PREDICTIVEMART ANALYTICS, A BIG MARTSALES PREDICTION SYSTEM

Assistant Prof. Mr. Sanjeev Soni
Department of Information Technology
ABES Engineering College Ghaziabad.
Dr. A.P.J. Abdul Kalam Technical University Lucknow
Sanjeev.soni@abes.ac.in

Mansi Rao
Department of Information Technology
ABES Engineering College Ghaziabad.
Dr. A.P.J. Abdul Kalam Technical University Lucknow
Mansi.20b0131110@abes.ac.in

Ishika Goel
Department of Information Technology
ABES Engineering College Ghaziabad.
Dr. A.P.J. Abdul Kalam Technical University Lucknow
Ishika.20b0131036@abes.ac.in

Kshitiz Verma
Department of Information Technology
ABES Engineering College Ghaziabad.
Dr. A.P.J. Abdul Kalam Technical University Lucknow
Kshitiz.20b0131174@abes.ac.in

Muskan Agrawal
Department of Information Technology
ABES Engineering College Ghaziabad.
Dr. A.P.J. Abdul Kalam Technical University Lucknow
Muskan.20b0131201@abes.ac.in

***Abstract*— PredictiveMart Analytics, which is a trailblazer in the field of retail sales forecasting, uses state-of-the-art machine learning to provide major marts with previously unheard-of levels of accuracy and insight. This ground-breaking research uses sophisticated algorithms to search through enormous datasets of past sales, consumer behavior, and outside variables in order to find hidden correlations and patterns. Regression models, time series analysis, and neural networks are some of the tools that PredictiveMart Analytics uses to produce incredibly accurate predictions of future sales trends. Big marts get a strategic edge in a number of important areas thanks to this all-encompassing approach to data analysis. Realizing precise inventory management allows for less waste and optimal stock levels. Real-time data-driven dynamic pricing strategies guarantee a competitive edge and optimize profit margins. A new level of reactivity and agility is added to overall business planning, enabling major marts to take advantage of every opportunity and adjust to market changes with ease.**

The arrival of PredictiveMart Analytics represents a sea change for the retail industry. Big marts are no longer in the dark regarding potential sales growth. Rather, they are outfitted with an advanced, data-driven system that facilitates well-informed decision-making, enhances operational effectiveness, and eventually boosts profitability. With the ability to transform the whole industry, this ground-breaking technology will put data power at the core of retail success. PredictiveMart Analytics is essentially a game-changer rather than merely a forecasting tool. It provides huge marts with a glimpse into the future, enabling them to predict and manage changes in the market, make the most use of their resources, and ultimately prosper in the dynamic realm of retail.

Keywords—Predictive Analysis, , Sales Forecast, Random Forest Grid, Linear Regression, Standard Scaler, Hyper Parameter Tuning.

I. INTRODUCTION

For retail to be successful over the long term, precise sales forecasts are essential due to the ever-changing consumer and market conditions. In this environment, PredictiveMart Analytics stands out as a shining example of innovation, offering a state-of-the-art way to rethink sales forecasting that is especially suited for giant marts. This intelligent system makes use of cutting-edge machine learning algorithms to evaluate past sales information, interpret consumer behavior, and account for outside factors. PredictiveMart Analytics' main goal is to produce incredibly accurate projections for future sales trends, giving major marts a tactical edge as they navigate the complexities of their business operations.

PredictiveMart Analytics is based on a comprehensive approach to data analysis that uses a variety of machine learning techniques to glean valuable insights from large datasets. Regression models are a useful tool for closely analyzing past sales data and identifying trends that improve forecasting accuracy. An other important element is time series analysis, which provides a dynamic perspective on how sales fluctuate over time by capturing temporal dependencies and patterns. Neural network integration increases the system's analytical power by revealing hidden correlations that may be missed by more conventional analytical techniques.

Beyond just being a tool for sales forecasting, PredictiveMart Analytics is important because it is a driving force behind a complete overhaul of huge marts' operational strategies. The system offers a strategic edge in inventory management by precisely forecasting demand trends, optimizing stock levels, and reducing the likelihood of overstock or understock. Big marts are able to connect their pricing strategy with market realities thanks to the sophisticated pricing optimization process, which takes into account consumer behavior insights and real-time market dynamics. Additionally, the data-driven insights cover general business planning and provide useful information for product launches, marketing campaigns, and expansion plans.

Essentially, PredictiveMart Analytics presents itself as a strategic partner for large retailers looking to prosper in a quickly changing retail landscape, rather than merely a technical fix. This technology provides organizations with a sophisticated and data-driven approach to sales forecasting, enabling them to make well-informed decisions, adjust to market fluctuations with ease, and improve their overall profitability and operational efficiency. We shall examine PredictiveMart Analytics methods in more detail, as well as the concrete advantages it offers massive mart operations in the sections that follow.

Problem Statement: Big Mart, a well-established retail chain with multiple stores, confronts the persistent challenge of accurately predicting sales and efficiently managing inventory across diverse product categories and store locations. The current inventory management system struggles with the accurate estimation of product demand, leading to either surplus stock or frequent stockouts. This discrepancy between predicted and actual sales results in increased costs due to excess inventory holding or lost revenue opportunities due to unmet customer demands. Therefore, the central problem addressed in this project is the inadequacy in sales forecasting and its subsequent impact on inventory management and profitability.

I. LITERATURE REVIEW

The emergence of machine learning (ML) techniques is causing a revolutionary change in the field of sales forecasting. In order to show the many uses and significant advantages of machine learning (ML) in forecasting future sales patterns, this paper summarizes the most important discoveries from previous studies.

Domain-Specific Forecasting: Research such as [1, 2, 4, 5] highlights the importance of customizing models for certain domains. Clothing sales, electric vehicle sales, fashion markets, and hypermarkets (Big Marts) all profit from tailored forecasting techniques that capture distinct market dynamics.

Temporal Dynamics and Time Series Analysis: Understanding the temporal dimensions of sales is a major area of study interest ([4, 5, 13, 14, 15]). It becomes clear that time series analysis is an essential tool for creating precise forecasts that take into account both short- and long-term sales trends.

Predictive analytics and data mining: Studies such as [6] emphasize the usefulness of data mining in revealing customer preferences and behavior. Businesses can enhance inventory management and customize marketing campaigns by comprehending demand patterns for particular products or categories.

Geographic Relevance and Market Dynamics: Forecasting models that are relevant to a given region are necessary due to regional variances and market idiosyncrasies ([1, 11]). Research such as [1] conducted in Baghdad and [11] in the Brazilian free market highlight how crucial it is to modify models to suit local conditions in order to make precise forecasts.

Hybrid models and ensemble methods: As discussed in [4, 8, 9], combining several forecasting methodologies through hybrid models has the potential to increase sales estimates' accuracy and resilience. This method overcomes the shortcomings of each algorithm by utilizing its strengths.

Interdisciplinary Viewpoints: Research like [10, 12] clearly demonstrate the potential of interdisciplinary approaches. Examining sales forecasting in cloud systems and fashion markets emphasizes the necessity of using a variety of approaches to address issues in different sectors of the economy.

Evaluation and Model Performance: Although not stated specifically, these studies most likely use evaluation measures such as Random Forest Grid and Linear Regression to evaluate the performance of the models and prediction accuracy.

In conclusion, machine learning has shown itself to be a potent instrument for transforming sales forecasting. Through the customization of models to particular domains, temporal dynamics analysis, and utilization of a variety of approaches, enterprises may realize substantial advantages concerning precision, adaptability, and well-informed decision-making. We may anticipate ever more complex and multidisciplinary methods to improve sales forecasting's predictive capacity as research advances.

III. PROPOSED SYSTEM

The proposed research provided the following procedures for forecasting the sales of different categories using the sales data from the retail business. Figure 1 shows the recommended schematic of the system's architecture. The process's several phases are listed below.

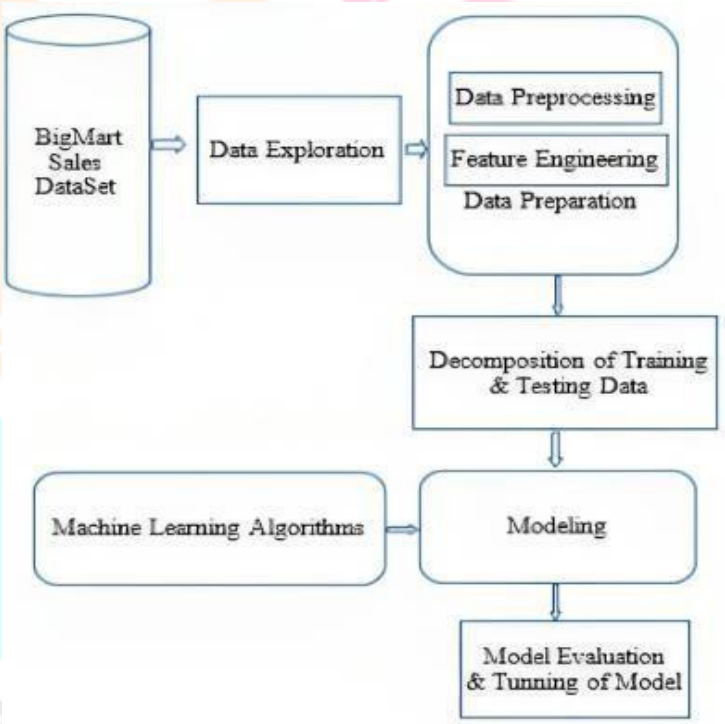


Fig-Architecture of sales predication framework

A. Hypothesis Generation:

The most crucial phase in the data analysis procedure is this one .This stage involves creating a lot of assumptions based on reading the problem description. The construction of the assumptions favours the intended outcome.

Examine the problem statement that follows: "The dataset acquired includes sales data from 10 stores in various cities for 1559 goods in 2013."

The objective is to develop a prediction model that will estimate the sales volume of each product at a given retailer." Consequently, this stage's main objective is to get knowledgeable about a product's attributes and the stores that might potentially impact sales. Because forecasting is based on products and stores, it is a little more difficult. The theories fall into two categories: "Store Level Hypotheses" and "Product Level Hypotheses." A number of essential elements should be taken into account, including brand, packaging, display area, usability, promotional offers, and shop exposure. Advertising and other product-level assumptions may also have an impact on sales.

While a number of factors, including population density, shop capacity, rivals, location, customer behavior, advertising, environment, and other store-level presumptions, might affect sales. For instance, branded goods sell better than non-branded goods. This increases the degree of trust that customers have inthe brand, which boosts sales.

Similar to this, we should anticipate better sales from establishments that are kept up well and run by kind, modest staff members since there will be more walk-in customers. Consequently, we have generated fifteen conjectures that aid inour understanding of the circumstances.

B. Data Exploration

Our goal is to improve accuracy whenever we consider a business problem by applying and upgrading different models.

We will emphasize, nevertheless, that there will come a point at which we will be unable to increase the accuracy of the model. These kinds of problems are solved by data investigation. Investigating the data set and learning as much as you can about the information that is available and speculative is the first step in the data exploration process. Based on our study, the dataset contains six characteristics that were hypothesized and found, three features that were theorized but not found, and nine features that were postulated but not found. The most accurate illustration of this is the figure in Figure 2. The dataset under evaluation has some missing values in the columns labeled "Outlet Size" and "Item Weight". Data values that are missing will be imputed during the data preparation step. We have separated the variables in our dataset into two groups: numerical variables and category variables. Table 2 provides a broad description of the numerical variables. The table allows for two fundamental conclusions to be made.

1. Item Visibility has a minimum value of 0. Problems arise when the number 0 indicates that the product is being sold even if it cannot be seen. Visibility in this instance should thus be greater than 0.

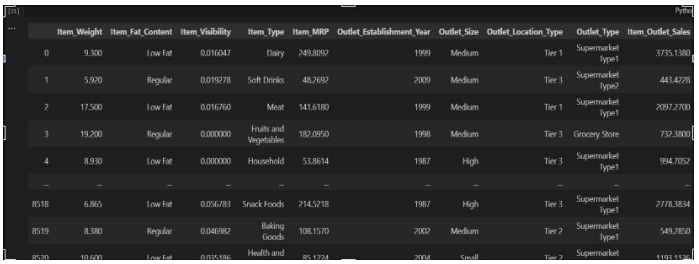
2. The alternatives for installing a plug It covers the years 1985 through 2009. During the pre-processing stage, these variables are converted to represent the age of a shop. The dataset has 10 distinct stores and 1559 unique commodities depending on thevarious categories.

After examining the frequency of different categories and analyzing the dataset, we arrived at the following results.

1. In the Item Fat Content category, certain "Low Fat" values are labeled "LF" and "low fat," whereas other "Regular" valuesare labeled "regular."

2. The Item Type category has sixteen subcategories.

These commodities include anything from food to beverages, meat to canned food, household goods to prescription drugs, and so on, suggesting that each store stocks a wide range of things.



	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet_Establishment_Year	Outlet_Size	Outlet_Location_Type	Outlet_Type	Item_Outlet_Sales
0	9.300	Low Fat	0.016047	Dairy	249.8092	1999	Medium	Tier 1	Supermarket Type1	3735.1380
1	5.920	Regular	0.019278	Soft Drinks	48.2692	2009	Medium	Tier 3	Supermarket Type2	443.4228
2	17.500	Low Fat	0.016760	Meat	141.6180	1999	Medium	Tier 1	Supermarket Type1	2097.2700
3	19.200	Regular	0.000000	Fruits and Vegetables	182.0950	1998	Medium	Tier 3	Grocery Store	732.3800
4	8.930	Low Fat	0.000000	Household	53.8614	1987	High	Tier 3	Supermarket Type1	994.7052
...	...	...	...	...	...	...	...	...	...	...
8518	6.865	Low Fat	0.056783	Snack Foods	214.5218	1987	High	Tier 3	Supermarket Type1	2778.3834
8519	8.380	Regular	0.046982	Baking Goods	108.1570	2002	Medium	Tier 2	Supermarket Type1	549.2850
8520	10.600	Low Fat	0.035186	Health and	85.1224	2004	Small	Tier 2	Supermarket	1193.1136

Fig- Data with different categories.

C. Data Pre-processing

This stage often involves managing any dataset outliers and imputes missing values. Both the Outlet Size and Item Weight columns in the dataset are empty. While outlet size is a categorical variable, item weight is a numerical variable.

Item Weight, which is then used to replace the missing data, is imputed using the sample mean weight of that specific item. Since Outlet Size is a categorical variable, we are unable to calculate the average; instead, we utilize the mode approach to fill in the blanks. Therefore, the missing numbers in the outlet size are found by determining the size mode depending on the kind of outlet.

When a Big Mart sales prediction system uses models like as standard scaler, hyperparameter tuning, Random Forest, and Linear Regression, data pre-processing is essential to guaranteeing the precision and effectiveness of the predictive models. First, in order to avoid skewing forecasts, the data must be cleaned up by removing outliers and missing values. This calls for the use of methods like robust statistical approaches for outlier detection and imputation for missing data. Next, in order to prepare categorical variables for numerical analysis, encoding methods such as label encoding and one-hot encoding are applied.

The next step is feature scaling, which uses the standard scaler to normalize numerical features and make sure they are all on the same scale. This keeps no one feature from controlling the training of the model. To further improve the predictive potential of the models, feature engineering—the process of creating new features from preexisting ones—is also essential. For example, developing new features such as average sales per item category or total sales per store might yield insightful data. Principal Component Analysis (PCA) is one dimensionality reduction technique that may be used to decrease the computational complexity of a dataset without sacrificing its vital information.

In order to assess model performance, the dataset is ultimately divided into training and testing sets. Cross-validation techniques are then employed to further guarantee robustness. In the end, this painstaking workflow for preparing data makes predictive models easier to apply later, which leads to more precise sales projections in the Big Mart setting.

The dataset goes through further preparation stages suited to the needs of each model after it has been divided into training and testing sets. In Random Forest and Linear Regression, the most pertinent features that contribute to the prediction task are found using feature selection techniques like Recursive Feature Elimination (RFE) or feature importance analysis. This reduces the dimensionality of the dataset and may enhance model performance and training time.

The next critical stage is hyperparameter tuning, which entails determining the ideal set of hyperparameters for each model using grid search or randomized search approaches. In order to determine the optimal configuration, this procedure entails systematically adjusting the hyperparameters and assessing the model's performance.

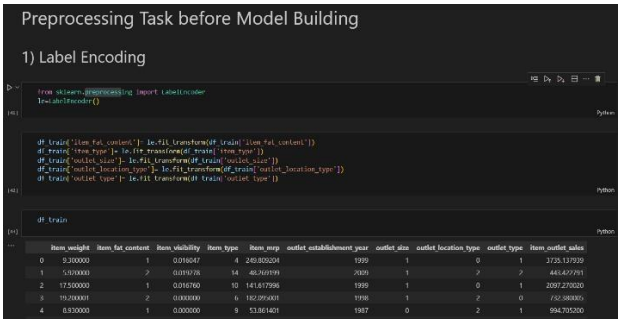


Fig- Data Pre-Processing.

D. Feature Engineering

In the process of exploring the data, we found a few oddities. In order to prepare our data for analysis, this step handles the nuances as well as creating new variables from existing variables. We were unaware that, throughout the data exploration phase, Item Visibility had a minimum value of 0. We resolved the issue by imputed the product's average visibility and treating it as a missing quantity.During the hypothesis development process, we included visibility as one of the store level hypotheses, based on the notion that sales increase with a product's visibility in the shop. Comparing a product's exposure at one location to its average visibility throughout all stores. Consequently, a new variable called Item Visibility MeanRatio is created to store these values. We discovered during the data exploration phase that Item Type contains 16 distinct categories; if these categories are grouped together by a comparable property, they could be helpful in the analysis. This functionality is best shown by the Item Identifier variable, which holds the Unique ID of every item. In this variable, the initials FD, DR, or NC stand for Food, Drinks, and Non-Consumables, respectively. Consequently, we expand the table by adding a new column called Item Type. Each product was classified as either Food, Drinks, or NonConsumables using the two columns Item Type and Item Identifier. The Outlet Establishment Year field contains the year the company originally opened. From this, we create a new variable that indicates the duration of the store's operation. The ages so varied from 4 to 28. We noticed that the Item Fat Content variable showed some variation in the way the categories were displayed. To remedy this, swap out Regular for Regular and Low Fat for LF. Non-consumables, nevertheless.

E. Segregation of Training and Testing Data

Feeding data into the algorithm and teaching it to spot patterns is one of the most important steps in machine learning.

To ascertain the algorithm's level of understanding, another dataset must be provided to it once it has learnt the pattern. For training and testing purposes, it is standard procedure to split the available data into two groups at a 4:1 ratio. To get better outcomes, though, we might have to change this ratio. Moreover, there may be variations in the ratio between regression techniques.

F. Model Building and Prediction

The procedure of Model Building and Prediction in a Big Mart Sales Prediction System entails using the Random Forest Grid and Linear Regression to produce an accurate and effective predictive model.

Because Random Forest with grid search can handle complex correlations and nonlinearities in the data, it is chosen during the model development phase. A thorough investigation of possible hyperparameter combinations is carried out using grid search, making it possible to choose the ideal values that will enhance prediction performance. In the meanwhile, the modeling of linear correlations between attributes and sales is made easier and more transparent with the use of linear regression. The Standard Scaler plays a critical role in maintaining numerical feature standardization and preventing scale disparities from allowing one feature to dominate the model training process.

Both models use the information gathered from the preprocessing stage to predict sales with accuracy. The efficacy of any model may be determined by comparing the forecasts with real sales data, which will help with future iterations and enhancements to the predictive algorithms. The sales forecasting system is made more reliable by this methodical approach to model construction and prediction, which also provides decision-makers with useful information for enhancing profitability and streamlining operations in the Big Mart context.

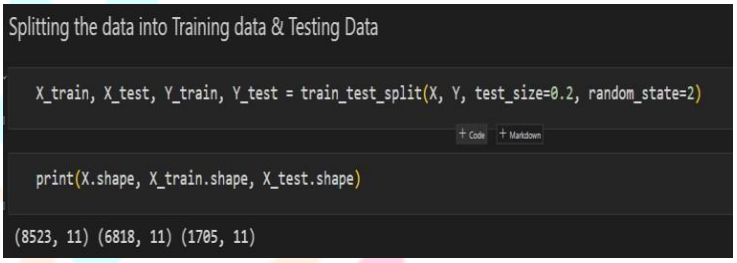


Fig – splitting the dataset in training and testing data.

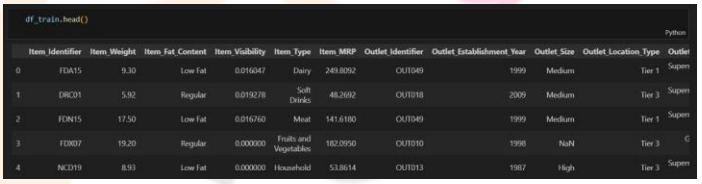


Fig – training the dataset

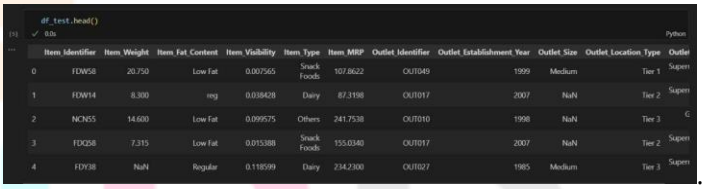


Fig- Testing thre dataset

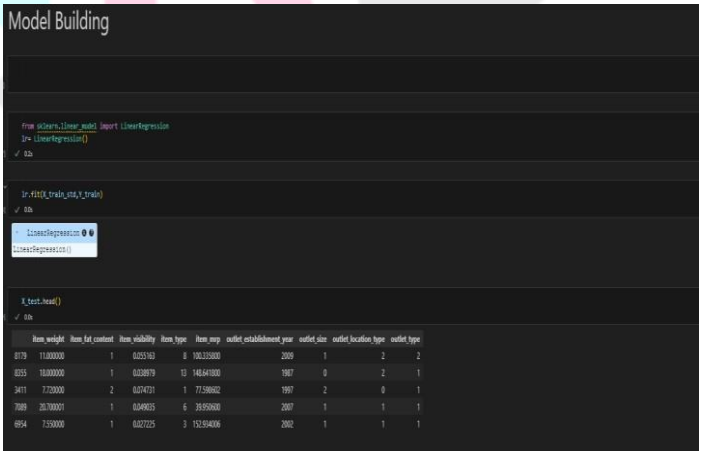


Fig- Model Building

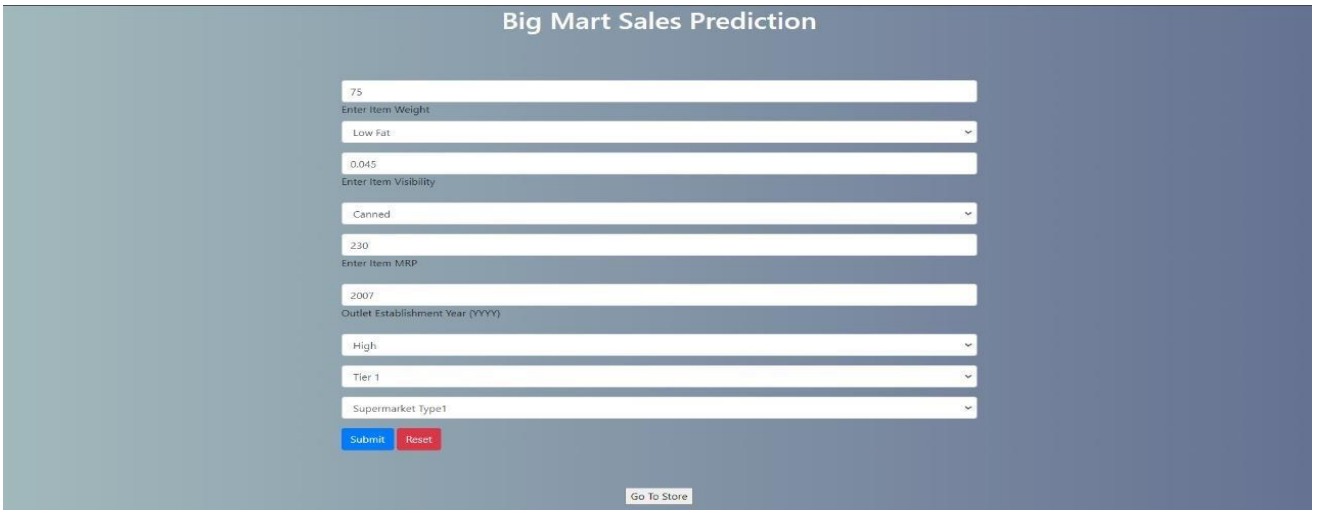


Fig- UI showing the inputs for prediction

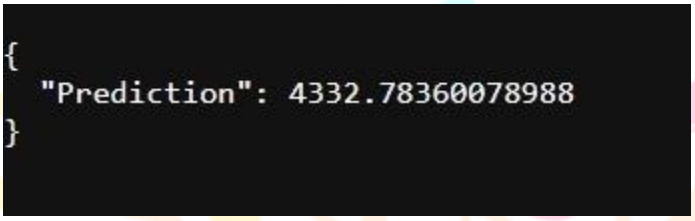


Fig- Output of the final Prediction

G.DATASET DETAILS

This BigMart sales dataset has 1559 goods distributed over 10 sites in different cities. There are 5681 trains and 8523 overall test records. The train dataset contains both input and output variables. The dataset has 12 characteristics that serve as this item's identifiers: This product code is exclusive to it. object Weight: The mass of the object. Item Fat Content: Indicates whether or not the product is low in fat. Item Visibility: The portion of a store's overall display space devoted to a particular item. Item Type: This describes how the product is categorized. The maximum selling price (list price) of a product is known as its MRP (maximum retail price). The product's distinct store ID is found on the outlet. Installation of the Outlet Year: The first year of the store's opening.

The store's square foot size is referred to as the outlet size. Outlet city: The name given to the store's location within a certain city. Whether it's a tiny food shop or a full-fledged supermarket, there's something for everyone. Product sales at a certain shop are known as item outlet sales. This. This is the variable that requires foresight.

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishment_Year	Outlet_Size
0	FDA15	9.30	Low Fat	0.016047	Dairy	249.8092	OUT049	1999	Medium
1	DRC01	5.92	Regular	0.019278	Soft Drinks	48.2692	OUT018	2009	Medium
2	FDN15	17.50	Low Fat	0.016760	Meat	141.6180	OUT049	1999	Medium
3	FDX07	19.20	Regular	0.000000	Fruits and Vegetables	182.0950	OUT010	1998	NaN
4	NCD19	8.93	Low Fat	0.000000	Household	53.8614	OUT013	1987	High

Fig- Dataset details

IV. RESULT AND DISCUSSION

The model's accuracy, the hardware's performance issue being addressed. By fine-tuning of the Random Forest during training and prediction, and the difficulty of the Grid and Linear Regression model's hyperparameters, which include the number of trees, learning rate, and regularisation parameters, the proposed system can be made more effective. This can aid in shortening the training period and enhancing the model's dimensionality accuracy to enhance the presentation. power of the model, pertinent factors including product exposure, shop size, and promotional events were created. Used scaling and normalizing methods to guarantee consistency in feature contributions.

4.1 Experimental Setup:

Data Collection: - Acquired a large dataset from Big Mart that included product details, sales history, and pertinent contextual information. Maintained data quality by addressing outliers, resolving missing values, and performing preprocessing procedures.

Feature Engineering: - To improve the predictive

Model Selection: - To determine the best model for sales forecast, a variety of machine learning techniques, including [list algorithms], were evaluated. Took into account elements like predicted performance, scalability, and interpretability.

Training-Validation Split: - Used a [insert percentage] training-validation split to test the model's performance on a different dataset after training it on historical data. Utilized cross-validation methods to guarantee generalization and robustness.

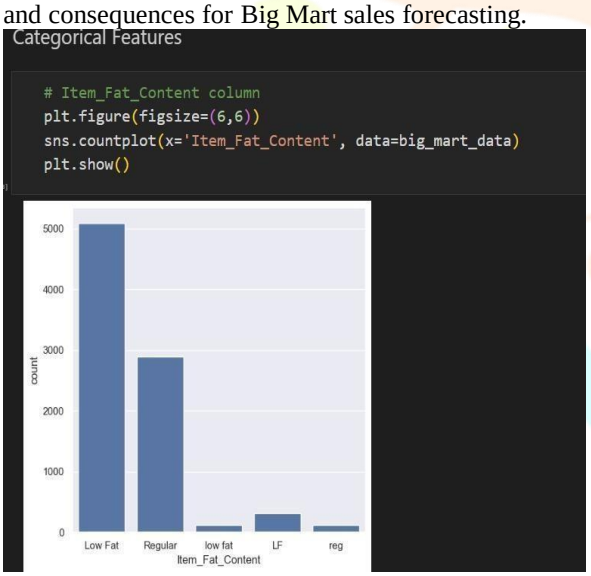
4.1 Performance Matrix to Evaluate the Proposed Methodology:

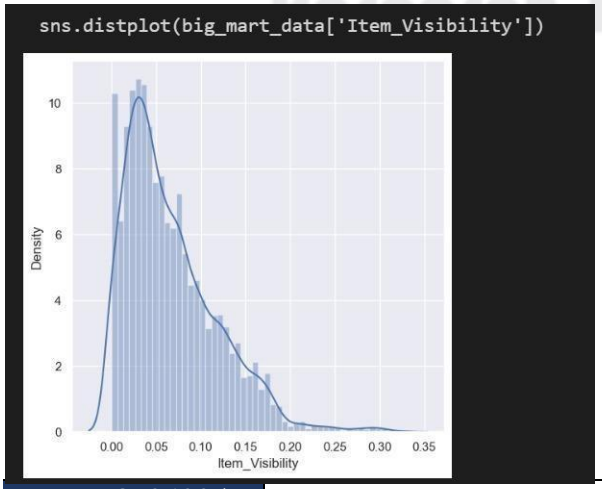
Accuracy Metrics: - Determined the model's accuracy, precision, recall, and F1-score in order to assess its predictive power in detail. To comprehend true positive, true negative, false positive, and false negative cases, the confusion matrix was examined.

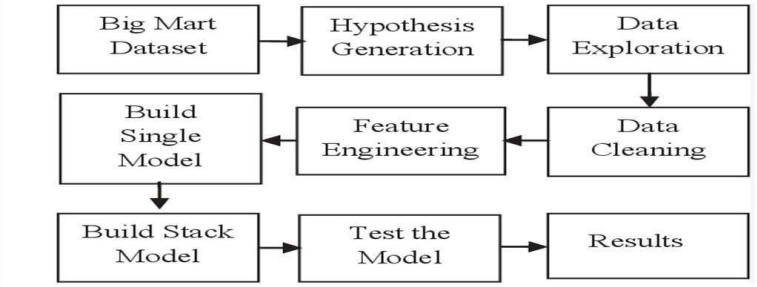
Regression Metrics: - Measured the model's predicted accuracy in terms of sales values using regression measures like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

Sensitivity Analysis: - Performed sensitivity study to see how each characteristic affects the results of predictions. Determined the critical elements and their relative relevance impacting sales projections.
Scalability Testing: - By gradually expanding the amount of the dataset, the suggested methodology's scalability was assessed. Evaluated resource needs and computational efficiency for managing bigger datasets.

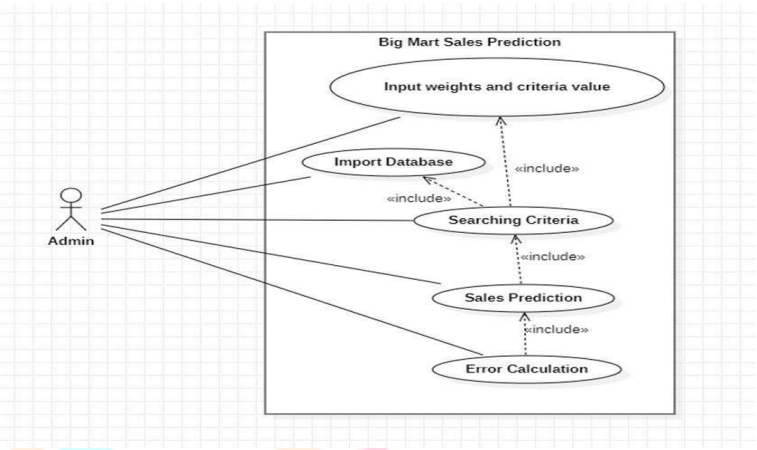
Overall Discussion: -
Model Robustness: - The results show how reliable the suggested technique is in forecasting Big Mart sales.
Interpretability: - The performance measures used boost interpretability by giving a clear picture of the model's advantages and shortcomings.
Temporal Dynamics: - Time-series analysis guarantees that temporal trends are captured and utilized by the model, which is necessary for precise sales forecasts.
Comparative Advantage: - Comparative study demonstrates the useful benefits of the suggested technique by proving its superiority over baseline models.
Scalability and Efficiency: - Scalability testing verifies the model's effectiveness in managing bigger datasets, demonstrating its applicability for dynamic, real-world settings.
Recommendations for Future Work: - Talk about possible improvements, such investigating sophisticated modeling strategies, utilizing outside data sources, or customizing the process for certain business scenarios. This dual explanation of the experimental design and performance assessment offers a thorough rundown of the approach, highlighting its advantages and consequences for Big Mart sales forecasting.







Project Flow Diagram



Use Case Diagram

FUTURE WORK: - By adding more data sources to the XGBoost model, like weather patterns and economic indicators, its accuracy can be increased. It entails extracting new features from the current dataset that could enhance the model's capacity for prediction. This can involve modifying already-existing features, fusing various features, or developing brand-new features from scratch using domain expertise.

V. CONCLUSION

5.1 Restate the Research Question or Hypothesis: Clarity is ensured and the study's main emphasis is reinforced by restating the research question or hypothesis. Regarding "PredictiveMart Analytics: A Big Mart Sales Prediction System," the clarification that might be provided is as follows:
Research Question: How can PredictiveMart Analytics effectively predict Big Mart sales, and what is the impact of its implementation on business operations?
5.2 Discuss the Implications: **Business Impact**: The business operations of Big Mart will be greatly impacted by the successful deployment of PredictiveMart Analytics. More accurate sales forecasts lead to better stockout prevention, better inventory control, and higher total profitability. **Decision-Making:** Decision-makers are empowered to make well-informed decisions, which optimize the allocation of resources and marketing strategies, thanks to the precise and timely projections. **Competitive Advantage**: The effectiveness and precision of the system provide Big Mart a competitive edge in the retail sector and establish the company as a data-driven, forward-thinking one.
5.3 Address Limitations: Transparent Acknowledgment: Recognize any obstacles that arose throughout PredictiveMart Analytics' creation and testing. This might involve places where the model has to be further refined, possible biases, or limitations in the data that are available. **Continuous**

Improvement: In future revisions of the system, address highlighted restrictions to demonstrate your commitment to continual improvement.

5.4 Suggest Future Research: **Advanced Modeling Techniques:** To further improve prediction accuracy, investigate the incorporation of deeper learning or more sophisticated machine learning approaches.**Incorporation of External Factors:** To capture a wider range of affects on sales, look into include external aspects like societal trends or economic statistics. **Industry-Specific Adaptations**: Take into account customizing the system for distinct retail industry sectors in order to meet particular needs and obstacles. **User Feedback and Iterative Development:** Get user input to guide iterative development and make sure PredictiveMartAnalytics continues to meet changing business requirements.

5.5 Reiterate the Significance: Contribution to the Field: Stress the importance of PredictiveMart Analytics as a useful tool for sales forecasting and how it has advanced predictive analytics in the retail industry. **Practical Applications:** Emphasize the system's useful uses and how its effective deployment satisfies industry demands for data-driven decision-making. **Long-Term Impact**: Stress how PredictiveMart Analytics might have a long-term influence on how retail sales prediction techniques are developed in the future.

Overall Conclusion: To sum up, PredictiveMart Analytics is a formidable instrument that has the ability to completely change the way Big Mart makes sales predictions. The research topic has been effectively addressed, along with the practical consequences, limits, and proposed areas for future research. Furthermore, the study has emphasized the value of its contributions to academia and industry. PredictiveMart Analytics is expected to have a significant impact on how predictive analytics is used in the retail industry as it develops.

VI. REFERENCES

[1] Anwer, Mussab Osamah, and Sureyya Akyuz. "Sales Forecasting of a Hypermarket: Case Study in Baghdad Using Machine Learning. "2022 30th Signal Processing and Communications Applications Conference (SIU).IEEE,2022.

[2] Thivakaran T.K., M. Ramesh, "Exploratory Data analysis and forecasting of Big Mart dataset using supervised and ANN algorithm" Measurement: Sensors23 2022.

[3] Aguilar-Palacios, Carlos, Sergio Munoz-Romero, and Jose Luis Rojo-Alvarez. "Casual Quantification of Cannibalization 2021.

[4] Li, Yuanjiang, et al. "Clothing sale forecasting by a composite GRU-Prophet model with an attention mechanism." IEEE Transactions on Industrial Informatics 17.12 2021.

[5] Naveenraj, R., and R. Vinayaga Sundharam. "Prediction Of Big Mart Sales Using Machine Learning" International Journal of Modernization in Engineering, Technology and sciene, Volume:03, Issue:09 2021.

[6] Suma, V., and Shavige Malleshwara Hills. "Data mining based prediction of demand in Indian market for refurbished electronics." Journal of Soft Computing

[7] Paradigm (JSCP)2.02 2020.

[8] Aguilar-Palacios, Caros, et al. "Forecasting Promotional Sales within the Neighbourhood." Ieee Access 7 2019.

[9] Sun, Shaolong, et al. "A clustering-based nonlinear ensemble approach for exchange rates forecasting." IEEETransactions on Systems, Man, and Cybernetics: Systems50.6 2018.

[10] Wang, Yi, et al. "An ensemble forecasting method for the aggregated load with subprofiles." IEEE Transactions on Smart Grid 9.4 2018. Vol-9 Issue-3 2023 IJARIIE- ISSN(O)-2395-4396 20425 ijariie.com 2174.

[11] Baldan, Francisco et al. "A forecasting methodology for workload forecasting in cloud systems." IEEE Transactions on Cloud Computing 6.4 2016.

[12] Xavier, E.M., et al. "Requirements to Leverage the Electricity Distributors Sales and Revenues in the Brazilian Free Market." IEEE Latin America Transactions14.10 2016.

[13] Behesti-Kashi, Samaneh, et al. "A survey on retail sales forecasting and prediction in fashion markets." Systems Science & Control Engineering 3.1 2015. 29

[14] Duan, Zhaoyang, Brittni Gutierrez, and Lizhi Wang "Forecasting plug-in electric vehicle sales and the diurnalrecharging load curve." IEEE Transactions on Smart Grid5.1 2014.

[15] Gil-Alana, Luis Alberiko, Carlos Pestana Barros, and Albert Assaf. "Retail sales: persistence in the shortterm and long-term dynamics." IMA Journal of Management Mathematics 25.3 2014.

[16] Ren, Shuyun, Tsan-Minf Choi, and Na Liu. "Fashion sales forecasting with a panel data-based particle-filter model."IEEE Transactions on Systems, Man, and Cybernetics: System 45.3 2014.