5. Project Report: E – Commerce Sales and Return Rate Reduction Analysis

# Objective:

To analyze historical sales data and identify patterns responsible for high product return rates. The project also aims to build a return prediction model to assist in reducing overall return percentages and improving product quality and customer satisfaction.

# Project Components:

1. Dataset Used:

- Source: dataset.csv

- Key Columns: Product\_ID, Sales\_Quantity, Return\_Status, Customer\_Rating, Region, Category, etc.

2. Data Cleaning & Preprocessing:

- Removed duplicates and missing entries.

- Standardized categorical values.

- Handled imbalanced return labels using resampling techniques.

# Machine Learning Implementation:

Model: Logistic Regression & Random Forest

Code Files:

- return prediction.py – Model training and accuracy metrics.

- return rate analysis.py – Exploratory Data Analysis (EDA), Visualization.

Steps:

- Feature selection based on correlation heatmaps and domain knowledge.

- Model trained on labeled data with 80/20 train-test split.

- Accuracy achieved: [Insert exact metric here]

- Important features: Low customer ratings, specific product categories, certain regions.

# Visual Insights & Dashboards:

- Interactive dashboards in PDF format (glass sales dashboard, sales dashboard).

- KPIs: total revenue, sales trends, high-risk return products.

- Filters: Category, Region, Time Period.

- Chart snapshots highlight top 10 return-prone products and monthly return trends.

# High-Risk Products Identified:

- Based on high\_risk\_products.csv, products with return rate > 25% were flagged.

- These were mostly from specific categories such as [Insert category].

# Key Takeaways:

- Low customer ratings and high delivery time → higher return rates.

- Predictive model flags risky transactions in advance.

- Dashboards support real-time decision-making.

# Tools & Technologies Used:

- Languages: Python (pandas, sklearn, matplotlib, seaborn)

- Tools: Power BI / Tableau, Jupyter Notebook

- Techniques: Logistic Regression, EDA, Data Cleaning, Feature Engineering

# Outcome:

Improved understanding of return behavior. Predictive model assists in identifying high-risk orders early. Dashboards offer actionable insights for business stakeholders.