Customer Insights and Behavioral Patterns Analysis Using PySpark

The Catalysts

1 Project Title

Customer Insights and Behavioral Patterns Analysis Using PySpark

2 Project Idea

The project aims to analyze customer data to extract valuable insights that can help improve business strategies. The analysis will focus on understanding customer demographics, purchasing behavior, and satisfaction levels. The project will utilize PySpark for data processing and analysis, and the results will be visualized using a BI tool like Tableau.

3 Technology Summary and Database

- **PySpark**: For data processing and analysis.

- **Tableau**: For data visualization.

- **Jupyter Notebook**: For writing and executing PySpark code.

- **GitHub**: For version control and collaboration.

- **Database**: Customer Purchases Behaviour Dataset. Dowloaded from kaggale.
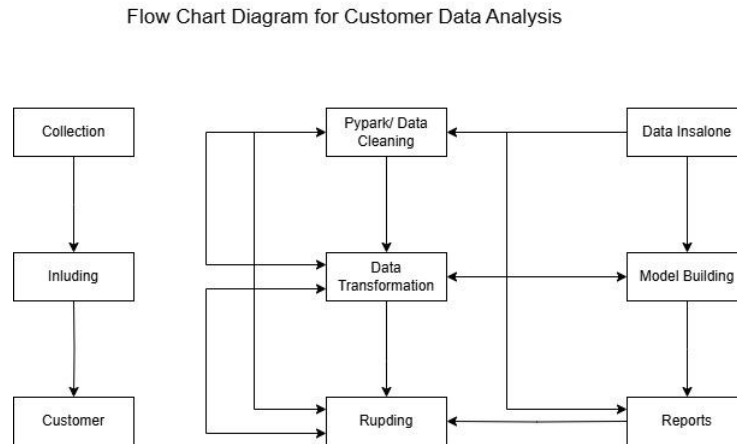
4 Architecture Diagram

Figure 1: Data Flow Diagram for Customer Data Analysis Project

5 Architecture Summary

- **Customer Data**: The raw customer data is stored in a CSV file.

- **PySpark Processing**: The data is loaded into PySpark for cleaning, transformation, and analysis.

- **Data Analysis**: Various analytical operations are performed to derive insights.

- **Data Visualization**: The results are visualized using Tableau.

- **Insights & Reports**: The final insights and reports are generated and shared.

6 Project Goals

1. Analyze customer demographics to identify key segments.

2. Determine purchasing patterns and frequency.

3. Assess customer satisfaction levels and identify factors affecting satisfaction.

4. Evaluate the impact of promotions on purchase behavior.

5. Generate actionable insights to improve marketing strategies.