

E-COMMERCE & RETAIL B2B CASE STUDY

Submitted By

1. Chayanika Hazra
2. Bhagyashree Bose
3. Bharat Bhushan Bedi

Contents

1. Problem Statement
2. Goal
3. Approach
4. EDA
5. Model Evaluation
6. Insights and Suggestions

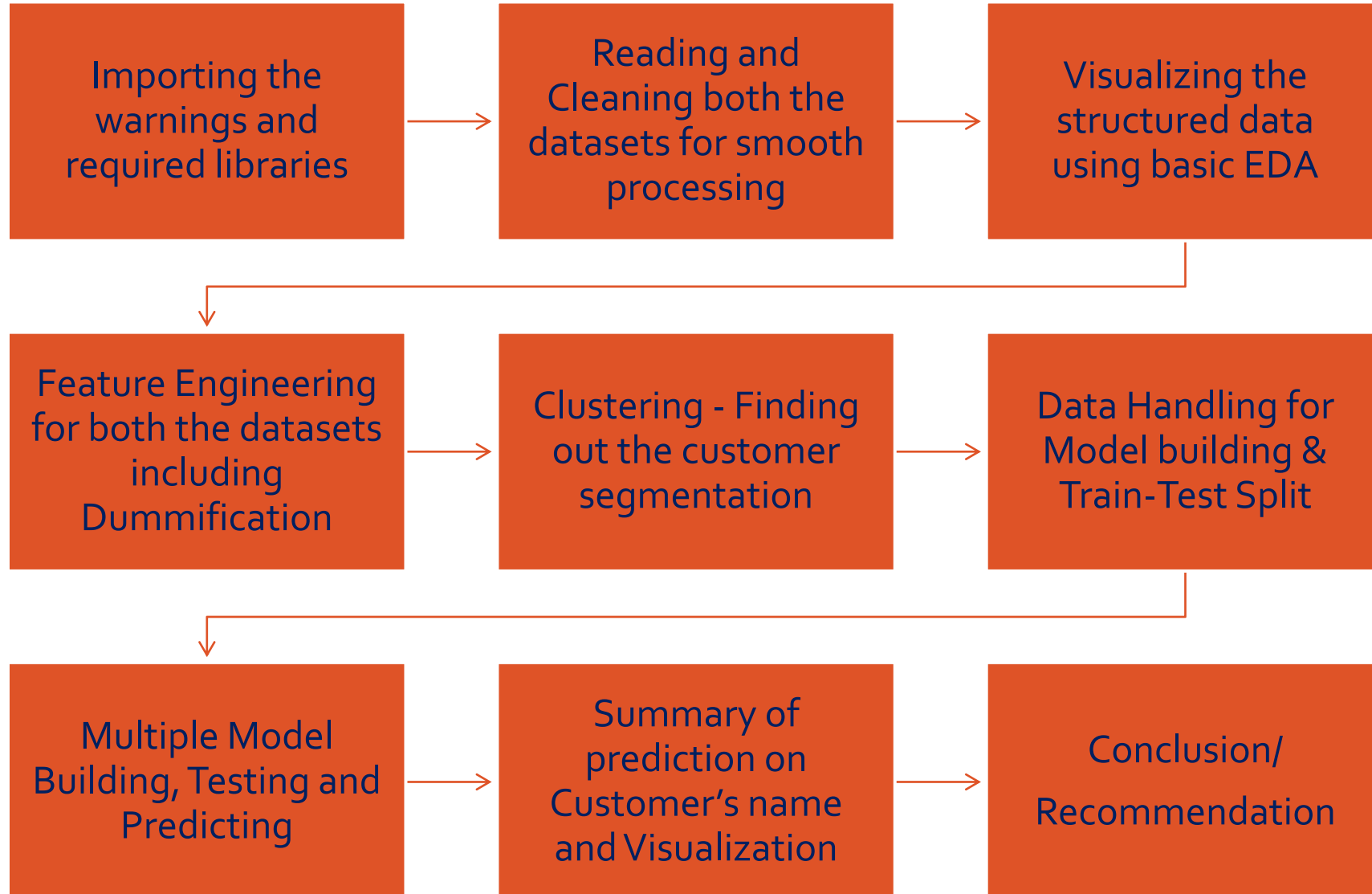
Problem Statement

- Schuster is a multinational retail company specializing in sports goods and accessories, engaging in substantial business with numerous vendors under credit arrangements. However, some vendors fail to adhere to payment terms, leading to late payments. While Schuster imposes significant late fees, this approach is detrimental to long-term relationships. Employees spend time chasing late payments, resulting in non-value-added activities and financial costs. Schuster aims to analyze customer payment behavior and predict the likelihood of late payments on open invoices.
- Whenever a transaction occurs, the accounting team generates an invoice detailing the goods, invoice value, creation date, and payment due date per the contract's credit terms. Given the frequency of transactions, multiple invoices are active for each vendor at any time.

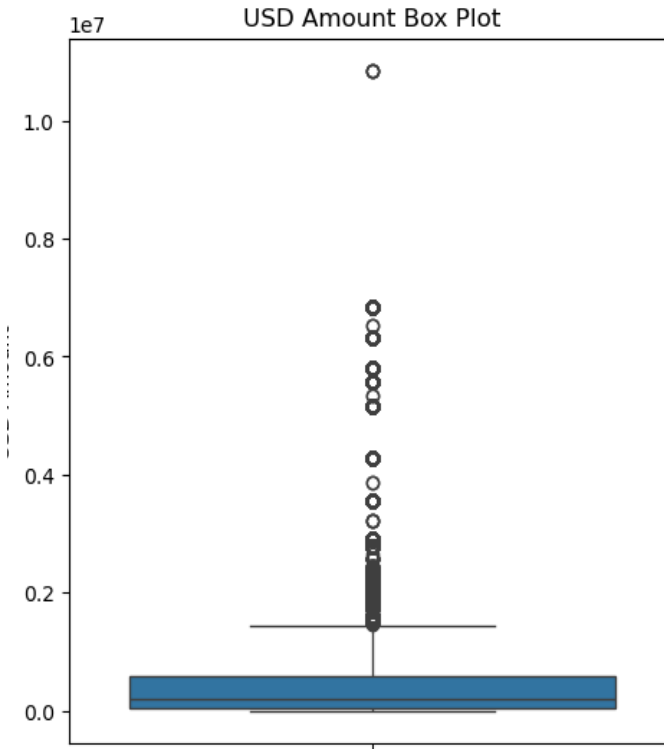
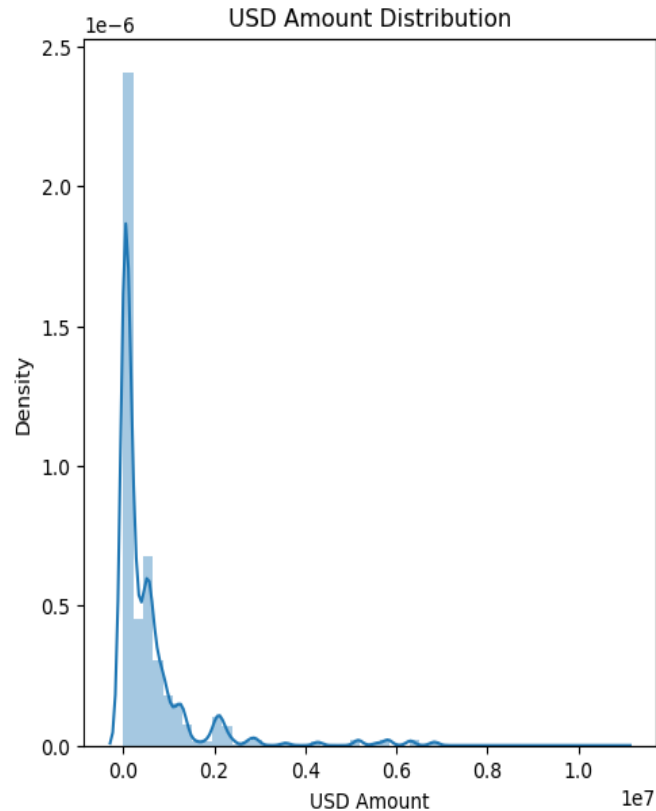
Goal

- The company seeks to understand customer payment behavior through historical patterns and segmentation. It aims to predict the likelihood of delayed payments on open invoices, enabling collectors to prioritize follow-ups and secure timely payments.

Approach

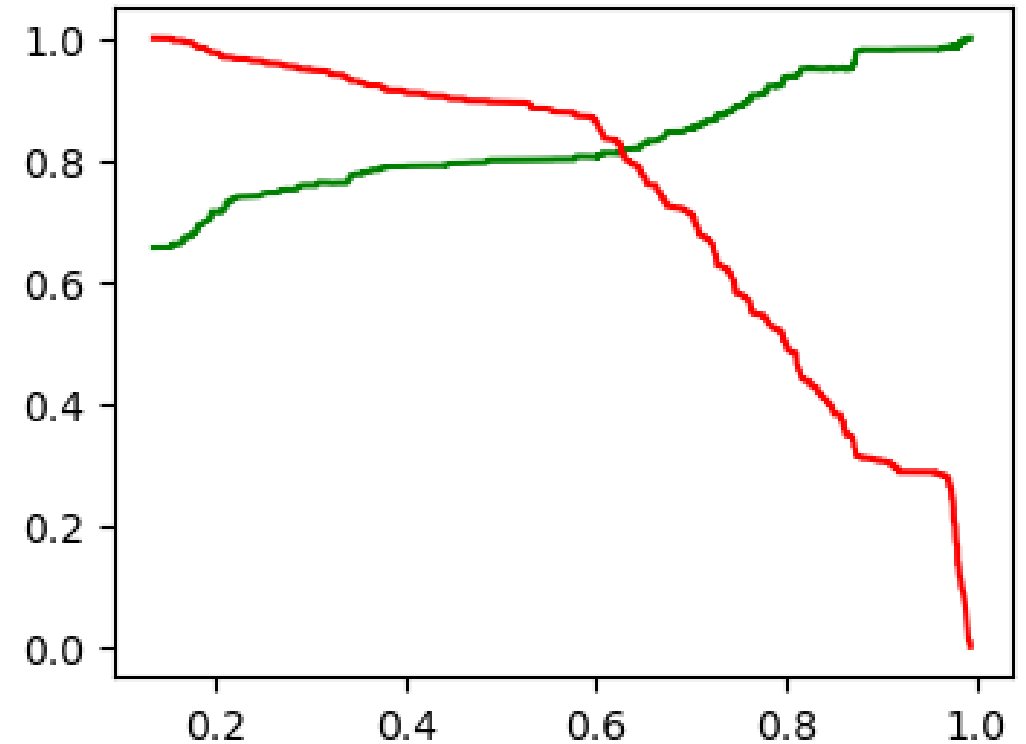
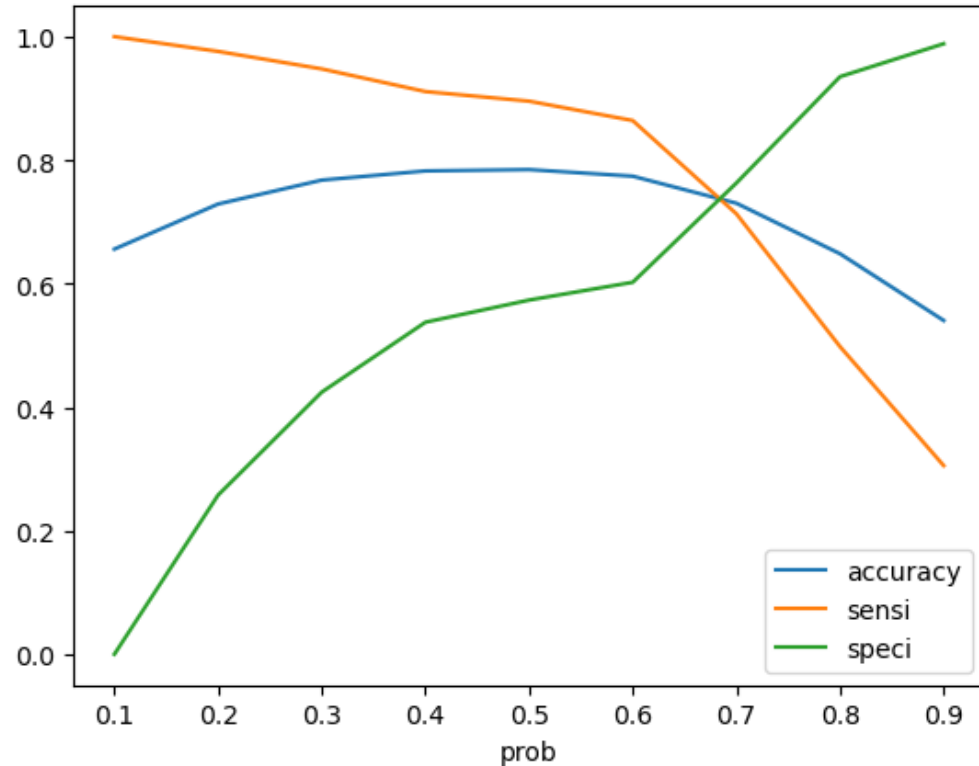


EDA

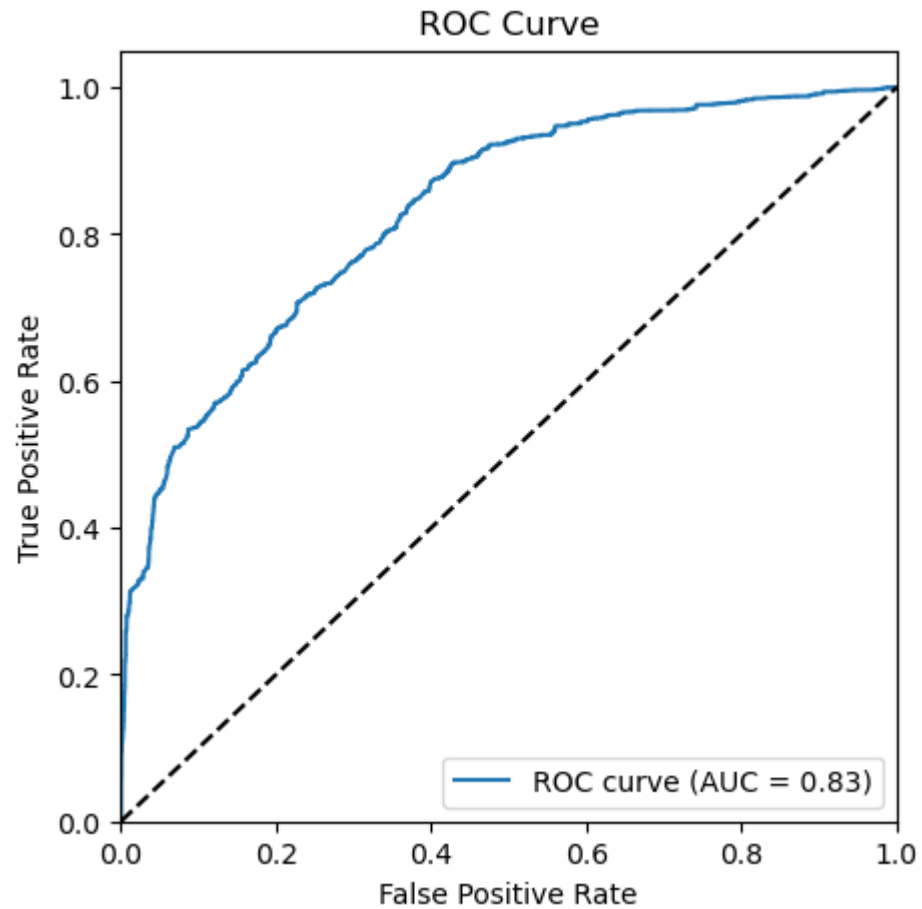


- **Highly Skewed Distribution:** The USD amount is highly right-skewed, with most transactions concentrated at lower values, indicating that small transactions dominate.
- **Presence of Outliers:** The box plot reveals numerous outliers (values far beyond the whiskers), suggesting the existence of very high transaction amounts that deviate significantly from the majority.
- **Tight Spread for Majority:** The box plot shows a compact interquartile range (IQR), meaning most transactions fall within a narrow band of values, while a few extreme values inflate the range.

Model Evaluation



Based on the above two diagrams, the decided optimum cut-off is 0.6.



- The AUC Score seemed to be really good and ideally acceptable therefore we moved ahead and made predictions on the Test Set.

Recommendations

- **Invoice Class Policies:** Credit Note Payments exhibit the highest delay rates compared to Debit Note or Invoice-type payments. It is recommended that the company implement stricter payment collection policies for such invoice classes to mitigate delays and improve cash flow efficiency.
- **Goods vs. Non-Goods Invoices:** Invoices for goods have significantly higher payment delay rates than those for non-goods. To address this, stricter payment policies should be enforced for goods-related invoices to reduce delay instances.
- **Low-Value Payments:** A significant portion of transactions involve lower payment amounts, which also show higher delay rates. The company should prioritize these transactions by introducing penalties for late payments. For example, penalties could be applied as a percentage of the billing amount, with higher penalty rates for smaller invoices. This policy should, however, be considered as a last resort.
- **Customer Segmentation Insights:** Customers were segmented into three clusters:
 - Cluster 0: Medium payment duration
 - Cluster 1: Prolonged payment duration Cluster 2: Early payment duration Customers in Cluster 1 (prolonged payment duration) demonstrated significantly higher delay rates compared to other clusters. It is recommended that the company prioritize focused attention and customized engagement strategies for Cluster 1 customers to reduce delay rates.
- **High-Probability Customers:** Companies or customers with the highest probability of delay, as well as those with substantial total and delayed payment counts, should be prioritized. These customers represent a significant risk to timely payment collection and require dedicated follow-up and tailored policies to ensure compliance.