Credit Risk Prediction Project Report

1. Business Problem

Credit card companies seek to proactively identify customers who are at high risk of defaulting on their credit repayments. This prediction helps them reduce financial risk and optimize lending strategies.

2. Methodology

- Data Cleaning: Removed ID columns, filled missing values.
- Feature Engineering: Created Total_Delays, Avg_Payment_Ratio, Payment_Consistency, Credit_Utilization.
- Preprocessing: Scaled numeric features, encoded categorical ones.
- Modeling: Trained Logistic Regression, Random Forest, and SVM models using pipelines.
- Evaluation: Accuracy, Precision, Recall, F1 Score, ROC AUC, Confusion Matrix, ROC Curve.

3. Insights

- Total_Delays and Credit_Utilization are strong predictors of default.
- Random Forest offered the highest performance across most metrics.
- ROC AUC was used as a key metric to compare model effectiveness.

4. Conclusion

The project provides a comprehensive approach to identifying high-risk customers using data-driven modeling, with visual and quantitative support for selecting the best model.