Journal of Food Technology Research
2025 Vol. 12, No. 4, pp. 302-314
ISSN(e): 2312-3796
ISSN(p): 2312-6426
DOI: 10.18488/jftr.v12i4.4625

An extension of the theory of planned behavior with machine learning to predict household food waste in India

Veena Ghuriani[1+]
Pooja Baweja[2]
Sakshi Suman[3]
Nisha Poddar[4]
Haritma Chopra[5]

[1]*Department of Computer Science, Maitreyi College, University of Delhi, Delhi, India.*
[1]*Email: vghuriani@maitreyi.du.ac.in*
[2,3,4]*Department of Botany, Maitreyi College, University of Delhi, Delhi, India.*
[2]*Email: pbaweja@maitreyi.du.ac.in*
[3]*Email: sakshisuman.2699@gmail.com*
[4]*Email: nishapoddar0777@gmail.com*
[5]*Department of Chemistry, Maitreyi College, University of Delhi, Delhi, India.*
[5]*Email: drharitmachopra@gmail.com*

(+ Corresponding author)

Article History
Received: 20 June 2025
Revised: 3 October 2025
Accepted: 1 December 2025
Published: 29 December 2025

Keywords
Explainable artificial intelligence
Food security
Food waste
Machine learning
Theory of planned behavior.

ABSTRACT

Food wastage is a complex social, economic, and environmental issue worldwide. Approximately one-third of edible food products are wasted globally. There is an urgent need to raise awareness about food wastage. The Theory of Planned Behavior (TPB), which encompasses attitude, subjective norm, and perceived behavioral control, has been extensively used to study food-related behaviors. In this study, we propose an extension of TPB by adding three attributes: self-responsibility, eco-consciousness, and food and home management, aiming to develop a more comprehensive framework for predicting food wastage behavior. A cross-sectional survey was conducted among individuals of different age groups, measuring the impact of various attributes on food wastage. A total of 341 responses were collected across India. The data were analyzed using statistical methods and machine learning (ML) algorithms such as Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF). The prediction accuracy of these ML classifiers for TPB and its extension was compared. Additionally, explainable artificial intelligence (XAI) techniques were employed to assess feature importance and identify significant attributes influencing food wastage. Results indicated that the prediction accuracy of all ML classifiers improved by 12% after incorporating self-responsibility, eco-consciousness, and food and home management attributes, thereby validating the extension of TPB. The integration of XAI highlighted key attributes associated with food wastage and increased the transparency of complex predictive models.

Contribution/Originality: The study employs an AI-driven pipeline to extend the Theory of Planned Behaviour by incorporating self-responsibility, eco-consciousness, and home management for assessing food wastage in Indian households. Additionally, it utilizes explainable AI to evaluate the significance of various attributes contributing to food wastage.

1. INTRODUCTION

The emergence, development, and persistence of human civilization are heavily dependent on the supply and availability of food (Godfray et al., 2010). While the United Nation's Universal Declaration of Human Rights asserts every individual's right to food (Martindale, 2010), approximately one-third of the food produced for human consumption is lost or wasted globally. This significant wastage occurs across every stage of the food supply chain, from production (agricultural land, industrial manufacturing) and processing to consumption (Mirabella, Castellani, & Sala, 2014; Stefan, Van Herpen, Tudoran, & Lähteenmäki, 2013; Williams & Wikström, 2011). Consequently, nearly 800 million people worldwide lack sufficient access to food, meaning roughly one in every nine individuals suffers from hunger (Racz, Vasiljev Marchesi, & Crnković, 2018). Given the resource-intensive nature of food production, these losses and wastes indirectly contribute to substantial environmental impacts, including water and air pollution, soil erosion, deforestation, and greenhouse gas emissions (Mourad, 2016).

In ancient times, food wastage was considerably lower due to localized food production and consumption patterns. The volume of food prepared was precisely tailored to immediate requirements, with any surplus either repurposed or consumed. The phenomenon of increased food wastage coincided with urban development and the proliferation of nuclear family structures (Bhatia & Sharma, 2023). Presently, a significant proportion of food waste originates at the household level, underscoring the critical necessity of effective prevention measures at this terminal point of the supply chain (Parfitt, Barthel, & Macnaughton, 2010). The squandering of food by end-users at this final stage renders all expended energy – across production, preparation, processing, and transportation – ultimately in vain.

Consumers are the largest individual contributors to food wastage (Griffin, Sobal, & Lyson, 2009). The new generation has a passion for wasting food by excessively buying and instead of eating it, throwing it away (Ghinea & Ghiuta, 2019). It is estimated that most of the food losses occur between production and transport of food to retail sites in the food supply chains in developing countries (Grandhi & Singh, 2016). Food wastage not only leads to an increase in the percentage of humans deprived of food, but also triggers climate change by emitting greenhouse gases amounting to approximately 17 million CO_2 equivalent tonnes (Pandey, 2021).

The impact of food wastage on the environment can be reduced by minimizing the amount of food wasted at the household level. Food wastage is influenced by various factors, and predicting its amount in a household is a challenging task (Quested, Marsh, Stunell, & Parry, 2013). According to the Food and Agriculture Organization, food waste can be defined as "the wholesome edible material intended for human consumption, arising at any point in the Food Supply Chain (FSC) that is instead discarded, degraded, or affected by pests"(Food and Agriculture Organization of the United Nations (FAO), 1981). A consequence of food wastage is the diminished availability and elevated cost of healthy food options. Food wastage not only has enormous environmental impacts and ethical issues but also can impact the national economy by incurring significant financial losses.

Qualitative research serves as a crucial first step in identifying the primary drivers of household food waste reduction. The theory of planned behavior is one well-known theory based on the factors: attitude towards behavior, subjective norm, and perceived behavioral control, which study the cognitive requirements of human behavior (Ajzen, 2012). Stefan et al. (2013) investigated the role of the Theory of Planned Behavior (TPB) in conjunction with food-related activities. Their study specifically explored the interdependence of intention and attitude in reducing household food waste. Recently, researchers have applied TPB to study consumer behavior towards various aspects like management, conservation, and disposal of food. Studies reveal that food waste behavior is impacted by attitudes toward food conservation, societal pressure to prevent waste, and presumed efforts in managing leftovers (Graham-Rowe, Jessop, & Sparks, 2015; Russell, Young, Unsworth, & Robinson, 2017).

According to the TPB model, the most significant part of behavior is intention, which indicates the level of motivation to carry out the desired behavior. The stronger the intention for a behavior, the more likely it is to be carried out. Intention is largely governed by attitude, subjective norms, and perceived behavioral control (Ajzen, 2012). Subjective norm is the societal pressure to behave in a particular way, attitude is whether the behavior is perceived as positive or negative by an individual, and perceived behavioral control is the judgment of an individual's ability to behave in a particular way.

The accuracy of the traditional TPB has been widely debated. Although it provides a useful baseline, it has been suggested to include additional predictor variables to fully capture moral, environmental, or practical factors that influence food wastage decisions, particularly in culturally diverse contexts such as India. TPB can be strengthened by incorporating additional attributes. Studies have advocated including the attribute of self-responsibility, which closely aligns with perceived behavioral control already supported by TPB. Bamberg and Möser (2007) have suggested that self-responsibility and eco-consciousness are personal traits towards environmentalism. They influence the TPB model's intentions and behaviors. Schlegelmilch, Bohlen, and Diamantopoulos (1996) have demonstrated a strong correlation between environmental consciousness and green purchasing decisions, resulting in less food wastage. Fielding, McDonald, and Louis (2008) had observed improvement in the model's predictive power by extending the TPB to include identity and ecological concern. Graham-Rowe et al. (2015) have advocated that food waste can be reduced by incorporating attributes of home planning and management. These attributes help in every aspect of household decision-making. Schanes, Dobernig, and Gözet (2018) have advocated that minimizing food waste in the household reduces the impact on the environment. The carbon footprint can be reduced by lowering greenhouse gas (GHG) emissions if the food is prevented from going to landfills. This, in turn, could save money and slow down the release of methane, which is harmful (Bhatia, Jha, Sarkar, & Sarangi, 2023). There is an urgent need to reduce the wastage of food so that it can reach the underprivileged and ensure food security around the world.

Artificial Intelligence is an emerging field in computer science that has found its application in almost every sphere of human life. Machine learning algorithms are being used to analyze data in various spheres. Panda and Dwivedi (2019) have proposed the use of machine learning algorithms to reduce food wastage in the mess of educational institutions. They have used Decision Tree and Naive Bayes algorithms to predict food wastage and reported that the Naive Bayes algorithm gives better results compared to Decision Tree. The food waste in the mess of two hostels was analyzed using AI and concluded that changes in behavioral aspects and management of the mess can help in reducing the wastage (Sama, Makkar, Prokshitha, Dhaloria, & Sharma, 2021). Anggraeni, Silaban, Anggreainy, and Cahyadi (2021) studied the Bayesian Networks and Agent-Based Modelling techniques to understand the interdependencies between the various attributes that lead to food wastage. They have studied the relationship between food waste generation and innovation techniques to be adopted in the retail sector. Grainger et al. (2018) used machine learning algorithms to study the household waste, they considered the drivers size of household, fussy eaters, status of employment, ownership of home and the local authority for food waste. They concluded that large households are the major factor for food wastage. Chowdary et al. (2024) employed XGBoost to predict food wastage levels by analyzing diverse datasets, including socio-economic indicators and environmental conditions, which can be applicable to Indian households facing similar challenges in food waste management. Ozcil (2024) explored the utility of machine learning in mitigating food waste by enabling precise food demand forecasting and efficient inventory optimization. This approach holds significant potential for adoption within Indian households, particularly for perishable goods, to bolster food security and promote sustainability.

In the present work, additional factors related to food wastage have been investigated, and an extension of the theory of planned behavior (Ext. TPB) has been proposed. This proposed extension encompasses factors related to self-responsibility, eco-consciousness, food and home management that influence food wastage in households. The work aims to demonstrate that the inclusion of these factors in TPB provides a better predictive framework for food wastage behavior compared to the traditional TPB model. The present work not only focuses on people's perceptions of food wastage, a global issue, but also assesses the importance of different attributes using XAI techniques. This study offers key behavioral insights for targeted interventions in Indian households.

Besides the behavioral aspect, the study provides an AI-driven pipeline that can be helpful in creating a framework globally for reducing food waste. The pipeline integrates machine learning models with explainable AI techniques such as SHAP, which can facilitate the development of intelligent applications for real-time monitoring of food consumption and disposal. The proposed pipeline can be used as a decision-support tool for developing food waste monitoring systems. This fusion of technology with behavioral study can help achieve the goal of sustainable food consumption with optimal resource utilization and less food wastage.

2. MATERIALS AND METHODS

2.1. Survey Technique

A cross-sectional survey was conducted to evaluate respondents' awareness regarding food wastage. The study utilized an online questionnaire comprising forty-five questions divided into eleven sections, followed by questions related to demographic studies. The questionnaire solicited opinions on household food wastage concerning attitudinal and behavioral aspects. The eleven sections were created based on the proposed Ext. TPB. G*Power statistical software (3.1.9.D package), with sample size (n), effect size (d), power, and significance criteria (α) as parameters, was used to determine the sample size (Faul, Erdfelder, Buchner, & Lang, 2009). Calculations performed using this software indicated a minimum required sample size of 295 to achieve a power of 0.95, with α=0.05 and d=0.50.

The online questionnaire was circulated to people from all age groups, covering rural and metro cities, students, homemakers, and working professionals. Participants were allowed to respond from March to April 2024.

2.2. Survey Design

For each statement within the questionnaire concerning Household Food Wastage Reduction, participants were required to select their response on a 7-point Likert scale. This scale ranged from 1 (Strongly Disagree) to 7 (Strongly Agree), with intermediate points defined as 2 (Disagree), 3 (Somewhat Disagree), 4 (Neutral), 5 (Somewhat Agree), and 6 (Agree) (Van Herpen et al., 2019).

There were no right or wrong answers, and the answers given were kept completely anonymous. The final section consisted of demographic information (e.g., gender, age, family status, marital status, occupation, area of residence, educational level, annual household income, etc.). Data was collected through an online self-administered survey, wherein the questionnaire was shared with the participants using various social media platforms. The data was collected across India, assuming that the respondents provided accurate responses to the survey questions.

2.3. Data Analysis using Machine Learning

The data analysis using machine learning classifiers was carried out using Python. The pipeline employed in the study uses Logistic Regression (LR), Naive Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF) classifiers.

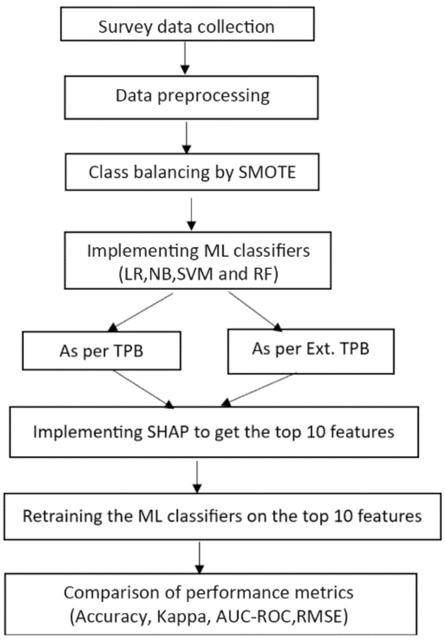

Figure 1. Proposed pipeline.

To address class imbalance, improve interpretability, and optimize predictive performance, the pipeline as depicted in Figure 1 was applied:

1. Data preprocessing: The missing values were replaced by the median values of respective columns.
2. Class balancing with Synthetic Minority Over-sampling Technique (SMOTE):

The study utilized five classes to categorize food wastage: Less or nothing, half a portion, about one portion, 2-3 portions, and more than 3 portions. Given the imbalanced nature of these classes in survey data, SMOTE was applied to the training dataset to achieve balance (Blagus & Lusa, 2013). This crucial step helps mitigate model biases and facilitates more effective learning.

3. Initial Model Training:

After class balancing, each classifier LR, NB, SVM, and RF was initially trained on the class-balanced data after applying SMOTE.

- LR - This is a classification machine learning algorithm. It is used to predict discrete values based on independent variables. It performs linear regression for classification problems (Subasi, 2020) according to Equation 1, where x is the input variable and y is the predicted value.

$$y = 1 / 1 + e^{-x} \qquad (1)$$

- NB - NB is based on Bayes' Theorem (Friedman, Geiger, & Goldszmidt, 1997), assuming that each input variable is independent of any other feature. This method is used for calculating conditional probabilities, i.e., the probability of the occurrence of an event assuming that another event has occurred, as given in Equation 2.

$$P(A|B) = P(B|A) * P(A) / P(B) \qquad (2)$$

- SVM - SVM represents a category of supervised learning algorithms that can be utilized for either classification or regression purposes. The core concept behind SVM is to identify a hyperplane that optimally distinguishes the various classes present in the training data. This is achieved by locating the hyperplane that maintains the largest margin, defined as the distance from the hyperplane to the nearest data points of each class. Once the hyperplane is established, new data can be classified by assessing which side of the hyperplane it lies on. SVMs are especially beneficial when the dataset contains numerous features and/or features a clear margin for separation among the data points.
- RF - RF is a well-known supervised learning method, suited for classification and regression both. It uses several decision trees on different dataset subsets and averages them to increase the dataset's predictive accuracy. Rather than depending on a single decision tree, the RF predicts the outcome based on the majority vote of predictions from each tree. RF is capable of handling datasets with high dimensionality. It keeps the overfitting problem at bay and improves the model's accuracy.

4. Hyperparameter Tuning and Cross-Validation

In order to maximize performance, the hyperparameters of each classifier were fine-tuned and k-fold cross-validation (k=10) was applied. The 10-fold cross-validation was applied to all the models for training and validation to ensure robustness and generalization. The dataset was partitioned into 10 equal-sized subparts. The models are trained on some subparts out of 10 and tested on the remaining subparts. This step helps to avoid overfitting of classifiers.

5. Feature selection using SHapley Additive exPlanations (SHAP)

In our work, we employed SHAP (Holzinger, Saranti, Molnar, Biecek, & Samek, 2020), one of the techniques of explainable AI. Explainable AI helps to interpret feature importance. The importance of the features of the Ext. TPB was calculated by computing SHAP values on the trained models and these values were represented as summary plots. We computed the top 10 features out of the available 45 features.

6. Retraining the Models

The machine learning classifiers were retrained on the balanced dataset retuned by SMOTE. The top 10 features were selected and hyperparameters were tuned again.

7. Model Evaluation

The machine classifiers were evaluated by various evaluation metrics including accuracy, root mean square error (RMSE) (Chicco, Warrens, & Jurman, 2021), kappa statistics (McHugh, 2012), and area under the receiver operating characteristic curve (AUC-ROC) (Van Calster et al., 2008) providing a comprehensive measure of classification effectiveness.

- Kappa statistics is used to measure the level of agreement between two raters. It varies from 0 to 1, with 0 indicating no agreement and 1 being a perfect agreement.
- AUC-ROC value is a graphical representation of the performance of ML classifiers. The higher the AUC-ROC value, the better the performance of the model.
- The RMSE is the root mean square error, which indicates how accurate the predicted values given by an ML classifier are. A good ML model will have a lower RMSE value.

The performance of the classifiers LR, NB, SVM, and RF was compared to predict food wastage as per TPB and Ext. TPB. A comparison of Kappa statistics, AUC-ROC, and RMSE values was used as the selection criteria for determining the best model among these.

3. RESULT AND DISCUSSIONS

3.1. Demographic Analysis

The descriptive analysis of demographic variables revealed several key characteristics of the participants. The majority were female (73%), never married (54.5%), and possessed an undergraduate education (58.7%). Furthermore, 76% resided in urban areas, and 25% reported an annual household income of one lakh INR or more. Regarding household composition, 30% of participants lived in households with four members, and 34.6% had one child.

Significant differences in food wastage patterns were observed across certain demographic groups. Male participants reported significantly different monthly food wastage compared to females (p=0.05). Given that females constituted the majority of respondents (73%) and typically manage meal planning and cooking, the average food wastage observed was notably low at 18.50%. Individuals who were never married wasted significantly less food than their married counterparts (p=0.05). Additionally, participants over 80 years old wasted significantly more food than those aged 41-60 years (p=0.05).

Conversely, no significant differences in food wastage were found based on residents' location, educational qualification, or household income (Table 1).

Table 1. Demographic characteristics of the respondents.

Demographic characteristic	Population distribution
Gender	Female - 249 (73%)
	Male - 89 (26.1%)
	Prefer not to say - 3 (0.9%)
Marital status	Married-104 (30.5%)
	Never Married-186 (54.5%)
	Widowed - 1 (0.3%)
	Divorced - 2 (0.6%)
	Living with partner - 1 (0.3%)
	Other - 46 (13.5 %)
	Prefer not to say - 1 (0.3%)
Household size	1 member - 53 (15.6%)
	2 members - 10 (2.9%)
	3 members - 84 (24.5%)

Demographic characteristic	Population distribution
	4 members - 103 (30.4%)
	5 members - 47 (13.9%)
	6 or more members - 42 (12.4%)
	Prefer not to say - 2 (0.6%)
Education	Higher School - 37 (10.9%)
	Undergraduate - 200 (58.7%)
	Postgraduate - 66 (19.4%)
	Diploma - 4 (1.2%)
	Doctorate - 33 (9.7%)
	Prefer not to say - 1 (0.3%)
Household location	Urban - 262 (76.8%)
	Rural - 79 (23.2%)
Number of children in household	None - 109 (32%)
	1 Child - 118 (34.6%)
	2 children - 81 (23.8%)
	3 children - 17 (5%)
	4 children - 11 (3.2%)
	5 children - 2 (0.6%)
	6 and more - 3 (0.9%)
Household income in INR (monthly)	(Rs.0 – Rs.19,999) 40 – 10.7%
	(Rs.20,000 – Rs.39,999) 61 – 21.4%
	(Rs.40,000 – Rs.59,999) 53 – 17.9%
	(Rs.60,000 – Rs.79,999) 45 – 14.3%
	(Rs.80,000 – Rs.99,999) 33 – 7.1%
	(Rs.1 Lakh and above) 97 – 25%
	Prefer not to say 12 – 3.6%

The survey revealed strong agreement among respondents on several key aspects of food wastage. Participants generally reported producing little to no waste (Mean = 6.33, Std. Dev. = 1.02), actively trying to repurpose food (Mean = 6.24, Std. Dev. = 0.89), and holding a strong belief that wasting food is immoral given global hunger (Mean = 6.50, Std. Dev. = 0.90).

Furthermore, respondents expressed significant discomfort with food wastage, indicating it disturbs them (Mean = 6.52, Std. Dev. = 0.78) and that they perceive it as a waste of money. The act of discarding food also elicited feelings of guilt (Mean = 6.45, Std. Dev. = 0.87). Many participants attributed their aversion to food waste to their upbringing, having been raised to believe in not wasting food (Mean = 6.27, Std. Dev. = 0.88), and linked food waste reduction to personal improvement (Mean = 6.07, Std. Dev. = 1.14).

A sense of duty and seriousness regarding food waste was also evident. Respondents largely viewed reducing food wastage as a duty (Mean = 6.10, Std. Dev. = 1.50) and a genuine and serious concern (Mean = 6.26, Std. Dev. = 0.91). They acknowledged the environmental implications of food wastage (Mean = 6.12, Std. Dev. = 0.98), recognizing the need for effective food waste management (Mean = 6.37, Std. Dev. = 0.86) due to its impact on food shortages in other regions (Mean = 6.03, Std. Dev. = 1.14). There was a collective agreement that every individual can contribute to global food waste reduction (Mean = 6.22, Std. Dev. = 1.25).

Statistical analysis further corroborated these perceptions, indicating that the additional attributes of self-responsibility, eco-consciousness, and food and home management play a significant role in predicting food wastage behavior.

Most respondents reported discarding less than one portion of purchased food or none, as shown by the survey results. Table 2 provides a comprehensive breakdown of the wastage observed across different food categories and their respective portion sizes.

Table 2. Percentage of food wasted in different food groups.

Categories of food	Percentage of food wasted				
	Less or nothing	Half portion	About 1 portion	2-3 portions	More than 3 portions
Fruits & vegetables	70.90%	14.10%	8.80%	4.70%	1.50%
Rice & breads	76.50%	14.70%	5.30%	2.90%	0.60%
Pulses	80.00%	11.80%	5.30%	1.80%	1.20%
Dairy products	87.40%	6.80%	3.20%	1.80%	0.90%
Non-vegetarian food	88.80%	7.60%	1.80%	0.60%	1.20%
Bakery food	85.60%	8.80%	3.80%	1.20%	0.60%

As revealed in Table 2, fruits & vegetables have the highest wastage. Although 70.9% of respondents reported minimal to no waste for these, 14.1% waste at least half a portion, and 8.8% discard approximately 1 portion. Since fruits and vegetables are highly perishable, they are wasted more than other food items. For rice and bread, wastage is notably minimal, with 76.5% of respondents reporting little to no waste. However, 14.7% still waste at least half a portion. This often occurs because these foods are commonly over-prepared and not always stored properly.

Pulses (80%), dairy, and non-vegetarian food (88.8%) show the highest percentages of people reporting "Less or Nothing" wasted. This minimal waste for dairy and non-vegetarian items is likely due to their relatively high cost, which encourages careful consumption. Pulses, being non-perishable, also contribute less frequently to waste. While bakery food wastage is considerable, it remains lower than that of fruits and rice. Specifically, 85.6% of respondents report minimal bakery waste, though 8.8% still discard at least half a portion. The generally low wastage of bakery products can be attributed to their tendency to be purchased in smaller quantities and consumed immediately.
Overall, the data presented in Figure 2 illustrates that food wastage across various food groups predominantly falls into the "Less or Nothing" category.

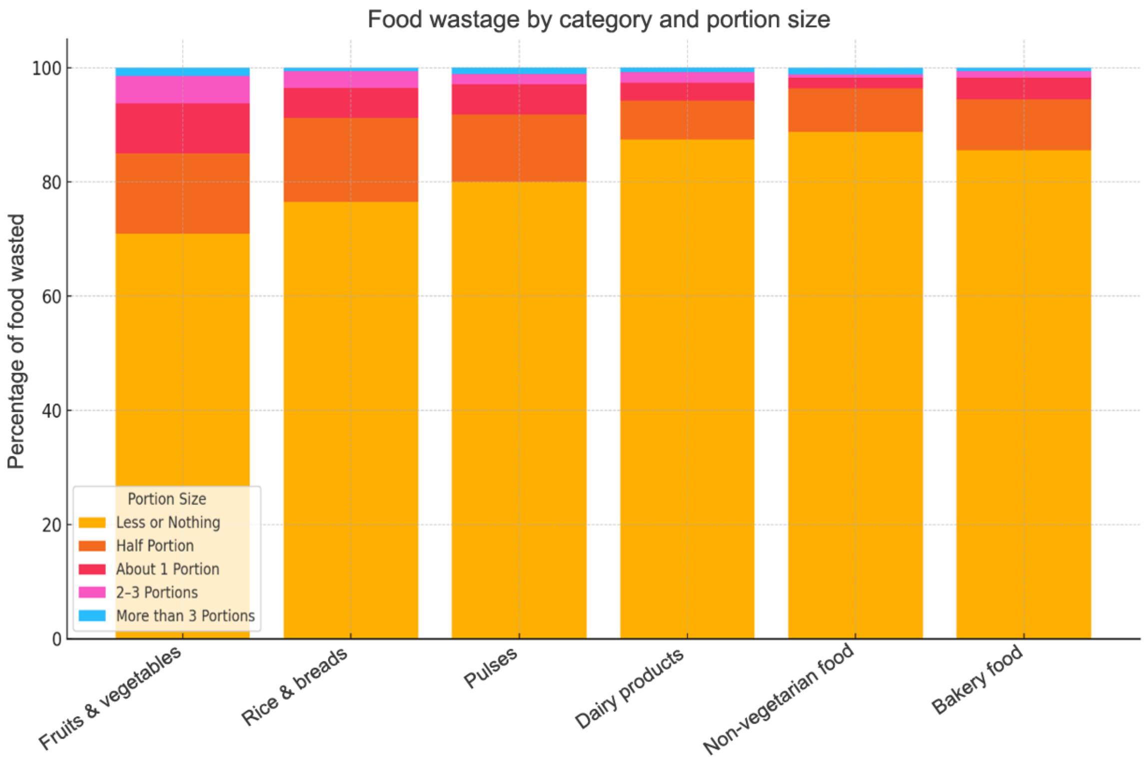

Figure 2. Graph between food category and wastage percentage.

3.2. Machine Learning Classifiers

The study of food wastage from an Indian perspective has been conducted using supervised machine learning classifiers. The dataset was compiled from a survey comprising 341 responses and 45 factors influencing household food wastage. The machine learning algorithms were evaluated through k-fold cross-validation. This method was chosen because the dataset size was small (341 responses), and therefore, a train-test split was not employed. After the k-fold validation, the machine learning algorithms were used to predict food wastage.

Initially, Logistic Regression (LR), Naive Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF) algorithms were applied to the traditional TPB attributes (Intentions, Personal Attributes, Financial Attributes, Environmental Attributes, Food Safety Attributes, Perceived Behavioral Control, Subjective Norms) from the survey data. Subsequently, these same classifiers were run on the Extended TPB (Ext. TPB) attributes, which additionally included Moral Norm, Personal Norm, Good Provider Identity, Household Planning Habits, Environmental Knowledge and Awareness, Awareness of Consequences, and Ascription of Responsibility.

A comparison of the algorithms' performance revealed that the prediction accuracy of LR, NB, SVM, and RF was lower when applied to the traditional TPB attributes compared to the Ext. TPB, where the additional attributes significantly improved predictive power. These results are summarized in Table 3.

Table 3. Performance matrix of different machine learning algorithms.

Machine learning classifiers	Correctly classified instances		Kappa statistics		Root mean squared error (RMSE)		Weighted AUC-ROC	
	Ext. TPB	TPB	Ext. TPB	TPB	Ext. TPB	TPB	Ext. TPB	TPB
LR	58%	46.90%	0.4036	0.1996	0.3013	0.3287	0.812	0.72
NB	50.14%	41.29%	0.3125	0.18	0.3552	0.3609	0.733	0.688
SVM	**84.67%**	**81.71%**	**0.862**	**0.7709**	**0.193**	**0.222**	**0.876**	**0.673**
RF	80%	78.65%	0.645	0.523	0.308	0.421	0.824	0.762

Note: Best performance is highlighted in bold.

It has been observed that the prediction accuracy in terms of correctly classified instances of SVM is 84.67% for Ext. TPB, whereas it is 81.71% for TPB. The kappa statistic for SVM is 0.862 for Ext. TPB, which is higher than 0.7709 for TPB. The RMSE of SVM for Ext. TPB is 0.193, which is lower than for TPB (0.222). Also, the weighted AUC-ROC for Ext. TPB is 0.876, higher than that for TPB (0.673). The evaluation metrics of TPB and Ext. TPB are graphically represented in Figure 3.

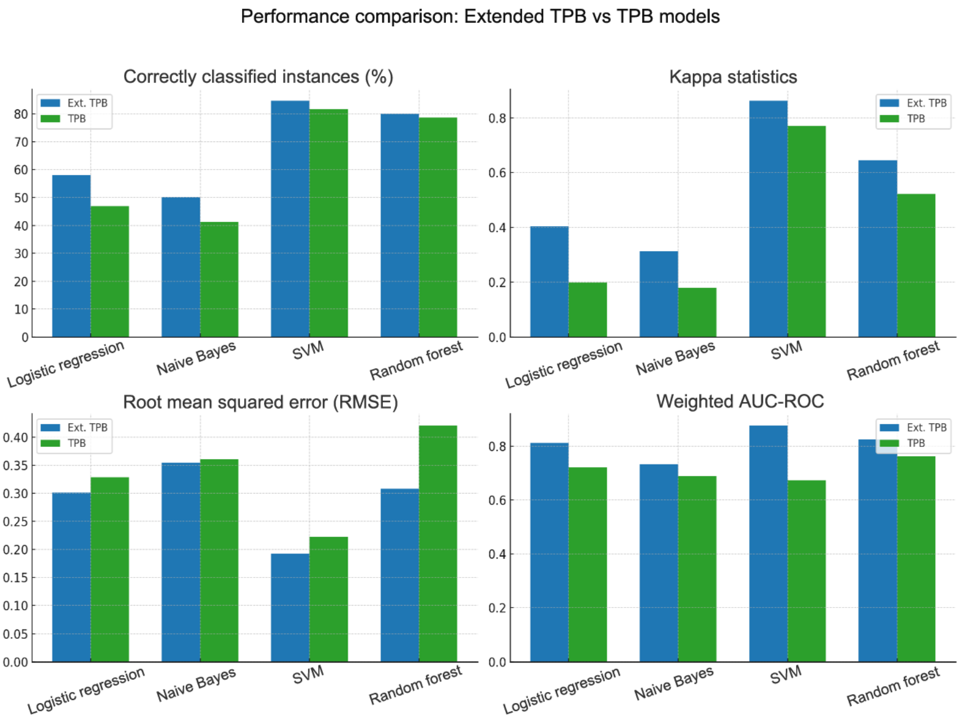

Figure 3. Comparison of evaluation metrics for TPB and Ext. TPB using machine learning classifiers.

Expanding upon established literature, which has augmented the Theory of Planned Behavior (TPB) with constructs such as guilt, behavioral control, injunctive norms, and knowledge-related factors to better understand food wastage (Aydin & Aydin, 2022; Chen, 2023), our study achieved a 12% improvement in prediction accuracy over the traditional TPB. This outcome supports prior research, such as Attiq, Habib, Kaur, Hasni, and Dhir (2021), who identified cognitive factors (e.g., awareness of consequences, environmental knowledge) as positively linked to food waste reduction behavior. A key distinction of our work is its focus on the Indian context, where food consumption is uniquely shaped by cultural practices and socio-economic diversity.

To interpret classifier predictions, the SHAP technique of XAI was applied to identify the top contributing attributes from the 45 factors influencing food wastage. Prior to SHAP application, section-wise averages of attribute values were calculated. Figure 4 illustrates the mean SHAP values, revealing perceived behavioral control as the most significant attribute. This is closely followed by household planning habits, subjective norms, personal norms, intention, personal attitude, financial attitude, good provider identity, environmental awareness, and ascription of responsibility. Notably, household planning habits, an attribute specifically included in the Ext. TPB plays a pivotal role in predicting food wastage. These SHAP results reinforce the comprehensiveness of the Ext. TPB and demonstrate the utility of XAI techniques in uncovering interpretable attributes related to food wastage.

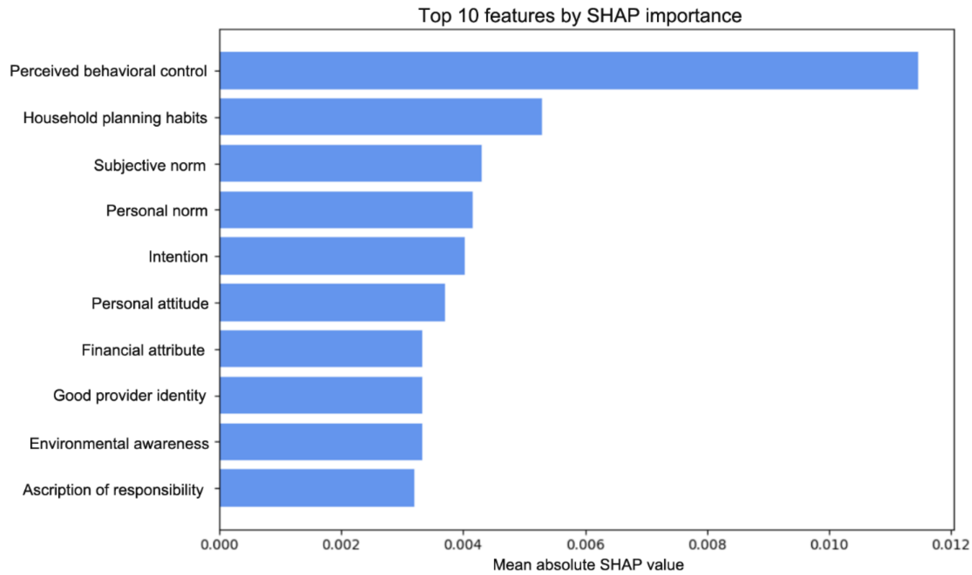

Figure 4. Top 10 attributes of food wastage given by SHAP.

This study demonstrates the need to go beyond the traditional TPB to understand something as complex as household food wastage. The machine learning models performed noticeably better with the inclusion of additional behavioral and contextual factors. This suggests that people's food waste behavior is influenced by more than just intentions or perceived control it's shaped by a broader mix of financial, environmental, and personal factors. A broader perception is required to obtain a more realistic and effective understanding of food wastage.

This is particularly valuable in behavioral research, where transparency is essential for translating model outputs into actionable insights. The SHAP analysis revealed that food management practices, perceived behavioral control, subjective norms, and personal norms had the highest positive SHAP values across most predictive models. Conversely, variables such as attitude and environmental concern exhibited lower SHAP values, indicating less predictive influence in this context.

The limitation of the work is the survey data which does not cover all the Indian states. Although the online questionnaire was circulated pan-India, it still lacks representation from some states. Hence, it does not fully cover the socio-cultural, economic, and environmental diversity across the country. The regional factors like local food and cuisines, cultures, and customs influence household food waste behavior across different states. Future work should expand data collection from all Indian states, spanning North, South, East, and West; this will help in generalizing the findings and incorporate regional relevance.

The findings of this study can help policymakers and food agencies to create awareness campaigns to prevent food waste. The awareness programs should emphasize the importance of household planning, which is one of the attributes responsible for food wastage in households as given by SHAP. Training sessions should be designed to educate people about meal planning, reusing food leftovers, and sensitizing them about food scarcity across the globe. Mobile apps and SMS-based reminders can be designed to help people understand their food requirements and plan food purchases and meal preparation accordingly. The interpretation given by SHAP can be used for data-driven policymaking.

4. CONCLUSION

In the present work, it has been observed that in Indian households, most respondents are sensitive to the global food shortage and do not believe in wastage; hence, they try to reuse excess food in other ways. Respondents believe that food waste reduction is a duty and an approach towards sustainability and eco-friendly practices.

The present work provides an integrative approach to predicting food wastage in Indian households by proposing an extended Theory of Planned Behavior (Ext. TPB), supported by machine learning algorithms and explainable AI techniques. Machine learning models, particularly Support Vector Machines (SVM), outperformed other classifiers on Ext. TPB. Moreover, the incorporation of SHAP (SHapley Additive exPlanations) offers insights into the relative importance of each attribute, helping to bridge the gap between predictive performance and interpretability. The attributes of food management practices, perceived behavioral control, subjective norms, and personal norms are among the most influential factors in reducing food waste. These findings emphasize the need for good household planning skills and not just attitudinal change.

These insights can be helpful to policymakers and environmentalists in targeting campaigns for household planning and individual accountability for reducing food waste. Future studies can consider broader geographic and demographic samples to enhance predictive performance. Additionally, incorporating ensemble learning and other XAI frameworks can be explored to improve prediction accuracy and provide more behavioral insights.

Funding: This study received no specific financial support.
Institutional Review Board Statement: The college does not have an Institutional Review Board currently. The research was undertaken adhering to the ethical principles of informed consent, voluntary participation and anonymity of responses. The questionnaire did not collect any medical or sensitive data about the participants. The authors had obtained a No Objection certificate no. MT/PO/2025/770 dated 8th September, 2025 from the college.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Design and conceptualization, Pooja Baweja (PB), Haritma Chopra (HC); data collection, circulation, result interpretation, Sakshi Suman (SS), Nisha Poddar (NP); machine learning analysis, Veena Ghuriani (VG). All authors have read and agreed to the published version of the manuscript.

REFERENCES

Ajzen, I. (2012). The theory of planned behavior. In P. A. Van Lange, A. W. Kruglanski, E. T. Higgins (Eds.), The theory of planned behavior. In (Vol. 1, pp. 438-459): SAGE Publications Ltd. https://doi.org/10.4135/9781446249215.n22.

Anggraeni, M. C., Silaban, C. A., Anggreainy, M. S., & Cahyadi, E. (2021). *Role of artificial intelligence in the management of food waste*. Paper presented at the 2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS), IPOH, Malaysia. https://doi.org/10.1109/AiDAS53897.2021.9574167

Attiq, S., Habib, M. D., Kaur, P., Hasni, M. J. S., & Dhir, A. (2021). Drivers of food waste reduction behaviour in the household context. *Food Quality and Preference*, *94*, 104300. https://doi.org/10.1016/j.foodqual.2021.104300

Aydin, H., & Aydin, C. (2022). Investigating consumers' food waste behaviors: An extended theory of planned behavior of Turkey sample. *Cleaner Waste Systems, 3*, 100036. https://doi.org/10.1016/j.clwas.2022.100036

Bamberg, S., & Möser, G. (2007). Twenty years after Hines, Hungerford, and Tomera: A new meta-analysis of psycho-social determinants of pro-environmental behaviour. *Journal of Environmental Psychology, 27*(1), 14-25. https://doi.org/10.1016/j.jenvp.2006.12.002

Bhatia, A., & Sharma, S. (2023). Identifying determinants of household food waste behavior in urban India. *Cleaner Waste Systems, 6*, 100105. https://doi.org/10.1016/j.clwas.2023.100105

Bhatia, L., Jha, H., Sarkar, T., & Sarangi, P. K. (2023). Food waste utilization for reducing carbon footprints towards sustainable and cleaner environment: A review. *International Journal of Environmental Research and Public Health, 20*(3), 2318. https://doi.org/10.3390/ijerph20032318

Blagus, R., & Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics, 14*(1), 106. https://doi.org/10.1186/1471-2105-14-106

Chen, M.-F. (2023). Integrating the extended theory of planned behavior model and the food-related routines to explain food waste behavior. *British Food Journal, 125*(2), 645-661. https://doi.org/10.1108/BFJ-07-2021-0788

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, e623. https://doi.org/10.7717/peerj-cs.623

Chowdary, K. S., Praneetha, L. K., Holika, S., Priya, D. B., Kumar, S. V. P., & Kishore, K. V. K. (2024). *Prediction of Food Wastage using xg boost*. Paper presented at the 2024 8th International Conference on Inventive Systems and Control (ICISC), Coimbatore, India. https://doi.org/10.1109/ICISC62624.2024.00059

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Fielding, K. S., McDonald, R., & Louis, W. R. (2008). Theory of planned behaviour, identity and intentions to engage in environmental activism. *Journal of Environmental Psychology, 28*(4), 318-326. https://doi.org/10.1016/j.jenvp.2008.03.003

Food and Agriculture Organization of the United Nations (FAO). (1981). *Food loss prevention in perishable crops*. Rome, Italy: FAO.

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine Learning, 29*(2), 131-163. https://doi.org/10.1023/A:1007465528199

Ghinea, C., & Ghiuta, O.-A. (2019). Household food waste generation: Young consumers behaviour, habits and attitudes. *International Journal of Environmental Science and Technology, 16*(5), 2185-2200. https://doi.org/10.1007/s13762-018-1853-1

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., . . . Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. *Science, 327*(5967), 812-818. https://doi.org/10.1126/science.1185383

Graham-Rowe, E., Jessop, D. C., & Sparks, P. (2015). Predicting household food waste reduction using an extended theory of planned behaviour. *Resources, Conservation and Recycling, 101*, 194-202. https://doi.org/10.1016/j.resconrec.2015.05.020

Grainger, M. J., Aramyan, L., Piras, S., Quested, T. E., Righi, S., Setti, M., . . . Stewart, G. B. (2018). Model selection and averaging in the assessment of the drivers of household food waste to reduce the probability of false positives. *PloS One, 13*(2), e0192075. https://doi.org/10.1371/journal.pone.0192075

Grandhi, B., & Singh, J. A. (2016). What a waste! A study of food wastage behavior in Singapore. *Journal of Food Products Marketing, 22*(4), 471-485. https://doi.org/10.1080/10454446.2014.885863

Griffin, M., Sobal, J., & Lyson, T. A. (2009). An analysis of a community food waste stream. *Agriculture and Human Values, 26*(1), 67-81. https://doi.org/10.1007/s10460-008-9178-1

Holzinger, A., Saranti, A., Molnar, C., Biecek, P., & Samek, W. (2020). *Explainable AI methods-a brief overview*. Paper presented at the International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers.

Martindale, W. (2010). Waste: Uncovering the global food scandal. *International Journal of Sustainable Engineering*, *3*(2), 144-145 https://doi.org/10.1080/19397030903573317

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276-282. https://doi.org/10.11613/BM.2012.031

Mirabella, N., Castellani, V., & Sala, S. (2014). Current options for the valorization of food manufacturing waste: A review. *Journal of Cleaner Production*, *65*, 28-41. https://doi.org/10.1016/j.jclepro.2013.10.051

Mourad, M. (2016). Recycling, recovering and preventing "food waste": Competing solutions for food systems sustainability in the United States and France. *Journal of Cleaner Production*, *126*, 461-477. https://doi.org/10.1016/j.jclepro.2016.03.084

Ozcil, I. E. (2024). Reducing food waste using machine learning models: Forecasting and optimization approaches. *Open Access Journal of Data Science and Artificial Intelligence*, *2*(1), 1-3. https://doi.org/10.23880/oajda-16000136

Panda, S. K., & Dwivedi, M. (2019). *Minimizing food wastage using machine learning: A novel approach*. Paper presented at the Smart Intelligent Computing and Applications: Proceedings of the Third International Conference on Smart Computing and Informatics.

Pandey, A. (2021). Food wastage: Causes, impacts and solutions. *Science Heritage Journal*, *5*(1), 17-20. http://doi.org/10.26480/gws.01.2021.17.20

Parfitt, J., Barthel, M., & Macnaughton, S. (2010). Food waste within food supply chains: Quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1554), 3065-3081. https://doi.org/10.1098/rstb.2010.0126

Quested, T. E., Marsh, E., Stunell, D., & Parry, A. D. (2013). Spaghetti soup: The complex world of food waste behaviours. *Resources, Conservation and Recycling*, *79*, 43-51. https://doi.org/10.1016/j.resconrec.2013.04.011

Racz, A., Vasiljev Marchesi, V., & Crnković, I. (2018). Economical, environmental and ethical impact of food wastage in hospitality and other global industries. *Jahr: Europski Časopis za Bioetiku*, *9*(1), 25-42. https://doi.org/10.21860/j.9.1.1

Russell, S. V., Young, C. W., Unsworth, K. L., & Robinson, C. (2017). Bringing habits and emotions into food waste behaviour. *Resources, Conservation and Recycling*, *125*, 107-114. https://doi.org/10.1016/j.resconrec.2017.06.007

Sama, L., Makkar, A., Prokshitha, P., Dhaloria, D., & Sharma, B. K. (2021). *AI based management of food wastage*. Paper presented at the Proceedings of the International Symposium of Intelligent Control.

Schanes, K., Dobernig, K., & Gözet, B. (2018). Food waste matters-A systematic review of household food waste practices and their policy implications. *Journal of Cleaner Production*, *182*, 978-991. https://doi.org/10.1016/j.jclepro.2018.02.030

Schlegelmilch, B. B., Bohlen, G. M., & Diamantopoulos, A. (1996). The link between green purchasing decisions and measures of environmental consciousness. *European Journal of Marketing*, *30*(5), 35-55. https://doi.org/10.1108/03090569610118740

Stefan, V., Van Herpen, E., Tudoran, A. A., & Lähteenmäki, L. (2013). Avoiding food waste by Romanian consumers: The importance of planning and shopping routines. *Food Quality and Preference*, *28*(1), 375-381. https://doi.org/10.1016/j.foodqual.2012.11.001

Subasi, A. (2020). *Practical machine learning for data analysis using Python*. Boston, MA: Academic Press.

Van Calster, B., Van Belle, V., Condous, G., Bourne, T., Timmerman, D., & Van Huffel, S. (2008). *Multi-class AUC metrics and weighted alternatives*. Paper presented at the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence). https://doi.org/10.1109/IJCNN.2008.4633979.

Van Herpen, E., Van Geffen, L., Nijenhuis-De Vries, M., Holthuysen, N., Van Der Lans, I., & Quested, T. (2019). A validated survey to measure household food waste. *MethodsX*, *6*, 2767-2775. https://doi.org/10.1016/j.mex.2019.10.029

Williams, H., & Wikström, F. (2011). Environmental impact of packaging and food losses in a life cycle perspective: A comparative analysis of five food items. *Journal of Cleaner Production*, *19*(1), 43-48. https://doi.org/10.1016/j.jclepro.2010.08.008