Data-Driven Credit Risk Assessment

Predicting Customer Default for Smarter Lending Decisions
Capstone Project Presentation

JATIN PURI [23f3000916] [MAY 2025 TERM]

About the Business

- FinSecure Corp is a mid-sized financial services firm specializing in consumer lending products.
- Operates in the B2C segment offering credit cards.
- Target audience: Middle-income urban consumers.

Core Challenge:

- The company is facing **rising customer default rates**, which negatively impacts profitability and increases financial risks.
- There is a critical need to evolve from traditional, intuition-based risk assessment to a more accurate, **data-driven approach**.

The PDM Approach

Problem (P)

- Develop a reliable model to predict defaults and reduce loses.

Data (D)

- **Source**: Secondary dataset from UCI Machine Learning Repository.
- **Scope**: 30,000 customer records with 24 features.

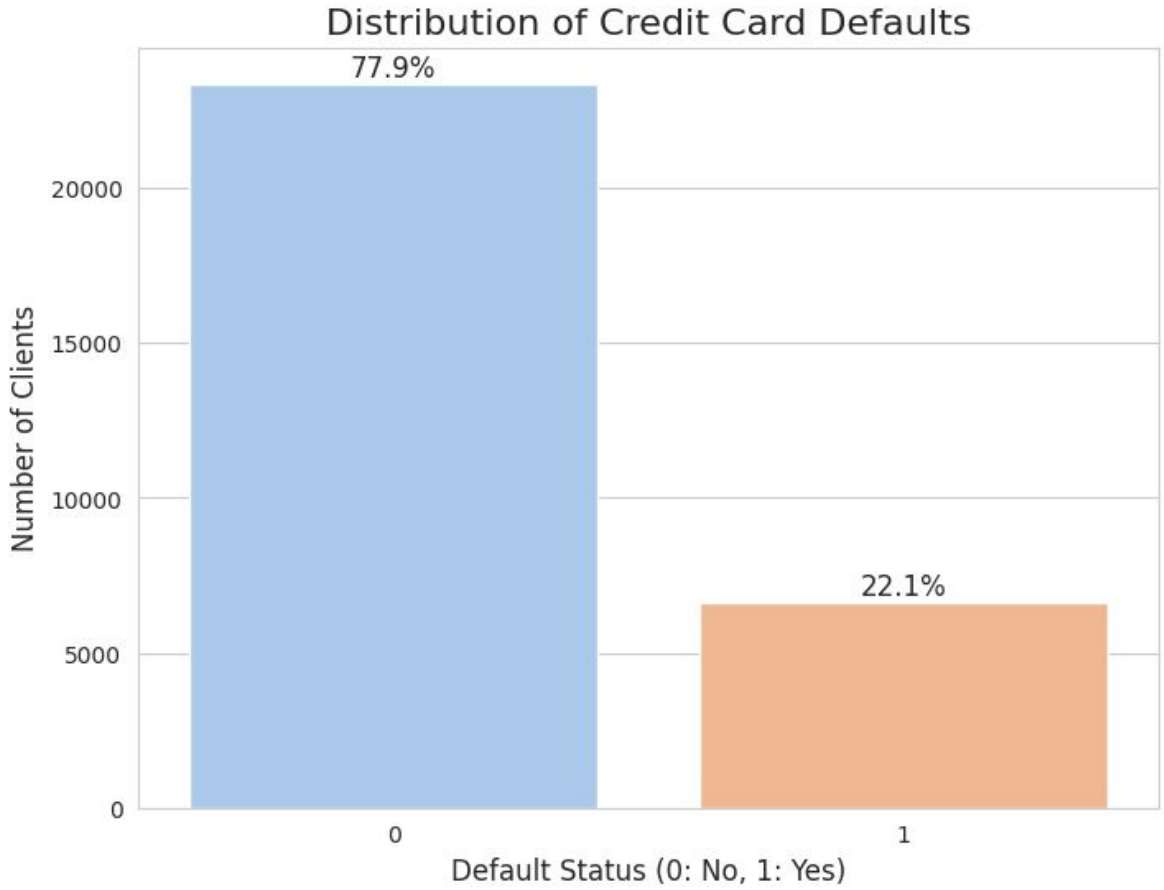
Methodology (M)

- **Analysis**: Exploratory Data Analysis (EDA) to find initial trends.
- **Modeling**: Compared three classification models to find the best fit.
    1. Logistic Regression (Baseline)
    2. Decision Tree
    3. Random Forest (Final Model)

Key Findings from EDA

The Scale of the Default Problem

Initial analysis of the data immediately confirmed the severity of the business problem.

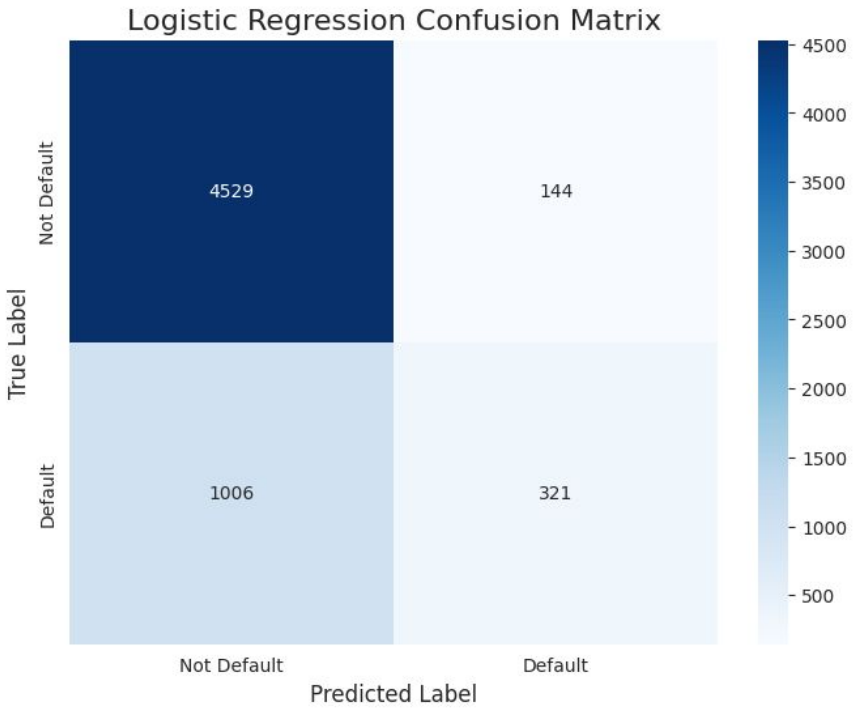


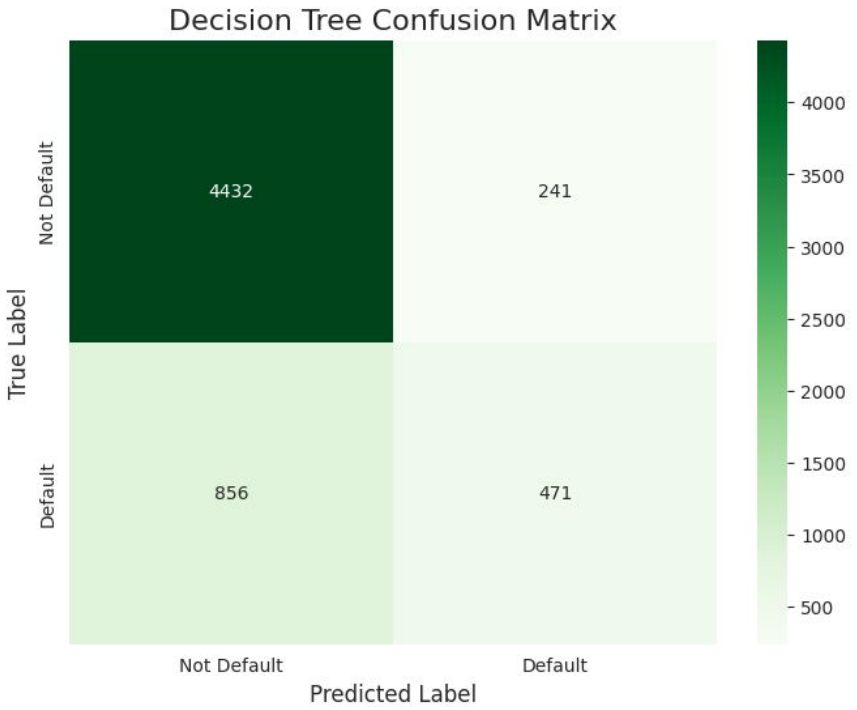
Key Insight:

The data revealed that 22.1% of all clients defaulted on their payments. This substantial number confirms that default risk is a major issue for FinSecure Corp.

Model Comparison

The first two models were evaluated to establish a performance baseline and identify areas for improvement.



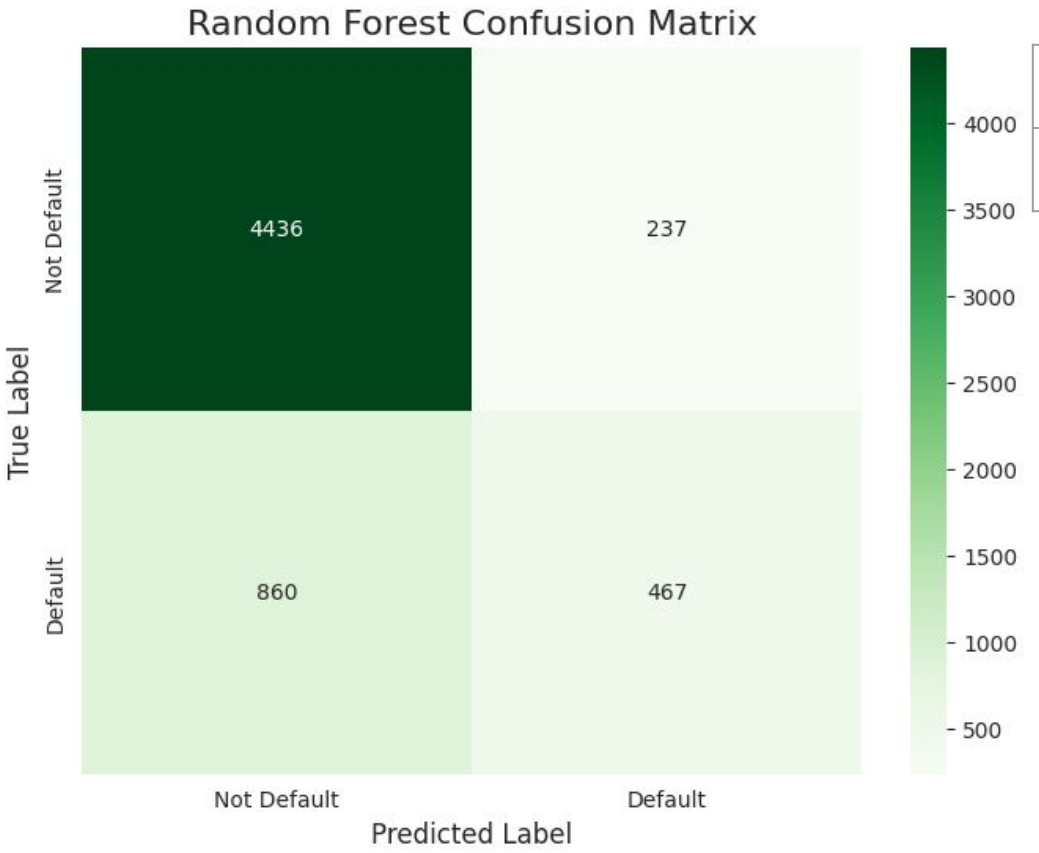


Accuracy: 81.15%	**Accuracy**: 81.72%
Recall: 23%	**Recall**: 36%

Finding: The Decision Tree significantly outperformed the baseline Logistic Regression model in its ability to find actual defaulters, with **Recall increasing from a low 23% to a much improved 36%.**

Final Model Selection: Random Forest

The Random Forest model provided the best overall performance, making it the most suitable choice for the business.

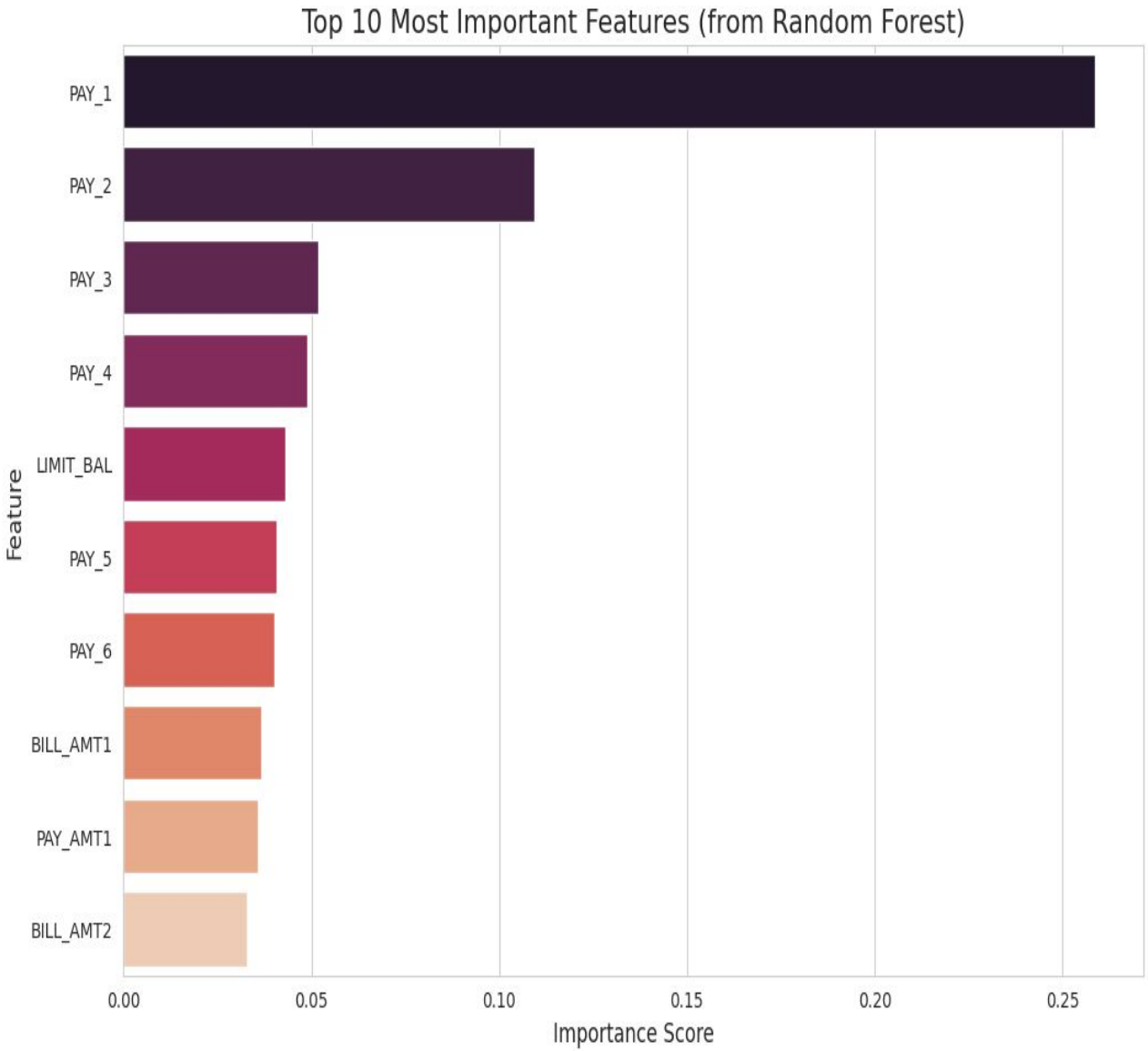


Accuracy	Precision	Recall	F1-Score
0.8190	0.66	0.36	0.47

Justification: While all models had similar accuracy, **Random Forest** was chosen because it achieved the best balance of performance, particularly a **high recall (36%)**. For the business, correctly identifying the correct defaulter is the most critical task to minimize loss.

The Single Most Important Driver of Default

The analysis uncovered one overwhelmingly significant factor that predicts default risk



The "Aha!" Insight: Recent payment behaviour (PAY_1) is by far the most important predictor. This finding shifts the business focus from "**who a customer is** " to "**what a customer is doing now.**"

The Strategic Insight

Credit Risk is Dynamic, Not Static

- The overwhelming importance of recent payment behavior proves that a customer's risk profile is not fixed; it changes based on their **current actions**.
- This challenges traditional underwriting models that rely heavily on static data like education or age.
- The key takeaway is that the **most recent data is the most valuable data**. Our model provides a clear, data-driven framework to operationalize this dynamic view of risk.

Actionable Recommendations

Based on the analysis, three specific, data-driven recommendations are proposed:

1. **Adopt the Random Forest Model for Risk Scoring**
    - **Action:** Integrate the model into the underwriting system to automatically score all new and existing customers.
    - **Impact:** Directly reduce defaults by flagging high-risk applicants.
2. **Prioritize Recent Payment Behavior in Decision-Making**
    - **Action:** Update credit policies to give the highest weight to `PAY_1` and `PAY_2` status in any risk assessment.
    - **Impact:** Shift from a static to a more accurate, dynamic risk strategy.
3. **Develop a Proactive, Tiered Alert System**
    - **Action:** Create automated alerts based on payment delays (Tier 1 for 1-month delay, Tier 2 for 2-month delay).
    - **Impact:** Intervene early to prevent customers from spiraling into serious delinquency.

Thank You

Open to Questions!