INTERPRETABLE PREDICTIVE ANALYTICS FRAMEWORK FOR INVENTORY FORECASTING IN SMALL RETAIL BUSINESSES

PALLAV RAJPUT

Research Proposal

AUGUST 2025

Abstract

The growth of large retailers and quick-commerce platforms has changed customer expectations, pressuring small businesses to adapt to their operations. There is a rise in customer demand for quick ordering. These Q-commerce stores, which are part of big enterprises poses direct competition to small businesses. Because they are not technically expert and do not have a framework to estimate overstocks and understock in the conditions where external factors affect the sales. This inefficiency causes waste and stockouts that reduce sales and customer trust, and they move to the competitors.

This research aims to develop a practical model that help small retailers in meeting the inventory needs with greater accuracy and better understanding. The study will use the Retail Store Inventory Forecasting Dataset that has around 73000 daily records, to build a framework that integrates both classical time series models and regression-based machine learning approaches. The process will involve cleaning the data, creating new features, and including outside factors such as holidays or competitor's pricing that often change sales patterns. The study aims to build a system that cuts waste and keep stock available, and the results must be understandable for shopkeepers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AIC …… Akaike Information Criterion

AI …… Artificial Intelligence

ARIMA …… AutoRegressive Integrated Moving Average

ARIMAX …… AutoRegressive Integrated Moving Average with Exogenous Variables

BI …… Business Intelligence

BIC …… Bayesian Information Criterion

CV …… Cross Validation

DL …… Deep Learning

DS …… Data Science

EC2 …… Elastic Compute Cloud

EDA …… Exploratory Data Analysis

GBM …… Gradient Boosting Machine

IQR …… Interquartile Range

KPI …… Key Performance Indicator

MAE …… Mean Absolute Error

MAPE …… Mean Absolute Percentage Error

MASE …… Mean Absolute Scaled Error

ML …… Machine Learning

MSE …… Mean Squared Error

OOS …… Out of Stock

RF …… Random Forest

RMSLE …… Root Mean Squared Logarithmic Error

RMSE …… Root Mean Squared Error

S3 …… Simple Storage Service

SARIMA …… Seasonal AutoRegressive Integrated Moving Average

SES …… Simple Exponential Smoothing

SHAP …… SHapley Additive exPlanations

XGBoost …… Extreme Gradient Boosting

……

1. Background

Inventory planning involves predicting product demand and sales. Customers now expect consistent product availability, which make it is critical for small retailers or shops who operate with limited storage, competition and fluctuating because of customer demand. According to (Govindarajan et al., 2019), small businesses can lose up to 15% of revenue due to poor forecasting and competition. While methods such as moving average or exponential smoothing remain common, they often overlook factors like promotions and discounts or rapid changes in demand.

My interest in this problem came while observing local retailers struggling in my community with inventory decisions and competition Another issue is that these small businesses usually have sales data across other of their stores, which is often irregular, making the traditional models less effective and requires adjustments to their old framework.

While large retail gaints use advanced ML-driven with high-end techniques, smaller retailers do not have affordable and practical solutions. Regression-based ML models like Gradient Boosting (Friedman, 2001), Random Forests (Breiman, 2001) demonstrate strong forecasting potential, but most of these studies rely on large-scale datasets, that leaves a gap in applicability to these small business environments. Also forecasting for small shop must consider local factors like festivals, weather conditions, compititor's pricing, since these have a much stronger effect on sales as compared to large retailer organizations (Feddersen and Cleophas, 2024). Methods such as exponential smoothing (Hyndman and Athanasopoulos, 2021) work for stable patterns but fails to identify effect of promotions or sudden change in demand.

To address this gap, I propose developing a predictive framework which can identify extenal influencing factors, can handling multiple variables and their linear relationships accurately and most importantly present the results in a form that is easily understandable for store managers. The research mainly focuses on two methods: regression based ML models time-series forecasting techniques, such as ARIMA (Box et al., 2015) and integrates by including the interpretability methods like SHAP (Lundberg and Lee, 2017), which is important for our research subjected to these non techical users. Mixing these approaches allows us to handle outside influences and find relation between factors. This interpretability is important where people who usually lack advanced technical skills, want to adopt the method.

2. Literature Review and Problem Statement

Prior research has evaluated both regressions based and traditional time series forecasting approaches to address this. While past studies highlight useful methods, still they have gaps concerning small-scale retail environments.

(Daruvuri et al., 2025) explored demand prediction using XGBoost (Chen and Guestrin, 2016) for price and sales forecasting, showing that Tree based ML techniques can capture non-linear patterns better than the simple regression. Which also showed that XGBoost performed better than the other models. However, the study relied on big retail organisations with detailed data that makes it less useful for small retail owners where we can try combined approaches, by adding factors like promotions or competitor pricing.

(Jenčová et al., 2025) applied ARIMA and Holt-winters models for retail finance forecasting successfully captured the seasonal trends. Their work was more focused on long-term market patterns with financial related indicators rather than product-based inventory management, which require daily predictions. Still, these findings confirm how important is seasonality, which our framework surely will cover.

(Krishna et al., 2019) tested multiple regression and ensemble techniques that included Linear Regression, polynomial regression, ridge, lasso, adaboost and Gradient Boosting for sales forecast, across multiple stores, where Gradient boost performed best and proved that ensemble models are powerful when sufficient data is there. The study focused on data with multiple large numbers of stores, but there is no study for smaller datasets.

Recent work also shows that tree-based techniques can also perform better than deep learning approaches like LSTM for retail demand prediction (Nasseri et al., 2023).

(Khan et al., 2020) combined predictive models like Decision Trees and SVM with BI dashboards, which enabled real-time monitoring making easier for organization to act quickly on business-oriented results. The approach was more effective in corporate environment; hence it requires advanced infrastructure that shops or stores requires.

(Gomes et al., 2025) compared equivalent temperature methodologies in energy demand forecasting. They combined weather data with regression models like Random Forest and ARIMA and found that including the weather data improved the accuracy of energy forecasting models. For our study, this shows that demand can be influenced by outside events. Adding variables like holidays/promotions helps making the predictions better.

(Madhukumar et al., 2022) evaluated a total of nineteen different regression models at university campus. Models included Linear regression, Decision Tree, SVM, Gaussian process regression models and many more, where Gaussian Process Regression gave the most accurate forecasts. Hence testing and comparing several models instead of relying on single proved to be a better alternative. The findings implies that when having store focused data like in our case with small stores, a precise model selection can still produce better results.

This also aligns the findings for parsimonious forecasting models that are simple yet effective (Petropoulos et al., 2021)

(Kalhori and Beliakov, 2025) studied how to make electricity demand forecasts better. They used a method known as the Choquet integral which gave better forecasts than simple methods that suggests the advanced ways of combining models can manage uncertainty better than older methods. For retailers, this we can combine different forecasting models in such a way that it could state both sales trends and external factors.

(Yeesin and Kajornkasirat, 2024) examined sales data from a 'Thai multi-store' retailer to compare traditional forecasting methods. Using ARIMA family models they evaluated performance through MSE and MAE, where SARIMA found to be the most effective model for handling seasonal fluctuations and regional demands highlighting how this model family can significantly improve forecasting accuracy in multi-store environments. The study is based on larger retail operations, which limits direct applicability for our smaller shops with less complex data.

(Bajoudah et al., 2023) studied sales forecasting in E-commerce store using data set from year 2018 to 2022 and compared the ARIMA and Prophet models and tested the ways to handle missing data by forward and backward filling. Where prophet was easier to use and fasters to code, the ARIMA gave better result. The study was more focused on online selling hence it does not reflect the challenges the small store owners face.

(Takahashi and Goto, 2021) used shipment data from baking factories and focused on only a single product that was 'bread'. They applied a method known as LINE to predict potential sales of OOS products. It performed better than t-SNE. This study was only focused on products with shorter life cycles, hence limiting them to the retail category.

(Edwardo and Ruldeviyani, 2020) predicted the mini market sales using deep learning methods and evaluating by MAE and RMSE. The study was more focused on predicting the sale targets. The study focused on larger mini market chains not on smaller independent retailers, as in our case.

(Zheng et al., 2024) ‘s study focused on pricing strategy for fresh vegetables in grocery store using techniques like ARIMA and linear equation. Combined approach increased the profit and reduced waste and found a negative relation between sales and up costing. We also have such data for competitor pricing that could help us understand the factors affecting if we have decreased sales. The study only focuses on fresh products like perishable goods; our case is not only focused on it.

(Unni et al., 2025) predicted E-commerce data and created an optimized XGBoost model to predict the sales failure. The model had a good result than LightGBM and Catboost, capturing nonlinear relations effectively. The model was accurate across different user groups. The study was only focused on E-commerce transactions limiting their applicability to small offline stores.

(Senthilkumar Renuka et al., 2025) developed a retail model focused on short term delivery (Q-commerce). The study compared machine learning techniques like Naïve Bayes, Decision Tree, Random Foresters and SVM. The results were better for Decision Tree and Random Forest whereas SVM reduced the false positive in classification. These retail stores work on quicker sales cycles, and therefore these studies are less applicable to our requirements.

(Li and Shi, 2025) predicted rural E-commerce data and proposed a model called TIART Tri-Stage Integrated Attention Residual Transformer. The model handled the noisy and high dimensional data and identified key drivers like family income and logistic performance. It achieved a good prediction accuracy that was better than XGBoost and other ML models. The study focused on policy planning for ruler development hence very different from our case.

Table 1. Literature Review Summary of Forecasting Models Addressing Research Gaps

Approach	What previous studies found	What we will do differently
Random Forest Regressor	Studies like (Krishna et al., 2019) and (Khan et al., 2020) found that Random Forest can capture non-linear demand patterns and handle noisy data. Models were done on large datasets and have not focused on interpretability.	We will use Random Forest for smaller data sets and use SHAP values to make results transparent and understandable for our non-technical store owners.
Gradient Boosting - XGBoost, LightGBM, CatBoost	(Daruvuri et al., 2025) and (Unni et al., 2025) proved the high accuracy of boosting models in retail and e-commerce forecasting. Models can identify relations which are complex but are mostly tuned for large datasets.	We can use boosting models to small store sales data and additionally use SHAP and residual analysis making predictions both accurate and interpretable.
ARIMA, SARIMA, SARIMAX.	The models were widely used in studies like (Yeesin and Kajornkasirat, 2024) and (Zheng et al., 2024) for capturing trends and seasonality Models perform well for regular demand but often fail when sudden promotions or weather changes occur.	We will test ARIMA family models on small store daily sales data along with external factors (promotions, weather, etc.) to see if they can remain useful in retail.
Prophet	(Bajoudah et al., 2023) showed Prophet as simple to implement and quick to uncover the trends and holidays, but it was sometimes less accurate than ARIMA.	In our study, Prophet will be used as an evaluation criteria for comparing the ease of use and speed of more complex models, especially for retailers with less technical knowledge.
Hybrid Studies	(Makridakis et al., 2018) and (Gomes et al., 2025) showed that combining models or adding external factors improves robustness, but their work mainly focused on large-scale or different domains like energy and finance.	Our workflow will have a consistent comparison of time series and regression ML models in the small retail inventory setting, filling a gap in current research.

Problem Statement

Small retail stores usually struggle with inventory forecasting as they operate with a limited stock, as many of them have budget limitations. There is also irregular demand due to events such as holidays, weather conditions etc. The large retailers have advanced analytics teams and very detailed datasets, whereas these stores usually manage inventory manually or use simple offline or digital workbooks, this often leads to overstocking of low demand goods or running OOS (Out-of-Stock) of high selling items. These challenges reduce their profit, cause a lot of waste (for perishable items) and unavailable items damage the customer's trust and miss the sales goals. Current models focus either on accuracy (deep learning, which is too complex for this study) or on interpretability (using ARIMA), but few approach can balance both in the context of small retailers which have less data. Hence, a practical and more interpretable forecasting framework is required that works with small data to help their businesses to manage their inventories.

This study addresses three gaps:

1. Most forecasting solutions are designed for large-scale datasets, leaving a requirement for small retailers. This study specifically focuses on such environments. Testing models on small retail-based datasets will ensure this gap will be filled and can reduce wasted stock while keeping other items in demand available, that will directly increase profits and customer trust.
2. Advanced models can achieve good accuracy, but they are not that transparent, which makes it difficult for non-technical users to adopt in everyday business. To overcome it, the framework integrates SHAP along with residual analysis, making forecasting interpretable and easy to understand. In simple language the framework will create user-friendly insights for retail owners.
3. Direct comparisons between statistical and machine learning models in small business settings are still rare, even though broader studies (Makridakis et al., 2018) finds that traditional statistical models perform better, and even better than modern ML methods in many cases. This study will run a consistent evaluation of classical and ML models along with highlighting the tradeoffs.

Table 2. Comparison of Classical Forecasting and Machine Learning Models.

Category	Models	Strengths	Limitations
Time Series Models	ARIMA, ARIMAX, SARIMA, Prophet	These ML models are best at finding seasonal trends and are interpretable.	May struggle with sudden demand shifts or irregular data, may require adaptations like stationarity.
Machine Learning Regression Models	Random Forest, Gradient Boosting Techniques	Handle complex and non-linear relationships, adapt to diverse features, high accuracy potential.	Less interpretable with risk of overfitting with small datasets, require more tuning.

The literature review for this study was done using IEEE Xplore, Google Scholar and Kaggle datasets by using keywords like 'retail forecasting' and 'time series forecasting for perdictive analysis'.

3. Research Questions

1. Which traditional time series methods like ARIMA models family, Prophet (Taylor and Letham, 2018) and the regression-based ML techniques like Random Forest, Gradient Boosting, could provide a good balance of easy understanding at best and a better accuracy in prediction for small retail inventory management?
2. How incorporating external factors like holidays/promotions etc. can improve forecasting compared to models using only historical sales data?
3. What are the trade-offs between using classical time series models and regression-based ML approaches when they are used to their environments with limited data and technical resources?

4. Aim and Objectives

Aim: To develop a forecasting framework that minimizes inventory waste and stockouts in small stores, like grocery shop businesses, and helps in reducing customer churn by improving customer satisfaction.

This will ensure that small retailers can retain both store efficiency and customer trust. By focusing on simplicity, the framework also makes it easier for non-technical users to understand and rely on predictions for a better inventory stocking.

Objectives:

- Clean and prepare the data before analysis to identify errors, missing values and outliers is retail data.
- To create features from past sales and integrating external influences or regressors. Finding most meaningful predictors and selecting effective variables.
- Bridging the gaps by comparing regression methods with classical forecasting methods to identify trade-off between complexity, accuracy and ease of implementation.
- Evaluate model performance and apply interpretability methods so that results are accurate and understandable for small businesses.

5. Significance of the Study

This study is significant because it tackles issues like predicting demand that have usually less data, mismanagement and errors that could result in missing revenue goals and reducing customer satisfaction. This research will introduce a framework that compares time series and regression-based ML approaches, offering both precise results and easy to understand models. Helping retailers to reduce losses and avoid items being out of stock and helps them to gain more customer trust.

Main Contributions of the Proposed Method Summarized: -

- Comparable Forecasting - Making it easy to compare or combine both types of model approaches, giving small retailers more reliable predictions to save capital waste and increase product availability.
- Transparent and Interpretable prediction - Using methods can create a hybrid framework, so store owners can understand the reason why sales are expected to rise or fall, so that they can make a confident decision.
- Uncovering External Factors - External factors are essential to determine like change in weather and holiday/promotions, as they put most impact on the retail sales like domain.

- Practical and Affordable solutions using only free options like open-source tools, and free analytic tools, and explain with output with proper guidelines so that it becomes usable for non-technical users.

6. Scope of the Study

The research focuses on the goal of daily inventory forecasting for small shops or stores and would be using 'Retail Store Inventory Forecasting' dataset with 73000 entries.

The study will incorporate external regressors like competitors pricing, weather etc. there might be a need of creating synthetic data in case if the data is unavailable.

Evaluation will be done using metrics like MAE, RSME and MAPE along with handing interpretability utilising SHAP technique.

For engineering the study will be done using simple tools, the work will be conducted in Python for coding and visualized using Tableau.
And the framework created would make the application accessible, affordable and prioritizing ease of understanding so that these retailers would not require advanced technical help, along with making proper guideline available for practical implementation.

Out of Scope
The Study won't be covering certain areas to focus on scope and maintain clarity and excludes using deep learning due to data and complexity. Real-time inventory systems and on ground trails with local business, extension to other domain with similar inventories like health clinics, energy management at residential societies can be left as future scopes.

7. Research Methodology

Inventory forecasting the core of the research and research methodology plays a very vital role in setting up the road map of whole evaluation. In the context of small retail business, the accurate forecasting is not the only goal, we also must consider the limitation of business. A good methodology in our context will make sure that process is consistent and not too technical to miss out the constrains related to inventory management.

7.1 Overall Workflow

The research will be carried out in a sequential workflow to build an inventory forecasting framework for small retail businesses. The process is initiated with the dataset from Kaggle that have around 73k of daily records. The data will be cleaned and pre-processed to prepare for the model training. There might be a need to merge the data into aggregated table of sales and products.

Then the EDA will be performed on the data to identify sizes and portions of different variables, and to identify for any anomalies before we train the data and begin with feature engineering.

Feature engineering will then be applied to construct features that could include day, week or biweekly data, particularly the lag and rolling features. Also, the external indicators would be taken care of like holidays, weather condition can competitors' data. Once we have the dataset prepared, both time series forecasting models and regression-based ML models will be used and compared.

The evaluation of these models would be done by using error matrices. In addition to accuracy of models, the interpretability will be handled. The workflow is planned so that it not only gives accurate results but also keeps the process clear enough for non-technical person to follow.

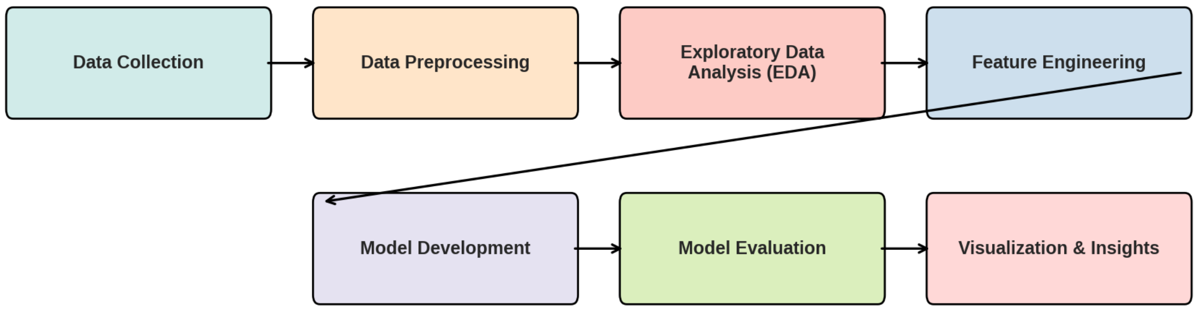


Fig 1. Basic Flowchart of Methodology

7.2 Dataset description

The dataset used here is the Retail Store Inventory Forecasting Dataset from Kaggle (Anirudh, 2023), which contains around 73000 entries of daily sales across multiple stores and products. Each entry includes:

- 'Date': That has daily records from starting to ending date.
- 'Store ID & Product ID': Are unique identifiers for stores and products repectively.
- 'Category': The product categories such as Electronics, Groceries, Clothing, etc.
- 'Region': The geographic region of the store.
- 'Inventory Level': Stock units available at the beginning of the day.
- 'Units Sold': Stock units sold during the day.
- 'Units Ordered': Quantity of products ordered to restock.
- 'Demand Forecast': The predicted demand based on past trends.
- 'Price': Selling price of the product on tha day
- 'Discount' Promotional discount applied
- 'Weather Condition': Report of daily weather impacting the sales.
- 'Holiday/Promotion': Indicators for holidays or promotions.
- 'Competitor Pricing': Pricing of compititor's product.
- 'Seasonality': Indicating seasonal patterns.

This data is sufficient as it is very specific to small retail and provides product level details of more than 70k daily entries, which is necessary for inventory level forecasting. The dataset also includes external factors such as competitors pricing, weather, seasons all of which strongly influence retail demand. And this kind of simple data allow the analysis of price, sales and the impact of the market trends and changes easily as compared to larger and complex datasets.

7.3 Data Preparation

The Kaggle's Retail Store Inventory Forecasting Dataset is already in a structured format without any gibberish or unformatted values. However, it is still important to look for missing values and outliers as a part of data cleaning and outlier detection. We can perform first EDA on raw data (using python or Tableau) to understand the data size and proportion of different variables and we can have a detailed EDA after cleaning the data. Then we can perform feature engineering with train and split Here are the steps with details: -

7.3.1 Data Cleaning:

The missing values in the data would be handled using simple imputation techniques, as it a time series field we can use forward fill. Median or mode importation can be used for categorical attributes. Extreme outliers will be checked with IQR (Interquartile Range) and if we found any uncertainty of their presence, they can be removed.

7.3.2 Data Transformation:

Categorical variables like store ID and product ID will be converted into numerical labels, while continuous features may be standardized depending on the models. As we are using the ARIMA model family and tree-based ML models, we can scale data using standardization for continuous variable.

7.3.3 Feature Engineering

Feature engineering is essential for our forecasting study, where historical data may be limited. To maximize information, the following transformations are applied:

- Lag Features: Past sales values are included as predictors, like sales from the previous day, last week, bi-weekly sales. It will help model to understand short-term based demands.
- Rolling Features: Moving averages and rolling standard deviations are generated between 7 day to 30 days of windows helping clearing out random spikes and highlight underlying trends to observe long term demand.
- External Regressors or Factors: We will include indicators for holiday/ promotion and even weather conditions, which usually cause change in demand .

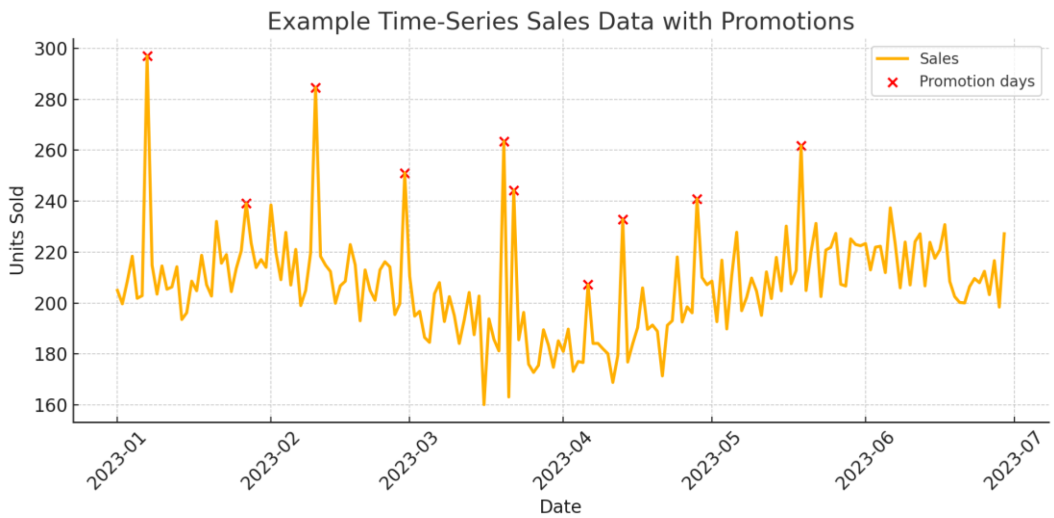
To reducing dimensionality (reducing complexicity) less important features like temporal indicators (dates based) will be removed after feature importance analysis to keep model simple and to save computational cost.

7.4 Exploratory Data Analysis (EDA)

EDA helps in gaining insights and deriving relationships among features. We can observe the trends and patterns by looking at various type of different graphs. The following types of EDA will be performed to understand the dataset:

- Univariate Analysis -uses single variable to check the distribution of sales, promotions and product-level demand using histograms, boxplots, line chart etc.
- Bivariate and Multivariate Analysis -use more than one variable to derive relationships between sales and other influencing factors such as promotions, holidays, store, or product IDs by using scatterplots, pair plots and correlation heatmaps.
- Visualization of Time-Series Trends: helps observing trends, seasonality and cyclic patterns through line plots and decomposition like shown in Figure 2.
- Outlier and Anomaly Detection: For outlier detection this could be performed on the cleaned data to identify unusual spikes or drops in sales with boxplots, IQR and any seasonal anomaly detection.
- Correlation: for determining corelation among numerical variables and external regressors like holidays, promotions, lag features.

Figure 2 derived from our own dataset (Retail Store Inventory Forecasting) shows how daily sales fluctuate with both seasonal patterns and sudden promotion spikes capturing the real challenges small retailers face in managing inventory explaining why the external factors must be included in inventory forecasting.



\

Fig 2. Daily Sales Trends with Seasonal Variation and Promotion Effects.

7.5 Model Development

Model development is the stage where the prepared data is turned into actual forecasting models. Here, both classical time-series methods and machine learning techniques will be used, tested and compared. To check which approach works best. The aim is to compare model's strengths and trade-offs to build a reliable framework for small retail business. Two sets of models will be developed and compared:

7.5.1 Time Series Forecasting Models

- ARIMA and SARIMA

These models are easy to understand and have been used in retail industry for a very long time. ARIMA (Autoregressive Integrated Moving Average) Is one of the most widely used time series models and it is well suited for our dataset to see short term patterns.

$$\boldsymbol{ARIMA\ Model}$$

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta q \varepsilon_{t-} q + \varepsilon_t$$

SARIMA extends ARIMA by adding seasonality which is essential to see weekly, biweekly or monthly cycles in sale, capturing demand during seasonal events. In (Yeesin and Kajornkasirat, 2024) SARIMA was the best-performing model for multi-store demand forecasting, outperforming ARIMA and SARIMAX on criteria like AIC and MAE.

$$\boldsymbol{SARIMA\ Model}$$

$$ARIMA(p, d, q) \times (P, D, Q)_s$$

- SARIMAX

SARIMAX moves one step forward by adding external regressors such as holidays and competitor pricing, all of which are included in our dataset. This is critical in small retail, where an external event like a local holiday or competitor discount can affect regular demand. By using SARIMAX, we test whether including these external influences could make forecasts more realistic. (Zheng et al., 2024) used ARIMA with exogenous factors to forecast vegetable pricing,

$$\boldsymbol{SARIMAX\ Model}$$

$$SARIMAX(p, d, q) \times (P, D, Q, s)$$

- Prophet

Prophet is a model which is developed by Facebook. It is very simple to implement and can explain forecasts best with holiday/promotion effects. This would also work as benchmark for speed. (Bajoudah et al., 2023) compared Prophet with ARIMA in e-commerce sales, found it easier but slightly inaccurate.

7.5.2 Regression Machine Learning Models

To handle non-linear interactions between the features in our data set, these machine learning models could be our best choice. Time series models mainly focus on seasonality, but these Regression Models can process variables like sales lag, discounts, competitors and other. The models will be trained to capture complex relationships among variables. Hyperparameters will be tuned using GridSearchCV to ensure optimal performance.

- Linear Regression (Baseline)

We can start with simplest supervised machine learning approach. This will act as a baseline to evaluate other models with other Regression models.

$$\boldsymbol{Linear\ Regression}$$

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \cdots + \beta nXn + \varepsilon$$

- Random Forest Regressor

An ensemble learning technique that generates the mode of predictions by building multiple decision trees. It doesn't require extensive feature transformation to capture non-linear effects. Since our dataset reflects inventory stocking related conditions that may have some noise, RF is also noise-resistant, which is useful.

$$\boldsymbol{Random\ Forest}$$

$$\hat{y} = (1/T)\ \Sigma\ (t = 1\ to\ T)\ h_t(x)$$

- Gradient Boosting Models (XGBoost, LightGBM, CatBoost)

These models are some of the best that could work for retail data. They just don't make predictions but fix the mistake in the next one, capturing very minor details.

XGBoost: Handles share features well and can model complex relationships between sales lags, discounts and promotions. (Unni et al., 2025) used optimized XGBoost for e-commerce sales failure prediction, achieving 93% accuracy,

LightGBM: The model is known to be very fast and efficient when dealing with many features. This is useful for our dataset if we create a lot of extra features for rolling averages and lagged sales.

CatBoost: Specially designed to handle categorical data (like Store ID or Product Category) without needing too much extra processing. This makes it simpler and less time-consuming to use.

$$\boldsymbol{Gradient\ Boosting}$$

$$F_{\mathrm{m}}(x) = F_{\mathrm{m-1}}(x) + \eta \cdot h_{\mathrm{m}}(x)$$

7.6 Evaluation Metrics

The Evaluation Metrice is how we can judge if the prediction or forecasting models are actually accurate and useful by telling the degree of accuracy attained and how close or far are the predictions or forecasting are from the actual.

The models will be compared on following metrics:

- Error Metrics: Forecasting accuracy will be numerically measured using following techniques:

MAE (Mean Absolute Error): Gives the average degree of error and it is the best for communicating results to a non-technical user.

$$\boldsymbol{Mean\ Absolute\ Error}\ (\boldsymbol{MAE})$$

$$MAE = (1/n)\ \Sigma\ (t = 1\ to\ n)\ |\ y_{\mathrm{t}} - \hat{y}_{\mathrm{t}}\ |$$

RMSE (Root Mean Squared Error): By focusing on large errors, RMSE can evaluate large mistakes like missing holiday/promotion. And we can compare performance with such criticalities.

$$\boldsymbol{Root\ Mean\ Squared\ Error\ (RMSE)}$$

$$RMSE = \sqrt{((1/n)\ \Sigma\ (t = 1\ to\ n)\ (y_t - \hat{y}_t)^2)}$$

MAPE (Mean Absolute Percentage Error)

When comparing errors across categories and scales, this is most helpful evaluation. Error is expressed in percentage.

$$\boldsymbol{Mean\ Absolute\ Percentage\ Error\ (MAPE)}$$

$$MAPE = (100/n)\ \Sigma\ (t = 1\ to\ n)\ |\ (y_t - \hat{y}_t)\ /\ y_t\ |$$

- Interpretability metrics: SHAP (Shapley Additive Explanations) values will be used to explain the importance of features in machine learning models, especially for Random Forest and Gradient Boost. This would help in transparency by explaining the factors that can drive our target variable.

$$\boldsymbol{SHAP\ Value}$$

$$\varphi_i = \Sigma\ (S \subseteq F \setminus \{i\})\ [\ (|S|!\ (|F| - |S| - 1)!)/|F|!\ ]\ \cdot\ [\ f(S \cup \{i\}) - f(S)\ ]$$

7.7 Practical Relevance

This ensures models will remain reliable for retail use. The analysis will be done in Python using libraries such as pandas, scikit-learn, seaborn, Statsmodels, XGBoost and Prophet. For clear and generate actionable insights, visualization tools such as Tableau will be used. The goal is to build an accurate, transparent, reproducible forecasting framework. We will expect the results that provides accuracy and also deliver actionable insights for our their need. This combination of framework would allow the small retailers to benefit from this hybrid forecasting regression technique without heavy technical requirements.

Figure 3 shows the step-by-step research workflow, from cleaning of data to building and testing the Models.

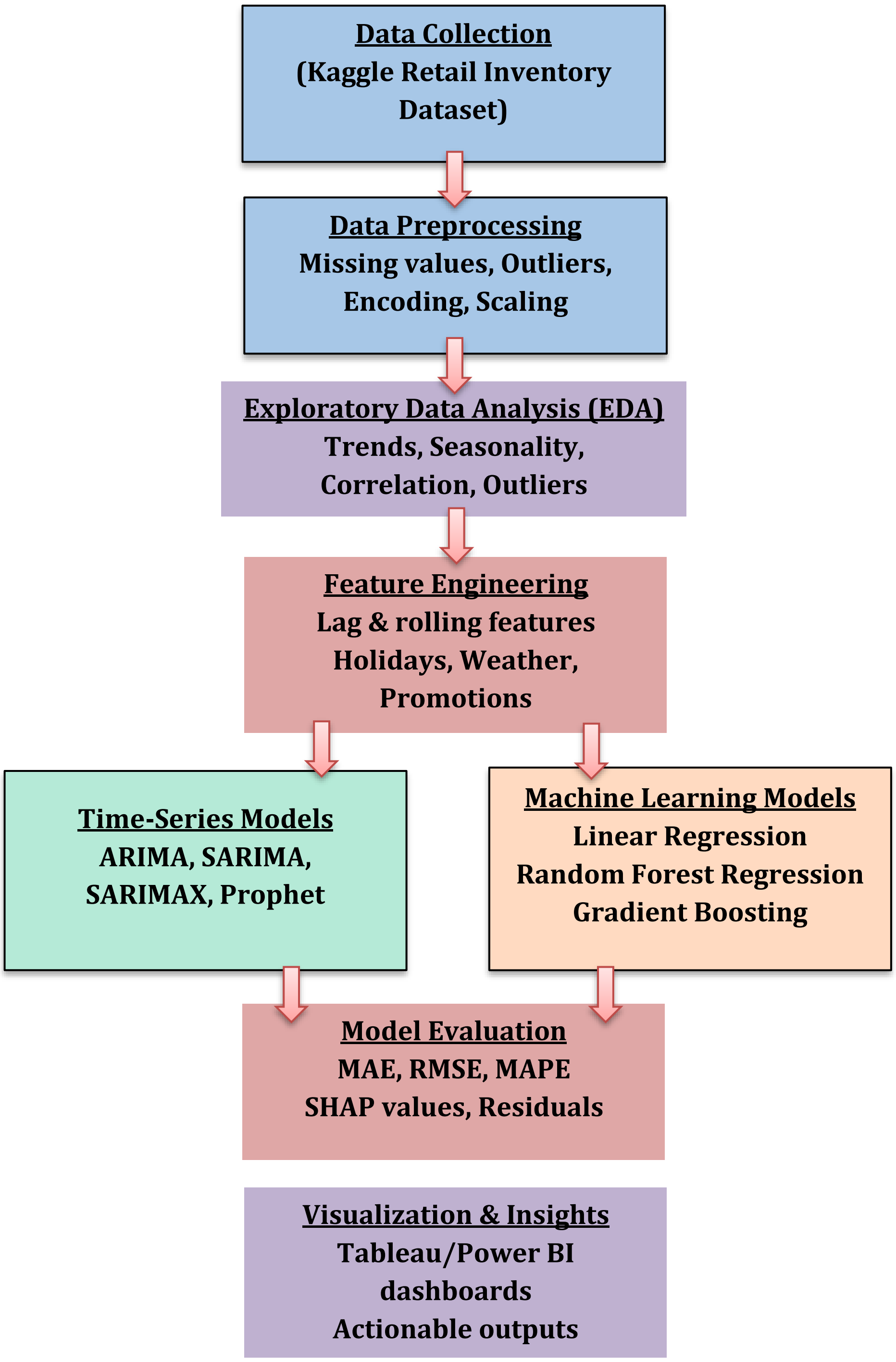


Fig 3. Detailed flow chart of methodology

8. Requirements and Resources

Table 3. Software Requirements

	Version / Type
Operating System	Windows 10/11, Linux (Ubuntu 20+), macOS 11+
Program	Python 3.9 or above, Tableau / PowerBI… (for visualization)
Python Library	Basic: (Pandas, NumPy), ML Models: (scikit-learn, Statsmodels, XGBoost, Prophet), Visualization: Matplotlib, Seaborn
Documentation	Jupyter Notebook, JupyterLab, Google Colab
Development Environment	Microsoft Word (with Mendeley Cite Addon) Mendeley (for Harvard-style referencing)

Table 4. Hardware Requirements

	Version / Type
Processor	Intel Core i5 (10th gen+), AMD Ryzen 5+, Apple M1 or latest
Storage	128 GB or above SSD, Google Drive, One drive (for cloud storage)
RAM	8 GB minimum, 16 GB recommended
GPU	NVIDIA GPU (GTX 1650+) with at least 4 GB VRAM.
Cloud Compute (Optional)	Lower processor CPU/GPU to run web-browser or CLI, to access Azure ML Studio or AWS Sagemaker or AWS EC2 instance + S3

Table 5. Dataset Requirements

	Requirement
Source	Retail Store Inventory Forecasting Dataset (Kaggle, 2023) ~73,000 daily records of multiple stores and products.
Key Attributes	Date, Store ID, Product ID, Category, Region, Inventory Level, Units Sold, Units Ordered, Demand Forecast, Price, Discount, Weather Condition, Holiday/Promotion, Competitor Pricing, Seasonality, already explained in Dataset Description 7.2.
Model Input	Feature engineering (lag & rolling features), model training, evaluation

9. Research Plan

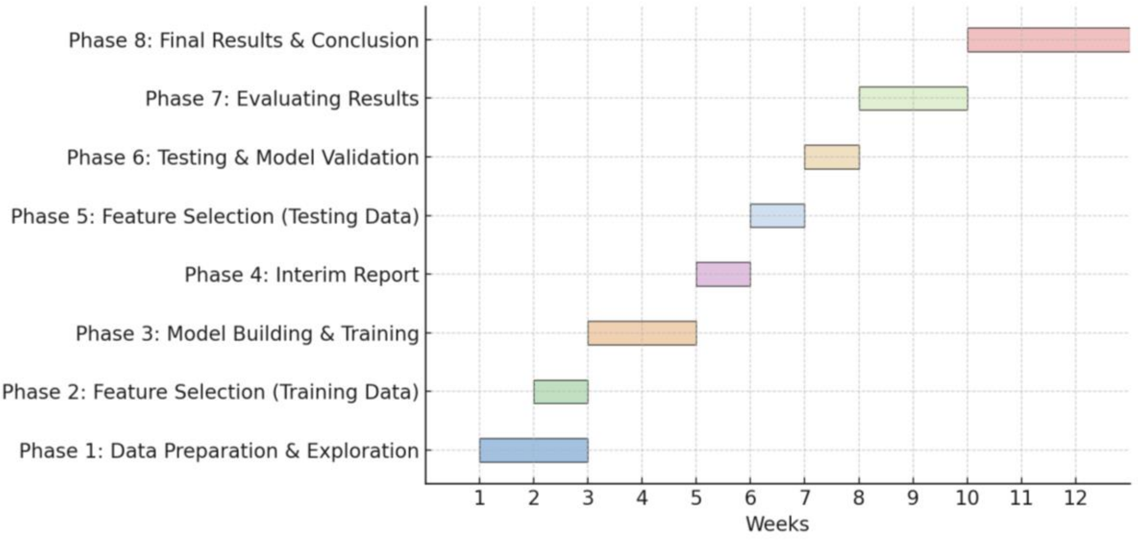


Fig. 4 Research Plan Gantt Chart

Phase 0 : The research was started with indentification of Research Interest, Literature review for topic submission and Research topic submission and its approval, that had already took almost 8 weeks. The current process is drafting the research proposal.

Phase 1 (Weeks 1-2): Data Preparation & Exploration : Data preprocessing and performing EDA providing a clean and structured data to identify the anomalies and meaningful patterns.

Phase 2 (Week 2): Feature Selection (train) : The significant features would be identified. And only most relevant features would be selected to avoid complexity.

Phase 3 (Weeks 3-4): Building and Training the Models : Both of the considered ML techniques will be trained and applied. And we will compare different approaches.

Phase 4 (Week 5): Interim Report :Documenting the progress up to this point. The report will consist of initial results or any changes made in methodology.

Phase 5 (Week 6): Feature Selection for Testing : Just like the training data, the test data also needs to go through feature selection so that only the most important variables are kept.

Phase 6 (Week 7): Testing and Model Validation : The trained models will be tested on unseen data to check accuracy.

Phase 7 (Weeks 8-9): Evaluating the Results : Model performance will be measured using MAE, RMSE and MAPE. And interpretability tools like SHAP will also be applied.

Phase 8 (Weeks 10-12): Final Results and Conclusions : Merging final results, insights and recommendations for small retailers, including a video explaining the project.

Risk & Contingency Plan

Dataset may may not fully capture acutal small retail conditions, since it comes from Kaggle and does not talk about it's origin.
Contingency: Add features that reflect more real world scenerios or look for a relevant dataset.

Resource Constraints - Running complex models can stress the weaker hardware. And it might take a very long time in running codes.
Contingency: Leveraging cloud services like AWS or Azure or optimizing code for efficiency.

Timeline Delays - Unexpected technical issues, data challenges or professional commitments may cause the project to fall behind schedule.
Contingency: Setting smaller goals and effective time management.

References

Bajoudah, A., Alsaidi, M. and Alhindi, A., (2023) Time Series Forecasting Model for E-commerce Store Sales Using FB-Prophet. In: *2023 14th International Conference on Information and Communication Systems (ICICS)*. pp.1–6.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., (2015) *Time series analysis: Forecasting and control*. Hoboken, NJ: Wiley.

Breiman, L., (2001) Random forests. *Machine Learning*, 451, pp.5–32.

Chen, T. and Guestrin, C., (2016) XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16)*. San Francisco: ACM, pp.785–794.

Daruvuri, R., Patibandla, K.K. and Mannem, P., (2025) Data Driven Retail Price Optimization using XGBoost and Predictive Modeling. In: *Proceedings of the International Conference on Intelligent Computing and Control Systems, ICICCS 2025*. Institute of Electrical and Electronics Engineers Inc., pp.220–225.

Edwardo, T.O. and Ruldeviyani, Y., (2020) Sales Prediction Analysis in Determining New Minimarket Stores. In: *2020 International Conference on Information Technology Systems and Innovation (ICITSI)*. pp.79–84.

Feddersen, A. and Cleophas, C., (2024) Hierarchical neural additive models for interpretable demand forecasts. *arXiv preprint*. [online] Available at: https://arxiv.org/abs/2404.04070.

Friedman, J.H., (2001) Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 295, pp.1189–1232.

Gomes, J.P.R., Santos, F.F.G. Dos, Mota, D.D.D., Assis, F.N. De, Vieira, D.A.G., Reis, T.C.B., Mendonca, M.D.O., Silva, G.R.L., Tavares, L.D., Lisboa, A.C. and Pereira, F.A.C., (2025) Calculation methodologies for equivalent temperature applied to energy demand forecasting. *IEEE Access*, p.1.

Govindarajan, V., Lev, B., Srivastava, A. and Enache, L., (2019) The gap between large and small companies is growing. Why? *Harvard Business Review*. [online] Available at: https://hbr.org/2019/08/the-gap-between-large-and-small-companies-is-growing-why.

Hyndman, R.J. and Athanasopoulos, G., (2021) *Forecasting: Principles and practice*. [online] OTexts. Available at: https://otexts.com/fpp3/.

Jenčová, S., Vašaničová, P., Košíková, M. and Miškufová, M., (2025) A Time Series Approach to Forecasting Financial Indicators in the Wholesale and Retail Trade. *World*, 61, p.5.

Kalhori, M.R.N. and Beliakov, G., (2025) A Fuzzy Measure and Choquet Integral-Based Approach in the Predictive Knowledge-Based Systems: Application to the Electricity Demand Forecasting. *IEEE Transactions on Fuzzy Systems*, 334, pp.1379–1390.

Khan, M.A., Saqib, S., Alyas, T., Ur Rehman, A., Saeed, Y., Zeb, A., Zareei, M. and Mohamed, E.M., (2020) Effective Demand Forecasting Model Using Business Intelligence Empowered with Machine Learning. *IEEE Access*, 8, pp.116013–116023.

Krishna, A., Akhilesh, V., Aich, A. and Hegde, C., (2019) Sales forecasting of retail stores using machine learning techniques. In: *IEEE International Conference on Machine Learning Applications*. [online] IEEE, pp.1–6. Available at: https://ieeexplore.ieee.org/abstract/document/8768765.

Li, H. and Shi, N., (2025) Predicting Priority Factors for Rural E-Commerce Development Using Tri-Stage Integrated Attention Residual Transformer. *IEEE Access*, 13, pp.61578–61593.

Lundberg, S.M. and Lee, S.-I., (2017) A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems (NeurIPS)*. [online] Available at: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.

Madhukumar, M., Sebastian, A., Liang, X., Jamil, M. and Shabbir, M.N.S.K., (2022) Regression Model-Based Short-Term Load Forecasting for University Campus Load. *IEEE Access*, 10, pp.8891–8905.

Makridakis, S., Spiliotis, E. and Assimakopoulos, V., (2018) Statistical and machine learning forecasting methods: Concerns and ways forward. *PLoS ONE*, 133, p.e0194889.

Nasseri, S., Kraus, M. and Feuerriegel, S., (2023) Applying machine learning in retail demand prediction: A comparison of tree-based ensembles and LSTM. *Applied Sciences*, 1319, p.11112.

Petropoulos, F., Fildes, R., Goodwin, P. and Boylan, J.E., (2021) Parsimonious forecasting: The case for simplicity. *European Journal of Operational Research*, 2912, pp.520–534.

Senthilkumar Renuka, S.R., C A, S., D R, S., Liviya L, M., Jayakumar, S.S. and Anuradha, R., (2025) Quick Commerce: Product Price Classification Using Machine Learning Algorithms. In: *2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)*. pp.1–6.

Takahashi, K. and Goto, Y., (2021) Forecasting Potential Sales of Bread Products at Stores by Network Embedding. In: *2021 5th IEEE International Conference on Cybernetics (CYBCONF)*. pp.114–119.

Taylor, S.J. and Letham, B., (2018) Forecasting at Scale. *The American Statistician*, [online] 721, pp.37–45. Available at: https://doi.org/10.1080/00031305.2017.1380080.

Unni, G.A., A, A., Sabu, H. and Krishnan, M.S., (2025) Optimized Learning for E-Commerce Sales Failure Prediction: An XGBoost Perspective. In: *2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)*. pp.746–751.

Yeesin, K. and Kajornkasirat, S., (2024) Application of Machine Learning in Multi-Store Retail Demand Forecasting: A Case Study of Tuenjai Panit Company. In: *2024 International Conference on Intelligent & Innovative Practices in Engineering & Management (IIPEM)*. pp.1–5.

Zheng, Y., He, Z. and He, H., (2024) Research on Pricing Strategy of Fresh Vegetable Commodities in Grocery Stores Based on Linear Regression and ARIMA Model. In: *2024 International Conference on Computers, Information Processing and Advanced Education (CIPAE)*. pp.183–189.