AI CASHFLOW OPTIMIZATION PORTFOLIO

Introduction

In today's cash-sensitive business environment, optimizing the Accounts Receivable (AR) cycle is critical. This project leverages **machine learning** to drive actionable insights from customer payment patterns and streamline AR operations. Its core objectives are to:

1. **Predict late payments** using supervised learning models.
2. **Segment customers** based on historical payment behavior through unsupervised clustering.
3. **Recommend intelligent collection strategies** that are both data-driven and operationally effective.

1. Data Generation

Purpose:
Synthetic data was created to simulate a realistic Accounts Receivable (AR) dataset for customers. This includes fields like invoice dates, due dates, amounts, and payment behavior.

Key Actions:

- Used numpy and pandas to generate random but structured financial data.
- Simulated invoice and payment dates, customer features, credit scores, and payment status.

Output:
A synthetic dataset representing customer transactions and behavior for cash flow analysis.

---

2. Data Cleaning & Preprocessing

Purpose:
To ensure the data is clean, consistent, and ready for analysis and modeling.

Key Actions:

- Converted date columns to date and time format.
- Calculated derived features like days-late and paid-late.
- Handled null or erroneous values.

Output:
A well-structured and cleaned dataset suitable for both supervised and unsupervised learning.

---

3. Exploratory Data Analysis (EDA)

Purpose:
Understand patterns, trends, and relationships in the data to gain initial business insights.

Key Actions:

- Visualized late payment distribution using histograms.
- Analyzed relationships between late payments and features like credit score or invoice amount.
- Identified segments of customers with higher risk of late payments.

Output:
Key insights about customer payment behavior, used to inform model building.

---

4. Late Payment Prediction (Supervised Learning)

Purpose:
Predict which customers are likely to pay late, enabling proactive cash flow strategies.

Key Actions:

- Feature engineering: created variables like invoice-to-due-days, credit-score, historical-late-rate.
- Trained models: Logistic Regression, Random Forest, XGBoost.
- Evaluated using accuracy, precision, recall, and ROC AUC.

Output:
A predictive model that classifies customers as likely on-time or late payers.

---

5. Customer Segmentation (Unsupervised Learning)

Purpose:
Group customers based on payment and financial behavior to tailor strategies.

Key Actions:

- Used KMeans clustering on features like avg-invoice-amount, late-rate, and credit-score.
- Determined optimal cluster number using the Elbow Method.
- Visualized clusters using PCA for dimensionality reduction.

Output:
Clusters of customers with similar behavior—helpful for targeted collection strategies.

---

6. Strategy Recommendation Engine

Purpose:
Provide AI-driven, data-backed recommendations to improve cash flow and reduce late payments.

Key Actions:

- Mapped clusters to action plans (e.g., early reminders for high-risk customers).
- Matched predicted late payers with appropriate collection strategies.
- Created a decision framework based on prediction and segmentation results.

Output:
An actionable recommendation system that supports business decision-making.
