IJSOM International Journal of Supply and Operations Management

An End-to-End CRISP-DM Machine Learning Pipeline for Forecasting Demand in FMCG Chain Stores

Amir Mohammad Khani [1], **Arman Rezasoltani** [2], **Maghsoud Amiri** [3], **Ali Husseinzadeh Kashan** [4]✉

1. Department of Industrial Management, Faculty of Management, University of Tehran, Tehran, Iran. E-mail: amir.mo.khani@ut.ac.ir
2. Department of Industrial Management, Faculty of Management, University of Tehran, Tehran, Iran. E-mail: armanrezasoltani@ut.ac.ir
3. Department of Industrial Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran. E-mail: amiri@atu.ac.ir
4. Corresponding author, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran. E-mail: a.kashan@modares.ac.ir

Article Info

Article type:
Research Article

Article history:
Received August 5, 2025
Received in revised form October 26, 2025
Accepted January 2, 2025
Available online February 12 , 2026

***Keywords*:**
demand forecasting
machine learning
FMCG supply chain
retail analytics
CRISP-DM framework

ABSTRACT

Objective: Accurate forecasting of customer demand is necessary to optimize the efficiency of a supply chain, maximize profits through reduced inventory costs, and increase customer satisfaction. This research presents a new machine learning methodology based on the CRISP-DM for customer order forecasting that is both interpretive and interpretable and validates it with a real-world application from the Ofogh Kourosh Company, which offers the largest number of physical retail locations in Iran.

Methods: The dataset analyzed for this research contained 844,275 sales transactions from 40 separate physical locations. Six advanced ensemble machine learning models were developed to forecast customer order demand. A beneficial factor of this research was the ability to automate hyperparameter tuning of the six predictive models using the Optuna framework. The performance of the predictive models was then evaluated using MAE, RMSE, MSE, and R^2 metrics.

Results: Based on R^2 score, LightGBM was the most accurate predictive model with an R^2 score of 0.536. Feature importance analysis from LightGBM demonstrated that the three factors that would most determine customer order demand were the percentage of discount, price, and store location.

Conclusion: This research contributes both theoretically and practically to the development of a forecast model that is regionally, culturally, and contextually relevant within the Iranian retail marketplace. Compared to the literature, this study uses actual transactional data with ML models to narrow the theory-practice gap. Future research should emanate from this development, incorporating external influences such as climate, advertising, and macroeconomic influences for even greater forecast accuracy.

Cite this article: Khani, A. M., Rezasoltani, A., Amiri M., & Husseinzadeh Kashan, A. (2026). An End-to-End CRISP-DM Machine Learning Pipeline for Forecasting Demand in FMCG Chain Stores. *International Journal of Supply and Operations Management*, *13*(1), 133-156. https://doi.org/10.22034/ijsom.2026.110857.3442

Publisher: Kharazmi University
DOI: https://doi.org/10.22034/ijsom.2026.110857.3442

1. Introduction

In today's hypercompetitive and data-intensive retail environment, demand forecasting has become a strategic priority for companies in fast-moving consumer goods (FMCG) industries (Basson et al., 2019; Yang & Zhang, 2019; Lee et al., 2023; Basavaraju & Valilai, 2025).Demand forecasting directly affects supply chain planning, inventory allocation, pricing, and ultimately customer satisfaction (Mitra et al., 2022).FMCG products include food, beverages, personal care items, and household consumables, and they usually are in the marketplace for a short time and sell out relatively quickly (Shakur et al., 2024; Olutimehin et al., 2024).Effectively managing their supply chains involves operational agility that responds to operational realities, but in order to achieve this goal there must be also prediction based on intelligence grounded in data science (Oyeyemi et al., 2024; Nweje & Taiwo, 2025).Classical approaches to forecasting using statistical methods, like ARIMA and exponential smoothing, have long benefited businesses. These models are nevertheless handicapped by their linearity assumptions, their inability to describe complex interactions, and their lack of scalability to big data (Fatima & Rahimi, 2024; Wang, 2025).With a continually shifting market moving towards omnichannel sales, frequent promotions, and changes in customer behavior, traditional forecasting techniques are no longer sufficient in grasping non-linear trends and unexpected upturns in demand(Fildes et al., 2021).This weak point has caused scholars and practitioners to focus on the promise of machine learning (ML) models, which can represent more complex relationships in the data more accurately, learn at scale using historical data, and respond to emerging trends in real-time (Feizabadi, 2020).

Machine learning yields an increase in the accuracy of forecasts and helps prevent disruptions in the supply chain. Anchuri (2024) conducted a comparison of 127 instances and discovered through their comparative study that a hybrid machine learning model that incorporates real-time Internet of Things (IoT) information would yield a decrease of 42% in stock-outs and a 28% decrease in holding inventory costs. Similarly, Perumallapalli (2025) argues that the implementation of machine learning allows retailers to dynamically adapt to unpredictable demands, especially when using time-varied parameter integration into the forecasting model. While these recent developments are creating opportunities for retailers, a significant limitation of the currently published research is that most literature relies upon datasets collected from large multinational retail businesses. Additionally, the lack of contextual studies on empirical research means the majority of research has come from artificial supply chain configurations or industry-specific datasets like electronics or pharmaceuticals (Douaioui et al., 2024). As a result, very few empirical studies exist within a developing country such as Iran, where differences in retail channels, consumer preferences, and volatility in the marketplace can be dramatic compared to the developed Western markets.

Many earlier studies are limited to a few algorithms, such as Support Vector Machine (SVM) and Random Forest, and do not perform a comprehensive hyperparameter tuning process. Therefore, the overall fit of the models resulting from these studies is generally low. Limited attention has been given to the application of ensemble methods or the optimization methods available through particular tools such as Optuna as ways to enhance the accuracy of model predictions by automatically searching for hyperparameter values (Zheng, 2024). There are significant opportunities to improve methods of conducting research, developing methodologies with increased methodological transparency and reproducibility across the entire research process, from initial data preparation through to experimental evaluation of predictive systems. Furthermore, as most prior research studies on forecast accuracy do not include an analysis of feature importance, decision makers therefore do not have the ability to assess the most important supply and demand determinants, including price sensitivity, day of the week effects, brand effects, promotional effects, and so on.

The purpose of this article is to address the above limitations and to develop such a demand forecasting system that applies advanced machine learning algorithms and an automated method for hyperparameter tuning via Optuna to conduct demand forecasting within the Iranian retail grocery sector and evaluate results based on large transaction data collected from Ofogh Kourosh stores (Calcutta, India), one of the leading FMCG retail chains in Iran. Analyzing feature importance within the model to determine the features that would cater to consumer demand. There are four)

sections of contribution from this article: (1) Developing a scalable and geographically specific ML demand forecasting architecture for the Iranian FMCG retail sector; (2) Utilizing state-of-the-art ML technologies and optimization algorithms to achieve the best possible outcomes; (3) Understanding the business drivers of demand fluctuation; (4) Providing a repeatable demand forecasting framework that adheres closely to the CRISP-DM model, thereby enhancing both the academic rigor of the study as well as providing practical industry relevance. The remaining sections of this article are organized as follows: Section 2 consists of a thorough literature review; Section 3 covers the methodology used and data collection; Section 4 includes results and analyses; and Section 5 provides insights and implications for future research and findings.

2. Literature Review

In Table 1, we synthesize and compare important studies related to demand forecasting in the fast-moving consumer goods supply chain that uses machine learning. It presents important dimensions, including year, application area, algorithms, data type, evaluation metric, and key findings of the studies, with two additional columns presenting each study's relevance to our research case and each study's similarity to the current Ofogh Kourosh case. Project Overall Demand Forecasting The descriptive and theoretical studies they reviewed included many cases from the informing, popular press, etc., which allows us to compare dimensions in literature when conducting our secondary research. The purpose of the table is unique and enables researchers to identify trends, gaps, and relevant methodologies.

Table 1. Literature Review

Paper Title	Year	Application Domain	Algorithms Used	Compared to Traditional Methods	Real Data	Evaluation Metrics	Key Findings	Usefulness for My Case Study
Forecasting Supply Chain Demand Using Machine Learning Algorithms	Carbonneau et al. (2009)	Chocolate, toner	SVM, others	✓	✓	MAPE, ranking	SVM most accurate	Generalizable to FMCG domain
Modeling Promotional Demand Variability	Abolghasemi et al. (2019)	Fluctuating demand	ARIMAX, SVR	✓	✓	MAPE, RMSE	The ARIMAX+ SVR hybrid model excelled in handling fluctuating demand.	Effective for volatile-demand products in Ofogh Kourosh stores.
Machine learning demand forecasting and supply chain performance	Feizabadi (2020)	Steel supply chain	ARIMAX, Neural Network	✓	✓	SC performance	ML improves upstream coordination	Partially relevant (methodology)
Demand Forecasting for E-Grocery	Golabek et al. (2020)	Online Grocery (FMCG)	LSTM (single and multivariate)	✓	✓	MAPE	LSTM outperformed RF and Regression models in predicting demand for food products.	A structure similar to FMCG products in Iranian retail stores

Table 1. Literature Review (Continued)

Predicting the Demand for Fmcg using Machine Learning	MebalP et al. (2021)	FMCG	Classification algorithms	✓	✓	Accuracy, F1	ML improves inventory and profits	Highly relevant: FMCG focus
Demand forecasting based machine learning algorithms on customer information	Zohdi et al. (2022)	Customer behavior	ANN, ELM, KNN, DT	✓	✓	MAE, MSE, R^2	ANN and ELM performed best	Useful for customer-centric demand modeling
Demand forecasting accuracy in the pharmaceutical supply chain: a machine learning approach	Yani and Aamer (2022)	Pharmaceutical SC	RF, Simple Tree	✓	✓	Accuracy gains (10–41%)	ML boosts forecast precision	Methodology useful, but sector differs
Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical Review	Douaioui et al. (2024)	Literature review	Wide range of ML/DL	✗	✓	Mixed	Reviews trends & methods in ML/DL	Useful for model selection
Development of machine learning based demand forecasting models for the e-commerce sector	Fırat et al. (2024)	FMCG e-commerce	MLP, MQRNN, RF	✓	✓	MAPE	MQRNN outperformed others	Applicable for short-term FMCG sales
Explainable MCDFN for Demand Forecasting	Jahin et al. (2024)	FMCG supply chain	MCDFN (CNN+LSTM+ GRU)	✓	✓	MAPE, ShapTime	Best MATE ~20%, feature explanation with ShapTime	High interpretability ; suitable for store managers' decision-making
Demand Forecasting in FMCG – Ofogh Kourosh	2025	FMCG retail (Iran)	LightGBM, RF, XGBoost, CatBoost, Extra Trees, Gradient Boost	✓	✓	MAE, RMSE, MSE, R^2	LightGBM achieved the highest R^2 (53.6%), discount and price were the top predictors	It's the exact study

Table 2 summarizes the most common constraints identified in the literature on machine learning-based demand forecasting in the FMCG supply chain. Despite the insights provided by numerous studies, a number of methodological or contextual limitations restrict their applicability in the context of Iranian retail chains. Its drawbacks comprise a deficiency of real-world retail data, intermittent application of sophisticated ML methods, inadequate feature

importance analysis, and an absence of transparent and reproducible modeling. The recognition of these weaknesses serves as a clear choice in support of the applied research focus at this time and its methodological implementation.

Table 2. Common Limitations and Methodological Weaknesses in Prior Studies

Limitation/Weakness	Description
Lack of real data from Iranian retail chains	Studies focused on a particular industry, online shops, and select countries other than Iran and did not reflect the local features or physical store structure in Iran.
Limited use of modern algorithms and parameter optimization	Most papers simply compared simpler algorithms (SVM, ARIMAX, Random Forest) without optimized parameterizations or similar optimization techniques, like Optuna.
Lack of analysis on key demand-driving features	The models usually measured accuracy alone, which is not very advantageous without also assessing the relative importance of any variables, such as discounts, seasons, brands, or days of the week.
No transparent and repeatable approach in preprocessing/modeling	Not all studies carefully described steps such as data preparation, treatment of missing values, feature selection, or criteria used to select the final model.

Table 3 provides the main research gaps identified from the critical review of existing research and how the current research systematically addressed each gap. Most prior studies in the domain of FMCG demand forecasting identify shortcomings with restricted access to local, real-world retail data and limited algorithmic variation with no explainability or thorough investigation. The current research has contributed to carrying out a more robust and applicable forecasting framework with local and real-world retail data, using a large-scale dataset published by Ofogh Kourosh chain stores about products, implementing six different advanced ensemble machine learning algorithms with hyperparameter tuning, and conducting thorough model assessments with recognizable variable importance. Therefore, this study may identify and fill those gaps in the literature in the below sections.

Table 3. Research Gaps in the Literature and How They Are Addressed in This Study

Gap	How It's Addressed in This Study
Lack of studies using real data from Iranian retail chains	Use of actual sales data of 844,000 transactions at Ofogh Kourosh retail outlets
Need for comparison of multiple modern ML algorithms with advanced tuning	Testing 6 of the most powerful algorithms (LightGBM, CatBoost, Extra Trees, etc.) with parameter tuning via Optuna
No analysis of feature importance	Important feature analysis through LightGBM model (discount, price, day of week, etc.)
Need for separate performance analysis on training and test data	Individual MAE, RMSE, MSE, and R 2 results per model on training and test datasets

Table 4 summarizes key innovations of the proposed study when compared to past research on FMCG demand forecasts. In contrast to numerous previous works based on small datasets and generic algorithms, this project integrates various complex machine learning models and uses automatic hyperparameter optimization with Optuna. It, too, uses real transactional data of Iranian retail chain stores (Ofogh Kourosh) and uses a transparent, structured, and replicable process grounded on the CRISP-DM methodology. Moreover, the study isolates and explains the significance of major characteristics (including discount percentage and pricing) in determining consumer demand in a novel way. These advances determine the scientific and practical significance of the study.

Table 4. Methodological and Practical Innovations of the Present Study

Innovation	Description
Integration of multiple advanced ML models with Optuna optimization	Use of powerful algorithms + automated tuning to enhance accuracy
In-depth analysis of demand-driving features (feature importance)	Identification of the key roles of discounts, pricing, seasonality, brand, and store in the model
Use of large-scale real data from Iranian physical retail stores	Practical and generalizable for real-world Iranian retail chains
Transparent and repeatable structure based on CRISP-DM framework	From data mining to final model evaluation with comprehensive details

Table 5 provides a summary of the main advances made by the present study, both theoretically and realistically. Most prior literature is either contextually unadapted or methodologically shallow, whereas this study provides a context-specific, data-grounded model applied to the Iranian FMCG retail market. Methodologically, it combines advanced machine learning methods, feature importance analysis, and automatic hyperparameter search. In terms of managerial insights, the model can offer practical information to aid pricing, discounting, inventory, and supply chain planning. The following contributions make this study relevant to both researchers and practitioners.

Table 5. Scientific and Practical Contributions of the Present Study

Domain	Contribution
Localized Knowledge	Development of a practical, data-driven model for the FMCG supply chain in Iran
Advanced Methodology	Introduction of a hybrid framework combining machine learning modeling, feature analysis, and automated parameter tuning
Managerial Application	Support for decision-making in pricing strategies, discount policies, inventory planning, and supply management

3. Data and Methodology

This section describes the prime components of the model proposed for forecasting demand in fast-moving consumer goods and establishes the stages for executing the research based on the data mining approach. Because of the data-driven nature of the problem, the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology is used as the main analytical framework. CRISP-DM was developed by a consortium of experts and companies involved in the data mining field, and provided a standardized, flexible framework for conducting data mining projects (independent of application domain, industry and technologies used) (Tripathi et al., 2021). They don't tell us how to do this but rather what is the typical outputs from the six phases of a data mining project. Thus, the CRISP-DM methodology is a cyclical process (Schröer et al., 2021), with six phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. In Figure 1, an overview of the stages of this research based on the CRISP-DM methodology is presented (Wiemer et al., 2019).

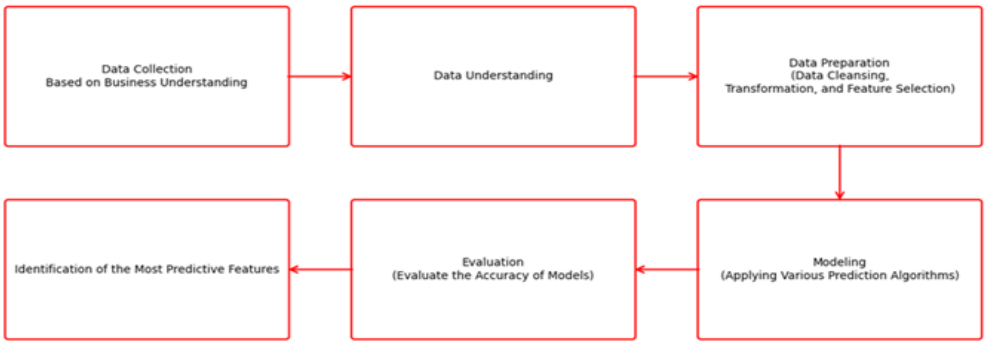

Figure 1. The Flow Diagram of the Methodology

The flow diagram in Figure 2 shows the order of procedure undergone to create and test the demand forecasting model, including collecting raw retail sales data, refining the model, testing the model, and extracting managerial insight using advanced machine learning methods.

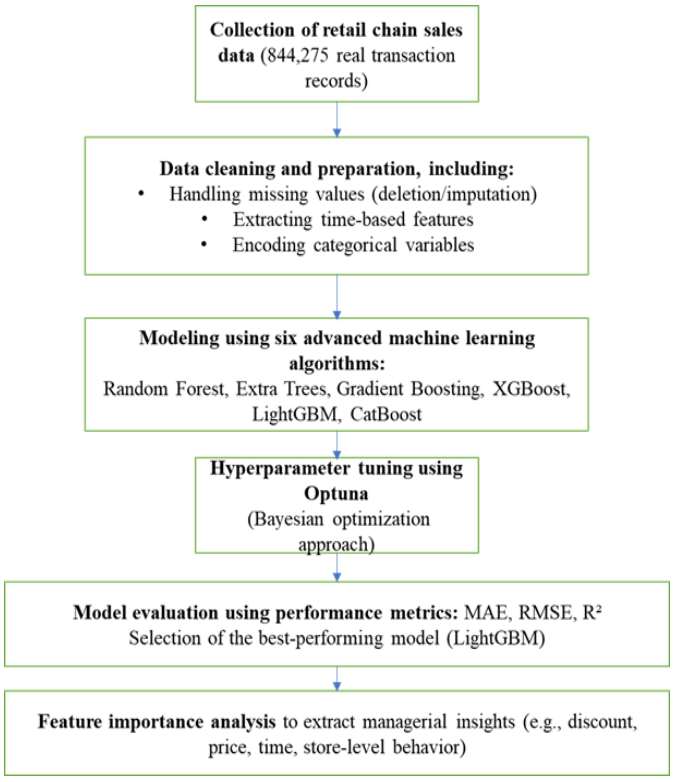

Figure 2. Problem-Solving Steps Based on the Proposed Research Framewo

3.1 Data Collection

Using actual product retail transaction data from the Ofogh Kourosh retail network as the source for this research, to develop an understanding of customer purchasing behavior, we used a large quantity of data on several independent variables, such as the product characteristics, store characteristics, and characteristics of the economic environment

surrounding the sale.The dataset will allow us to view sales trends over time and predict future demand or develop future sales forecasts; therefore, it has the potential to allow us to analyze demand behavior and sales patterns as well as forecast total sales volume of FMCG products through a time-series style method of analysis. The dataset also allows the researcher to analyze sales of FMCG products based on a time-series method. Since the dataset contains extensive amounts of transactional data from Ofogh Kourosh stores, as well as a variety of product categories, store locations, and the various economic conditions, the dataset has a high level of potential for generating high-quality forecasting models.The Ofogh Kourosh retail dataset consists of product–store–day unit transactions recorded between 21 March 2021 and 20 March 2022 and includes 844,275 observations from various product categories and store locations in the product.types.csv file of the dataset.

3.2 Data Understanding

In this study we use a dataset comprising 844.275 records of actual sales transactions of fast-moving consumer goods in Ofogh Kourosh chain stores. The data spans one year and has a reasonable variety of product types, brands, store locations, and selling contexts that will allow us to conduct our analysis and modeling; hence, it is appropriate for research purposes. Table 6 shows the main variables used in the forecast model. The variables included some categorical variables such as product ID, category, and store and quantitative variables such as discount, price, and volume, all of which support accurate predictions of sales quantities.

Table 6. Understanding Variables (Features) Used in This Research

Feature Name	Description	Type	Number of Unique Values	Sample Values
Gregorian_Date	The Gregorian calendar date when the transaction occurred	Categorical	365	2021-03-21, 2021-03-22, 2021-03-23, ..., 2021-03-30
ItemID	Unique identifier assigned to each product item	Categorical	195	Item_175, Item_63, Item_62, ..., Item_48
Category	Product category or type	Categorical	20	Cat_15, Cat_12, Cat_11, ..., Cat_1
Brand	Manufacturer or brand of the product	Categorical	37	Br_36, Br_9, Br_11, ..., Br_13
Supplier Name	Name of the supplier providing the product	Categorical	27	Sup_23, Sup_9, Sup_8, ..., Sup_4
Store	Store ID or location code where the product was sold	Categorical	40	Store_20, Store_35, Store_7, ..., Store_40
Sale QTY	Number of product units sold in the transaction	Numerical	317	2, 7, 17, 5, 3, 10, 8, 14, 4, 9
Discount Perc.	Percentage of discount applied to the product price	Numerical	152,180	0.1, 0.0629, 0.05, 0.0694, 0.0505, 0.0515, 0.0589, 0.0501
Shelf Life	The product's shelf life (in days) from production to expiry	Numerical	26	365, 180, 293, 303, 150, 295.68, 120, 200, 273, 300
Volume	Volume or size of the product unit	Numerical	53	514.25, 425.0, 2592.0, 3800.0, 1518.75, 1536.0, 1664.0, 1406.25, 231.5, 1984.0
Price	Selling price per unit of the product	Numerical	11,357	30000.0, 35000.0, 99000.0, 135000.0, 75000.0, 98999.99, 35000.00002

The company's sales information system provided the data set used in this analysis and, given the volume and quality of the data, is a good basis for developing demand forecasting models and analyzing sales behavior at both the store and product levels.

3.3 Data Preparation

During the Data Preparation phase, the initial assessment of the data included checking for the presence of missing values. The assessment found that three columns had missing values: "Discount Perc." (682 missing values), "Shelf Life" (7.961 missing values), and "Price" (869 missing values). Because of the amount of data as well as the value of the data, removing the rows where the missing values occurred could have resulted in loss of relevant information. For this reason, we utilized a two-step imputation method: missing values would first be filled with the mean of the feature by ItemID to maintain a level of similarity among similar items (Rezvan et al., 2022) and second, if there were any missing values afterward, the overall mean of the column would be used to fill the remaining columns. This would also ensure that the statistical makeup was preserved and the data structure was kept intact from bias in the modeling process (Xu et al., 2020).

Next, we proceeded to conduct data preprocessing operations. The "Gregorian_Date" column was altered to a Python-readable datetime format to enable the extraction of more specific temporal features. The four main temporal features that we extracted were day of the month, month of the year, year, and day of the week. These temporal features are necessary to help with the analysis of consumer behavior and seasonally impacted sales. For example, the day of the week variable captures consumer behavior on holidays and weekends. In this context, these days have shown greater changes in consumer behavior and thus influence FMCG purchases.

Next, categorical features, such as "ItemID", "Category", "Brand", "Supplier Name", and "Store", were encoded to numerical codes to allow for machine learning algorithms to use these features. The purpose of label encoding is to help the model learn relationships among categories without having to deal with alphanumeric characters. During feature selection all columns except "Gregorian_Date" and "Sale QTY" were used as inputs for the model development process. The "Sale QTY" was formatted into a vector, appropriate for use with machine learning models. Additionally, Pearson correlation coefficients were used to assess whether there may be high correlations between features and hence possible multicollinearity. Based on the results, there were no features with any correlation coefficient greater than 0.7 with each other (Khani et al., 2023). Hence, no features were removed from the model development process due to multicollinearity and all features were included. Figure 3 presents a correlation heatmap depicting the relationships among all features.

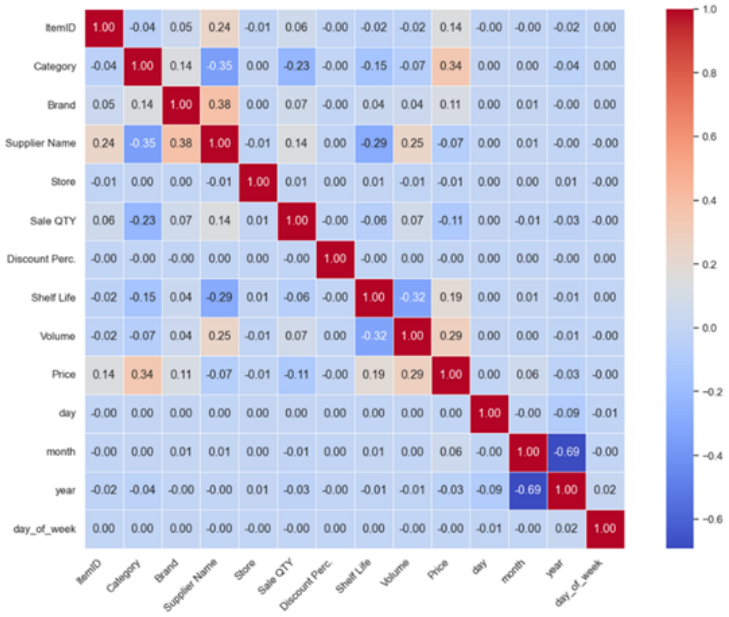

Figure 3. Correlation Heatmap of All Features

Overall, these data preparation steps resulted in a clean, consistent, and reliable dataset, which was used for accurate modeling of fast-moving consumer goods sales.

3.4 Modeling

Our goals during the modeling phase of the project were to create & validate models that accurately predicted FMCG (Fast Moving Consumer Goods) demand in grocery/chain stores using six highly advanced algorithms from multiple categorical regression types found in both industry practice and research: Random Forest, Gradient Boosting, Extra Trees, XGBoost, LightGBM, and CatBoost, and describe each algorithm, provide justification for selecting each, and then conclude with a summary of our findings. Random Forest Regressor is an established ensemble method of prediction based on the methodology of 'decision trees.' Random Forest uses bootstrapping & then generates an average of the predictions made by a sample of different decision trees. Random Forest helps to ensure stability and reduces overfitting to some extent (Hennart et al., 2022). Random Forest provides good results on noisy data with many predictor variables (Ghosh & Cabrera, 2021). Unlike Random Forest, Gradient Boosting Regressor builds new trees to fix the previous tree's mistakes until it has built enough trees, making its final prediction. Gradient boosting is highly sensitive to hyperparameter tuning, but once fine-tuned, it provides some of the best predictive results available (Meaney et al., 2025).

The Extra Trees Regressor is similar to the Random Forest model, but it introduces more randomness in selecting split points. By increasing the diversity among trees, this approach often leads to reduced variance and faster training times (Afshar et al., 2022). The XGBoost (Extreme Gradient Boosting) algorithm is an optimized and significantly faster version of traditional Gradient Boosting (Rezasoltani, et al., 2025). It has gained a prominent position in scientific competitions and real-world applications by employing techniques such as pruning and regularization (Wiens et al., 2025). LightGBM is also a tree-based boosting algorithm created by Microsoft (Jafarnejad et al., 2025). It is faster than other models due to a number of computational optimizations, and the way it creates trees, a leaf-wise instead of a level-wise strategy (Rizkallah, 2025). Finally, CatBoost is a powerful and modern decision tree–based boosting algorithm specifically designed to handle categorical features. It delivers high performance without requiring extensive preprocessing of these features (Geeitha et al., 2024). Despite their strong predictive performance, the selected machine learning models also have inherent limitations. Tree-based ensemble models may suffer from reduced interpretability compared to simpler statistical approaches and can be sensitive to hyperparameter configurations, potentially leading to overfitting if not properly tuned. In addition, these models rely heavily on historical patterns and may not fully capture sudden structural changes in demand caused by external shocks or rare events (Rezasoltani et al, 2025).

To improve the performance of the various machine learning models, the hyperparameters for all of these ML models were tuned using the Optuna library (Lai et al. 2024). Optuna is an automated hyperparameter optimization framework, which is open source and allows efficient exploration of a broad range of hyperparameter values while minimizing computational resources (Lai et al. 2023). The other two typical methods of optimizing hyperparameters, namely, grid search and manual tuning, both of which are computationally intensive, often do not provide a global optimal solution. Optuna solves this issue by using Bayesian optimization with the Tree-structured Parzen Estimator (TPE) Algorithm to rank the models in an iterative manner and to further refine them (Akiba et al. 2019). The TPE is designed to balance exploration (taking a chance on new & untried parameters) with exploitation (refining and optimizing identified areas of parameter space) for an efficient and adaptive search (Watanabe, 2023). This hybrid search nature makes it possible for Optuna to rapidly converge onto a set of high-performing hyperparameter values with significantly fewer iterations than traditional search approaches. For this study, a custom-configured search space for each model based on the mathematical characteristics and structure of each model was created; this included parameters such as learning rate, number of estimators, and other parameters related to tree structure, etc.MSE was selected due to its sensitivity to large deviations, thereby disincentivizing other models from producing large forecasting errors, and therefore promoting better parameters that yield consistent and reliable predictions. After the

optimization procedure converged, the best performing parameters were selected and retrained on the whole training dataset before the final assessment. The search space configurations associated each algorithm are provided in the following table.Table 7 displays the specific hyperparameter search spaces for all of the machine learning algorithms developed for the experiment. These ranges were used in the Optuna optimization framework to guide the exploration and tuning of model configurations to achieve the best performance.

Table 7. Defined Search Spaces for Each Algorithm

Model	Hyperparameter	Search Range	Description
Random Forest	n_estimators	100 – 1000	Number of trees in the forest
	max_depth	5 – 50	Maximum depth of each decision tree
	min_samples_split	2 – 20	Minimum number of samples required to split an internal node
	min_samples_leaf	1 – 20	Minimum number of samples required at a leaf node
Gradient Boosting	n_estimators	100 – 1000	Number of boosting stages (trees)
	learning_rate	0.001 – 0.3	Shrinks the contribution of each tree
	max_depth	3 – 50	Maximum depth of individual trees
Extra Trees	n_estimators	100 – 1000	Number of randomized trees in the ensemble
	max_depth	3 – 50	Maximum depth of each tree
XGBoost	n_estimators	100 – 1000	Number of boosting rounds
	learning_rate	0.001 – 0.3	Step size shrinkage to prevent overfitting
	max_depth	3 – 50	Maximum depth of trees
LightGBM	n_estimators	100 – 1000	Number of boosting iterations
	learning_rate	0.001 – 0.3	Controls contribution of each tree
	max_depth	3 – 50	Maximum tree depth
	num_leaves	10 – 200	Maximum number of leaves per tree, controlling model complexity
CatBoost	iterations	100 – 1000	Number of boosting iterations
	learning_rate	0.001 – 0.3	Learning rate for gradient updates
	depth	3 – 12	Depth of individual trees

This optimization process allowed each model to be trained with the best possible parameters, maximizing their predictive accuracy.

3.5 Evaluation

In this analysis, four widely utilized and interpretable regression measurements were utilized to assess the accuracy and efficiency of the demand forecasting models. These included Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Coefficient of Determination (R^2)(Wu et al., 2024). The metrics allow the overall model performance to be evaluated on the training and test data separately while also identifying potential signs of overfitting on the training data. In this comparison, the forecast data was separated into 80% training data and

20% test data (Durap, 2023).Due to the large dataset size (844.275 samples), 168,855 samples were randomly sampled to serve as the test set, and the remaining 675.420 samples were to be used to train the models. This was done randomly using stratification to maintain the distribution so that the models results will generalize. One of the main metrics in this study will be the Coefficient of Determination (R^2). The R^2 value indicates how much of the variance in the dependent variable is explained by the independent variable. R^2 values range from 0 to 1; larger values (closer to 1) represent more explanatory power of the model (Mehregan & Khani, 2024). R^2 can also be interpreted as the "relative reduction in uncertainty with respect to Kullback-Leibler divergence from using explanatory variables." The equation for R^2 is expressed as (Le et al., 2023):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (1)$$

Where Y_i represents the actual observed values, Y_i denotes the predicted values by the model, $\bar{Y}$ is the mean of the actual values, and n is the number of samples. This formula indicates the proportion of the data's variance that is explained by the model. In addition to R^2, three other evaluation metrics were also used, described as follows:

Mean Absolute Error : MAE represents the average absolute difference between the actual and predicted values and is defined by the following formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \quad (2)$$

Mean Squared Error : MSE indicates the dispersion of errors in the model and is more sensitive to larger errors:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad (3)$$

Root Mean Squared Error : RMSE is the square root version of MSE and, like MSE, is sensitive to large errors. However, its unit matches that of the model output, making it easier to interpret.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \quad (4)$$

By incorporating these metrics, researchers could have an accurate view of how the model performs in predicting the response numerically and their applicability to new data, which is especially relevant as the model output is the actual product demand and thus a quantity that must be understood to be a real-valued, numerically meaningful quantity (real-valued interpretability). In the end, through the course of measuring and reporting all of these indicators through all of the models, we chose the final model that provided the highest accuracy and lowest prediction error.

3.6 Identification of the Most Predictive Features

In the last step of modeling, we used the feature importance capability in the LightGBM algorithm to analyze model performance and find the feature factor that influences the estimation of future product demand the most (Chen et al., 2024). The feature selection method based on LightGBM is a feature selection embedded method because, in the LightGBM method, the importance weights of features are extracted internally during model training (Hancock & Khoshgoftaar, 2020). LightGBM (Light Gradient Boosting Machine) is a very modern and fast gradient boosting algorithm, meaning that it has a lot of advantages when working with structured data (Ileri, 2025). One of these is that it allows you to run categorical features without requiring you to run complex transformations like one-hot encoding (Lu et al., 2022).

This ability not only makes processing much faster and with less memory, but it will also allow for better analysis of the nonlinear association(s) between categorical features and the target variable. During training, the LightGBM model builds a large number of decision trees to fit, sequentially and in a boosting manner, all of the data (Demirtürk et al., 2025). Each decision tree will contribute to the improvement and fit of the model among the features depending on which features contribute the most reduction in the cost function (thus minimizing the MSE) (Zhou et al., 2024). Therefore, the contribution of each feature towards a valid and reliable data split will rank the importance of each of these features among the trees. In the end, the features must be ranked from most important to least important based on the importance of each feature by LightGBM (Adler & Painsky, 2022). This ranking is useful for getting a deeper understanding of the data structure and model behavior and can also serve to optimize future models and inform operational decisions on managing and ordering inventory for high-demand items.

In addition to the internal feature importance scores from the LightGBM model, we also applied a model agnostic interpretability method called SHAP (SHapley Additive exPlanations) to better understand the impact of individual features on product demand prediction in a depthful and reliable manner. SHAP is a unified interpretability framework from cooperative game theory, in which the prediction of the machine learning model can be thought of as a game, and every feature is a "player" contributing to the final output (Khani et al., 2025). The goal of SHAP is to measure the marginal contribution (i.e. the difference in output based on the model prediction with and without the feature) for each feature to the output based on many combinations or coalitions of features. Therefore, SHAP can quantify the size of the feature's influence on the prediction of demand, as well as whether it positively or negatively impacted the predicted value.

In contrast to embedded feature importance methods which use only feature selection frequency from decision tree splits, SHAP furnishes consistent and theoretically justified estimates of the contributions of features to both local and global modeling, meaning it provides insight into not only which features are significant, but how and in what direction they contribute to the predictive power of a model. This property of SHAP is especially useful when used by managers in practice (e.g., retail settings) when knowing why the model is making these predictions equals the utility of high predictive accuracy. SHAP was applied in this paper using the optimized LightGBM model, specifically using the TreeExplainer package.

4. Results

The next section provides performance metrics for the machine learning models. This study used Python as the programming language and performed all model training and experiments on a local workstation with a 13th Gen Intel ® Core™ i7-13700H CPU (2.40 GHz), 16 GB RAM, and Python 3.12 version running on 64-bit Windows. The whole training and hyperparameter optimization pipeline took about 2.5 hours to run on the given platform. Table 8 shows the performance of the machine learning models using the evaluation metrics as a foundation.

Table 8. Performance of Machine Learning Models Based on Evaluation Metrics

Model	Train_MAE	Test_MAE	Train_RMSE	Test_RMSE	Train_MSE	Test_MSE	Train_R^2	Test_R^2
LightGBM	1.5721	1.6668	4.1773	4.4488	17.4497	19.7914	0.6572	0.5360
Random Forest	1.6431	1.8371	5.0149	4.9190	25.1496	24.1970	0.5059	0.4327
Gradient Boost	1.8143	1.8792	4.5920	5.1018	21.0864	26.0285	0.5858	0.3897
Extra Trees	1.5534	1.9139	4.0301	5.0046	16.2420	25.0460	0.6809	0.4128
XGBoost	1.9668	1.9575	5.7296	5.0952	32.8281	25.9610	0.3551	0.3913
CatBoost	1.8258	1.8721	4.8090	4.8884	23.1263	23.8962	0.5457	0.4397

When we analyze the results of our regression models, we looked at four terms; Mean Absolute Error , Root Mean Squared Error , Mean Squared Error , and the Coefficient of Determination (R^2), for both training and test sets to fully evaluate the models to see how accurate each model is, and how generalizable they are. This also allows us a means to measure the absolute accuracy of the models, as well as observe the consistency of the models or give us insight in overfitting.In the first iteration, the LightGBM model did not perform too poorly. Specifically, the model outputted a test MAE of 1.6668 and an RMSE of 4.4488. Additionally, combining a relatively low MSE value (19.7914), which suggests LightGBM predicts at a low prediction error, with an R^2 value of 0.5360 on the test set which suggests that LightGBM explained about 53.6% of the variation in demand (one of the highest of the models tested), tells us that LightGBM performed relatively well.The Random Forest model performed reasonable too, with MAE being 1.8371 and RMSE 4.9190. The MSE (24.1970), and R^2 (0.4327) value is relatively lower than the MSE and R^2 values reported in LightGBM; thus Random Forest could model and predict at a slightly lower accuracy than LightGBM.

There is some mild overfitting in this model, which is validated by a small difference in the R^2 when comparing the training and test results (0.5059 in training vs. 0.4327 in testing). The Gradient Boosting Model seems relatively accurate, but on the test set it drops off significantly from the training set accuracy (Train R^2 = 0.5858 and Test R^2 = 0.3897) and therefore, cannot generalize the data as effectively, suggesting that there may be some negative tuning of parameters. The RMSE on the test set (5.1018) also is one of the highest relative to the other models. The Extra Trees model provides some unique results. It achieves a lower training RMSE (4.0301) and a lower training MSE (16.2420), however on the test set, is slightly low with RMSE (5.0046) and R^2 (0.4128). Which suggests that the model learned approximately fairly well, but its applicability to generalization is the limiting factor.

The XGBoost model performed the worst of all the models, providing an R^2 of 0.3913 on the test set with a high MAE of 1.9575. Although it does not appear to have overly over-fit the data, (Train R^2 = 0.3551) it is overall less accurate than the other models, combined with a high training MSE (MSE = 32.8281).Lastly, CatBoost, which had a test MAE of 1.8721 with an RMSE of 4.8884, had better accuracy than XGBoost but worse accuracy than LightGBM and Extra Trees. The R^2 value of CatBoost is 0.4397, indicating that the model explains almost 44% of the variance in the dependent variable.

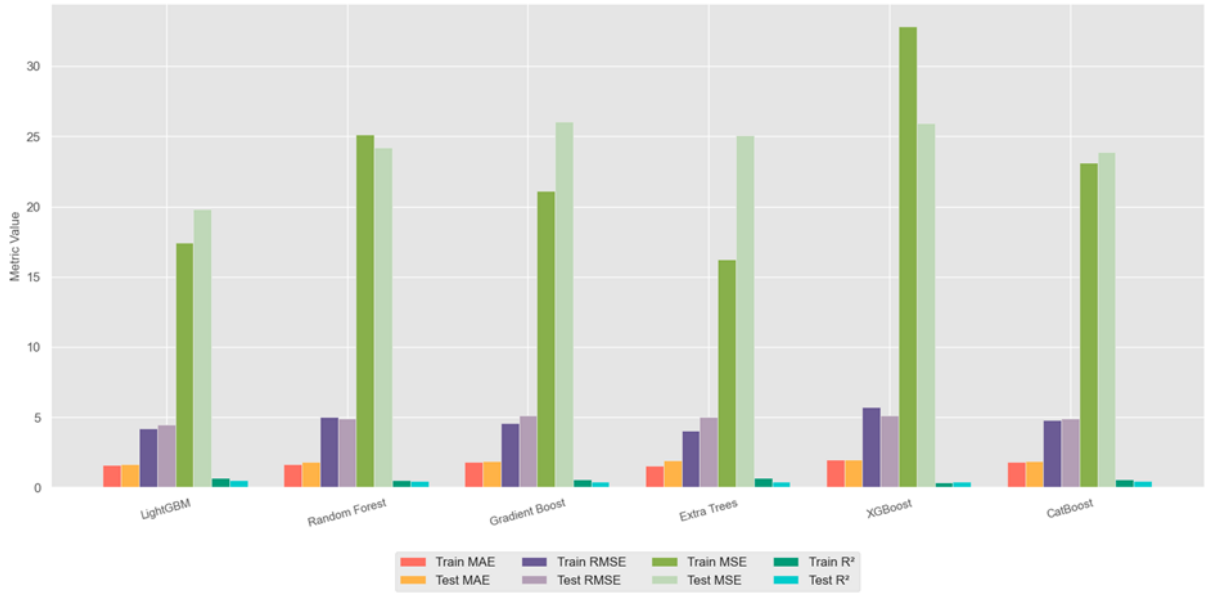

Figure 4. Train vs. Test Performance Metrics of the Models

Figure 4 provides a summary of the models' performance based on four evaluation metrics (MAE, RMSE, MSE and R^2). In summary, we find that the LightGBM model provides the best overall combination of accuracy/ excellent generalization/ acceptable error rate. Additionally, the Extra Trees model was good for learning, but there was a noticeable drop in accuracy when generalizing the performance on the test data. All of the other models displayed varying levels of error and fluctuation during the training and test metrics. This thorough analysis revealed that when clearly selecting a model for demand forecasting tasks, all of the performance indicators ought to be accounted for while also considering the trade off between accuracy and generalization as well as the level of error. Table 9 presents the best-hyperparameter values found for each algorithm.

Table 9. Best Optimized Hyperparameters for Each Algorithm

Model	n_estimators	learning_rate	max_depth	min_samples_split	min_samples_leaf	num_leaves	iterations	depth
LightGBM	889	0.03889	11	-	-	117	-	-
Random Forest	333	-	37	6	12	-	-	-
Gradient Boost	304	0.25486	5	-	-	-	-	-
Extra Trees	141	-	18	-	-	-	-	-
XGBoost	260	0.08017	4	-	-	-	-	-
CatBoost	-	0.15408	-	-	-	-	209	11

Figure 5 shows the feature importance using the LightGBM model. This chart shows the relative contribution of each input variable in predicting the demand of fast-moving consumer goods in chain stores. This ordering of feature importance has important practical managerial and strategic decision-making implications, which may affect cost-optimal pricing policies, discount policies, and supply and inventory planning.

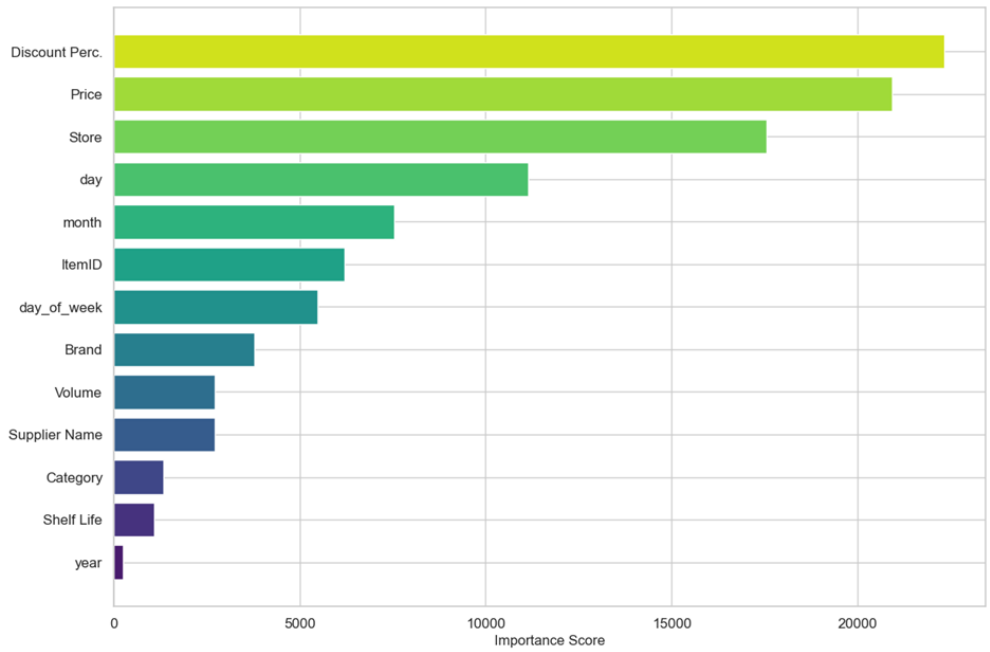

Figure 5. Feature Importance from LightGBM

The variable found to have the greatest impact on demand is referred to as Discount Percentage (Discount Perc.). This observation shows that variations in the level of discounts have had the most significant effect on the willingness of customers to buy. This knowledge can inform marketing and sales managers in setting more specific discount policies over specific periods to encourage sales. The second-most-important feature is the "Price", which emphasizes the sensitivity of consumers to price variations. Pricing strategies being optimized without raising opportunity cost or lowering profit margin can have significant effects on demand. An intelligent price and discount combination, being the two most dominant demand generators, acts as a strong market share expansionist. Features like Store name, Day of the month (day), Month (month), and Day of the week (day_of_week) are found at subsequent ranks. Such a strong emphasis on these characteristics highlights the significance of temporal and spatial patterns of customer purchasing behavior. As an example, the stores with more advantageous geographical positioning or accessibility have seen increased sales. Also more demand days of the week, or day/month (pre-holidays or early in the month when customers get paid) supports the usefulness of calendar-based variables in both supply and marketing planning.

Mid-tier functionalities involve Item ID, Brand, and Product Volume. This can imply that some products or brands are intrinsically more popular at any given time or at any price. That the size of a pack or the physical characteristics of a commodity can affect customer preference. Next comes features such as Supplier Name, Product Category, and Shelf Life. These are of minor value but contribute to the model and can be useful in supply chain solutions or product portfolio diversification strategies. Lastly, the Year variable was identified as the least important, which is understandable given that the data referred to a relatively narrow period of time and a year-to-year change has not been a decisive factor during this time frame. The factor that was concluded to be the most significant influencing factor when it comes to the demand is the Discount Percentage (Discount Perc.). This observation is a clear indication, that variations in the level of discounts have influenced the willingness of clients to procure them more than any other factor. This can inform the marketing and sale managers to come up with more focused marketing discount policies within specific periods in order to improve the volume of sales. The feature of the price ranks second in priorities showing the sensitivity of a consumer to the prices. The current move to optimize the pricing strategies which do not either raise the opportunity costs or cut the profit margins can have a heavy effect on demand. It is an effective strategy to exploit price and discount as the two primary influencers of demand in order to grow a market share with a clever mix.

The fields of Store name, Day of the month (day), Month (month) and Day of the week (day_of_week) are offered in the next levels. These features are of great significance, which means that the pattern of customer purchasing behavior depends on time and space. An example of this is that stores that are geographically placed or accessible have increased their sales. Also, the higher sale on particular days of the week or month, e.g., pre-holiday, or early in the month when a customer gets the salary, prove the effectiveness of calendar-related variables to be included in supply and marketing planning. Some of the middle-range features are Item ID, Brand, and Product Volume. It implies that some products or brands are simply more popular and that regardless of when a certain product is being offered or how much it is discounted there is a chance that packaging size or other physical characteristics of a particular product can affect the customer buying decision. After these, there are aspects such as Supplier Name, Product Category and Shelf Life. They are not as critical although contribute in the model and they can be useful in the supply chain studies or in a diversification of products offering. Lastly, the least significant variable was revealed to be the so-called Year variable, which is understandable given the fact that the values source is the short period of time, and annual fluctuation has not had a significant impact in such short period.

We conducted SHAP-based global interpretability analysis to better understand the impact of each feature on how the optimized LightGBM model makes predictions. Figure 6 presents the SHAP beeswarm plot, which illustrates the magnitude and direction of feature effects for all model predictions. As shown in Figure 6, Brand and Price were the most influential features driving demand variation—ItemID and Volume were again below Brand and Price. The horizontal spread of SHAP values for these features is wide, indicating variation across different products and stores, while Supplier Name and year tended to have very limited spread and minimal contribution to model outcome, suggesting they contribute marginally to demand behavior.

In SHAP analysis, the sign of the SHAP value indicates the direction of a feature's impact on the model prediction. A positive SHAP value means that the corresponding feature increases the predicted demand relative to the model's baseline prediction, while a negative SHAP value indicates a decreasing effect. The magnitude of the SHAP value reflects the strength of this contribution.

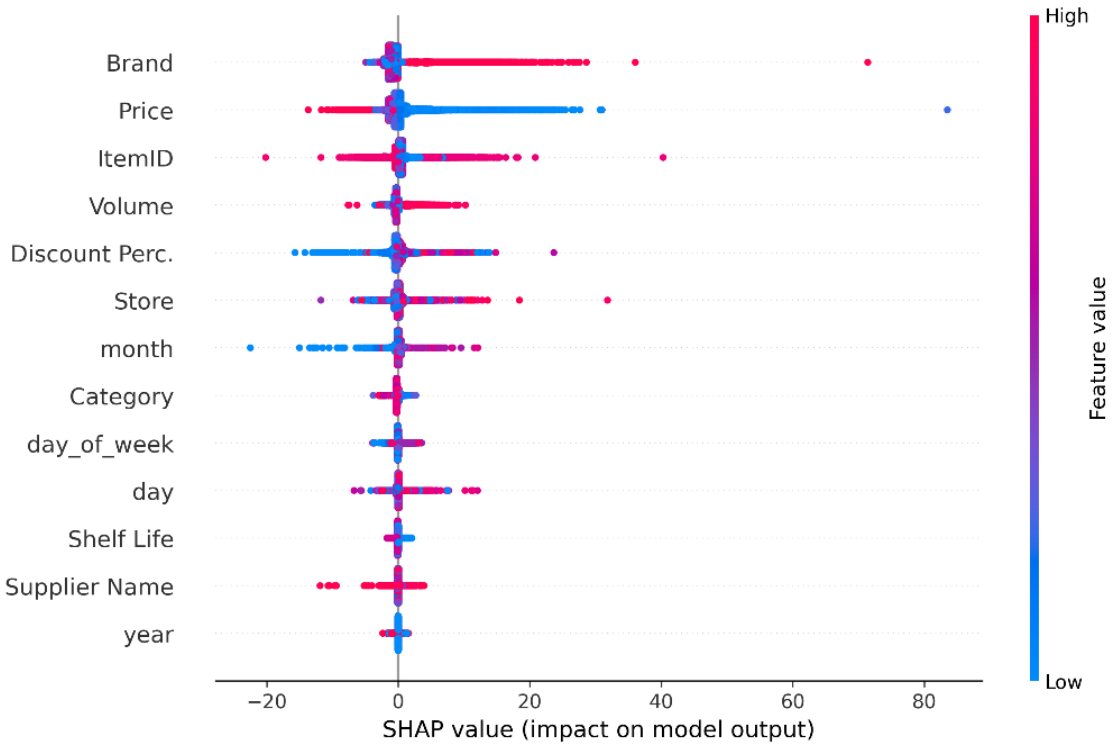

Figure 6. SHAP Beeswarm Plot Showing Global Feature Contributions to Demand Prediction

The color gradient shown in Figure 6 gives insights into how the order of magnitude across each feature impacted its implied expected contribution to predicted demand. For example, when Price is higher (red), it is typically correlated with a negative SHAP value, which means higher prices generally decrease model-predicted demand, which is consistent with established research on the effect of price elasticity when consumers are making purchasing decisions. By contrast, the higher the values of Discount Perc., or deeper promotional discount, the model-predicted unit sales volume increased, so more promotional discount resulted in positive SHAP contributions. Calendar-related features such as month and day_of_week also tend to show ordered and symmetrical effects, but less pronounced, which in this example, expressed seasonality and week patterns that drives purchasing activity.

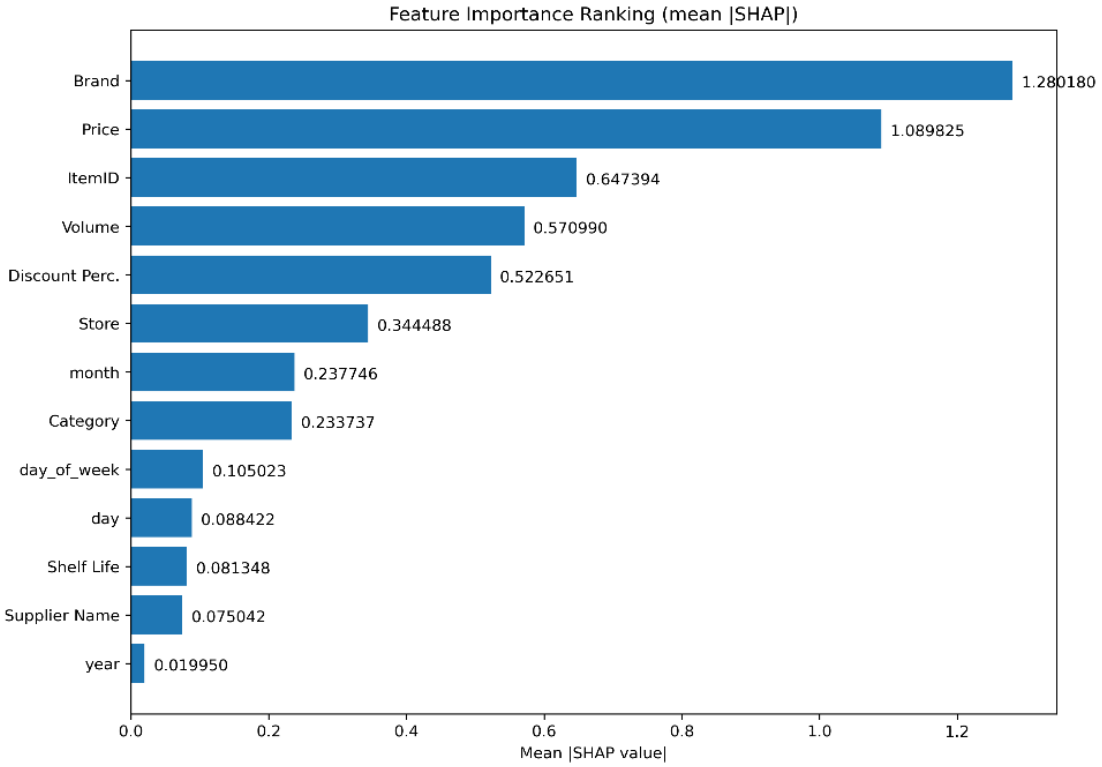

Figure 7. Feature Importance Ranking Based on Mean Absolute SHAP Values

Figure 7 displays the ranking of features by the mean absolute SHAP value, which portrays an interpretability hierarchy of feature importance. The results reiterate the overwhelming influence on demand comes from Brand, Price, ItemID, Volume, and Discount Perc. The ranking shows that certain contextual features such as Store and Category have a moderate influence as well. These results reinforce that product characteristics (identity, market positioning, packaging size) and pricing/promotional strategies represent the foundational determinants of purchasing behaviour within FMCG retail channels. Overall, the SHAP results add to the interpretability of the model not only by clarifying which features are important, but also by offering insight into how and in what direction the respective features influence demand to help with decision-making related to pricing strategy, promotion planning, assortment management, and tactical inventory management.

5. Discussion and Conclusion

This paper sets out to design and test a machine learning-based model to predict FMCG demand in retail chain stores in Iran. With the help of real-world datasets of nearly 844,000 transactions in Ofogh Kourosh retail stores, this study created a precise, transparent, and localized model to help supply chain decision-making. The section offers a detailed discussion of the findings, a comparison with other studies, theoretical and managerial implications, limitations, and suggestions for future studies.

5.1 Comparative Analysis of Findings

The results of this study found that machine learning models, specifically LightGBM, are highly accurate in forecasting demand. LightGBM performed best compared to Extra Trees, CatBoost, and XGBoost based on the R^2 score, which was about 0.536. LightGBM is able to capture complex, nonlinear relationships and interactions of various features. Random Forest and Gradient Boosting performed well, but the accuracy is not as high as current boosting algorithms.

LightGBM outperformed XGBoost primarily due to its leaf-wise growth strategy, efficient histogram-based learning, and native categorical handling, which together enhanced model efficiency and generalization on high-cardinality retail data.

A key innovation of this research was the feature importance analysis, which indicated that the four most important variables affecting demand variability were “discount amount”, “price”, “store”, and “day”. This is particularly important at the operational level, as it illustrates that consumers are controversially affected by pricing and discount decisions. The feature importance analysis indicated that brand, volume, and shelf life contributed significantly to the variability; however, the effect was lower than discount and price.

5.2 Comparison with Previous Studies

This study presents distinct differences and betterment compared to previous research studies. First, many previous studies used samples from developed nations or laboratory examples; however, this study used actual sales information from the Iranian market. The use of actual local sales data provides greater generalizability, validity, and comfort with regard to using this model for operational purposes within the Iranian market. Second, previous studies (e.g., Anchuri, 2024, and MebalP et al., 2021) provided no automated parameter tuning to validate their models. However, we employed Optuna and Bayesian search techniques to develop an automated parametric optimization procedure to statistically enhance the performance of our models. Third, unlike other comparable studies, this study did not focus solely on the accuracy of the final predictive model but also contributed to the ability to derive useful, actionable managerial insights by evaluating feature importance. The ability to extract actionable business intelligence from these predictive models represents an important benefit to business decision-making capabilities. Finally, this study utilized the CRISP-DM methodology to clearly define each phase of the modelling process, including data understanding, data preparation, and data evaluation.

5.3 Theoretical and Managerial Insights

The research provides evidence via a theoretical framework that machine learning, and in particular, boosting algorithms such as LightGBM or CatBoost, are able to identify complex patterns in sales data related to consumer behavior. The predictive feature analysis is also a critical aspect of developing causal or behavior-focused models to accurately identify relationships between marketing attributes (e.g., discount amount) and demand.In addition to rapidly developing theoretical and empirical understandings of consumer behavior, the research provides empirical evidence that supports concepts of pricing and responsiveness to promotion.The practical applications of the research from a manager's perspective are numerous. First, the research reinforces the criticality of discounts and their ability to drive demand via efficient management of prices, which will allow managers to develop pricing strategies with greater consideration of other variables. Second, the research indicates that incorporating day-of-week and location-based impacts will create better opportunities for planning promotional campaigns and distributing inventory appropriately. Third, due to the ability to clearly develop and implement a conceptual framework related to this research, managers should be able to readily apply the research findings to their operational systems. Conditions presented by the research to supply chain, IT, and sales will also provide relevant data to help them make decisions on a daily basis. For the Iranian retail market, the research findings that identify price and discount as primary predictors suggest the importance of a pricing strategy that is adaptable to regional market dynamics.

In addition to this, it should be noted that the predictive accuracy obtained from the LightGBM model ($R^2 \approx 0.536$) is useful only in the context of real world FAST-MOVING CONSUMER GOOD business operations where demand is always affected by the fact that the consumer exhibits variation in their purchasing patterns, competition at the store level is very high, and promotional efforts drive out-of-stock conditions. In situations of high-level variation in considered sales, being able to explain over half of this variance in out-of-stock context is of operational value in today's retail environment. This level of predictiveness is sufficient to make informed tactical supply chain management decisions (i.e., replenishment timing, reorder point planning, promotional planning, etc.) in a demand

forecasting context. In the real world, even a marginally improved accuracy of demand forecasting creates substantial reductions in out-of-stock conditions, reduction and holding costs for inventory, and better utilization of limited storage and shelf space. For the aforementioned reasons, the model justifies not only statistical justification, but a tangible managerial benefit in terms of improved demand visibility reducing the gap between supply chain tactics and consumer purchasing activity.

5.4 Research Limitations

The present study offered several valuable insights but was limited in numerous ways. First, the data was limited to just one retail chain (Ofogh Kourosh) and occurred during just one calendar year. This factor may hinder how the results can be generalized to other companies or time periods. Second, external factors like macroeconomic effects, special events (e.g., holidays or national celebrations), or large advertising campaigns were not considered in the dataset. These could have a significant effect on demand which was discounted from the modelling at this point.

Third, the modelling process utilized only structured data; unstructured or behavioral data (such as customer reviews, online searches, or social media activity) were not incorporated for this analysis. Fourth, while the models provided excellent forecasting performance, there was no scenario analysis to consider how the models would behave under atypical conditions or market shocks (e.g., pandemics or currency swings). Moreover, the dataset used in this study is limited to sales transactions from physical retail stores and does not incorporate online or omnichannel sales data. As consumer purchasing behavior increasingly shifts toward digital and hybrid channels, the absence of online sales information may restrict the generalizability of the findings to fully omnichannel retail environments.

5.5 Recommendations for Future Research

In order to overcome the above limitations, various avenues for future research are proposed:

1. Utilizing multi-year and multi-store data to examine the extent to which models are robust and stable across time and space.
2. Considering external data sources (e.g. holidays, promotional campaigns, weather, exchange rates) in forecasting models to increase estimation accuracy.
3. Utilizing deep learning forecasting models (e.g. LSTM, Transformers) when faced with complex time series.
4. Putting in place causal frameworks to investigate managerial decisions (e.g. amount of discount) in relation to demand variations.
5. Conducting sensitivity and scenario analyses to assess the model's predictions' reaction to market shocks, or changes in policies.
6. Building intelligent dashboards that can capture and convey the forecasting results to non-technical managers and automate data-informed managerial responses.

This allows taking advantage of real-world data, cutting-edge machine learning algorithms, and a clear organization, helping to create a localized and scientific framework to forecast FMCG demand in Iran. Not only did the model demonstrate satisfactory accuracy, but the feature analysis given also offered managerial insights. Hopefully, this framework can be deployed throughout the retail, logistics, and marketing industries and inspire future studies in demand forecasting and data-driven decision-making.

References

Abolghasemi, M., Gerlach, R., Tarr, G., & Beh, E. (2019). Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1909.13084

Adler, A. I., & Painsky, A. (2022). Feature Importance in Gradient Boosting Trees with Cross-Validation Feature Selection. *Entropy*, *24*(5), 687. https://doi.org/10.3390/e24050687

Afshar, F., Seyedabrishami, S., & Moridpour, S. (2022). Application of Extremely Randomised Trees for exploring influential factors on variant crash severity data. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-15693-7

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna. *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631. https://doi.org/10.1145.3292500.3330701

Anchuri, N. S. (2024). Machine Learning-Driven Demand Forecasting: A comparative analysis of advanced techniques and Real-Time integration. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, *10*(6), 1352–1361. https://doi.org/10.32628/cseit241061175

Barreñada, L., Dhiman, P., Timmerman, D., Boulesteix, A., & Van Calster, B. (2024). Understanding overfitting in random forest for probability estimation: a visualization and simulation study. *Diagnostic and Prognostic Research*, *8*(1). https://doi.org/10.1186/s41512-024-00177-1

Basavaraju, K., & Valilai, O. F. (2025). Developing a demand planning strategy for joint forecasting and employing analytical tool in an empirical case study. *Deleted Journal*, *7*(4). https://doi.org/10.1007/s42452-025-06740-9

Basson, L. M., Kilbourn, P. J., & Walters, J. (2019). Forecast accuracy in demand planning: A fast-moving consumer goods case study. *Journal of Transport and Supply Chain Management*, *13*. https://doi.org/10.4102/jtscm.v13i0.427

Carbonneau, R., Vahidov, R., & Laframboise, K. (2009). Forecasting supply chain demand using machine learning algorithms. In *Advances in intelligent information technologies series/Advances in intelligent information technologies (AIIT) book series* (pp. 328–365). https://doi.org/10.4018.978-1-60566-144-5.ch018

Chen, W., Yang, H., Yin, L., & Luo, X. (2024). Large-scale IoT attack detection scheme based on LightGBM and feature selection using an improved salp swarm algorithm. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-69968-2

Demirtürk, D., Mintemur, Ö., & Arslan, A. (2025). Optimizing LightGBM and XGBOOST algorithms for estimating compressive strength in High-Performance Concrete. *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13369-025-10217-7

Douaioui, K., Oucheikh, R., Benmoussa, O., & Mabrouki, C. (2024). Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical review. *Applied System Innovation*, *7*(5), 93. https://doi.org/10.3390/asi7050093

Durap, A. (2023). A comparative analysis of machine learning algorithms for predicting wave runup. *Anthropocene Coasts*, *6*(1). https://doi.org/10.1007/s44218-023-00033-7

Fatima, S. S. W., & Rahimi, A. (2024). A review of Time-Series Forecasting Algorithms for Industrial Manufacturing Systems. *Machines*, *12*(6), 380. https://doi.org/10.3390/machines12060380

Feizabadi, J. (2020). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, *25*(2), 119–142. https://doi.org/10.1080.13675567.2020.1803246

Fildes, R., Kolassa, S., & Ma, S. (2021). Post-script—Retail forecasting: Research and practice. *International Journal of Forecasting*, *38*(4), 1319–1324. https://doi.org/10.1016/j.ijforecast.2021.09.012

Fırat, A. T., Aygün, O., Göğebakan, M., Akay, M. F., & Ulus, C. (2024). Development of machine learning based demand forecasting models for the e-commerce sector. *Uluslararası Mühendislik Tasarım Ve Teknoloji Dergisi.*, *7*(1), 13–20. https://doi.org/10.70669/ijedt.1567739

Geeitha, S., Ravishankar, K., Cho, J., & Easwaramoorthy, S. V. (2024). Integrating cat boost algorithm with triangulating feature importance to predict survival outcome in recurrent cervical cancer. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-67562-0

Ghosh, D., & Cabrera, J. (2021). Enriched random forest for high dimensional genomic data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *19*(5), 2817–2828. https://doi.org/10.1109/tcbb.2021.3089417

Golabek, M., Senge, R., & Neumann, R. (2020). Demand Forecasting using Long Short-Term Memory Neural Networks. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2008.08522

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/s40537-020-00369-8

Ileri, K. (2025). Comparative analysis of CatBoost, LightGBM, XGBoost, RF, and DT methods optimised with PSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks. *International Journal of Machine Learning and Cybernetics*. https://doi.org/10.1007/s13042-025-02654-5

Jafarnejad, A., Rezasoltani, A., & Khani, A. M. (2025). Cost-sensitive machine learning for predicting production defects: A novel approach based on MetaCost. Research in Production and Operations Management, 16(2), 73–94. https://doi.org/10.22108/pom.2025.144489.1610

Jahin, M. A., Shahriar, A., & Amin, M. A. (2024). MCDFN: Supply chain demand Forecasting via an explainable Multi-Channel Data Fusion network model integrating CNN, LSTM, and GRU. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2405.15598

Khani, A. M., Kazazi, A., & Taqhavi Fard, M. T. (2022). Evaluating the quality of services of the cultural and social deputy of Tehran Municipality in the field of culture and art. Social Development & Welfare Planning, 13(50), 205–250. https://doi.org/10.22054/qjsd.2021.58035.2110

Khani,A. M. , Rezasoltani,A. , Arjmandpour,S. , Jafarnjad,A. and Hosseinian,S. H. (2025). The Impact of Total Quality Management and Visual Quality on Customer Satisfaction and Loyalty in the Apparel Industry: A Hybrid Approach Using PLS-SEM and SHAP. *Journal of Advertising and Sales Management*, *6*(2), 153-176. https://doi.org/10.22034/asm.2025.2064300.3407

Lai, J., Lin, Y., Lin, H., Shih, C., Wang, Y., & Pai, P. (2023). Tree-Based Machine Learning Models with Optuna in Predicting Impedance Values for Circuit Analysis. *Micromachines*, *14*(2), 265. https://doi.org/10.3390/mi14020265

Lai, L., Lin, Y., Liu, Y., Lai, J., Yang, W., Hou, H., & Pai, P. (2024). The Use of Machine Learning Models with Optuna in Disease Prediction. *Electronics*, *13*(23), 4775. https://doi.org/10.3390/electronics13234775

Le, H., Sang, V. N. T., Thuy, L. N. L., & Bao, P. (2023). The fuzzy Kullback–Leibler divergence for estimating parameters of the probability distribution in fuzzy data: an application to classifying Vietnamese Herb Leaves. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-40992-y

Lee, K. H., Abdollahian, M., Schreider, S., & Taheri, S. (2023). Supply Chain Demand Forecasting and Price Optimisation Models with Substitution Effect. *Mathematics*, *11*(11), 2502. https://doi.org/10.3390/math11112502

Lu, J., Li, J., Ren, J., Ding, S., Zeng, Z., Huang, T., & Cai, Y. (2022). Functional and embedding feature analysis for pan-cancer classification. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.979336

Meaney, C., Wang, X., Guan, J., & Stukel, T. A. (2025). Comparison of methods for tuning machine learning model hyper-parameters: with application to predicting high-need high-cost health care users. *BMC Medical Research Methodology*, *25*(1). https://doi.org/10.1186/s12874-025-02561-x

MebalP, A., S, H., SJ, J., & M, M. (2021). Predicting the Demand for Fmcg using Machine Learning. *International Journal of Engineering and Advanced Technology*, *10*(3), 169–171. https://doi.org/10.35940/ijeat.c2253.0210321

Mehregan, M. R., & Khani, A. M. (2024). Improving organizational performance: The role of supply chain 4.0 and financing in reducing supply chain risk. Journal of International Business Administration, 7(3), 39–59. https://doi.org/10.22034/jiba.2024.60005.2164

Mitra, A., Jain, A., Kishore, A., & Kumar, P. (2022). A Comparative Study of demand Forecasting Models for a Multi-Channel Retail Company: A novel hybrid Machine learning approach. *Operations Research Forum*, *3*(4). https://doi.org/10.1007/s43069-022-00166-4

Nweje, N. U., & Taiwo, N. M. (2025). Leveraging Artificial Intelligence for predictive supply chain management, focus on how AI- driven tools are revolutionizing demand forecasting and inventory optimization. *International Journal of Science and Research Archive*, *14*(1), 230–250. https://doi.org/10.30574/ijsra.2025.14.1.0027

Olutimehin, N. D. O., Nwankwo, N. E. E., Ofodile, N. O. C., & Ugochukwu, N. C. E. (2024). STRATEGIC OPERATIONS MANAGEMENT IN FMCG: A COMPREHENSIVE REVIEW OF BEST PRACTICES AND INNOVATIONS. *International Journal of Management & Entrepreneurship Research*, *6*(3), 780–794. https://doi.org/10.51594/ijmer.v6i3.935

Oyeyemi, N. O. P., Anjorin, N. K. F., Ewim, N. S. E., Igwe, N. a. N., & Sam-Bulya, N. N. J. (2024). The influence of supply chain agility on FMCG SME marketing flexibility and customer satisfaction. *International Journal of Applied Research in Social Sciences*, *6*(10), 2546–2563. https://doi.org/10.51594/ijarss.v6i10.1665

Perumallaplli, R. (2025). Machine learning approaches for improving supply chain efficiency and demand prediction. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5228503

Rezasoltani, A., Jafarnejad, A., & Khani, A. M. (2025). A voting-based hybrid machine learning model for predicting backorders in the supply chain. Journal of Decisions and Operations Research, 10(1), 194–213. https://doi.org/10.22105/dmor.2025.511401.1924

Rezvan, P. H., Comulada, W. S., Fernández, M. I., & Belin, T. R. (2022). Assessing alternative imputation strategies for infrequently missing items on multi-item scales. *Communications in Statistics Case Studies Data Analysis and Applications*, *8*(4), 682–713. https://doi.org/10.1080.23737484.2022.2115430

Rizkallah, L. W. (2025). Enhancing the performance of gradient boosting trees on regression problems. *Journal of Big Data*, *12*(1). https://doi.org/10.1186/s40537-025-01071-3

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process model. *Procedia Computer Science*, *181*, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Shakur, M. S., Lubaba, M., Debnath, B., Bari, A. B. M. M., & Rahman, M. A. (2024). Exploring the challenges of Industry 4.0 adoption in the FMCG sector: Implications for Resilient Supply Chain in Emerging economy. *Logistics*, *8*(1), 27. https://doi.org/10.3390/logistics8010027

Tripathi, S., Muhr, D., Brunner, M., Jodlbauer, H., Dehmer, M., & Emmert-Streib, F. (2021). Ensuring the robustness and reliability of Data-Driven Knowledge discovery models in production and manufacturing. *Frontiers in Artificial Intelligence*, *4*. https://doi.org/10.3389/frai.2021.576892

Wang, Q. (2025). A Hybrid Transformer-ARIMA model for forecasting global supply chain disruptions using multimodal data. *International Journal of Advanced Computer Science and Applications*, *16*(1). https://doi.org/10.14569/ijacsa.2025.0160153

Watanabe, S. (2023). Tree-Structured Parzen Estimator: Understanding its algorithm components and their roles for better empirical performance. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.11127

Wiemer, H., Drowatzky, L., & Ihlenfeldt, S. (2019). Data Mining Methodology for Engineering Applications (DMME)—A holistic extension to the CRISP-DM model. *Applied Sciences*, *9*(12), 2407. https://doi.org/10.3390/app9122407

Wiens, M., Verone-Boyle, A., Henscheid, N., Podichetty, J. T., & Burton, J. (2025). A tutorial and use case example of the eXtreme Gradient Boosting (XGBOOST) artificial intelligence algorithm for drug development applications. *Clinical and Translational Science*, *18*(3). https://doi.org/10.1111/cts.70172

Wu, R. M. X., Shafiabady, N., Zhang, H., Lu, H., Gide, E., Liu, J., & Charbonnier, C. F. B. (2024). Comparative study of ten machine learning algorithms for short-term forecasting in gas warning systems. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-67283-4

Xu, X., Xia, L., Zhang, Q., Wu, S., Wu, M., & Liu, H. (2020). The ability of different imputation methods for missing values in mental measurement questionnaires. *BMC Medical Research Methodology*, *20*(1). https://doi.org/10.1186/s12874-020-00932-0

Yang, D., & Zhang, A. N. (2019). Impact of information sharing and forecast combination on Fast-Moving-Consumer-Goods demand forecast accuracy. *Information*, *10*(8), 260. https://doi.org/10.3390/info10080260

Yani, L. P. E., & Aamer, A. (2022). Demand forecasting accuracy in the pharmaceutical supply chain: a machine learning approach. *International Journal of Pharmaceutical and Healthcare Marketing*, *17*(1), 1–23. https://doi.org/10.1108/ijphm-05-2021-0056

Zheng, Y. (2024). Application of Machine Learning Algorithms in Enterprise Supply Chain Demand Forecasting. *2024 2nd International Conference on Design Science (ICDS), 1-4.*, 1–4. https://doi.org/10.1109/icds62420.2024.10751682

Zhou, W., Yan, Z., & Zhang, L. (2024). A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-55243-x

Zohdi, M., Rafiee, M., Kayvanfar, V., & Salamiraad, A. (2022). Demand forecasting based machine learning algorithms on customer information: an applied approach. *International Journal of Information Technology*, *14*(4), 1937–1947. https://doi.org/10.1007/s41870-022-00875-3