Journal of Posthumanism
2025
Volume: 5, No: 6, pp. 571–588
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i6.2073

Towards a Sustainable Retail Food Chain: Artificial Intelligence Driven Dynamic Pricing and Promotions for Food Waste Reduction

Nawaf Abdullah[1], Abdul K Shaikh[2], Ahlam Almusharraf[3]

Abstract

Food wastage represents a substantial environmental, economic, and social challenge, contributing to resource reduction. In response to these issues, this study investigates the potential of AI-driven dynamic pricing and promotions to decrease food waste within the supply chain of food retail. By leveraging machine learning models, namely, Gradient Boosting, SVR, Decision Tree, KNN, Random Forest, and Neural Networks, the research seeks to optimize pricing strategies for perishable goods based on factors such as shelf life, inventory levels, and demand fluctuations. The model's performance was evaluated using MAE, R^2 scores and RMSE. Gradient Boosting emerged as the most effective model, achieving the lowest error rates (MAE: 0.113, RMSE: 0.536) and highest R^2 (0.828), indicating strong predictive power and accuracy in price adjustment. The results demonstrate that AI-driven dynamic pricing can accurately adjust prices in real-time, encouraging the sale of near-expiration items and thereby reducing food wastage. Future research may explore reinforcement learning approaches and expanded datasets to further refine pricing accuracy and expand the model's applicability across different retail contexts.

Keywords: *Artificial Intelligence, Machine Learning, Dynamic Pricing, Food Waste Reduction, Sustainable Retail Food Chain*

Mathematics Subject Classification: 68T01

Introduction

Prophet Muhammad, peace be upon him, said over 1400 years ago, "Do not waste water even if you were at a running stream" (Ibn Mājah et al., 2007). However, it is estimated that around 30 percent of all food produced for human consumption and intended for consumers is lost or wasted, which amounts to approximately 1.3 billion tons yearly. This goes to show the extent of an inefficiency that unfolds at several levels along the long and intricate global food supply chain throughout its entire lifecycle, from production in the field to consumption by the end user. It poses grave economic, environmental and social challenges since this system is far from efficient. The inefficiency in the system has to be tackled (Gustavsson et al., 2011).

Wastage occurs at different levels in a food chain. At the lower end of the scale, large losses occur at the early- and middle-stages of the supply chain in lower-income nations due to a lack of technological and infrastructural support (Parfitt et al., 2010). Whereas in higher-income regions, the major portion of wastage is realized at the retail and consumer levels, due to weak quality standards, inefficient practices and reckless consumer behaviours (Quested et al., 2013).

[1] Department of Information Systems Sultan Qaboos University Alkhod PC 123 Muscat Oman, Email: s139495@student.squ.edu.om
[2] Department of Information Systems Sultan Qaboos University, Alkhod PC 123 Muscat Oman, Email: shaikh@squ.edu.om
[3] Department of Business Administration, College of Business and Administration, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia, Email: aialmusharraf@pnu.edu.sa

Due to the international variations of the distribution of causes for wastage in a food chain, it is evident that a single notion of how to reduce this issue is inadequate for all nations. Each nation will have supply chain-specific or dynamic solutions that need to be adopted and applied depending on context.

In spite of exhausting attempts through the years to change past practices of wasting food, traditional interventions have struggled to properly adjust to the nuances of modern food supply chains and consumer markets. Since linear food supply and demand do not adhere to rigid patterns any longer, we require radical innovations that are dynamic and adaptable to such supply and demand fluidities. Dynamic pricing and promotions through AI is a promising new technology transformation that is changing existing practices (Kumar et al., 2017). The growing world population and the rising future food demand further compound the need to tackle food wastage.

The other big thing in modern commerce is the adoption of AI in dynamic pricing. AI's implementation is essentially tied to the mechanics of revenue management which is an increasingly important business model across all industries. Dynamic pricing is a form of revenue management that necessitates fluctuating prices in time as well as markets and internal variables like demand and stocks. AI is programmed to rapidly compute massive amounts of data, predicting demand and changing prices near real-time for companies (Ramezani et al., 2011). This is particularly important for those providing products that depreciate in value over time, especially in the case of travel, hospitality, and more recently, retail environments. Models that hold AI applications for dynamic pricing are based on sophisticated algorithms, including machine learning, that allow firms and companies to predict how consumers will react to various pricing methods. Such models can incorporate many external influences such as changes in the weather, economic conditions or cultural events, and previously they were hard to incorporate rapidly in pricing methods (Dasgupta & Hashimoto, 2004).

Aside from the dynamic pricing that many of us have already encountered, AI can also be used to make personalized offers and discounts to consumer profiles, purchasing histories and predicted future behaviors that will help to engage customers, increase perceived value to customers and thereby increase purchases. This can also be used to help increase consumer loyalty (DiMicco et al., 2001) the ability for AI to analyses market conditions in real time and respond accordingly gives a distinct advantage of competency for businesses in terms of their ability to be proactive with pricing This in turn can increase profit margins and potentially capture more market share.

The objective of this research is to identify and deploy AI-based tools to minimize food wastage along the global food supply chain through a series of research objectives, starting from an identification of challenges in the food supply chain that lead to food wastage, and designing and implementing an AI-based dynamic pricing and promotions model which incorporates variables such as demand, shelf life and the level of inventory. It also investigates how dynamic pricing influences consumer decision making such as purchasing more items closer to their expiry date, by observing the changing of product prices in the market due to various variables to affect consumers decision-making when buying products in stores. Finally, how well an AI based dynamic pricing and promotions algorithm will reduce food wastage along the supply chain can be evaluated and measured in order to understand the real efficiency of AI based algorithms to reduce food wastage along the supply chain and increase the overall sustainability of the food supply chain. These research objectives aim to contribute to the advancement of sustainable supply chain management through relevant evidence-based insights and practical interventions supported through the use of emerging technologies such as AI.

Literature Review

There are a number of problems contributing to inefficiency in the food supply chain, leading to large-scale wastage, especially in developing countries. Research summarizes a number of factors that contribute to this inefficiency. Poor technological infrastructure is a key issue driving high losses. for instance, one study of food wastage in India noted the absence of optimal food cold storage and transportation facilities was a major contributor of food wastage like fruit and vegetables spoil during the transportation, before reaching consumers (Chauhan, 2020). Another study likewise identified 16 different causes for food wastage across the supply chain, however the lack of effective scientific harvesting process and the unnecessary number of middlemen in the supply chain significantly increase food losses by these means that are identified as root causes in need of immediate redress to enhance efficiencies in the supply chain (M. & K., 2016). Lack of risk management in the supply chain can also be considered a key factor, as argued by this particular study, and it connects supply chain risk with food wastage, using a model that maps relationships among the different risks identified to argue the potential to reduce food wastage through tackling the identified risks such as non-skilled personnel, and poor IT system implementation (Mithun Ali et al., 2019). Consider some differences between the countries found in the developed world and the developing world; consumer behaviour at retail and consumption stages causes high levels of food wastage in the developed countries, while in the developing countries, most food loss at the first quarter of the supply chain exist due to the inefficiencies in the production, harvesting and in distribution stages and policies which suggest that strategic interventions can be used in these regions as the first phase for intervention for reducing food loss (Porter & Reay, 2016). Also, unlike non-perishable foods, the perishable food supply chain brought bigger wastage compared with other types. This type of food supply chain is facing a number of challenges, including absence of horizontal integration and inefficient pre-harvest management and poor governmental support, which require better infrastructure and regulation to facilitate the supply chain to become more efficient and sustainable in the future (Kumar et al., 2020).

Dynamic pricing is presented in various domains as one of the solutions to curb food wastage. An e-commerce retail research study, based on a field experiment with more than 100 million customers on the largest global retail platform Alibaba, showed that short-term price promotions drive twofold increases in the sales of the promoted products. However, these promotions also induced long-term strategic behaviors, such as increased expectations of future discounts and more aggressive searches for lower prices. These strategic behaviors were then modeled and applied to both promoter and non-promoter sellers, showing the far-reaching effects of dynamic pricing strategies on the aggregate market ecosystem (Zhang et al., 2018). In the retail sector, dynamic pricing has been used productively in revenue management for perishable inventory items. In the study, they examine the use of ‘buy one get one free’ and ‘50 percent off’ promotions and found that the size of the purchase as well as the timing of purchases can result in significant influences on customers’ buying patterns, aligning closely with inventory and revenue management goals (Kim et al., 2016). Price in the airlines is variable according to demand, length of time you reserve, and the seats. We have seen most of the airlines use dynamic pricing in order to make the most amount of money but mostly on a perishable inventory like airline seats. It’s common now to price airline seats, hotel rooms and car rental real-time to accommodate for supply and demand inefficiencies (Elmaghraby & Keskinocak, 2003). In the online sale and service of cars and other automotive products, we can use dynamic pricing algorithms to dynamically adjust the price of vehicles and services depending on demand, available inventory, and competing offers. The study suggests the use of dynamic pricing along with a direct-to-customer business model to help automotive manufacturers to respond to pressures from the market, helping the inventory and pricing (Biller et al., 2005).

Artificial intelligence (AI) is revolutionizing supply chain management through improved decision-making processes, higher efficiency, and reduced operation costs. Numerous studies showcase the diverse applications of AI to manage diverse sides of the supply chain, including procurement, logistics, and inventory management, to benefit organizations. One study investigated beforehand the importance of AI for supply chain financing, highlighting how AI holds the power teds considerably towards a higher level in mitigating risks, detecting fraudulent activities, and improving working capital efficiency. Based on the result of this study, AI can be utilized to analyze huge amounts of data for identifying inefficiency and mitigating risk upfront by informing more transparent and informed strategic decisions of financing activities by companies. This leads to lower-cost financing decisions beneficial for improving organizations' financial performance (Rajagopal et al., 2023). In logistics, the second broad section analyzed the allowance of innovative technologies including AI application benefits to optimize the logistics process. AI can not only assist with automating well-defined workflows but also enhance the decision-making capabilities of logistics planners. Despite the potential to replace humans in the logistics process, AI integration in logistics cannot be more critical to improving decision-making performance. Specific trade-offs between logistics versus AI can be defined: AI can manage a barrel of complex data streams occurring within the whole logistics process that enables interactions at a human-like rate and timing, such as safety and worker's welfare, and operational flexibility (Boute & Udenio, 2023). The third section explored how AI can optimize performance in contemporary industrial supply chains. The study found that AI was capable of improving the accuracy of capacity planning, optimizing the utilization of equipment and resources, and mitigating sudden changes in the demand for goods and services. These practice settings were critical to maintaining productivity, reducing overall costs and optimizing the supply chain while providing extremely high-quality goods and services to the end consumer demand and ensuring safe and secure operating environments (Alomar, 2022).

The current climate of dynamic pricing sees the use of different dynamic pricing AI models to facilitate smart self-adjustments in real-time because it allows one to deliver adaptive pricing which responds to the dynamic supply and demand, and also to consumer behavior, current offerings, and other factors that come into play. A review of some machine learning models for dynamic pricing with different configurations is occasioned as Gradient Boosting Machines (GBM) have been applied widely due to their ability to cope with diverse and complex datasets. They have notable applicability in the e-commerce domain; where a project applied GBM to enable it to determine the optimal pricing approach. Project outcomes demonstrated near-perfect optimized results, through training relying on transactional data falling within a historically identified period. The model was able to derive insights that enabled precision in determining the optimal price points for the seller of the product, all geared towards improving revenue performance. The GBM machine learning aspect delves into feature-engineering and hyper parameter optimization techniques for superior predictive outcomes (El Youbi et al., 2023). Reinforcement Learning (RL) is another dynamic pricing AI model showing considerable promise in instances where decisions must be taken in specific dynamic conditions where uncertainty is involved. It is instructive to note that two algorithms were compared for evaluative purposes: the Deep Q-Network (DQN) and Soft-Actor Critic (SAC). Findings indicate that SAC did perform better than DQN when applied to dynamic online market scenarios. This is because of the RL property where it explores new pricing decision while balancing it against its ability to exploit what is already known of the pricing system (Kastius & Schlosser, 2022). Decision Tree Regressors also hold appeal for their simplicity, and seem effective in modelling nonlinear scenarios. This machine learning model was applied to the real-estate sector to predict values with high accuracy, given the ability of the algorithm to process copious amounts of geographical and market trend data. In certain sectors where pricing is driven by a host of factors, simplicity and interpretability are credible goals, especially if pricing is dependent on a wide array of variables (Kumar et al., 2023). Finally, one can point to the application of Bayesian Learning approaches that take uncertainty seriously in that they were applied to optimize pricing while coping with the variability of consumer behavioral responses to same. This approach is useful when the number of historical records is limited or where market conditions are viewed as highly unpredictable (Han, 2010).

The literature shows that AI-enhanced dynamic pricing models have experienced significant advancements. However, it also highlights that a gap remains for dynamic analysis on demand forecasting, shelf life and inventory optimization. The current approaches all endeavour to determine the price shifts described by the elastic and inelastic parts of general demand, while also relying on consumer segmentation for how different consumer behavior affects the sensitivity to price. Although some attempts acknowledge the critical factor of product perishability, they are sparse and therefore do not demonstrate the true value that real-time data can add to price setting. For example, Herbon focused on the optimal price strategy for a carbonated drink with high consumer sensitivity to product freshness. He proposed a linear model but failed to extend it to utilise real-time data to further refine the price dynamic strategy. Similarly, some are interested in the application of AI for perishable product supply chain management but their models are not explicitly linked to dynamic price adjustments. For packaged food products, Yimenu recently proposed an AI-powered modelling for real-time shelf-life estimation and its prediction (Yimenu et al., 2019). Their model admitted the promising application in dynamic pricing but the topic was off the scope of the study. These studies highlight the crucial gap in the literature on how AI-powered pricing models can facilitate market penetration of perishable products with real-time input from shelf life predictions and enhance operational efficiency and consumer experience.

Research Methodology

This section presents an overview of the datasets, the proposed work, the architecture, and the algorithms used for the AI pricing and promotion model to reduce food wastage.

Dataset

A private dataset was extracted from a supermarket POS system and published in Kaggle. Its food inventory list contains 171 items collected in 2020 and May 2021. It contains 14 columns, 13 features, and a target variable. The features are id, date_collected, retailer_type, retailer_detail, food_type, food_detail, label_type, label_language, label_date, image_id, label_explanation, collection_lat, and collection_long. The target variable is approximate_dollar_value. The value range of "Approximate Dollar Value" is from approximately $0.73 to $26.99, with a mean of $5.50. This indicates a wide range of food item values, potentially reflecting different types of food or portion sizes. The coordinates are concentrated around a specific geographic location (mean collection_lat: 40.6946, mean collection_long: -73.9924) with a small standard deviation in these coordinates confirming the data collection occurred within the limited area of the supermarket. As Table 1 explains, most items are perishable and packaged food. The dataset has 51 missing values in the "Label Language" column and 170 in the "Label Explanation".

Food type	**packaged**		**perishable**		**ready-to-eat**		**shelf stable**	
	Count	**Value $ (AVG)**	**Count**	**Value $ (AVG)**	**Count**	**Value $ (AVG)**	**Count**	**Value $ (AVG)**
bakery/deli	35	7.07						
chain grocer			34	5.99				
coffeeshop	2	3.75						
counter service	4	3.5			5	7.35		
drugstore	8	4.79	33	2.19	11	3.31		
health food grocer			35	6.82			4	11.24

Table 1 Characteristics of the Food

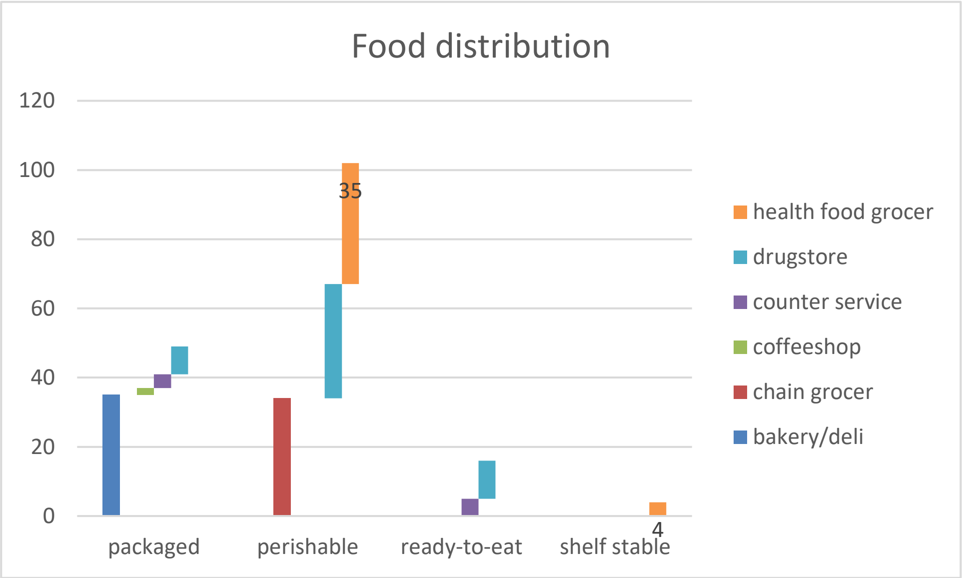

Figure 1 Dataset Distribution Graph

Model Building Process

The methodology used to build the AI dynamic pricing and promotion is presented here. The dataset is first preprocessed using various techniques to prepare it for analysis. Afterwards, the data is split into two sets: a training set for building the AI pricing model, and a test set for assessing its performance. The machine learning models are trained on the created training set and then tested on the test set using various metrics to evaluate their performance. Figure 2 illustrates the architecture of this process.

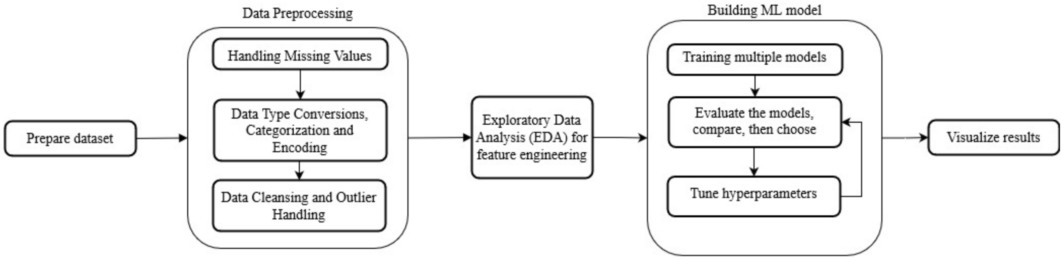

Figure 2 Data Analysis Process

Data Preprocessing

The data was preprocessed with data encoding, missing value imputation, transformation of skewed distributions, weighting the data, feature scaling, feature selection etc. Here's a breakdown of these techniques. As can be seen in figure 2 before, this is the complete preprocessing structure.

Handling Missing Values

For columns with missing values like "label_language" and "label_explanation", we need to decide whether to fill these gaps with default values using statistical methods to attribute them or drop them if they are not critical for our analysis (Iliou et al., 2015). To maintain dataset integrity without introducing bias, we attributed the missing values in 'label_language' and 'label_explanation' with 'Unknown' and 'No explanation', respectively.

Data Type Conversions, Categorization and Encoding

Following standard preprocessing practices, the 'date_collected' and 'label_date' columns were converted from strings to datetime objects, which facilitates more straightforward time-based analyses, ensures numeric fields are correctly formatted, and handles any anomalies or incorrect data types (García et al., 2016). Categorization and encoding of categorical variables like 'retailer_type,' 'food_type,' and 'label_type' are essential for preparing data for machine learning algorithms. Encoding techniques such as one-hot encoding or label encoding are important for transforming categorical data into a format that can be easily processed by machine learning models (Ramírez-Gallego et al., 2017). The choice between these techniques depends on the model requirements and the specific nature of the categorical data.

Data Cleansing and Outlier Handling

Data cleansing means identifying and removing data that is incorrect, corrupted, incorrectly formatted, duplicated, or incomplete from a data set. Data cleansing can include noise filtering and error correction, which is essential to the quality of the data for processing (Famili et al., 1997). We found some invalid data in the column "label_date" which were removed. Also, outliers in 'approximate_dollar_value' were managed using the Interquartile Range (IQR) method, a standard technique for reducing the influence of extreme values on the dataset's overall analysis. The literature supports this method for its effectiveness in normalizing data distributions (Iliou et al., 2015).

Exploratory Data Analysis (EDA)

Initial EDA was performed to gain understandings into the distribution and characteristics of key variables, such as “retailer type” and “food type”. This phase is critical for identifying patterns and potential areas of interest for deeper analysis and understanding data streams in preparation for data mining (Ramírez-Gallego et al., 2017). The following figures visualize the relationships and find some interesting insights. Figure 4 shoes that “Packaged” food type is the most common, followed by “perishable” food items. In contrast, the boxplot in figure 3 reveals the variation in “approximate dollar” values across different retailer types. “Bakery/deli” items have a higher range of values compared to others. Lastly, we can see in figure 5, the count of food waste occurrences by the day of the week shows that certain days like “Monday” and “Thursday” have higher counts, indicating possible patterns in food waste.

Feature Engineering

Feature engineering is a critical step in enhancing the predictive power of machine learning models by creating new features or modifying existing ones (Sreenivas & Srikrishna, 2013). The following features of engineering ideas can be helpful to analyze our dataset. Here are a few feature engineering ideas we can implement based on the existing data:

- Time to Expiry: Calculate the number of days from the date collected to the label date. This could give us insights into whether items closer to their expiry date are more likely to be wasted. Also, it may provide insights into shelf life and its impact on wastage.

- Weekday of Collection: Derive the day of the week from the date_collected. This could help in understanding if more food waste occurs on specific days of the week.

- Month of Collection: Extract the month from date_collected to see if certain months have higher wastage, potentially indicating seasonal trends.

- Year of Collection: Although the data range is not very long, extracting the year could be useful if the dataset is expanded in the future.

- Label Date Validity: Create a binary feature indicating whether the label date is valid (i.e., whether it is after the date collected). This could help identify labeling errors or unusual patterns.

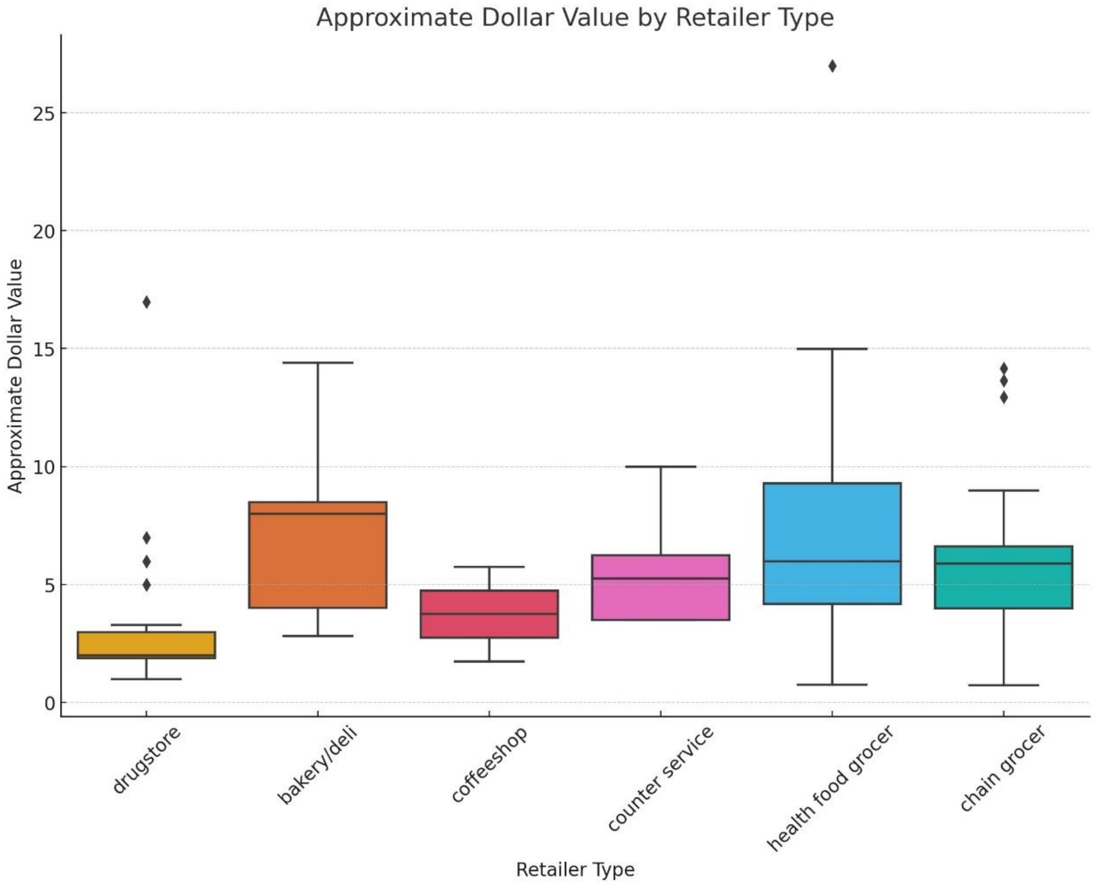

Figure 3 Approximate Dollar by Retailer Type

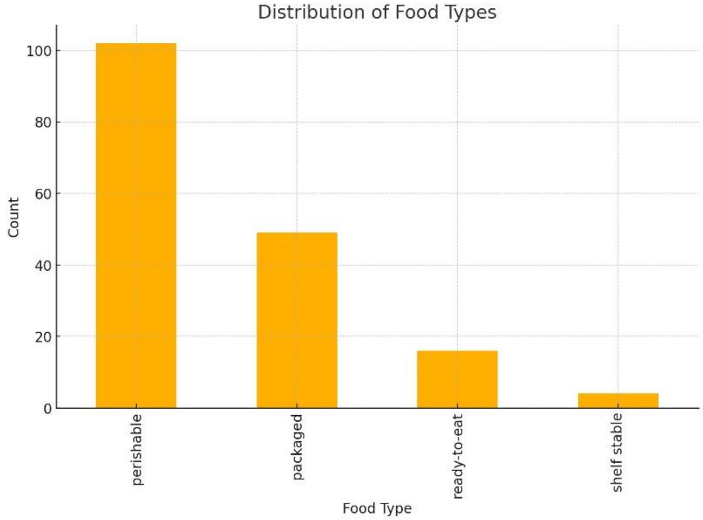

Figure 4 Distribution of Food Types

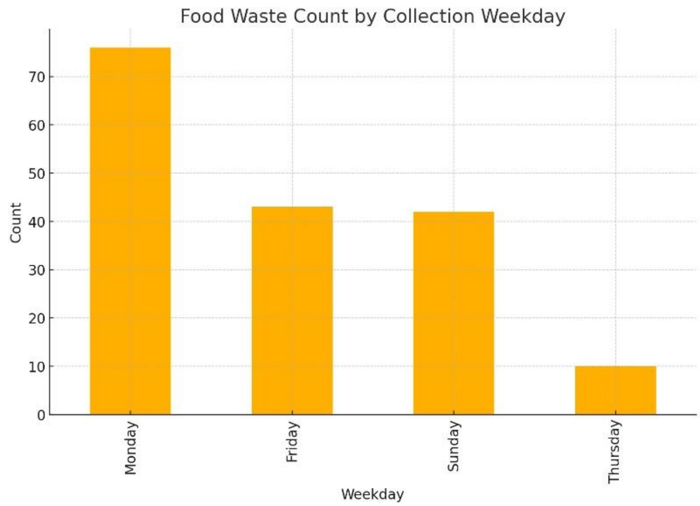

Figure 5 Food Waste Count

Building ML model

To create a dynamic pricing model that adjusts prices based on demand, shelf life, and inventory levels to reduce food wastage. First, we are investigating to find the best model for our scenarios so the following is the comparison between different algorithms.

Model Selection

Choosing the right models for dynamic pricing when trying to minimize food wastage is an important component of this research. The models must be robust, able to deal with massive datasets, and efficient at inferring the approximate dollar value of foods on different basis like shelf life, demand and stock levels. The following machine learning models were chosen because they have been validated for similar regression tasks and may be applicable for dynamic pricing models:

Random Forest (RF): RF is an ensemble learning method, which trains different decision trees and gets average prediction out of those trees. This model is famous for being robust, it can handle high dimensional and large data and reduces overfitting when compared to individual decision trees (Breiman, 2001). RF was selected for its capacity to manage diverse feature sets and its proven track record in various prediction tasks. The ensemble nature of RF helps in capturing complex relationships within the data, making it a strong candidate for dynamic pricing.

Support Vector Regression (SVR): SVR uses the Support Victor Machines (SVM) model which is specially built for regression problem. It attempts to approximate the best line with a threshold of some number with some error. SVR works on small to medium datasets and is especially efficient in high dimensional areas (Drucker et al., 1996). The flexibility of SVR in handling non-linear relationships through kernel functions (linear, polynomial, RBF) makes it a potential candidate for modeling the intricate dynamics of pricing based on various factors.

K-Nearest Neighbors (KNN): KNN is a nonparametric algorithm used in classification and regression. It find the K closest neighbors to some query point and averages them out. KNN is easy to set up and can be read but computationally expensive for big datasets (Altman, 1992). Despite its simplicity, KNN was considered for its effectiveness in capturing local data structures, which can be useful in pricing items based on similar historical data points.

Decision Tree (DT): DT is an interactive tree-like diagram of options and their implications. It is visualisable and interpretable which is useful for determining the model decision-making. But DT can overfit, so we should be aware when we are working with very large data (Quinlan, 1986). DT was selected for its simplicity and clarity in explaining the factors influencing pricing decisions, which is valuable for understanding and refining the dynamic pricing strategy.

Gradient Boosting (GB): GB is a way to merge models from other models sequentially, where the model below seeks to correct the previous models errors. GB works well for most types of predictive modelling since it is iterative and can boost the performance of weak learners by doing so (Friedman, 2001). GB was chosen for its superior performance in predictive accuracy and its robustness in handling complex, high-dimensional data. Its ability to minimize prediction errors iteratively makes it a strong candidate for dynamic pricing.

Neural Network (NN): NNs are based on human brain architecture and they can detect complex relationships in the data with layers of connected nodes. NNs are extremely accurate but still have a lot of computational load and tuning involved (Goodfellow et al., 2016). NN was considered for its potential to model complex, non-linear relationships in the data, which are often present in dynamic pricing scenarios.

Hyperparameter Tuning and Cross-Validation

Hyperparameter tuning is a must to get maximum performance from your model. That means choosing the optimal combination of hyperparameters that govern how the models learn. Hyperparameter tuning was done for each model by grid search and cross-validation. Grid search sequentially looks through various combinations of parameter values, then by cross-validating them finds out which one works best (Bergstra & Bengio, 2012).

Random Forest (RF): Key hyperparameters set were tree size estimators, tree depth and minimum number of samples needed to split a node. This way the model will not overfit or underfit the data (Breiman, 2001).

Support Vector Regression (SVR): Important hyperparameters were the kernel type (linear, polynomial, radial basis function), regularization parameter (C), and epsilon. This enables the model to accommodate different levels of non-linearity and modulate its tolerance to error margins (Smola & Schölkopf, 2004).

K-Nearest Neighbors (KNN): Neighbours (K) and distance units (Euclidean, Manhattan) were chosen optimally. This tuning is designed to smooth the bias-variance tradeoff of the model and to increase its generalization capability from training data (Altman, 1992).

Decision Tree (DT): The hyperparameters like maximum depth, minimum samples split, and minimum samples leaf were optimized. They also employed pruning to prevent overfitting by restricting the tree's complexity (Quinlan, 1986).

Gradient Boosting (GB): Learning speed, boosting stages (estimators) and depth maximum of each estimator were the hyperparameters we optimized. It's a technique called iterative booster which allows strong predictors to be produced from poor learners (Friedman, 2001).

Neural Network (NN): Hyperparameters pertaining to the network design, including layers and neurones per layer, learning rate, activation functions, and regularization parameters (dropout rates) were carefully calibrated. These adjustments enable the network to learn complex patterns without overfitting (Goodfellow et al., 2016).

This cross-validation was carried out to validate that the model's performance is generalized rather than limited to the training data. We used k-Fold cross-validation, where k is 5 or 10 and the data is partitioned into k subsets. The model is trained on k-1 of these subsets, with the last subset being validated. This is repeated k times, where each subset takes turns as the validation set (Stone, 1974). This method provides a robust assessment of model performance and reduces the likelihood of overfitting.

Evaluation Criteria

We scored each model across a number of parameters to get a true-to-goodness measure. These statistics indicate how good the models are, their margins of error, and their explanatory power.

Mean Absolute Error (MAE): MAE gives you an estimate of the average number of errors in the prediction and the real numbers which gives a good idea of how well the model predicts. It is the average of the absolute difference between the prediction and the value. Lower MAE means better model performance (Willmott & Matsuura, 2005).

Root Mean Squared Error (RMSE): RMSE is a measure that measures the variance between the estimated and measured value and it's more important for large errors. It is calculated by the square root of the mean of squared differences between predicted and measured values. RMSE can help you to know how error is distributed and penalize the larger errors harder (Chai & Draxler, 2014).

R^2 Score: The R2 score or coefficient of determination is how much of the variability in the dependent variable can be explained by the independent variables. A R2 score near 1 means that the model explained a high percentage of variance, and a score near 0 indicates lower explanatory power (Nagelkerke & others, 1991). This metric is essential for assessing how well the model explains the variability of the data.

These criteria were applied to evaluate both the predictive accuracy and the reliability of each model. The best performers were Gradient Boosting and Decision Tree, with R2 and low error rate respectively which make them ideal candidates for dynamic pricing for food wastage prevention (El Youbi et al., 2023).

Model Evaluation and Results

The model Decision Tree and Gradient Boosting both do well according to the analysis, with good R2 and low errors. Random Forest also does well, a bit less than Decision Tree and Gradient Boosting. And K-Nearest Neighbors (KNN) is performing slightly better. Meanwhile, Support Vector Regression (SVR) and Neural Network are also doomed in this scenario. Neural Network has a very low R2 value which means the model is not fitting to the data well. So, we'll use Decision Tree and Gradient Boosting as both are good in performance.

All of these metrics gave us clues as to whether the models were effective at forecasting the right price for perishable goods. The MAE indicator showed Gradient Boosting and Decision Tree close correlation with prediction and price. Lower MAE values translate into lower price errors, which is very important in dynamic pricing in order to get consumers to make purchases as soon as possible without affecting revenue. This was confirmed by the RMSE which for both Gradient Boosting (0.536) and Decision Tree (0.549) suggests low significant price deviation in predictions. The slightly higher RMSE of Random Forest and KNN (0.576 & 0.983 respectively) indicate that these models are introducing larger errors and are less effective for high-risk scenarios such as price on perishable goods.

The R2 scores were most helpful for evaluating how each model explained price variation with the input features. Gradient Boosting's R2 of 0.828 and Decision Tree's R2 of 0.820 demonstrate that such models can be able to account for the driver of price change such as demand shift and perishability. In contrast, the negative R2 scores for SVR and Neural Network suggest a mismatch with the data which could be caused by the fact that the models don't generalize well across all features.

Model	MAE	RMSE	R2
Random Forest	0.125646	0.576838	0.800906
SVR	0.954271	1.402496	-0.17694
KNN	0.455466	0.982745	0.422129
Decision Tree	0.118954	0.548928	0.819706
Gradient Boosting	0.113261	0.535695	0.828295
Neural Network	1.225096	1.866308	-1.08409

Table 2 Model Comparison

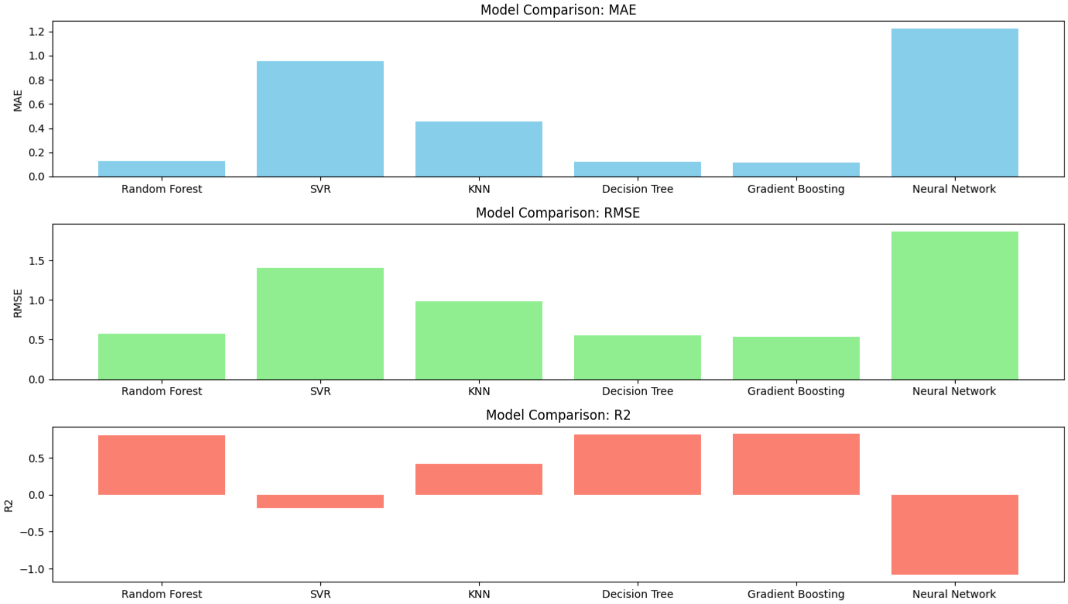

Figure 6 Model Comparison

So, the high R^2 and low error rates in Gradient Boosting and Decision Tree models ensured that these two models performed better than their counterparts. That also means these two models could be the best candidate for dynamic pricing applications that includes solving wastage issues, since they can provide real-time price changes based upon expiry date approaches, demand trend changes and stock levels. To put this in perspective, dynamic price changes allow retailers to better match consumer behaviour, in this case, to ensure that near-expiry products get sold instead of going to waste. Thus, dynamic pricing models that are better aligned with real-time market data can assist retailers in two important goals by addressing not only financial objectives but also the goals of more sustainable food supply chains.

Since the SVR and Neural Network models achieved the lowest results, perhaps more interpretable models – like Gradient Boosting and Decision Trees – will be better for this scenario. Additionally, the interpretability of a DT and GB can support retailers by providing actionable insights on how each feature is impacting decisions regarding what and when to price. Armed with these insights, retailers can become more operationally efficient and reduce wasteful fast-expiring stock. Although DT and GB models were very effective, future research could consider applying hyperparameter tuning to better optimise the models, as well as using real-time consumer behaviour data or external market-level factors. Further, a model could potentially be able to react in a more dynamic way to new data. For example, in reinforcement learning approaches, the agent not only chooses which actions will be the most effective, but also how to respond to changes in the environment like feedback from the sales department. We could envision a dynamic pricing system that autonomously makes decisions about what to price based on quantitative measures of sales and demand shifts, all of this potentially allowing us to improve model performance and impact waste reduction even further.

Discussions and Implications

This also demonstrates the potential use of AI-driven dynamic pricing in reducing food waste, which complements ongoing efforts to cut environmental impacts that result from extra steps added to the supply chain to mitigate ‘spoilage.’ The high level of accuracy of the Gradient Boosting model further suggests that dynamic pricing is a useful tool, one that can be used to target date-sensitive items and effectively induce customers to make timely purchases, before the item’s expiry date. This targeted intervention can go a long way in reducing the economic and environmental cost of food waste.

The environmental benefits of food waste reduction resulting from the use of this AI model are also significant. Each kilo of wasted food represents a significant loss of water, energy and farmland. Waste reduction is an important pillar of ecological modernisation and the concept of sustainable development in general (Porter & Reay, 2016). Therefore, the key objective of the application of AI in the case of dynamic pricing is to bring optimisation of the use and absolute reduction of the retailers’ and brands’ waste through adjustment of the prices of the items with a closing best-before date and through the channel of setting revisions to the price. According to numerous studies, cutting food waste may lead to a reduction of 8 to 10 per cent of all greenhouse gas emissions from the food sector (Pandey, 2021). Adjusting pricing according to real-time factors such as inventory or perishability has the capacity to improve food sales not only due to the price attractiveness but also due to the appreciation by the clients and help to support the concept of a circular economy. Such an economy involves consumption of all food products so that the quantity of waste can be minimised at all stages of supply.

The dynamic pricing produced by AI contributes to consumers’ long-term habits, promoting the behavioural change beyond the direct environmental effect of reduced surplus food waste. Having such items discounted near their expiry dates more often will prompt consumers to purchase them, which in turn helps them reduce their own food waste at home. As it comes to be seen as sustainable behaviour to buy near-expiry-date products, many more could begin engaging in this commercial activity, which in turn may amplify environment-friendly impact (Quested et al., 2013). Recent surveys have highlighted that ethical considerations are important in the purchase decisions of consumers when they consider where to shop. If the discounting practice of retailers incorporates more transparency, for example, by notifying the degree of discount in advance, saving consumers time from checking the price tag at the supermarket, or by pointing out the discounted items with visual cues, or flagging products that have near-expiry dates. This will encourage more consumers to participate in sustainable food waste reduction, promoting equitable consumption (Parfitt et al., 2010).

This highlights an area for further research: using reinforcement learning (RL) and other responsive AI models that adjust prices based on emerging consumer purchase patterns, which could provide an additional boost for waste reduction by addressing it inment learning has worked well in retailer pricing, as it adjusts prices iteratively to better suit consumer responsiveness and inventory status (Kastius & Schlosser, 2022). Running RL and similar models at scale in different retail contexts and regions, while still accounting for how local consumer behaviour and market conditions affect waste-reduction outcomes, could result in a scalable, ecological solution for food-waste mitigation.

To summarise, dynamic pricing with AI can help to reduce food waste resulting from high prices. First, a smoother price curve will keep at least some of the undesired customers away from restaurants during periods of high demand. Also, by setting prices that would prevent a high amount of unsold inventory, dynamic pricing will help to save natural resources and greenhouse gas emissions. Additionally, the application of dynamic pricing might lead to a green dietary shift in consumer behaviour. These results suggest that AI-based pricing strategies can make significant contributions to improve environmental sustainability in food supply chains. AI-driven dynamic pricing approaches should therefore be further developed and applied in different retail sectors.

References

Alomar, M. A. (2022). Performance Optimization of Industrial Supply Chain Using Artificial Intelligence. Computational Intelligence and Neuroscience, 2022, 1–10. https://doi.org/10.1155/2022/9306265

Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. The American Statistician, 46(3), 175–185. https://doi.org/10.1080/00031305.1992.10475879

Biller, S., Chan, L. M. A., Simchi-Levi, D., & Swann, J. (2005). Dynamic Pricing and the Direct-to-Customer Model in the Automotive Industry. Electronic Commerce Research, 5(2), 309–334. https://doi.org/10.1007/s10660-005-6161-4

Boute, R. N., & Udenio, M. (2023). AI in Logistics and Supply Chain Management. In R. Merkert & K. Hoberg (Eds.), Global Logistics and Supply Chain Strategies for the 2020s (pp. 49–65). Springer International Publishing. https://doi.org/10.1007/978-3-030-95764-3_3

Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature. Geoscientific Model Development, 7(3), 1247–1250.

Chauhan, Y. (2020). Food Waste Management with Technological Platforms: Evidence from Indian Food Supply Chains. Sustainability, 12(19), 8162. https://doi.org/10.3390/su12198162

Dasgupta, P. R., & Hashimoto, Y. (2004). Multi-attribute dynamic pricing for online markets using intelligent agents. Autonomous Agents and Multiagent Systems, International Joint Conference On, 2, 277–284.

DiMicco, J. M., Greenwald, A., & Maes, P. (2001). Dynamic pricing strategies under a finite time horizon. Proceedings of the 3rd ACM Conference on Electronic Commerce, 95–104.

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support Vector Regression Machines. In M. C. Mozer, M. Jordan, & T. Petsche (Eds.), Advances in Neural Information Processing Systems (Vol. 9). MIT Press. https://proceedings.neurips.cc/paper_files/paper/1996/file/d38901788c533e8286cb6400b40b386d-Paper.pdf

El Youbi, R., Messaoudi, F., & Loukili, M. (2023). Machine Learning-driven Dynamic Pricing Strategies in E-Commerce. 2023 14th International Conference on Information and Communication Systems (ICICS), 1–5. https://doi.org/10.1109/ICICS60529.2023.10330541

Elmaghraby, W., & Keskinocak, P. (2003). Dynamic Pricing in the Presence of Inventory Considerations: Research Overview, Current Practices, and Future Directions. Management Science, 49(10), 1287–1309. https://doi.org/10.1287/mnsc.49.10.1287.17315

Famili, A., Shen, W., Weber, R., & Simoudis, E. (1997). Data preprocessing and intelligent data analysis. Intelligent Data Analysis, 1(1–4), 3–23. https://doi.org/10.1016/S1088-467X(98)00007-9

FAO (Ed.). (2019). Moving forward on food loss and waste reduction. Food and Agriculture Organization of the United Nations.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of

Statistics, 1189–1232.

García, S., Luengo, J., & Herrera, F. (2016). Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. Knowledge-Based Systems, 98, 1–29. https://doi.org/10.1016/j.knosys.2015.12.006

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

Gustavsson, J., Cederberg, C., Sonesson, U., Van Otterdijk, R., & Meybeck, A. (2011). Global food losses and food waste. FAO Rome.

Han, W. (2010). A Dynamic Pricing Algorithm by Bayesian Q-learning. 2010 Second International Conference on Computer Modeling and Simulation, 515–519. https://doi.org/10.1109/ICCMS.2010.240

Herbon, A. (2014). Dynamic pricing vs. Acquiring information on consumers' heterogeneous sensitivity to product freshness. International Journal of Production Research, 52(3), 918–933. https://doi.org/10.1080/00207543.2013.843800

Ibn Mājah, M. ibn Y., Za'ī, A. Ṭāhir Z. 'Alī, Khattab, N., Khattab, H., & Abū Khalīl. (2007). English translation of Sunan Ibn Mâjah (Vol. 1). Darussalam.

Iliou, T., Anagnostopoulos, C.-N., Nerantzaki, M., & Anastassopoulos, G. (2015). A Novel Machine Learning Data Preprocessing Method for Enhancing Classification Algorithms Performance. Proceedings of the 16th International Conference on Engineering Applications of Neural Networks (INNS), 1–5. https://doi.org/10.1145/2797143.2797155

Kastius, A., & Schlosser, R. (2022). Dynamic pricing under competition using reinforcement learning. Journal of Revenue and Pricing Management, 21(1), 50–63. https://doi.org/10.1057/s41272-021-00285-3

Kim, K.-K., Lee, C.-G., & Park, S. (2016). Dynamic pricing with 'BOGO' promotion in revenue management. International Journal of Production Research, 54(17), 5283–5302. https://doi.org/10.1080/00207543.2016.1173250

Kumar, A., Mangla, S. K., Kumar, P., & Karamperidis, S. (2020). Challenges in perishable food supply chains for sustainability management: A developing economy perspective. Business Strategy and the Environment, 29(5), 1809–1831. https://doi.org/10.1002/bse.2470

Kumar, Prof. A., Rawat, H., & Singh, Y. (2023). Effective Machine Learning Algorithm to Predict House Price. International Journal for Research in Applied Science and Engineering Technology, 11(11), 1741–1744. https://doi.org/10.22214/ijraset.2023.56935

Kumar, V., Anand, A., & Song, H. (2017). Future of Retailer Profitability: An Organizing Framework. Journal of Retailing, 93(1), 96–119. https://doi.org/10.1016/j.jretai.2016.11.003

M., B., & K., A. (2016). Modeling the causes of food wastage in Indian perishable food supply chain. Resources, Conservation and Recycling, 114, 153–167. https://doi.org/10.1016/j.resconrec.2016.07.016

Mithun Ali, S., Moktadir, Md. A., Kabir, G., Chakma, J., Rumi, Md. J. U., & Islam, Md. T. (2019). Framework for evaluating risks in food supply chain: Implications in food wastage reduction. Journal of Cleaner Production, 228, 786–800. https://doi.org/10.1016/j.jclepro.2019.04.322

Nagelkerke, N. J. & others. (1991). A note on a general definition of the coefficient of determination. Biometrika, 78(3), 691–692.

Pandey, A. (2021). FOOD WASTAGE: CAUSES, IMPACTS AND SOLUTIONS. Science Heritage Journal, 5(1), 17–20. https://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

Porter, S. D., & Reay, D. S. (2016). Addressing food supply chain and consumption inefficiencies: Potential for climate change mitigation. Regional Environmental Change, 16(8), 2279–2290. https://doi.org/10.1007/s10113-015-0783-4

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

Quinlan, J. R. (1986). Induction of decision trees. Machine Learning, 1(1), 81–106. https://doi.org/10.1007/BF00116251

Rajagopal, M., Nayak, K. M., Balasubramanian, K., Abdul Karim Shaikh, I., Adhav, S., & Gupta, M. (2023). Application of Artificial Intelligence in the Supply Chain Finance. 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 1–6. https://doi.org/10.1109/ICONSTEM56934.2023.10142286

Ramezani, S., Bosman, P. A. N., & La Poutre, H. (2011). Adaptive Strategies for Dynamic Pricing Agents. 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, 323–328. https://doi.org/10.1109/WI-IAT.2011.193

Ramírez-Gallego, S., Krawczyk, B., García, S., Woźniak, M., & Herrera, F. (2017). A survey on data preprocessing for data stream mining: Current status and future directions. Neurocomputing, 239, 39–57. https://doi.org/10.1016/j.neucom.2017.01.078

Sreenivas, P., & Srikrishna, C. V. (2013). An analytical approach for data preprocessing. 2013 International Conference on Emerging Trends in Communication, Control, Signal Processing and Computing Applications (C2SPCA), 1–12. https://doi.org/10.1109/C2SPCA.2013.6749435

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society: Series B (Methodological), 36(2), 111–133.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30(1), 79–82.

Yimenu, S. M., Koo, J., Kim, B. S., Kim, J. H., & Kim, J. Y. (2019). Freshness-based real-time shelf-life estimation of packaged chicken meat under dynamic storage conditions. Poultry Science, 98(12), 6921–6930. https://doi.org/10.3382/ps/pez461

Zhang, D. J., Dai, H., Dong, L., Qi, F., Zhang, N., Liu, X., Liu, Z., & Yang, J. (2018). How Do Price Promotions Affect Customer Behavior on Retailing Platforms? Evidence from a Large Randomized Experiment on Alibaba. Production and Operations Management, 27(12), 2343–2345. https://doi.org/10.1111/poms.12964.