Ecommerce And Retail B2B Case Study

Done By : Swapnil Johri

Background

Company Overview : Schuster is a multinational retail company specializing in sports goods and accessories.

Challenges Faced

- Significant business conducted with numerous vendors under credit arrangements.
- Frequent late payments from vendors impacting cash flow and operational efficiency.
- Manual efforts required to chase payments, resulting in time and resource inefficiencies.

Objective

- Analyze customer transaction data to understand payment behaviors.
- Segment customers based on historical payment patterns.
- Predict the likelihood of delayed payments against open invoices to optimize collection strategies.

01

Implementation

Data Preparation

Data Cleaning:

- Remove duplicate rows.

- Handled missing values:
  - Replaced missing document numbers with "Unknown".
  - Dropped rows with missing receipt dates.

Data Transformation:

- Converted receipt date to datetime format.
- Extracted year and month from receipt date.

Customer Type	Customer_Name	Customer Account No	Transaction Number	Transaction Date	Payment Term	Due Date	Transaction Currency
3rd Party	GIVE Corp	49144.0	1.0021E+11	2021-12-21	Immediate	2021-12-21	AED
Related Party	AL J Corp	23152.0	1.0022E+11	2022-01-02	30 Days from Inv Date	2022-03-03	USD
Related Party	AL J Corp	23152.0	1.0022E+11	2022-03-24	30 Days from Inv Date	2022-04-23	USD
Related Party	AL R Corp	23312.0	1.0022E+11	2022-04-01	15 Days from Inv Date	2022-01-19	AED
Related Party	ALLI Corp	7530.0	1.0022E+11	2022-03-03	30 Days from EOM	2022-04-30	AED

Due Date	Transaction Currency	Local Amount	Transaction Class	AGE	USD Amount	INV_CREATION_DATE
2021-12-21	AED	-3,088	CREDIT NOTE	105	-3088	2021-12-21 12:53:00
2022-03-03	USD	2,000	INVOICE	33	2000	2022-02-01 14:09:00
2022-04-23	USD	2,000	INVOICE	-18	2000	2022-03-24 17:46:00
2022-01-19	AED	2,415	INVOICE	76	2415	2022-01-05 11:49:00
2022-04-30	AED	3,800	INVOICE	-25	3800	2022-03-03 22:30:00

Feature Engineering

Derive Payment Terms Feature:

Calculate payment terms as the difference between due date and invoice creation date for both received payment data and open invoices data.

Calculate Customer Payment Statistics - Received Payment Data:

- Group received payment data by customer name.
- Compute mean and standard deviation of payment time.

CUSTOMER_NAME	avg_payment_time	std_payment_time
3D D Corp	32.372093	15.387431
6TH Corp	78.444444	7.264832
A3 D Corp	26.600000	2.190890
ABC Corp	90.000000	0.000000
ABDU Corp	70.941406	30.821187

Perform Clustering on Customer Segments:

- Utilize K Means clustering to segment customers based on payment behavior.

CUSTOMER_NAME	avg_payment_time	std_payment_time	segment
3D D Corp	32.372093	15.387431	2
6TH Corp	78.444444	7.264832	0
A3 D Corp	26.600000	2.190890	2

Add Features to Training Data:

Merge customer payment statistics with received payment data based on customer name.

Integrate average and standard deviation of payment time as features for machine learning models.

Model Building And Evaluation

Model Selection:
Trained and evaluated various machine learning models, including:

Logistic Regression
Random Forest
Gradient Boosting

Model Evaluation:

- Random Forest model achieved the highest accuracy of 90.72% with balanced precision and recall for both classes.
- Logistic Regression model showed poor performance with an accuracy of 65.86% due to low precision and recall for class 0.
- Gradient Boosting model performed well with an accuracy of 81.28%, showing a good balance between precision and recall for both classes.

Random Forst Classification Report:

	precision	recall	f1-score	support
0	0.89	0.83	0.86	9532
1	0.91	0.95	0.93	18650
accuracy			0.91	28182
macro avg	0.90	0.89	0.89	28182
weighted avg	0.91	0.91	0.91	28182

Logistic Regression Classification Report:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	9532
1	0.66	1.00	0.79	18650
accuracy			0.66	28182
macro avg	0.33	0.50	0.40	28182
weighted avg	0.44	0.66	0.53	28182

Gradient Boosting Classification Report:

	precision	recall	f1-score	support
0	0.80	0.59	0.68	9532
1	0.82	0.93	0.87	18650

02

Visualization

Customer Segmentation Based on Payment Patterns

Insight: Segment 0 (Purple): These customers exhibit high variability in payment behavior.

Average Payment Time: Approximately -500 days (early payments).

Standard Deviation: Around 400 days (inconsistent payments).

1) **Segment 1 (Pink):** Customers in this segment consistently make late payments.

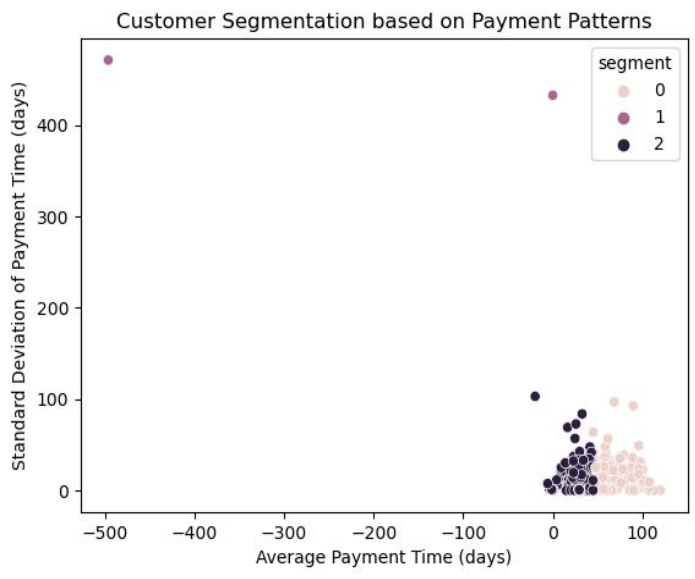
Average Payment Time: Approximately 75 days (late payments).

Standard Deviation: Low, around 25 days.

2) **Segment 2 (Brown):**Largest segment with consistent, on-time payments.

Average Payment Time: Near zero (on-time payments).

Standard Deviation: Low variability.



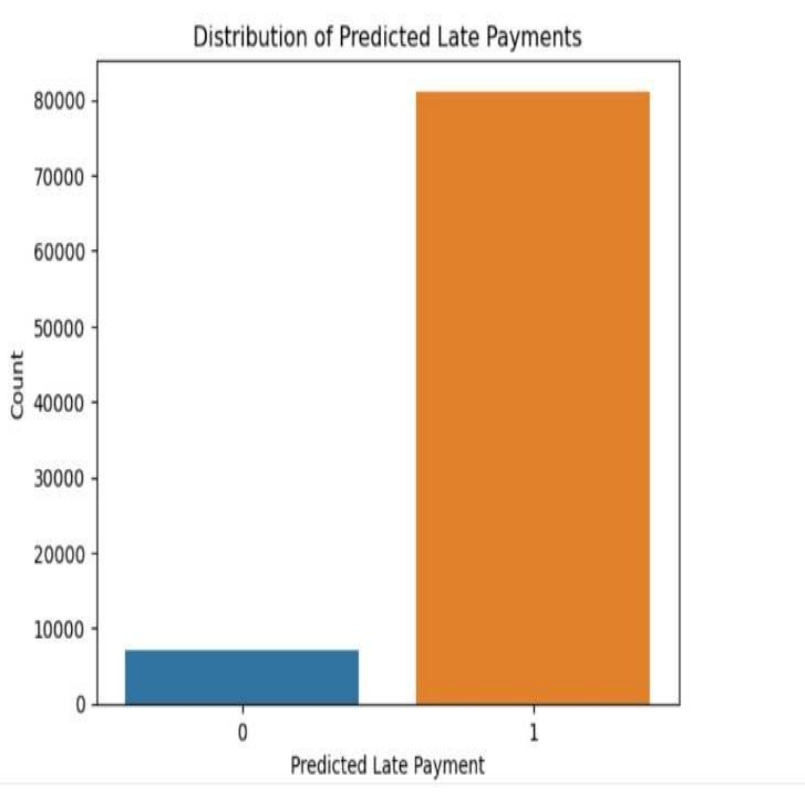
Distribution of Predicted Late Payments

1) **Late Payment Risk:** The majority of predictions fall into the “1” category, indicating a high risk of late payments. This suggests that a substantial portion of the population is likely to delay payments.

2) **On-Time Payments:** The “0” category, representing on-time payments, has a much smaller count. This implies that a smaller proportion of individuals consistently make payments on time.

3) **Imbalanced Distribution:**

   a) Skewed distribution highlights the need for targeted strategies to manage late payments.
   b) Actionable steps include personalized reminders, flexible payment options, and credit counseling.



List of all customer that has more than 80% chances of Late Payment.

562 Customers are at High Risk of Late Payments.

Number of Customers at High Risk of Late Payments: 562

Corp	AL W Corp	ASIA Corp	BLUE Corp	DANU Corp
6TH Corp	AL Y Corp	ASSI Corp	BOSC Corp	DAR Corp
A3 D Corp	AL Z Corp	ASWA Corp	BOTT Corp	DARS Corp
AALA Corp	AL-B Corp	AT K Corp	BOUT Corp	DARW Corp
ABC Corp	ALAM Corp	ATEL Corp	BRAN Corp	DAS Corp
ABU Corp	ALBU Corp	ATLA Corp	BRIT Corp	DASH Corp
ABWA Corp	ALDA Corp	AVEN Corp	BRUS Corp	DCLA Corp
ACQU Corp	ALEZ Corp	AXEL Corp	C EA Corp	DEBE Corp
ACTO Corp	ALFA Corp	AYAN Corp	CAIR Corp	DELH Corp
ADAM Corp	ALGH Corp	AZHA Corp	CANA Corp	DFS Corp
ADEL Corp	ALIC Corp	B8TA Corp	CAPS Corp	DGL Corp
AFAQ Corp	ALL Corp	BACC Corp	CARO Corp	DH S Corp
AHAD Corp	ALLI Corp	BAHA Corp	CARR Corp	DIFF Corp
AHIA Corp	ALMA Corp	BAHR Corp	CART Corp	DIGI Corp
AHLA Corp	ALNE Corp	BAM Corp	CASH Corp	DIWA Corp
AHME Corp	ALRA Corp	BAND Corp	CDAL Corp	DOLC Corp
AIN Corp	ALRY Corp	BARC Corp	CDC Corp	DOME Corp
AISH Corp	ALSH Corp	BASE Corp	CELI Corp	DORA Corp
AJIA Corp	AMAT Corp	BASM Corp	CG B Corp	DR M Corp
AJMA Corp	AMER Corp	BAWA Corp	CGT Corp	DUBA Corp
AL D Corp	AMI Corp	BAZL Corp	CHAL Corp	DUFR Corp
AL E Corp	AMOU Corp	BEAU Corp	CHAN Corp	DURR Corp
AL F Corp	ANFA Corp	BECC Corp	CHAU Corp	DUSI Corp
	ANSA Corp	BEHB Corp	CHER Corp	DUVA Corp