LOAN DEFAULT PREDICTION & EXPECTED LOSS ESTIMATION

October 2025

Rohan Veer

PROJECT OVERVIEW

- **Key insights and recommendations**

- Data used for prediction

- Models applied and results

- **Expected loss calculation**

- Key insights and recommendations

- Strategic recommendations for risk management

BUSINESS OVERVIEW

• Personal loan defaults have risen above expectations.

• Predicting defaults helps banks allocate sufficient capital reserves.

• Objective: Build a model to estimate Probability of Default (PD) and Expected Loss (EL).

OUR APPROACH

- Started by understanding the business objective: estimate Probability of Default (PD).

- Explored borrower data to identify key financial indicators affecting default risk.

- Selected two models for testing: Logistic Regression and Random Forest.

- Compared model results using Accuracy and ROC-AUC metrics.

- Integrated Probability of Default with Expected Loss calculation.

MODEL PERFORMANCE

Model	Accuracy	ROC-AUC
Logistic Regression	0.9955	0.99557
Random Forest	0.9945	0.99969

Both models show excellent performance, with Logistic Regression slightly ahead in interpretability and consistency.

EXPECTED LOSS CALCULATION

Example:
Borrower details → Income: ₹60,000; Loan: ₹20,000; FICO: 620
Predicted PD = 99.7%
Recovery Rate = 10%
Expected Loss = 0.997 × (1 – 0.10) × 20,000 = ₹17,954

Formula:

Expected Loss (EL) = PD × (1 – Recovery Rate) × Loan Amount

High debt and low FICO lead to higher default risk.

INSIGHTS & RECOMMENDATIONS

- FICO score and total debt are the strongest predictors of default.

- Integrate this model into JP Morgan's risk assessment framework.

- Use Expected Loss output to guide capital allocation and pricing.

- Regularly retrain models with new borrower data for accuracy.

- Consider adding macroeconomic variables for long-term improvement.

THANK YOU!

Rohan Veer