Store-Sales Prediction

DETAILED PROJECT REPORT

ABINESH B

BIG MART SALES

"***Predicting tomorrow's trends*** *with today's data. Sales*
forecasting made smarter!"

Detailed Project Report

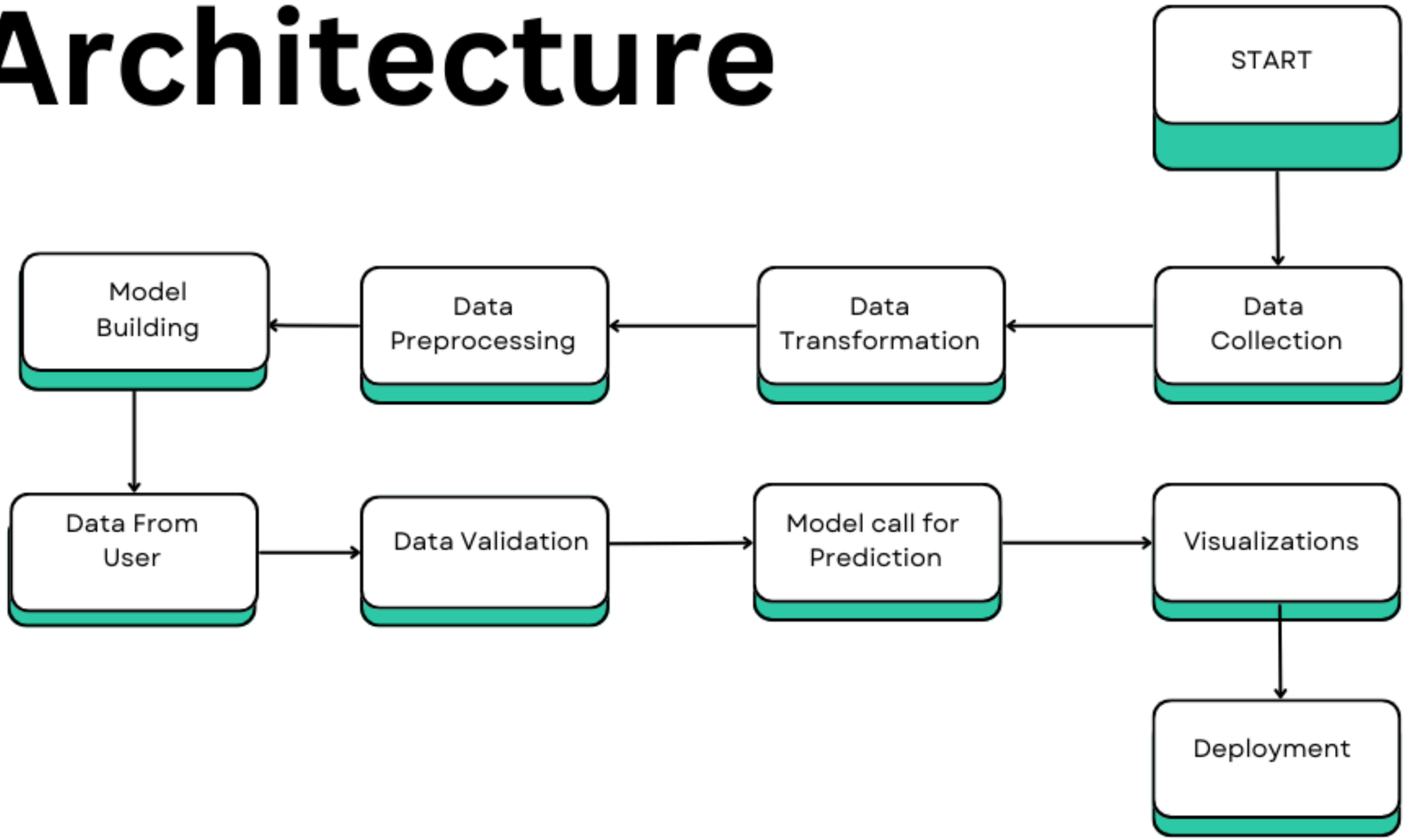
Objective

To develop a predictive model that can accurately forecast future sales based on historical sales data, customer behavior, and product-specific details. The model will help retailers optimize inventory levels, reduce overstock or stockout situations, and improve overall sales performance.

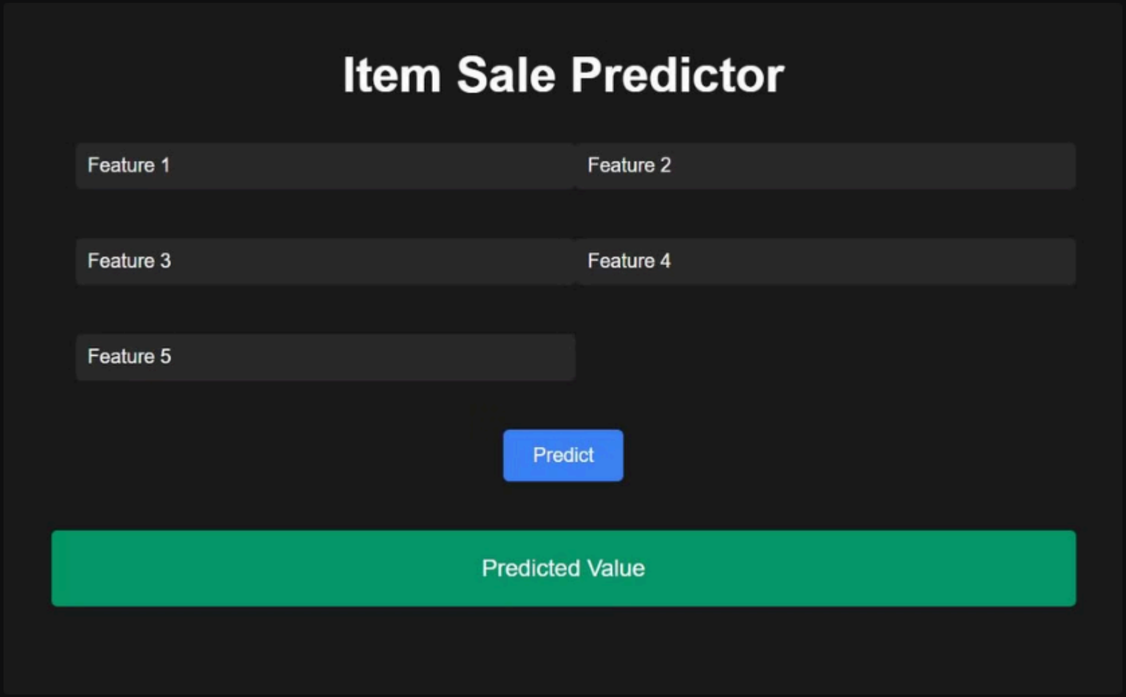
Workflow Overview

1. The journey begins with meticulous data collection, encompassing historical sales records, customer behavior patterns, and product-specific details. This raw data serves as the foundation for building a predictive model.
2. Data preprocessing is crucial, involving handling missing values, removing duplicates, and ensuring data consistency for analysis. This step ensures data quality and prepares it for model development.
3. Exploratory data analysis (EDA) delves into the data, uncovering trends and correlations between variables. This step provides valuable insights into customer behavior and product performance.
4. Machine learning models are then chosen and trained using the preprocessed data. Techniques like Random Forest or Gradient Boosting can accurately predict future sales based on historical patterns.
5. The model's performance is evaluated using metrics like MAE, RMSE, or R^2. Hyperparameter tuning refines the model for optimal accuracy and efficiency.
6. The trained model is then deployed for real-time sales forecasting, integrating seamlessly into inventory management systems or business decision-making processes.

Architecture

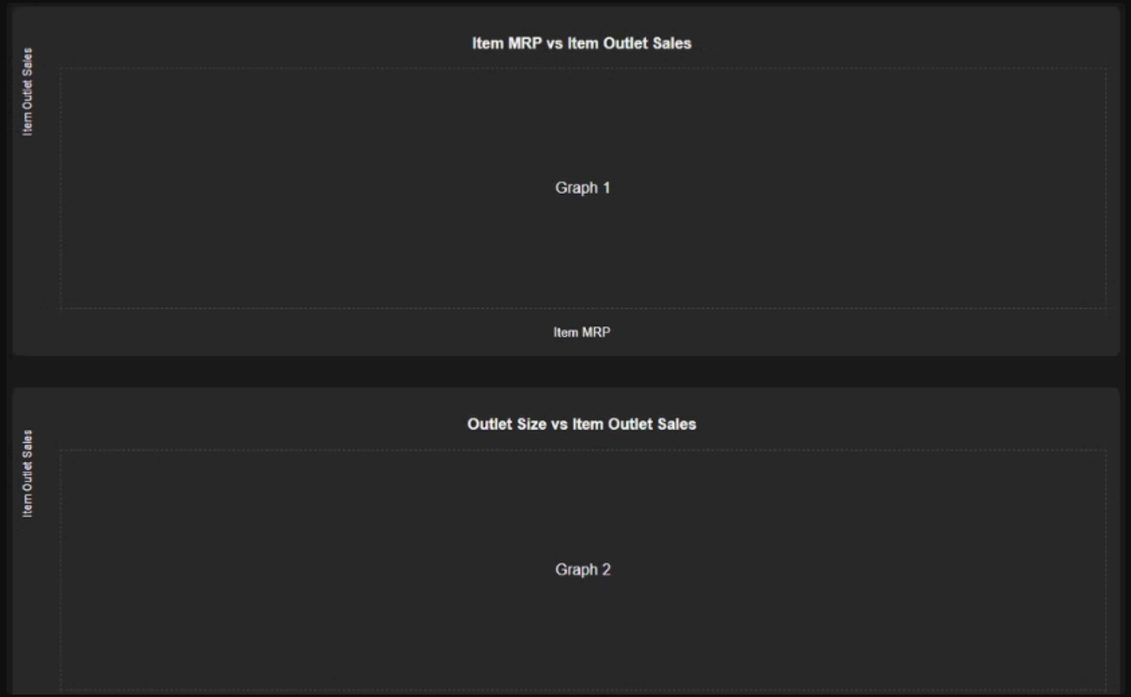


Wireframe



Input Parameters

User-friendly input fields for date range, store location, product category, and other relevant factors.



Prediction Dashboard

Interactive dashboard with visualizations for predicted sales, trends, and performance metrics.

View of Dataset

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishment_Year	Outlet_Size
0	FDA15	9.300	Low Fat	0.016047	Dairy	249.8092	OUT049	1999	Medium
1	DRC01	5.920	Regular	0.019278	Soft Drinks	48.2692	OUT018		
2	FDN15	17.500	Low Fat	0.016760	Meat	141.6180	OUT049		
3	FDX07	19.200	Regular	0.000000	Fruits and Vegetables	182.0950	OUT010		
4	NCD19	8.930	Low Fat	0.000000	Household	53.8614	OUT013		
5	FDP36	10.395	Regular	0.000000	Baking Goods	51.4008	OUT018		
6	FDO10	13.650	Regular	0.012741	Snack Foods	57.6588	OUT013		
7	FDP10	NaN	Low Fat	0.127470	Snack Foods	107.7622	OUT027		

Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishment_Year	Outlet_Size	Outlet_Location_Type	Outlet_Type	Item_Outlet_Sales
Low Fat	0.016047	Dairy	249.8092	OUT049	1999	Medium	Tier 1	Supermarket Type1	3735.1380
Regular	0.019278	Soft Drinks	48.2692	OUT018	2009	Medium	Tier 3	Supermarket Type2	443.4228
Low Fat	0.016760	Meat	141.6180	OUT049	1999	Medium	Tier 1	Supermarket Type1	2097.2700
Regular	0.000000	Fruits and Vegetables	182.0950	OUT010	1998	NaN	Tier 3	Grocery Store	732.3800
Low Fat	0.000000	Household	53.8614	OUT013	1987	High	Tier 3	Supermarket Type1	994.7052
Regular	0.000000	Baking Goods	51.4008	OUT018	2009	Medium	Tier 3	Supermarket Type2	556.6088
Regular	0.012741	Snack Foods	57.6588	OUT013	1987	High	Tier 3	Supermarket Type1	343.5528
Low Fat	0.127470	Snack Foods	107.7622	OUT027	1985	Medium	Tier 3	Supermarket Type3	4022.7636

tem_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishme
9.30	0	0.016047	4	249.8092	9	
5.92	1	0.019278	14	48.2692		
17.50	0	0.016760	10	141.6180		
19.20	1	0.000000	6	182.0950		
8.93	0	0.000000	9	53.8614		

Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishment_Year	Outlet_Size	Outlet_Location_Typ
4	249.8092	9	1999	1	
14	48.2692	3	2009	1	
10	141.6180	9	1999	1	
6	182.0950	0	1998	1	
9	53.8614	1	1987	2	

Dataset Features

Initial Set:

- 1.Item Weight
- 2. Item Fat Content
- 3. Item Visibility
- 4. Item Type
- 5. Item MRP
- 6. Outlet Identifier
- 7. Outlet Establishment Year
- 8. Outlet Size
- 9. Outlet Location Type
- 10. Outlet Type
- 11. Item Outlet Sales

Input Features:

- Item MRP
- Item Weight
- Outlet Type
- Outlet Location Type
- Outlet Size

Target Variable:

- Item Outlet Sales

Streamlit View

Item Sale Predictor

Item Weight

100.00 − +

Item MRP

249.81 − +

Outlet Size

High

Outlet Location Type

Rural Area

Outlet Type

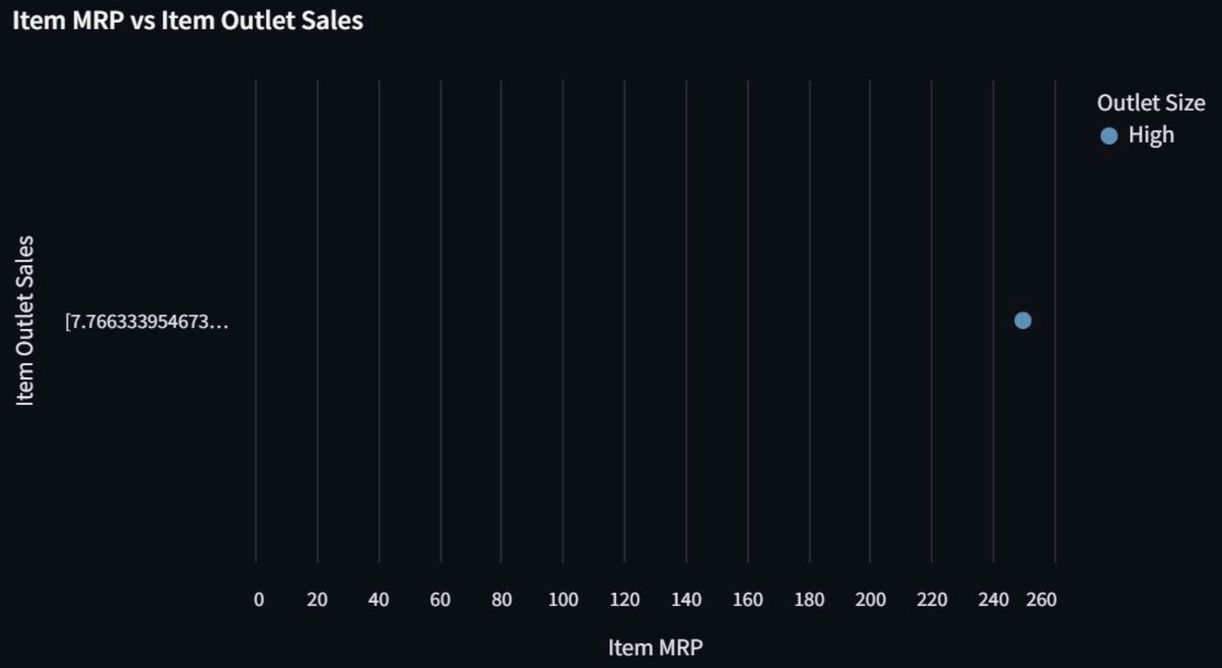
○ Convenience Store ○ Mini Mart

○ Neighborhood Market ◉ Hypermarket

Predict

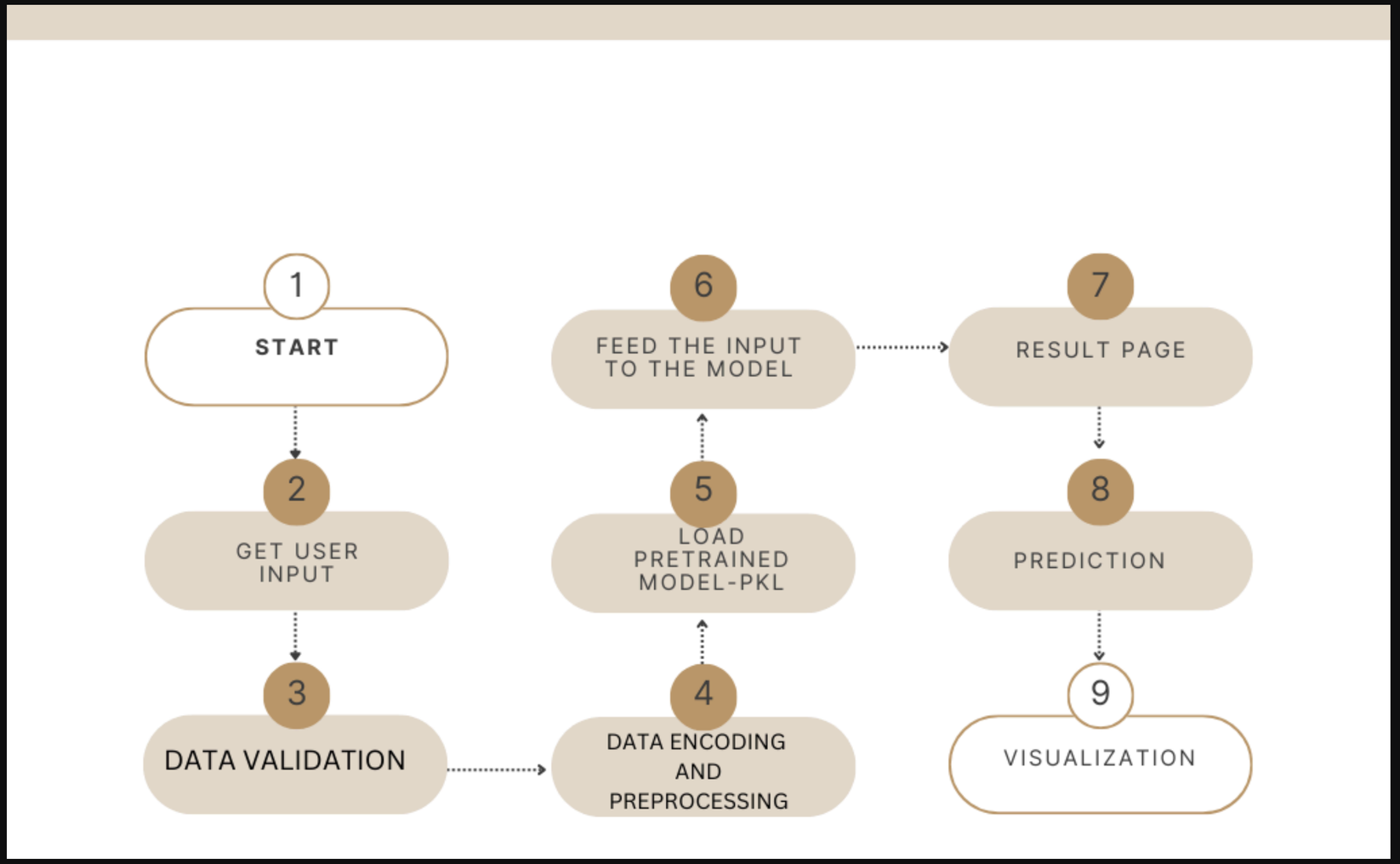
Predicted Value: 2358.80

Visualizations



Streamlit view

Deployment Process



Technology Stack

1.Python

The core language for data analysis, model development, and deployment.

2.Scikit-learn

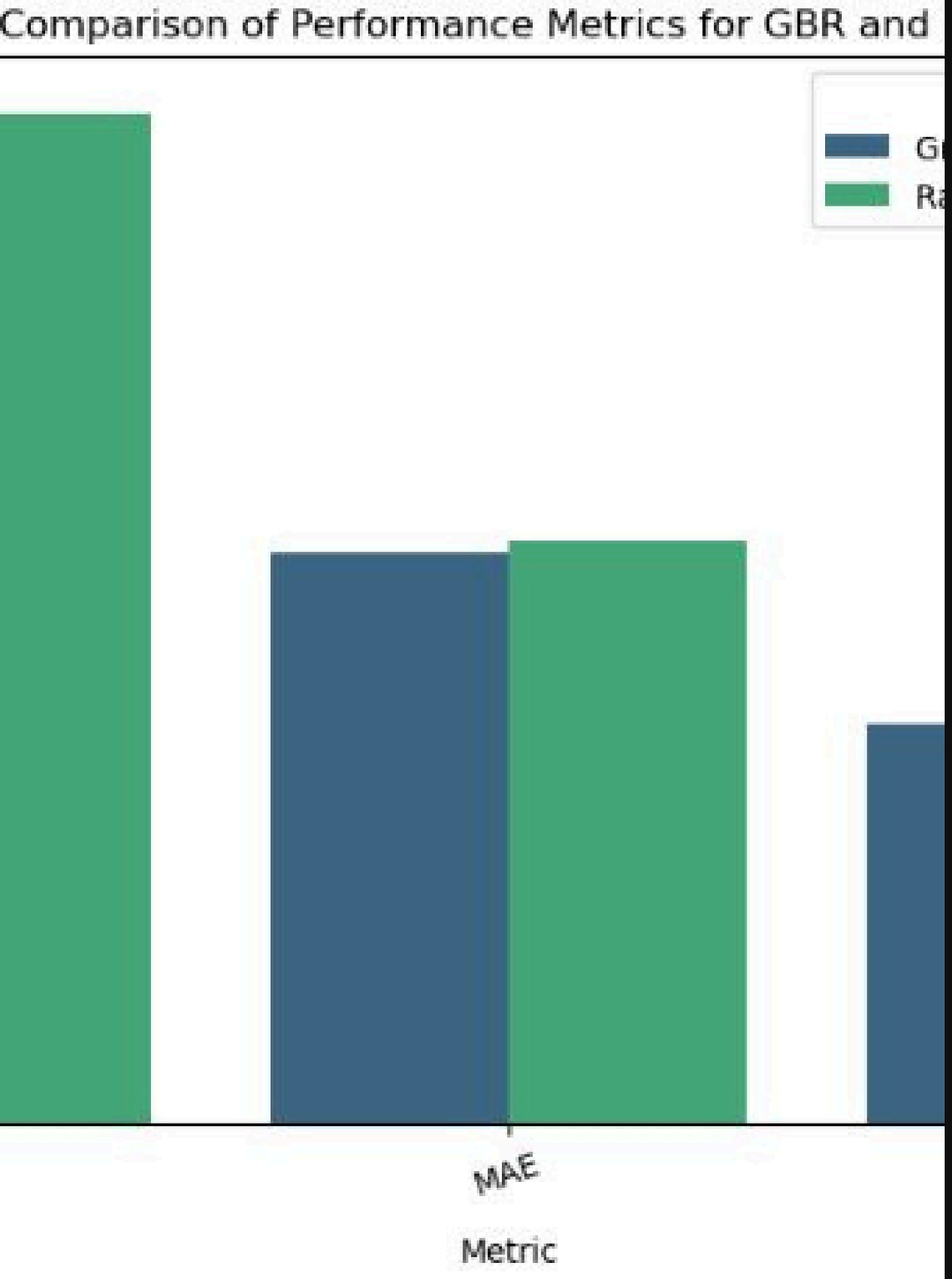
A machine learning library offering a wide range of algorithms for model development and evaluation.

3.Pandas

A library for data manipulation and analysis, providing powerful tools for cleaning and transforming data.

4.Streamlit

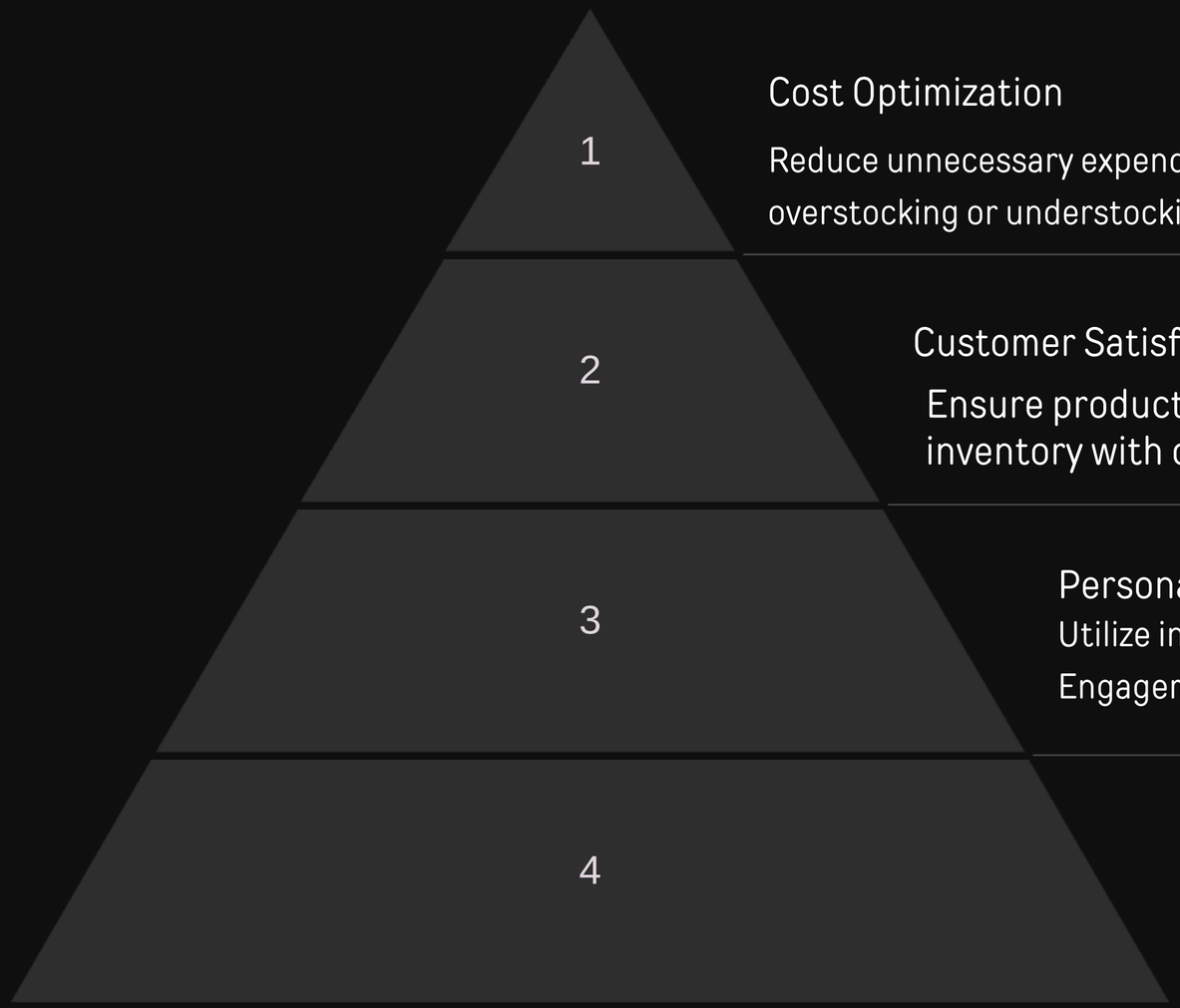
A framework for building interactive web applications for data visualization and model deployment.



Model Training and Evaluation

The Gradient Boosting Regressor excels in prediction accuracy and capturing complex data patterns, making it ideal for sales forecasting. In contrast, the Random Forest Regressor provides reliable but less precise results.

Benefits



Cost Optimization

Reduce unnecessary expenditures by accurately forecasting demand and avoiding overstocking or understocking.

Customer Satisfaction

Ensure product availability and better shopping experiences by aligning inventory with customer needs.

Personalized Marketing

Utilize insights to create targeted campaigns and promotions, increasing customer Engagement and sales.

Improved Financial Planning

Aid in budgeting and financial forecasting by providing clear sales trends and performance metrics.

Links

Architecture: https://docs.google.com/document/d/1F9u9pBJIDl2mmgxNM9dzoeAPZh6N4YYB/edit?usp=drive_link&ouid=104108091560678380977&rtpof=true&sd=true

High Level Design:

https://docs.google.com/document/d/1KR9JKajBTmq4RQK2HpE7kqCtl_BPSvZr/edit?usp=drive_link&ouid=104108091560678380977&rtpof=true&sd=true

Low Level Design:

https://docs.google.com/document/d/117aE25rMz5SKa--zKKajTCR7DL2XcQrE/edit?usp=drive_link&ouid=104108091560678380977&rtpof=true&sd=true

Wireframe:

https://docs.google.com/document/d/1ECqnbh8OwZfz_carylqwCCTM0rP1N1x5/edit?usp=drive_link&ouid=104108091560678380977&rtpof=true&sd=true

Code:

https://github.com/Abinesh2418/Sales-Store-Prediction