Introduction to E-Commerce and and Retail B2B Case Study

This case study explores the rapidly evolving landscape of e-commerce and its impact on the retail B2B sector. It delves into the strategies, challenges, and opportunities faced by businesses as they navigate the digital transformation in the retail industry.

~Prithibi Mondal

Problem Statement & Project Goals

1 Problem Statement

Schuster, a multinational retail company in sports goods and accessories, faces challenges with late vendor payments despite heavy late fees and manual follow-ups. To optimize its credit management process, Schuster aims to leverage data science to analyze and predict vendor payment behavior.

2 Project Goals

Schuster aims to leverage data science to analyze historical transaction data and predict vendor payment behavior. The goals include understanding payment patterns, segmenting vendors for tailored credit management, developing a predictive model for late payments, and deriving actionable insights to optimize collections and credit management processes.

3 Desired Outcomes

The desired outcomes of the project include gaining a deep understanding of vendor payment behaviors through data analysis, effectively segmenting vendors based on payment patterns to enhance credit management strategies, developing a reliable predictive model for forecasting late payments, and deriving actionable insights to improve operational efficiency, reduce financial risks, and strengthen vendor relationships through optimized credit practices.

Methodology

1. **Data Collection**

   Gather relevant data from Upgrad

2. **Data Preprocessing**

   Clean, transform, and aggregate the data to prepare it for analysis.

3. **Exploratory Analysis**

   Identify patterns, trends, and insights from the data to inform the modeling process.

4. **Model Building**

   Develop predictive models for vendor payment behavior.

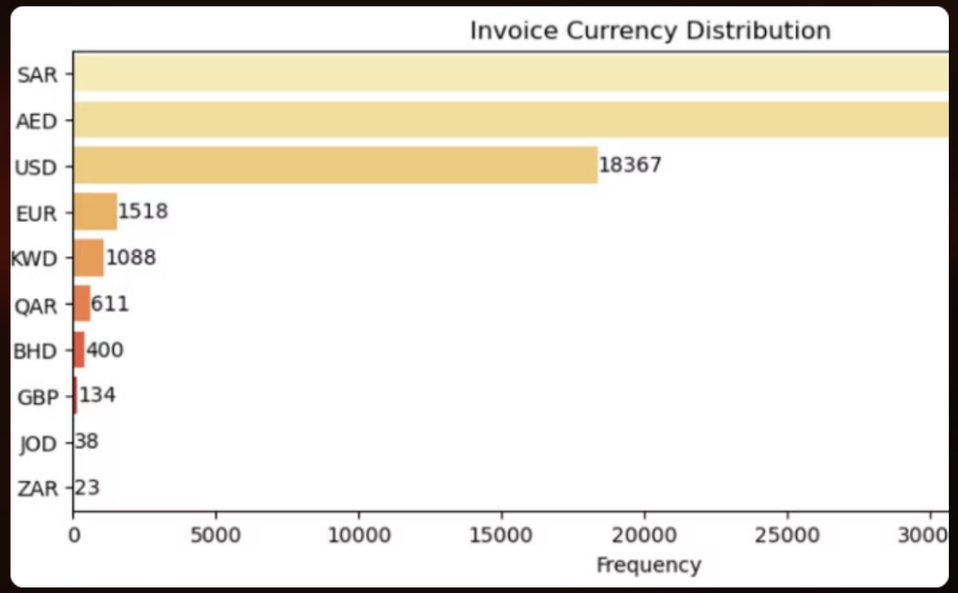
5. **Model Evaluation**

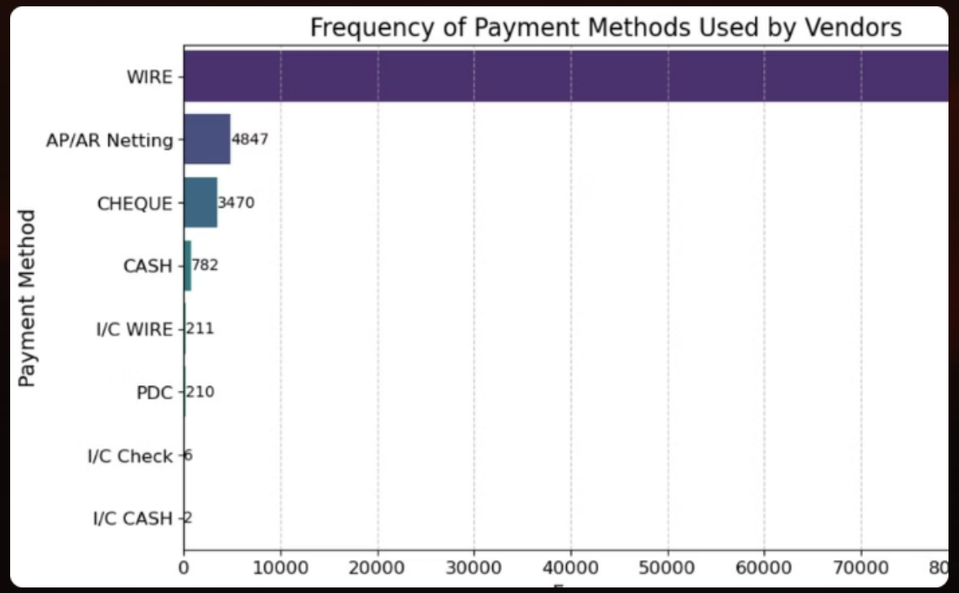
   Assess model accuracy and robustness.

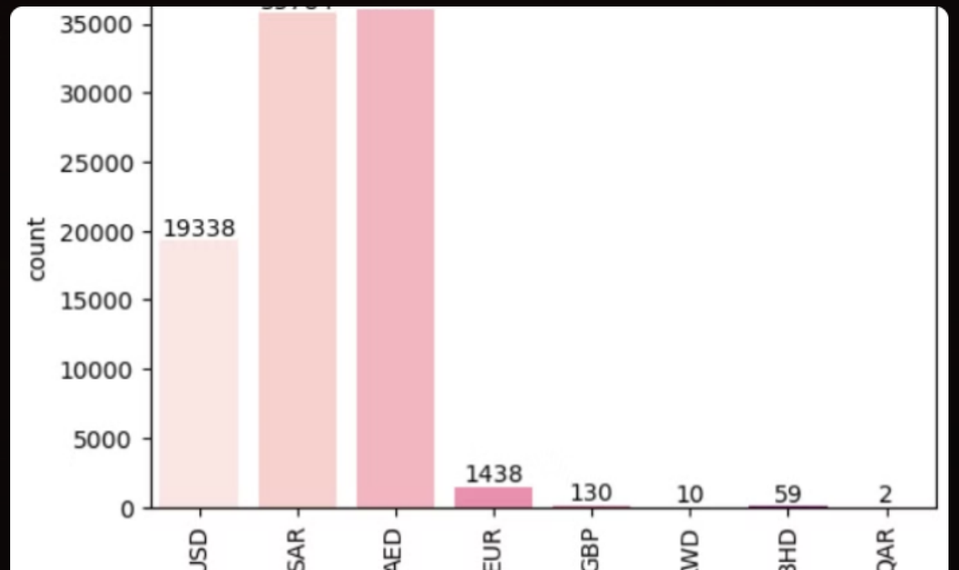
6. **Business Insights**

   Derive actionable insights to optimize credit management and collections strategies.

Exploratory Data Analysis







Model Development

Customer Segmentation

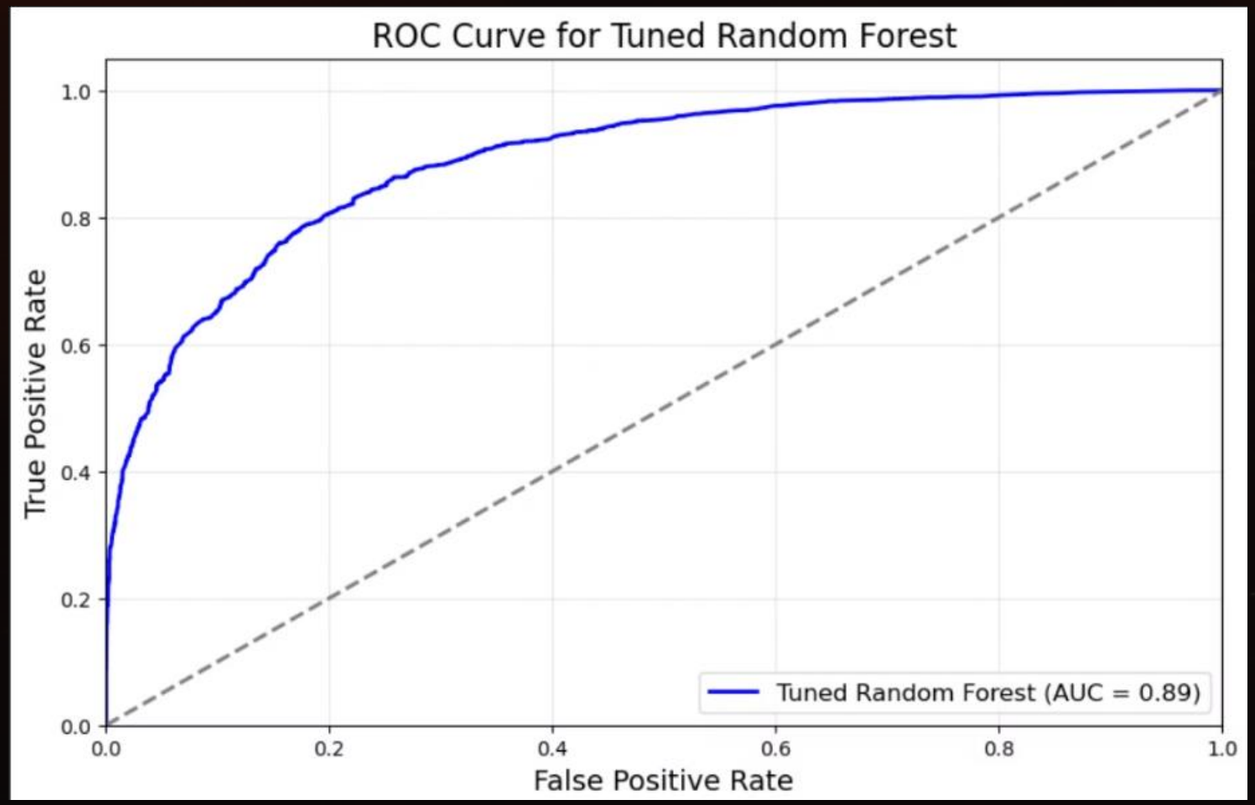
Utilize Random forest and Loegestic Regression to group customers based on their behavior and characteristics.

Predictive Modeling

Apply advanced techniques, such as logistic regression and random forest, to forecast customer churn and purchase likelihood.

Optimization Strategies

Develop models to optimize pricing, inventory management, and marketing campaigns for improved profitability.



Model Evaluation

Accuracy

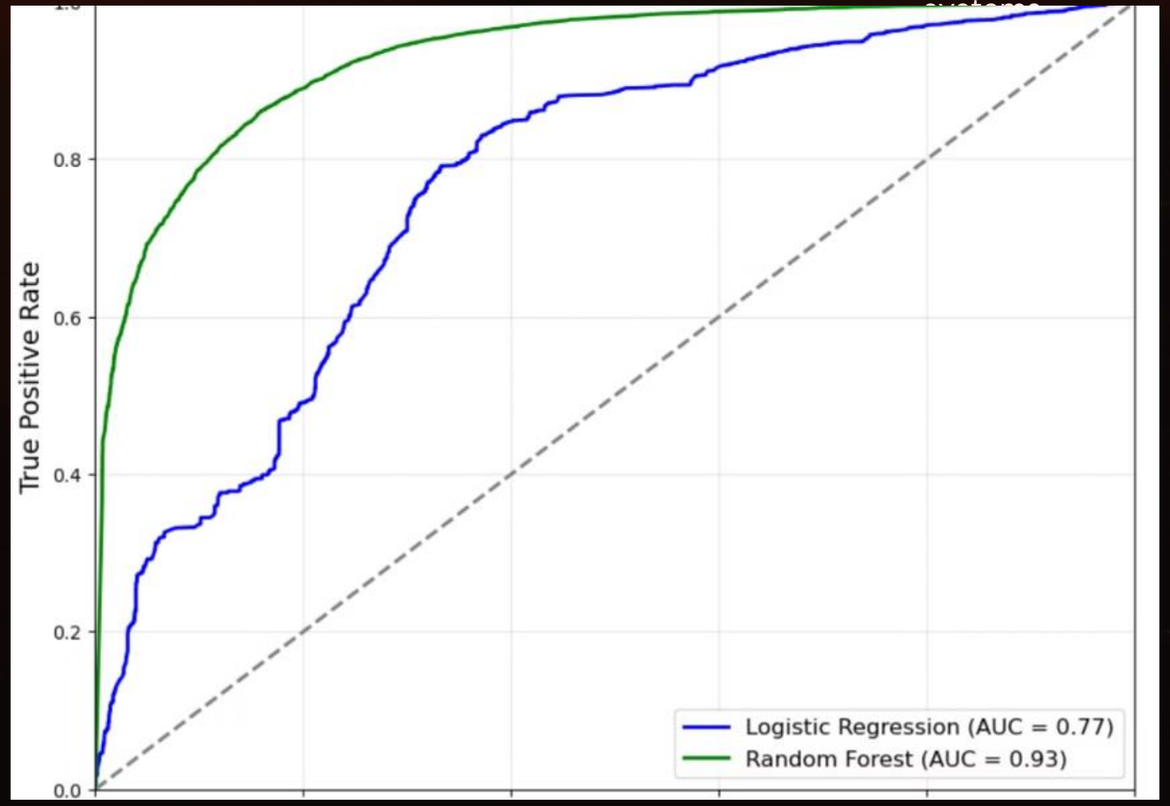
Assess the predictive accuracy of the models using appropriate metrics.

Robustness

Evaluate the models' ability to handle variations in data and real-world scenarios.

Scalability

Ensure the models can be seamlessly integrated into the company's existing processes and



Random Forest vs. Logistic Regression

Random Forest

Ensemble learning method that combines multiple decision trees to improve accuracy and reduce overfitting.

Logistic Regression

Widely used technique for binary classification problems, such as predicting customer churn.

Comparison

Random Forest often outperforms Logistic Regression in terms of predictive accuracy, but Logistic Regression may be preferred for its interpretability.

Logistic Regression:

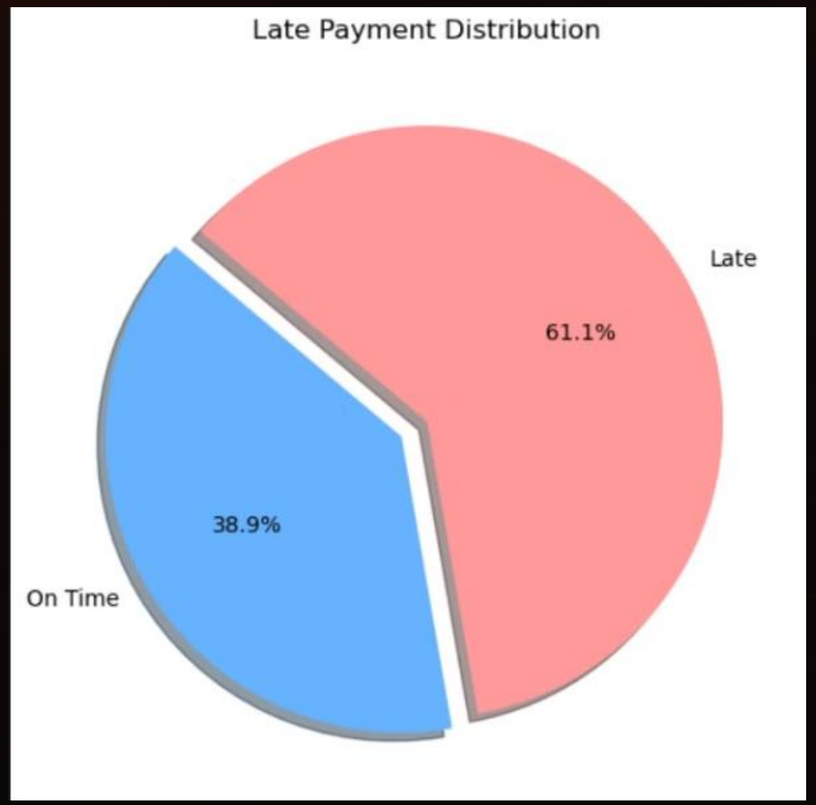
	precision	recall	f1-score	support
0	0.58	0.69	0.63	6378
1	0.82	0.74	0.78	12179
accuracy			0.72	18557
macro avg	0.70	0.72	0.71	18557
weighted avg	0.74	0.72	0.73	18557

Random Forest:

	precision	recall	f1-score	support
0	0.76	0.84	0.80	6378
1	0.91	0.86	0.89	12179
accuracy			0.85	18557
macro avg	0.84	0.85	0.84	18557
weighted avg	0.86	0.85	0.86	18557

Business Insights & Recommendations

- **Invoice Amount (USD Amount):** Higher invoice amounts tend to have a higher likelihood of late payments.
- **Payment Terms (PAYMENT_TERM_NUM):** Longer payment terms are associated with higher late payment risk.
- **High-Risk Customers:** Identifying customers with consistently late payment behavior allows the company to prioritize follow-up actions and mitigate financial risks.
- **Overall Late Payment Ratio:** The late payment ratio across all transactions is approximately 61%
- **Adjusting Credit Terms:** Review and adjust credit terms for customers with a history of late payments.
- **Proactive Follow-Ups:** Implement automated reminders and follow-up actions based on the model's predictions.



Conclusion and Key Takeaways

- Utilizing data science enables deeper insights into vendor payment behaviors.
- Predictive models aid in anticipating and mitigating late payment risks.
- Actionable insights support proactive credit management strategies.
- Ultimately, optimizing collections processes enhances operational efficiency and financial health.