

# 'PW Skills' Project for Internship

***Project Title:*** Credit Card Default Prediction

***Technology:*** Machine Learning Technology

***Domain:*** Banking

***Project Difficulty Level:*** Intermediate

## ***Executive Summary:***

The Credit Card Default Prediction project aims to develop a predictive model that accurately identifies customers who are likely to default on their credit card payments. By leveraging historical credit card data and applying machine learning techniques, the project aims to assist credit card companies in identifying high-risk customers, enabling them to take proactive measures to minimize financial losses.

## ***Introduction:***

Defaulting on credit card payments poses significant financial risks for credit card issuers. By accurately predicting credit card defaults, companies can implement effective risk management strategies, such as limiting credit lines, offering targeted repayment plans, or issuing warnings to high-risk customers. This project aims to develop a robust credit card default prediction model to assist credit card companies in making informed decisions regarding risk assessment and mitigation.

## ***Objectives:***

- Develop a predictive model that accurately predicts credit card defaults.
- Identify key factors contributing to credit card defaults.
- Enable credit card companies to proactively manage high-risk customers.
- Minimize financial losses associated with credit card defaults.

## ***Methodology:***

a) ***Data Collection:*** Gather a comprehensive dataset containing historical credit card transaction data, customer demographics, credit limits, payment history, and other relevant variables.

b) ***Data Preprocessing:*** Cleanse and preprocess the data to remove duplicates, handle missing values, and perform feature engineering, including scaling, normalization, and one-hot encoding.

**c) Feature Selection:** Identify the most significant features using techniques such as correlation analysis, feature importance ranking, and domain knowledge.

**d) Model Development:** Apply machine learning algorithms, such as logistic regression, decision trees, random forests, gradient boosting, or neural networks, to develop a predictive model for credit card default prediction. Tune hyperparameters to optimize model performance.

**e) Model Evaluation:** Evaluate the model's performance using appropriate evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Utilize techniques like cross-validation and train-test splits to assess the model's generalization ability.

**f) Model Interpretation:** Interpret the model's results to identify the key factors influencing credit card defaults. Analyze feature importances, coefficients, or SHAP values to gain insights into the underlying patterns and factors contributing to defaults.

## *Project Deliverables:*

- Detailed analysis and documentation of the dataset used.
- Preprocessed and cleaned dataset ready for modeling.
- Trained credit card default prediction model with optimized hyperparameters.
- Model evaluation metrics and performance analysis.
- Insights into key features influencing credit card defaults.
- Recommendations for risk management strategies based on the model's findings.

## *Timeline:*

- Data collection and preprocessing: 1 Day
- Feature selection and engineering: 1 Day
- Model development and tuning: 2 Days
- Model evaluation and interpretation: 4 Days
- Documentation and report preparation: 1 week

## *Risks and Challenges:*

- Availability and quality of historical credit card data.
- Dealing with imbalanced datasets where defaults are relatively rare.
- Overfitting or underfitting of the predictive model.
- Interpretability of complex machine learning models.

## *Conclusion:*

The Credit Card Default Prediction project aims to provide credit card companies with a powerful tool to identify customers at high risk of defaulting on their credit card payments. By utilizing advanced machine learning techniques, the project seeks to improve risk assessment and management, ultimately minimizing financial losses associated with credit card defaults.