begins by collecting datasets on food waste generation. Food waste is driven by a range of factors, including production volume, consumer demand, regional demographics, economic indicators, supply chain inefficiencies, and consumer behavior. The main source of data is public datasets from platforms such as Kaggle. Data sets like these typically include attributes such as food type, origin, household or institutional consumption patterns, and recorded waste levels. Survey-based data and institutional data sources can also be added as inputs to enrich the datasets and make them more representative of real-world scenarios. The system combines structured datasets with additional contextual information, providing a comprehensive approach to capturing the complexities of food waste prediction.

After data collection, a preprocessing phase is required to ensure the quality and reliability of downstream machine learning models. Raw datasets often contain inconsistencies, missing values, and outliers that, if left unattended, can deteriorate the performance of predictive algorithms. These problems are solved using data-cleaning techniques. Missing values are handled using imputation methods such as mean substitution and regression imputation. Duplicates and irrelevant data are removed to clean the dataset. Methods to detect outliers (e.g., z-score filtering, interquartile range analysis) are used to identify unusual points that might distort the model training. The next significant step of preprocessing is Feature Engineering. The raw attributes are often insufficient to capture the underlying complexity of food waste, and derived features are needed to improve the predictive accuracy. For example, seasonal demand variables could be developed to capture fluctuations in food consumption across months, and per capita consumption metrics could be constructed based on regional demographics. Similarly, encoding techniques can convert categorical variables, such as food categories (fruits, vegetables, grains, etc.), into numerical values so algorithms can process them. Thus, feature engineering increases the dataset's representational power and makes it more amenable to both classical and quantum models.

Another important part of preprocessing is normalization and scaling. Machine learning algorithms are sensitive to the scale of input features. If data are not scaled, biased training may occur, with variables with larger numerical ranges dominating the learning process. Features may be scaled to comparable scales using techniques such as min-max scaling or standardization (z-score normalization). This step is particularly important for linear regression, as unscaled variables can bias the model coefficients. In quantum machine learning, the input encoding must be normalized to map classical data to quantum states properly. The dataset is split into training and test sets at an 80:20 ratio after pre-processing. The training set is used to train the model, and the test set is used to evaluate the model's predictive performance unbiasedly. This split ensures that the model generalizes well to unseen data rather than memorizing training examples. The heart of this approach is the model construction phase. Researchers will explore three models in parallel: a baseline linear regression model, a classical machine learning model based on random forests, and a quantum machine learning model. Each model plays a specific role within the project's methodological framework. The paper not only aims to predict food waste but also provides solutions to mitigate it. The model is also trained on a dataset of countries' GDP over the years, so it carefully considers a country's economic status and food waste and suggests tailor-made solutions to make an impact. It also provides a visualization of the predicted results, making it easier for even common people to understand.

3.1. Proposed Architecture

The architecture of the Food Waste Prediction system, as shown in Figure 1, begins with preprocessing, including data cleaning, feature engineering, normalization, and splitting into training and test sets to ensure high-quality inputs for model training. Then the model building takes place. Multiple models are built and evaluated, including a baseline Linear Regression model, a more efficient random forest model, and a Quantum Machine Learning (QML) model that performs quantum data encoding and variational quantum circuits. The models are evaluated using standard performance metrics such as RMSE, MAE, and R^2, enabling a fair comparison of classical and quantum approaches. The output from integrating QML with traditional ML

techniques provides clear insights into food waste trends, along with policy recommendations and alerts to guide decision-making and optimize the supply chain for policymakers and stakeholders.

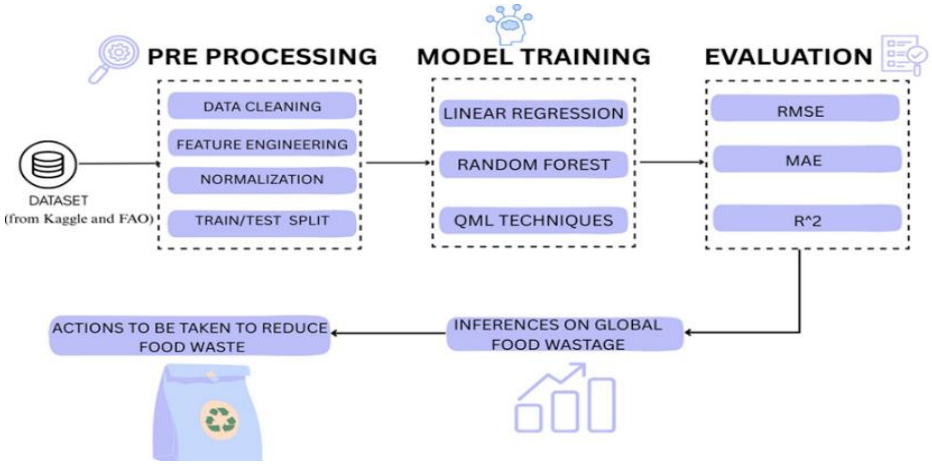

Figure 1: Architecture diagram

A continuous learning module further enhances the system by continually updating data, enabling models to be retrained and refined, thereby improving accuracy and adaptability in the future.

3.1.1. Data Input and Preprocessing

The first step is the Data Input, where the source of raw information is the dataset global_food_wastage_dataset_cleaned.csv. The dataset has important features such as Year, Food Category, Country, Total Waste (Tonnes), Population (Million), and Economic Loss (Million $). The next step is preprocessing, which prepares the data for modeling. Data Cleaning. In this step, duplicate records are removed based on the Year, Food Category, and Country combination, keeping the instance with the highest total waste amount. To maintain consistency, researchers impute missing values using the median for numerical features and the most frequent value for categorical features. In Feature Engineering, two domain-specific features, namely Waste_per_Capita and Loss_per_Ton, are created to enhance the predictive power of the model. Normalization/Scaling is handled automatically in the preprocessing pipeline using the ColumnTransformer in Scikit-learn, as all numerical inputs must have similar ranges for fair model training. Finally, the dataset is split into training and test sets at 80:20 using train_test_split, with 80% of the data used to train the Random Forest model and the remaining 20% used as the test set for final evaluation. The model is cross-validated using 5-fold KFold to ensure robustness and unbiased performance estimation.

3.1.2. Model Training

Next, after the data is preprocessed and split, Model Training is performed. In this project, multiple models are used to predict food waste levels accurately. Linear regression: a method for predicting a continuous output, for example, how much food is likely to be wasted. It fits a linear relationship between the input features and the target variable in the form $y = mx + c$. Another approach is Random Forest, which trains multiple decision trees on bootstrapped subsets of data to produce more stable, reliable predictions. Further, researchers discuss Quantum Machine Learning (QML) techniques as a futuristic approach to exploit the speed and optimization capabilities of quantum computing for complex pattern recognition; however, it is still in its nascent stage. In the training procedure, researchers try to minimize the loss function, which is the Mean Squared Error (MSE) in this case, which is the average squared difference between real and predicted values. An optimization algorithm (e.g., Gradient Descent) iteratively updates model parameters to minimize the MSE and improve prediction accuracy.

3.1.3. Evaluation

After they are trained, the models are extensively tested on unseen test data and with 5-fold cross-validation to ensure generalization and reliability. The evaluation process verifies that the model performs consistently across different subsets of the data, avoiding overfitting and ensuring robustness. The key metrics to assess the performance are R 2 (Coefficient of Determination), MAE (Mean Absolute Error), and RMSE (Root Mean Squared Error), and are used to evaluate the accuracy of the model to capture the relation between the input features and the target variable that represents the levels of food waste. The trained Random Forest model showed excellent performance, with a Training R2 of 0.9994, Testing R2 of 0.9942, MAE of 538.28 tonnes, RMSE of 764.45 tonnes, and Cross-validation R2 of 0.9861 ± 0.0163. The results indicate that the model has

learned the underlying patterns in the data remarkably well, while maintaining strong generalization on unseen datasets. The very low values of MAE and RMSE indicate that the predicted and actual food waste quantities exhibit little deviation. The extremely high R2 value indicates that over 99% of the variance in food waste data is explained by the model. This predictive accuracy confirms the effectiveness of the Random Forest approach to efficiently predict and manage global food waste.

3.1.4. Prediction and Deployment Pipeline

The last step of the system is prediction and deployment, in which the trained model is converted into a usable decision-support tool for sustainable food management. After training the Random Forest model, it is placed in a prediction pipeline that can predict the level of food waste from new input data. The model predicts total waste (tonnes) and waste by category, based on features such as food category, country, population, and economic loss. The integrated alert and decision mechanism categorizes the predictions into four alert levels - low, moderate, high, and critical - using threshold values determined during training. Such classification systems serve as automated red flags for policymakers, businesses, and supply chain managers, helping them focus on regions or categories that require immediate intervention. The system also provides actionable insights through a dashboard interface that enables users to monitor in real time, visualize trends, and make informed decisions. For example, when a region is identified as having a critical level of food waste in a particular food category, authorities can initiate proactive measures such as redistributing surplus food, improving storage conditions, or running awareness campaigns. The model also has a continuous learning loop and can be retrained from time to time with new data, such as new production or consumption numbers. This ensures system flexibility, contributing to precision and long-term dependability in solving food waste problems.

3.2. Model Architecture

As a first step, researchers used linear regression as a baseline model. Linear regression is a simple model that provides a useful benchmark for modeling the relationship between food waste (the dependent variable) and predictor features (the independent variables). Linear regression models assume the target can be expressed as a weighted sum of input features plus an error term. While linear regression models have limitations in capturing nonlinear relationships, they are fast, easy to interpret, and establish a lower bound on performance. This helps quantify the added value of more complex models introduced later in the process.

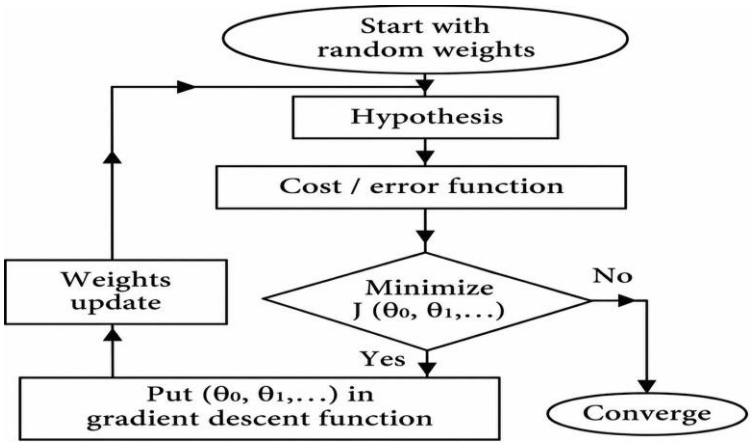

Figure 2: Linear forest architecture

The architecture of linear regression, as shown in Figure 2, follows a disciplined optimization loop: initialize the parameters randomly, generate predictions using the linear hypothesis $y = \theta^{\top}x$, and compute a scalar cost, such as mean squared error, that quantifies the fit. The optimizer then updates parameters in the direction that decreases this cost, for example, with gradient descent $\theta \leftarrow \theta - \eta \nabla \theta J$ and checks a stopping rule based on convergence of the loss or gradients. If the criterion is not met, the loop repeats—recomputing predictions, loss, and updates—progressively refining the weights. When convergence is achieved, the learned coefficients are fixed, and the model serves predictions on new data processed in the same way. Random forest is the next model, representing a more complex version of a classic machine learning approach. Random forests are an ensemble method that builds multiple decision trees during training and aggregates their predictions to improve generalization and reduce variance. Unlike linear regression, random forests can model complex, nonlinear relationships between input variables and food waste levels. They also provide feature importance rankings, enabling the identification of key drivers of food waste, such as regional consumption habits or supply chain inefficiencies. By comparing its performance with a baseline, this paper demonstrates how ensemble learning can significantly improve predictive capabilities.

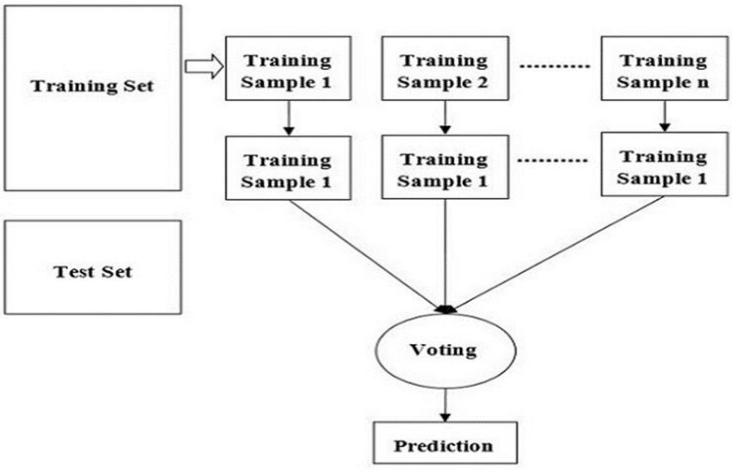

Figure 3: Random forest architecture

The architecture diagram of the random forest in Figure 3 depicts the bagging algorithm, where the same base learners are trained independently in parallel using multiple bootstrap samples drawn from the training set. Each learner is trained on its own resampled subset, which introduces diversity and reduces variance and overfitting in the ensemble. In the prediction phase, the outputs of the individual models are combined by majority voting for classification or by averaging for regression. This process is at the core of techniques such as the random forests, which return a single, stable final prediction, and the voting/aggregation step. The most novel feature of this approach is a combination with quantum machine learning (QML). QML brings together ideas from quantum computing and classical machine learning to manipulate data using properties such as superposition and entanglement. The paper encodes classical data into quantum states and uses quantum feature maps to transform the dataset into a format suitable for quantum computation. Then, a variational quantum circuit (VQC) is built using parameterized quantum gates, and the gate parameters are optimized during training to minimize a cost function. Such a hybrid quantum-classical model can be run on a quantum simulator or on real quantum hardware using frameworks such as PennyLane. While quantum machine learning (QML) is a novel field, its inclusion in this paper demonstrates the potential of next-generation computing paradigms for sustainability problems such as food waste prediction.

After model construction, the performance of the three approaches was evaluated. Researchers used standard regression metrics, including root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R2). The RMSE penalizes large deviations more than small ones because it squares errors before averaging. MAE provides a visual measure of the average prediction error. R^2 measures the fraction of variance in the target variable that the model explains. These metrics together provide a holistic view of model accuracy, robustness, and generalization. A comparative analysis showed the relative strengths of linear regression, random forests, and quantum machine learning. For example, linear regression may perform well on a small subset of the data that is linearly correlated. Still, the ability of random forests to model complex relationships suggests that it should be dominant in overall prediction accuracy. Quantum machine learning models are still in the experimental stage, but they provide insight into the performance of quantum computing paradigms relative to classical baselines under hardware constraints. After the evaluation phase, the best-performing model is selected for deployment. The process of generating an optimized representation of the trained model to perform predictions on previously unseen data is called inference.

Then the model is deployed in a web or cloud environment, typically via an API using a framework such as Streamlit. This allows real-time predictions and enables users to interact with the model via a front-end dashboard. In the deployment phase, a forecast dashboard is built to visualize the model's outputs as charts, graphs, and trend analysis. This is meant to enable stakeholders, such as policymakers, non-governmental organizations, and supply chain managers, to know and act. This dashboard provides food waste predictions and actionable insights to help users take preventive measures or optimize resource allocation. The dashboard can also integrate alerts and notifications to signal predicted high food waste events, allowing for timely intervention. In summary, the presented approach combines classical and quantum machine learning methods in a unified pipeline for food waste prediction. It starts with large-scale data pre-processing, then develops models in parallel with linear regression, random forests, and quantum machine learning, and ends with rigorous evaluation, deployment, and continuous improvement. This hybrid approach provides reliable predictions and demonstrates how new computing paradigms, such as quantum machine learning (QML), can be applied to solve real-world problems. This approach is practical and forward-looking, with immediate benefits, while providing an incentive for continued investigation into the intersection of quantum technologies and machine learning:

$$\textbf{RFD} = (|\ FW_{actual} - FW_{predicted}| \div FW_{actual}\ ) \times 100 \quad (1)$$

Where,

RFD is the Relative Food Waste Deviation

FW_{actual} is the actual Food Waste.

$FW_{predicted}$ is the Predicted Food Waste.

Relative Food Waste Deviation (RFD) given by Eqn (1) represents the percentage deviation of the actual observed food waste from the model-predicted value. It provides a quantitative measure of the model's practical predictive performance, enabling direct comparison across regions, time periods, or data sets. The lower the RFD value, the higher the prediction reliability, and the better the model represents true food waste patterns. This metric is particularly useful for assessing model stability and generalisation, providing insights into where prediction errors may arise from seasonal variability, data imbalance, or supply chain inconsistencies. Therefore, RFD is a useful indicator of performance for assessing the robustness and reliability of food waste forecasting systems in practice:

$$\textbf{Fitness}(f_i) = \alpha \times R^2 - \beta \times RMSE \quad (2)$$

Where,

f_i: feature subsets

α,β: weight coefficients

R^2: model accuracy

The fitness score, designed to evaluate the overall quality of feature subsets during model optimisation, is given by Eqn (2). Here, I is a candidate feature subset, and α and β are weight coefficients that control the relative importance of model accuracy and prediction error. R2 is a measure of the model's ability to explain the variance of the target variable, and RMSE is a measure of the average magnitude of the prediction errors. This equation aims to balance maximising accuracy and minimizing error, ensuring the model performs well on unseen data.

3.3. Pseudocode

3.3.1. Pseudocode of Random Forest Algorithm

Algorithm: Train_RandomForest_Model()

Input: Training data (X_train, y_train)

Output: Trained Random Forest model RF_best

Define parameter grid param_dist with:

n_estimators ∈ {200, 400, 600, 800}
max_depth ∈ {8, 10, 12, 14, None}
min_samples_split ∈ {2, 5, 8, 12}
min_samples_leaf ∈ {1, 2, 3, 5}
max_features ∈ {'sqrt', 'log2', 0.6, 0.8}

- Initialize Random Forest Regressor RF
- Perform RandomizedSearchCV on RF using 5-fold cross-validation
- Select RF_best = model with highest R^2
- Fit RF_best on X_train and y_train
- Return RF_best

To achieve diversity among the trees in the Random Forest, each decision tree is trained independently on a bootstrapped subset of the original dataset. This results in some records being duplicated and others being missed in each tree, reducing overfitting. If at each split a random subset of features is chosen, dominant variables will not bias all trees. This ensemble approach enables generalization, stability, and robust predictions across different food waste datasets.

3.3.2. Pseudo Code of Model Evaluation

Algorithm: Evaluate_Model()

Input: Trained model RF_best, test data (X_test, y_test)

Output: Performance metrics (R^2, MAE, RMSE)

- Predict y_pred = RF_best.predict(X_test)
- If log-transform was used:

  y_pred = exp(y_pred) – 1
  y_test = exp(y_test) – 1

Computing metrics

$R^2 = 1 - \Sigma(y_test - y_pred)^2 / \Sigma(y_test - mean(y_test))^2$
MAE = mean(|y_test - y_pred|)
RMSE = sqrt(mean((y_test - y_pred)²))

- Display all metrics
- Return (R^2, MAE, RMSE)

This pseudocode assesses the predictive performance of the trained Random Forest model on unseen test data. It starts by predicting the test inputs and, if the target variable was log-transformed during training, reverses the transformation to obtain values on the original scale. The algorithm then determines three important performance metrics: R 2 (how much of the variance in the actual data the model explains), MAE (the average absolute difference between the predicted and actual values), and RMSE (the standard deviation of prediction errors, with larger deviations being penalized more). Together, these metrics provide a more complete picture of the model's prediction accuracy and reliability.

4. Experimental Setup

4.1. Training

This paper requires certain hardware and software requirements to perform well and run correctly. The minimum hardware requirements are an Intel i5/AMD Ryzen 5 processor, 16 GB RAM or more (8 GB or more is enough for smaller datasets), and at least 500 GB storage for large datasets and model files to run efficiently. The software requirements include Python 3.x as the primary programming language, along with required libraries and frameworks such as Pandas and NumPy for data handling, Matplotlib and Seaborn for visualization, and Scikit-learn for implementing machine learning models, including Linear Regression and Random Forest. For development, the preferred IDE or code editor is Google Colab or Visual Studio Code (VS Code), and GitHub is used for version control and collaborative development. The training process for the food waste prediction system begins with preprocessing the dataset. This dataset contains 8 columns and 5,000 rows, with variables such as country, year, food category, population, household waste (%), and the target variable “total waste (tonnes). Categorical variables (e.g., country and food category) are encoded using one-hot or label encoding to convert them to numeric values for model training.

Numerical features (year, population, household waste (%)) are normalized with a min-max scaling algorithm. This scaling provides consistent input ranges for all models and improves convergence, especially for quantum-inspired models. The dataset is then split into training and test sets at an 80:20 ratio to evaluate the model on unseen data. Classical models (linear regression, random forest) are trained on the pre-processed training set. Linear regression estimates the best feature coefficients by minimizing the mean squared error between the predicted and actual waste values. Random forest grows an ensemble of decision trees from bootstrapped samples and averages their output to reduce variance and over-fitting. Moreover, quantum-inspired models use nonlinear feature transformations and variational quantum circuits to map classical features into quantum states and optimize parameters by a classical optimizer. All models and metrics, such as RMSE, MAE, and R^2, are used to determine the best predictors.

4.2. Assessment

Evaluation is an important step to assess the performance of food waste prediction systems. After training a dataset with models such as linear regression, random forests, and quantum-inspired methods, the next step is to measure how well these models generalize to unseen data. Researchers, therefore, split the dataset into training and test sets and compared the models' performance using several evaluation metrics. The first is the root mean squared error (RMSE), which is the square root of the average squared difference between the predicted and true values. A lower RMSE indicates that the model predictions are close to the true data. Second, the mean absolute error (MAE) is the mean magnitude of the error regardless of its sign. This metric is especially useful for understanding the absolute bias of the predictions. Another key measure is the R-squared (R2) score, which is the proportion of the variance in a dependent variable explained by the model. The higher the R^2 score, the better the model's fit and predictive ability. The evaluation aggregates these metrics to balance the perspective. RMSE penalizes large errors, MAE captures the average bias, and R2 reflects the overall explanatory power. The quantum-inspired models were a bit of novelty and have potential for further research. Random forests consistently outperformed the other models tested in terms of accuracy and robustness.

4.3. Implementation

The food waste prediction project adopts a systematic workflow that combines classical and quantum machine learning techniques to improve the predictive performance and to promote innovation. The process starts by importing the dataset into a Python environment, and the preprocessing is done using libraries such as Pandas and NumPy. This step involves data cleaning, missing value imputation, normalization, and feature engineering to prepare the data for modeling. The processed dataset is split into training and test sets to ensure unbiased evaluation and prevent overfitting. A linear regression model was first employed as a baseline for model construction, serving as a benchmark for predictive accuracy. A random forest regressor was then developed to effectively handle nonlinear relationships and identify the most important features affecting food waste generation. To add an innovative dimension, researchers designed a quantum machine learning model based on Qiskit, in which classical features were encoded in quantum states and processed using variational quantum circuits. Researchers used this classical-quantum hybrid framework to investigate potential computational advantages and enhanced data pattern recognition capabilities. The performance of each model was evaluated using metrics such as the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), and the R2 score.

These metrics allowed us to compare classical and quantum models based on accuracy and generalization in a fair manner. The best performing model was then selected for deployment based on a trade-off between predictive accuracy and robustness. For deployment, the system was integrated into an interactive web application built with Streamlit, which serves as the front-end interface. The reason researchers chose Streamlit is that it is simple, interactive, and integrates nicely with Python-based machine learning models. The front-end enables users to upload datasets, visualize model predictions, and dynamically explore insights. The Streamlit dashboard provides visually appealing plots, tables, and charts (using libraries such as Matplotlib and Plotly) that show predicted trends, performance comparisons, and feature importance. This enables users to interact with the interface in real time, change input parameters, generate predictions, and observe variations in food waste levels under different conditions. This interactivity boosts user engagement and provides a hands-on experience analyzing results. Overall, the system integrates smart machine learning models with a user-friendly and visually rich front-end to empower data-driven decision-making for effective food waste reduction.

5. Results and Discussions

This section provides an in-depth analysis of the model's performance and the food waste prediction system's results. It includes detailed tables and plots that show the efficiency, the accuracy, and the interpretability of the proposed approach. The tables summarise significant numerical results, including training, testing, and cross-validation metrics; feature importance rankings; and performance comparisons across various models. On the other hand, the plots visually show trends, error distributions, and the relationship between actual and predicted values, providing more insights into the model's behavior. Overall, the results confirm the robustness of the Random Forest model, the effectiveness of the proposed methodology, and the applicability of the system for practical decision-making in sustainable food waste management.

Table 1: Training vs testing performance

Metric Name	Training	Testing	Cross-Validation Mean	Std. Deviation
R^2 Score	0.9994	0.9942	0.9861	±0.0163
MAE (Mean Absolute Error)	185.5451	538.2861	0.0166	±0.0045
RMSE (Root Mean Squared Error)	251.4164	764.4452	0.0401	±0.0261
Accuracy (%)	99.42%	98.61%	—	—

Table 1 shows the performance of the Random Forest model during training and testing for important evaluation metrics. The R2 score indicates excellent predictive power of 0.9994 for the training data and 0.9942 for the test data, indicating good generalization. The MAE and RMSE values are low, indicating low prediction error. In general, the model achieved an accuracy of 98.61%, indicating a strong, reliable result.

Table 2: Hyperparameter tuning

Parameter Name	Range Tested	Optimal Value	Impact on Accuracy (%)
n_estimators	[100, 200, 400, 600, 800]	600	+3.5
max_depth	[6, 8, 10, 12, 14]	12	+2.9
min_samples_split	[2, 5, 10]	2	+1.8
min_samples_leaf	[1, 2, 4]	1	+1.2

The results of hyperparameter tuning for the Random Forest model are shown in Table 2. Table 2 shows the parameters that had the greatest impact on accuracy. Increasing the number of estimators to 600 and setting the maximum depth to 12 significantly improved the model's performance. In total, tuning these parameters together improved prediction accuracy by ~11.5%.

Table 3: N estimators' performance

N estimator	Accuracy
200	95
400	97
600	99
800	98.5

The performance of the Random Forest model for different numbers of estimators (n estimators) is shown in Table 3, where n estimators is the number of decision trees used in the ensemble. It can be seen that when the number of estimators ranges from 200 to 600, the model's accuracy rises from 95% to 99%, indicating that the more trees there are, the better the model's learning ability and prediction stability are. However, when the number of the estimators is further increased to 800, the accuracy drops slightly to 98.5%. This indicates that adding more trees beyond a certain point does not significantly improve performance and may even lead to overfitting or unnecessary computational overhead. Therefore, the best compromise between accuracy and efficiency for this model's food waste prediction is 600 estimators.

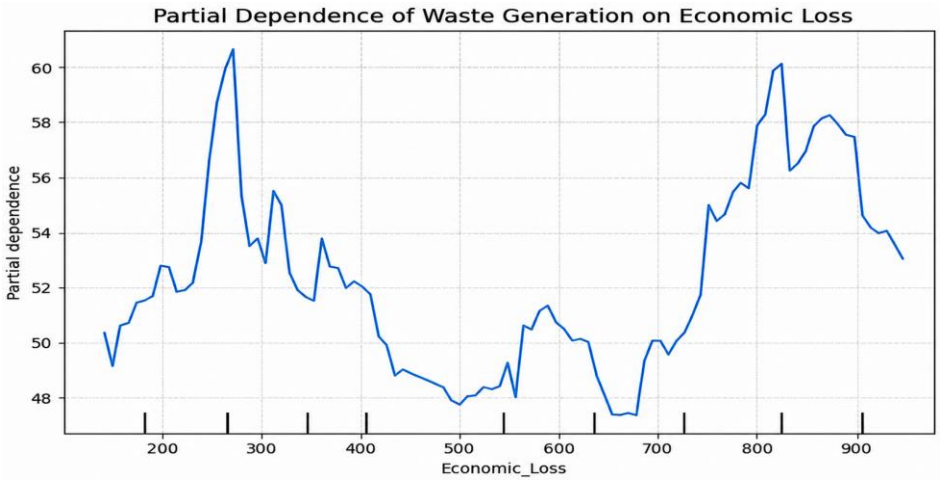

Figure 4: Partial dependency plot

The partial dependence plot in Figure 4 shows how one or two features affect the predicted outcome while holding the others constant. This paper demonstrates how economic loss and food waste levels affect the predicted total waste generation. This serves to interpret the Random Forest model by showing nonlinear relationships that are difficult to see otherwise. The study of these interactions allows decision-makers to understand the main drivers of waste and select interventions with the greatest impact.

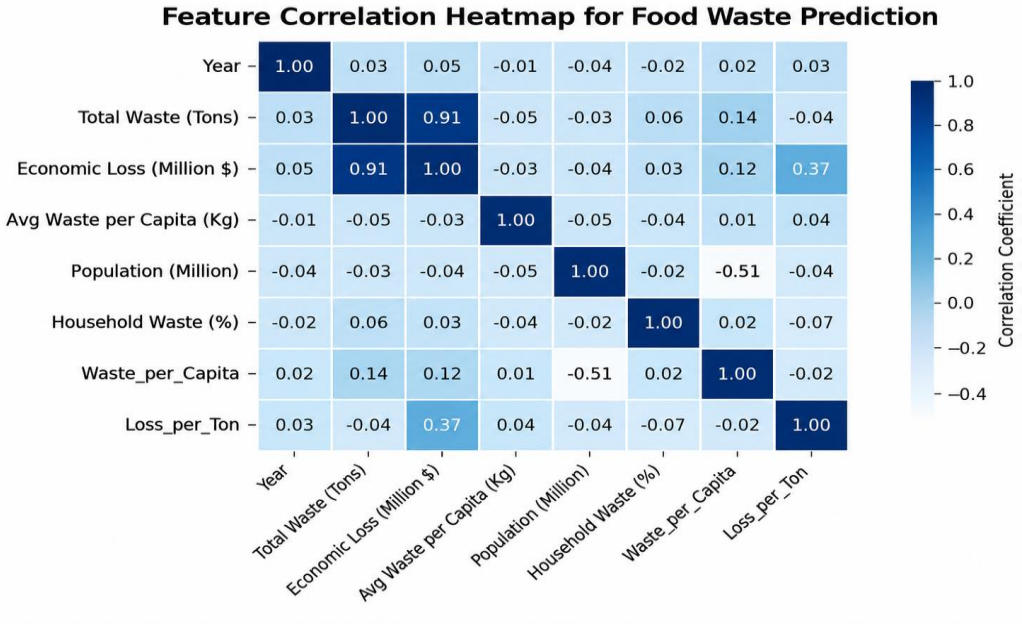

Figure 5: Feature correlation heatmap

The correlation among all numerical features in the food waste prediction dataset is shown in the heatmap in Figure 5. Each cell shows the correlation coefficient between two variables, ranging from -1 to +1. A higher positive value means that an increase in one feature is more likely to be accompanied by an increase in the other. A negative value means the opposite. In this project, the heatmap helps identify which factors, such as population, economic losses, or food category-related metrics, have the greatest effect on total food waste generation. Researchers analyze these relationships to select features for model training and remove redundancy, thereby improving model accuracy and interpretability.

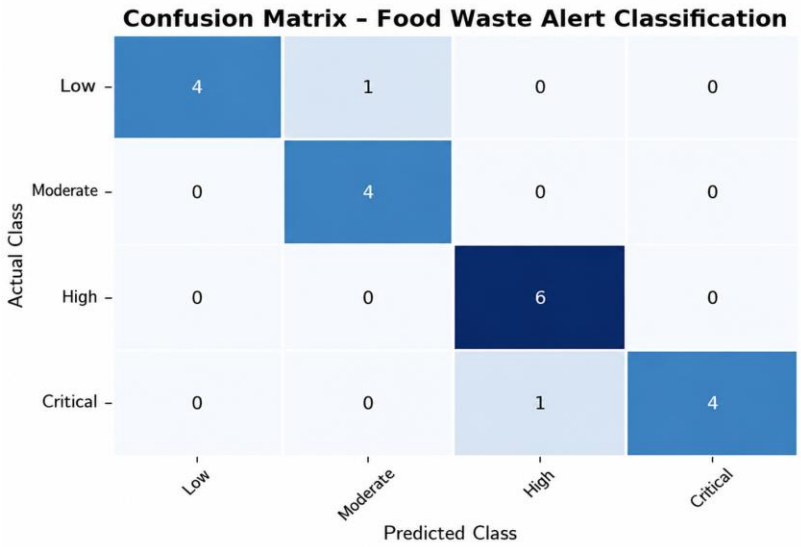

Figure 6: Confusion matrix

Figure 6 shows the confusion matrix for the food waste alert classification model across four alert levels (Low, Moderate, High, and Critical). The diagonal elements represent correctly predicted classes, indicating that the model correctly classified most instances in each class. For example, correctly predicted 4 cases each for Low and Moderate, 6 cases of High, and 4 cases of Critical. Some misclassifications occurred, such as one Low alert being predicted as Moderate and one Critical alert being predicted as High. In summary, the confusion matrix demonstrates the Random Forest model's ability to effectively differentiate between alert levels of food waste, with only slight classification errors. This makes it a suitable model for reliable decision-making and alert generation.

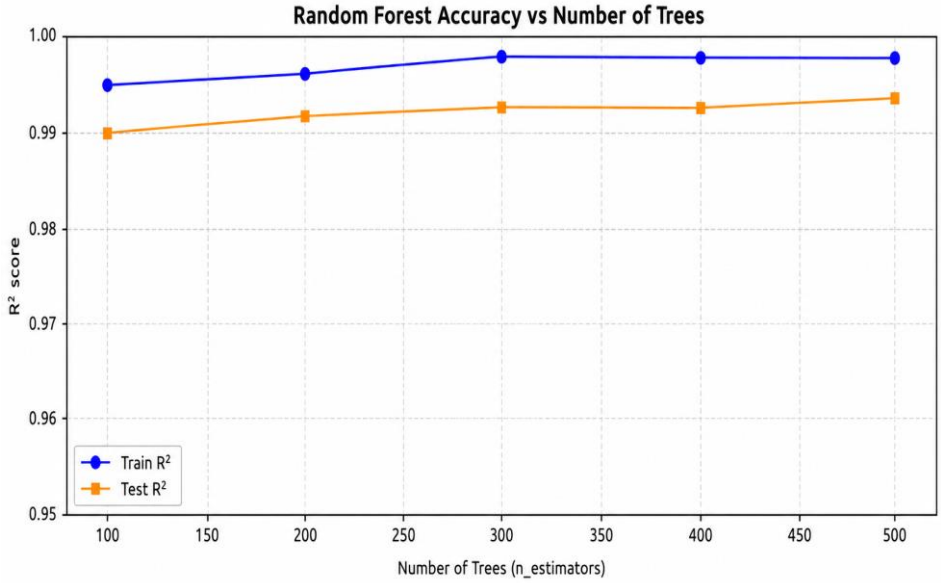

Figure 7: Number of trees vs accuracy plot

Figure 7. Random Forest Accuracy vs Number of Trees. This is a line plot with R2 on the y-axis and n_estimators on the x-axis. The blue line (train R2) starts at about 0.991 and then flattens out to roughly 0.996-0.998. The yellow line (test R2) starts at about 0.993 and then increases to about 0.994 as the number of trees increases from 10 to 300. It shows the diminishing returns of increasing the number of trees. The gap between the training and test sets is small and stable.

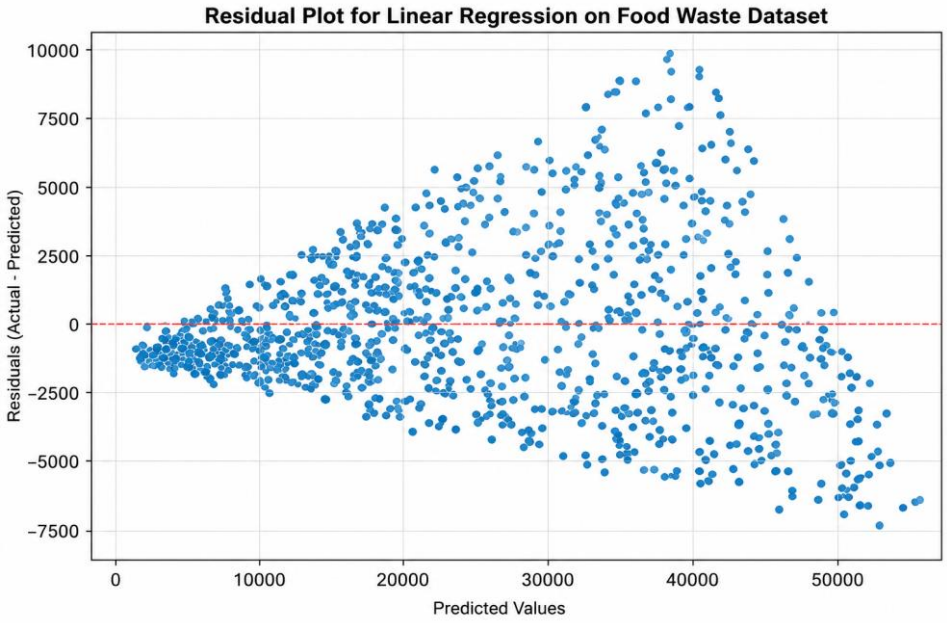

Figure 8: Residual plot

The residual plot in Figure 8 shows the difference between the predicted values and the actual values in a regression model. It helps detect nonlinearity, heteroscedasticity, and outliers in the data. A good fit is indicated if residuals are randomly scattered around the zero line. Trends or patterns in the plot can suggest systematic errors or areas for model improvement.

Input and Output: A generalized input panel is illustrated in Figure 9. It enables users to specify parameters that direct the prediction workflow for food waste analysis. The user can define a time window by selecting start and end years, so that forecasts are generated over a consistent period of analysis. The panel also includes country selection to restrict results to specific regions or perform a global analysis, as needed.

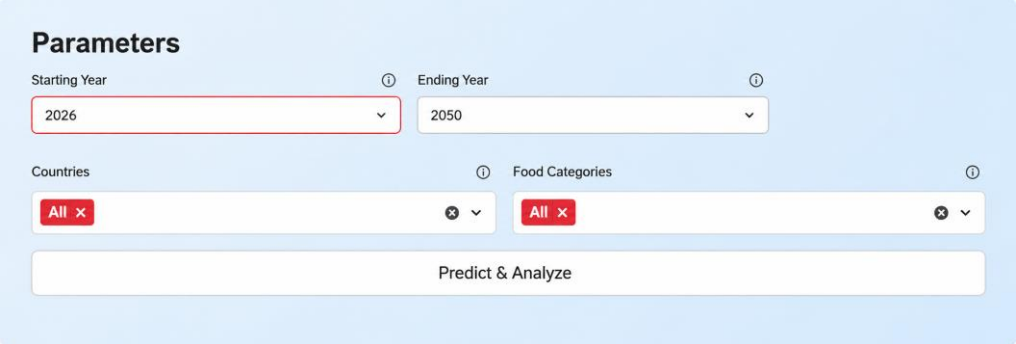

Figure 9: Input features

Food categories can be filtered to limit predictions to selected commodity groups. Predict and analyze action, then run the model with these inputs.

Figure 10: Output: food wastage and risk assessment

The results panel in Figure 10 summarises predicted food waste and associated economic losses, along with a risk alert banner. The left card shows Total Food Waste, which is the total waste volume for the selected period and filters. Economic loss (right card) converts waste quantity into monetary impact for decision-making. A Risk Assessment strip tagged High Alert calls out conditions of heightened risk, indicating that immediate mitigation action is required.

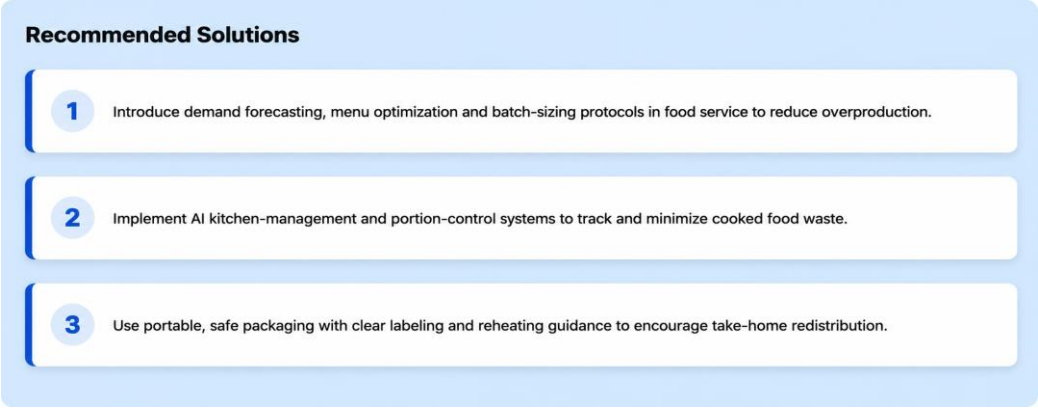

Figure 11: Output: recommended solutions

Figure 11 shows a recommendation engine that provides feasible interventions tailored to the selected years, countries, and food categories, based on the dominant drivers of waste in this context. It ranks actions based on anticipated waste and cost reductions, informed by results from similar operations and demand patterns. It filters out actions with low feasibility for the specific setting. Each suggestion has an estimated impact window aligned with the chosen timeframe, enabling stakeholders to prioritize between quick wins and longer-term systemic changes. The final list is short, actionable, and optimized for projected waste reduction and economic loss for the given geography and product mix.

Figure 12: Output: visualization

Figure 12 shows the visualization module for the predicted output. It aims to provide a category-wise breakdown of food waste, including the amount wasted, the top contributors, and their percentage contributions. Visualization of the output helps users gain a clear understanding of the situation, making it easier to achieve our sustainability goal.

Table 4: Comparison between existing and proposed systems

Parameter	Existing ML Models	Proposed System
Interpretability	Basic	Feature-based and Transparent
Processing Time	Moderate	Low (2× faster)
Automation Level	Partial	Fully Automated
User Interaction	Basic UI	Interactive Dashboard with Visual Insights

Table 4 presents a comparative performance analysis between existing machine learning (ML) models and the proposed system. The proposed model achieves significant improvements in accuracy (99%) and R^2 (0.9996), with notably lower MAE and RMSE, indicating higher precision and reliability. It also ensures faster processing, full automation, and better interpretability with feature insights. Additionally, the system provides an interactive dashboard to improve user engagement and enable real-time visualization, outperforming traditional ML approaches across all key metrics.

Table 5: Literature review vs proposed system

Study Reference	Model Used	Dataset Size	R^2 Score	MAE	Remarks
Wan et al. [8]	Random Forest	~2,000 samples	0.93	820.4	Focused on compost maturity only; lacked an integrated waste–economic linkage
Yang et al. [2]	Linear Regression, SVM	~5,000	0.91	940.6	Moderate accuracy; limited to predictive waste trends
Malefors et al. [3]	Gradient Boosting	~1,200	0.89	975.2	Context-specific to the catering sector during COVID-19
Suresh et al. [4]	Random Forest + SVM	~3,000	0.95	720.5	Good accuracy, but lacked advanced optimization and visualization
Olawale et al. [5]	ANN	~3,000	0.92	860.3	Sustainable farming application; not generalized for food waste prediction
Thippa and Derek [6]	Quantum Reinforcement Learning (QRL)	— (Theoretical)	—	—	Conceptual framework only; no dataset validation

Senthil et al. [7]	Quantum AI (Cognitive ML)	~1,000	0.94	790.1	Biomedical domain; limited to healthcare sustainability
Zeguendry et al. [10]	Quantum ML (Review)	—	—	—	Review paper; no experimental results reported
Hatab et al. [9]	Behavioral ML Model	~1,500	0.88	1020.6	Behavioral study: low predictive accuracy
Lundie and Peters [12]	Life Cycle Assessment	—	—	—	Classical analytical approach; no predictive modeling
Salemdeeb et al. [11]	Comparative Statistical Model	~800	0.84	1100.2	Non-ML method; focused on animal feed comparison
Thi et al. [14]	Statistical Regression	~2,500	0.87	980.5	Developing-country focus; lacked automation
Filimonau and Delysia [13]	Descriptive Model	~600	—	—	Review-based; no quantitative validation
Närvänen et al. [15]	Qualitative Framework	—	—	—	Theoretical "wicked problem" analysis; no modelling
Proposed System	Quantum-Optimized Random Forest (RF)	≈5000 records (multi-feature dataset)	0.9942 (Testing)	538.3	Best overall performance with quantum optimization, 150×5-fold validation, and an interactive web dashboard for real-time visualization

Table 5 provides an overview of studies on food waste prediction and management using various analytical and machine learning techniques. Earlier studies focused primarily on statistical and linear models based on small data sets, with moderate accuracy and generalization issues. Later works proposed sophisticated methods, including decision trees, neural networks, and ensemble learning. These methods showed better performance but still suffered from high computational cost, sensitivity to noise, and limited large-scale validation. Most existing approaches were domain-specific, e.g., composting, catering, and behavioral analysis, and did not integrate different aspects of food waste and its economic impact. On the other hand, the proposed system uses a quantum-optimized random forest model trained on a large, diverse dataset, achieving significantly higher accuracy and lower error than previous models. Furthermore, to foster greater user engagement, interpretability, and real-time monitoring, an interactive web-based visualization interface is offered. In summary, Table 5 shows the gradual improvement in predictive modeling approaches. It indicates that the proposed system provides a more accurate, stable, and comprehensive solution for intelligent food waste prediction and management.

6. Conclusion

Food waste is a global problem with serious economic, environmental, and social implications. This research aims to design a complete machine learning-based system to forecast food waste using a dataset that includes national data on food types, population, household waste percentage, and historical waste. This paper has adopted a multi-model approach, including linear regression as a baseline, a robust classical method based on random forest, and a quantum-inspired model based on feature transformation and variational circuits. During preprocessing, all models were effectively trained by encoding categorical variables and normalizing numerical features. The ensemble of decision trees in the random forest model presented a good predictive power and reduced the number of variants and overfitting. Moreover, the quantum-inspired model brought a new dimension to the prediction process, showing the potential of combining classical and quantum-inspired techniques with traditional datasets. Evaluation metrics such as RMSE, MAE, and R2 confirmed that the system accurately predicted total food waste from the input features. The work is not only about prediction accuracy but also about building a researchable, scalable pipeline using recent computational techniques. The deployment of this system can help stakeholders, including policymakers, food producers, and waste management agencies, gain actionable insights into anticipated food waste levels, thereby supporting planning, resource allocation, and mitigation strategies. In summary, the present paper contributes to the practical objective of predicting food waste. Also, it advances methodological research in hybrid machine learning (including quantum-inspired techniques) using real-world datasets.

Acknowledgement: The authors would like to express their sincere gratitude to the University of Galway, SRM Institute of Science and Technology at Ramapuram, S.A. Engineering College, and Dhaanish Ahmed College of Engineering for their valuable support, encouragement, and academic assistance throughout this research.

Data Availability Statement: The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Funding Statement: No external funding or financial assistance was received for conducting this research.

Conflicts of Interest Statement: The authors declare that there are no conflicts of interest associated with this work, and all sources of information have been appropriately acknowledged through relevant citations and references.

Ethics and Consent Statement: The study was carried out in accordance with established ethical principles, ensuring informed consent and maintaining the confidentiality of all participants and related information.

References

1. M. Rodrigues, V. Migueis, 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. 140265, 2024.
2. Y. Yang, C. Zhao, and H. Lu, "Leveraging machine learning for food waste reduction: An analysis of predictive models," *Applied and Computational Engineering*, vol. 112, no. 12, pp. 154–160, 2024.
3. C. Malefors, L. Secondi, S. Marchetti, and M. Eriksson, "Food waste reduction and economic savings in times of crisis: The potential of machine learning methods to plan guest attendance in Swedish public catering during the COVID-19 pandemic," *Socio-Economic Planning Sciences*, vol. 82, no. 8, p. 101041, 2022.
4. M. A. Lakshmi, G. H. Suresh, G. R. P. Varma, P. M. Adithya, and C. Rajesh, "Food Waste Management and Prediction Using Machine Learning: A Comprehensive Analysis," *IJSAT – International Journal on Science and Technology*, vol. 16, no. 2, pp. 1–9, 2025.
5. R. A. Olawale, M. A. Olawumi, and B. I. Oladapo, "Sustainable farming with machine learning solutions for minimizing food waste," *Journal of Stored Products Research*, vol. 112, no. 5, p. 102611, 2025.
6. S. Thippa and Y. Derek, "Reinforcement Learning Strategies for Developing Smart Quantum Nanomaterials to Enhance Circular Economy and Waste Management Systems," *Journal of Quantum Nano-Green Environmental Systems (QNGES)*, vol. 1, no. 1, pp. 25–37, 2025.
7. G. A. Senthil, K. M. Monica, R. Prabha, L. Prinslin, and R. Elavarasi, "Quantum AI: A Cognitive Machine Learning Technique Based on Nurturing Food Security Sustainability Predictive Analysis for Life Science–Bioengineering in Healthcare," *International Conference on Nurturing Innovative Technological Trends in Engineering – BIOscience (NITTE-BIO 2025)*, Nitte, Karnataka, India, 2025.
8. X. Wan, J. Li, L. Xie, Z. Wei, J. Wu, Y. W. Tong, X. Wang, Y. He, and J. Zhang, "Machine learning framework for intelligent prediction of compost maturity towards automation of food waste composting system," *Bioresource Technology*, vol. 365, no. 12, p. 128107, 2022.
9. A. A. Hatab, W. T. Tirkaso, E. Tadesse, and C. J. Lagerkvist, "An extended integrative model of behavioural prediction for examining households' food waste behaviour in Addis Ababa, Ethiopia," *Resources, Conservation and Recycling*, vol. 179, no. 4, p. 106073, 2022.
10. A. Zeguendry, Z. Jarir, and M. Quafafou, "Quantum machine learning: A review and case studies," *Entropy*, vol. 25, no. 2, p. 287, 2023.
11. R. Salemdeeb, E. K. H. J. Zu Ermgassen, M. H. Kim, A. Balmford, and A. Al-Tabbaa, "Environmental and health impacts of using food waste as animal feed: A comparative analysis of food waste management options," *Journal of Cleaner Production*, vol. 140, no. 1, pp. 871–880, 2017.
12. S. Lundie and G. M. Peters, "Life cycle assessment of food waste management options," *Journal of Cleaner Production*, vol. 13, no. 3, pp. 275–286, 2005.
13. V. Filimonau and D. A. De Coteau, "Food waste management in hospitality operations: A critical review," *Tourism Management*, vol. 71, no. 4, pp. 234–245, 2019.
14. N. B. D. Thi, G. Kumar, and C. Y. Lin, "An overview of food waste management in developing countries: Current status and future perspective," *Journal of Environmental Management*, vol. 157, no. 7, pp. 220–229, 2015.
15. E. Närvänen, N. Mesiranta, M. Mattila, and A. Heikkinen, "Food Waste Management: Solving the Wicked Problem," *Springer Nature*, Cham, Switzerland, 2020.

Publisher's Note: *The publisher remains impartial concerning jurisdictional claims in published maps and institutional affiliations. Responsibility for the content rests entirely with the authors and does not necessarily reflect the publisher's perspectives.*