Predictive Analysis of Late Payments in B2B E-commerce & Retail

Data-driven insights for better payment collection strategies.

by

Shayon Deb
Siddharth Joshi
Sanket Kamble

Content:-

- Business Problem
- Data Overview
- Data Analysis & Insights
- Predictive Model Performance
- Recommendations
- Business Implications
- Conclusion & Next Steps

Business Problem

1

Late Payments

Impact cash flow and operational efficiency.

2

Payment Delay Factors

Crucial for improving collection strategies.

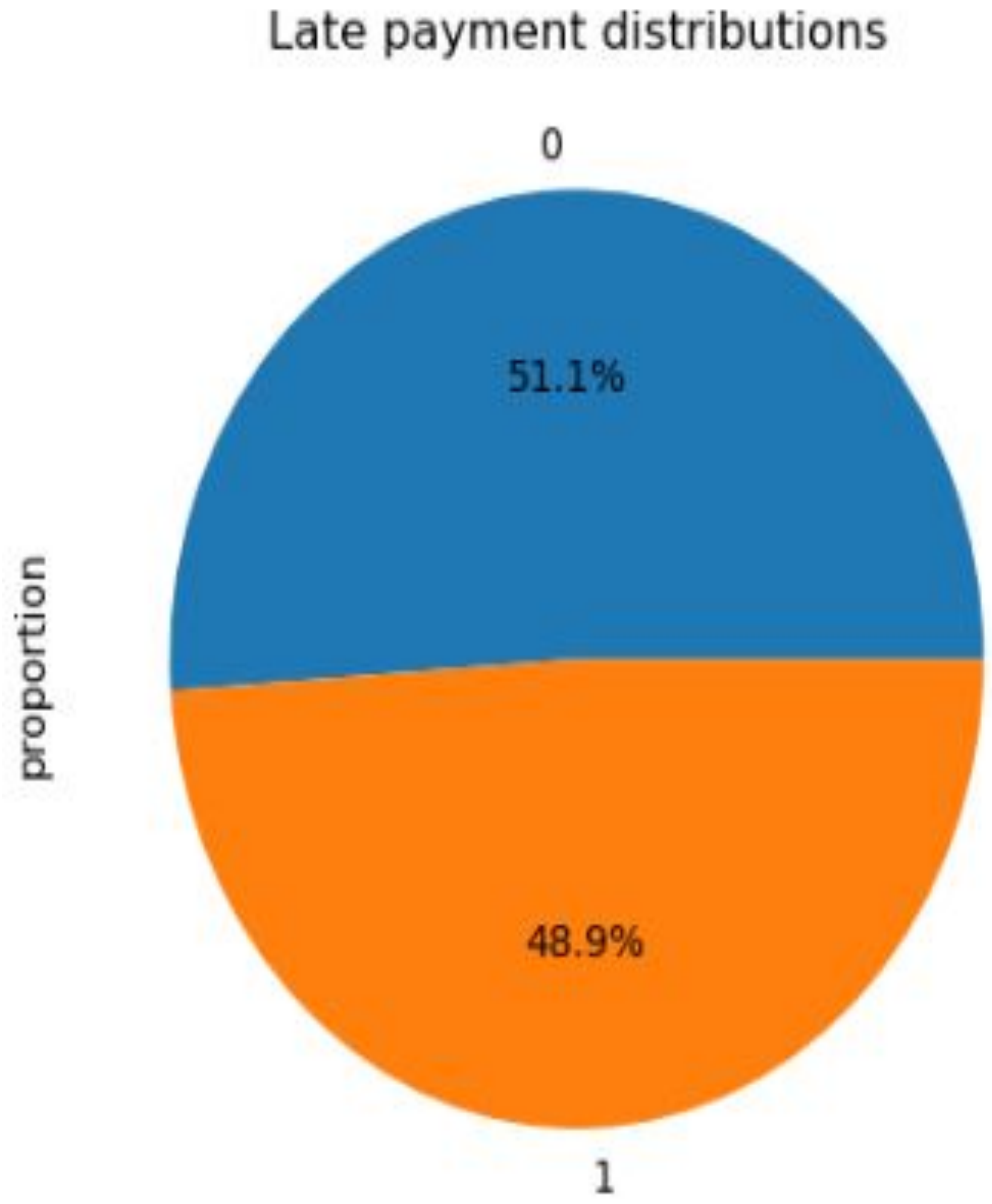
3

Predictive Model

Identify high-risk invoices and recommend corrective actions.

Data Overview

1. Dataset source: Kaggle, E-commerce & Retail B2B Case Study
2. Contains transaction details, invoice types, payment status, and customer information
3. Key variables: invoice type, payment status, payment amount, customer segmentation



Data Analysis & Insights

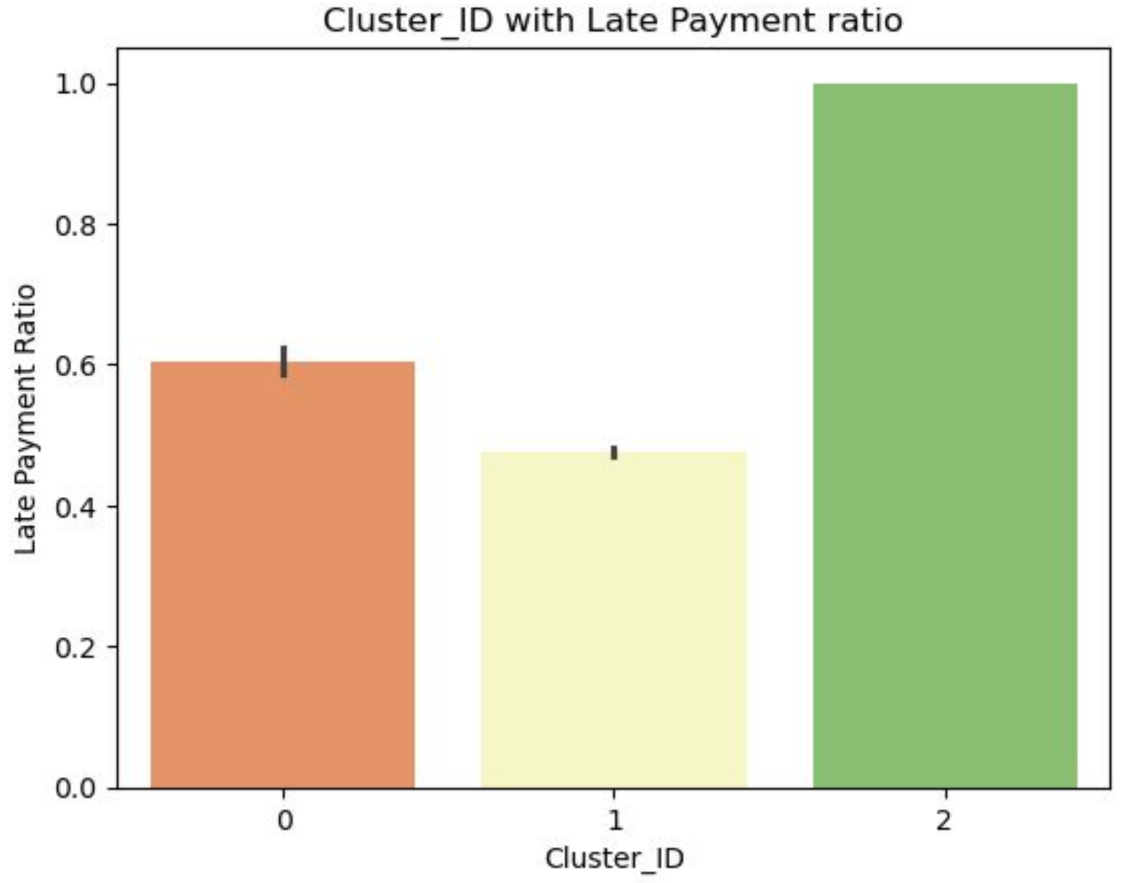
Payment Delay Trends

51.1% on-time, 48.9% delayed

Invoice Type Impact

Credit Note invoices faced highest delays

Data Analysis & Insights



cluster ID 2 has significantly higher ratio of default than clusters 0 and 1

Invoice Amount Impact

Lower-value payments had higher late payment rates

Customer Segmentation

Early payment cluster showed high delay rates

Predictive Model Performance

Train Data

Best hyperparameters: {'n_estimators': 150, 'min_samples_split': 2, 'min_samples_leaf': 1, 'max_depth': 30}
Best f1 score: 0.9393260434851571

	precision	recall	f1-score	support
0	0.96	0.91	0.94	22349
1	0.96	0.98	0.97	42618
accuracy			0.96	64967
macro avg	0.96	0.95	0.95	64967
weighted avg	0.96	0.96	0.96	64967

Test Data

	precision	recall	f1-score	support
0	0.91	0.86	0.88	9529
1	0.93	0.96	0.94	18315
accuracy			0.92	27844
macro avg	0.92	0.91	0.91	27844
weighted avg	0.92	0.92	0.92	27844

- From the above classification report, it can be seen that the f1-score for train and test set is 0.97 and 0.94, which implies that this is a good model. Hence moving forward with this as final model for prediction.

1. Model used: Random Forest Classifier, hyperparameters optimized

2. Performance metrics: Training F1 Score 0.97, Testing F1 Score 0.94

3. Model effectively predicts late payments, helps mitigate risk

Recommendations

Payment Policy Review

Stricter terms for Credit Notes and Goods-type invoices

Low-Value Payment Strategies

Tiered penalties and early payment discounts

Customer Segmentation Approach

Close monitoring of "early payment" cluster

Automated Reminders

Alert system for predicted late invoices

Business Implications

Improved Cash Flow

Early identification of risky invoices

Enhanced Customer Relations

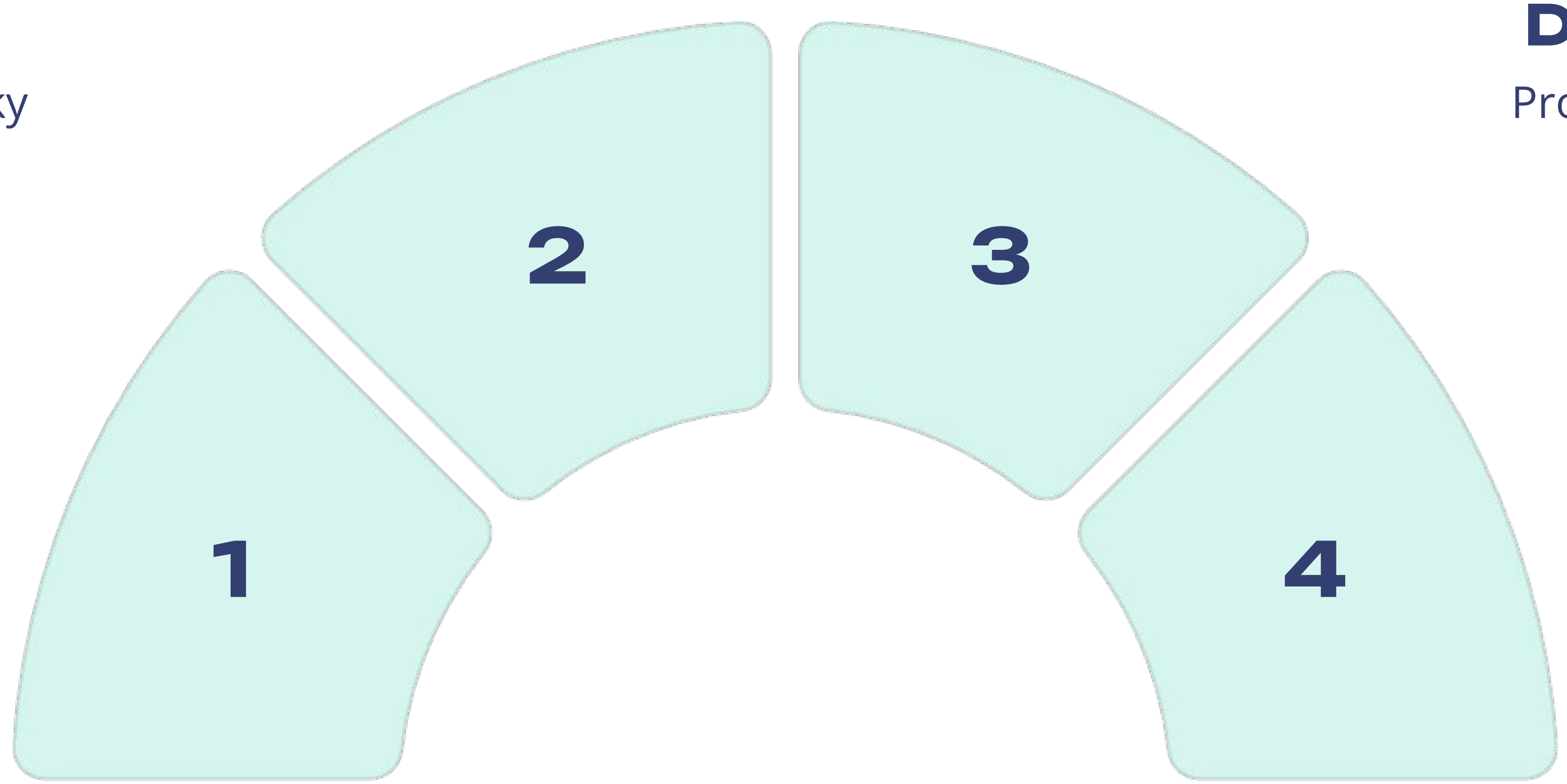
Personalized follow-ups for better collection efficiency

Strategic Policy Adjustments

Refined penalties and discounts for timely payments

Data-Driven Decision Making

Proactive strategies based on predictive analytics



Conclusion & Next Steps

1. Actionable insights to minimize late payments
2. Implement recommendations for improved collections
3. Deploy models in real-time payment monitoring
4. Test and iterate policies for continuous improvement

Thank You!