

## **Assignment 1**

### **Business Understanding: Banking – Loan Approval Model**

#### **1. Business Objective**

- The main business objective is to improve the loan approval process in the bank.
- By analysing customer financial data, the bank wants to predict whether a customer will repay the loan or default.
- This helps the bank reduce financial risk and avoid bad loans.
- Another objective is to speed up the loan approval process using data-driven decisions.
- The bank also wants to improve customer satisfaction by providing quick and fair decisions.

#### **Business Success Criteria**

- Reduction in loan default rate
- Increase in successful loan repayments
- Faster loan approval processing time
- Improved risk assessment accuracy
- Increased profitability and reduced Non-Performing Assets (NPAs)

#### **2. Assess Situation**

*(For Banking – Loan Approval Prediction)*

##### **Inventory of Resources**

- Historical loan application data
- Customer financial records
- Credit history data
- Banking experts and loan officers
- IT infrastructure and database systems

##### **Requirements**

- Accurate prediction of loan repayment capability
- Easy-to-understand results for bank managers
- Secure handling of sensitive customer data

## **Assumptions**

- Past loan data reflects customer repayment behaviour
- Customer information is accurate and complete
- Historical patterns can predict future defaults

## **Constraints**

- Data privacy regulations (e.g., banking rules)
- Limited or incomplete customer data
- Imbalanced dataset (fewer default cases)
- Regulatory compliance requirements

## **Costs and Benefits**

### **Costs:**

- Data collection and storage
- Model development and testing
- Skilled data science team
- System maintenance

### **Benefits:**

- Reduced financial losses
- Better credit risk management
- Faster decision-making
- Improved customer trust
- Competitive advantage in banking sector

## **3. Determine Data Science Goals**

### **Data Science Objective (Technical View)**

Build a predictive model to classify loan applications as:

- o Approved (Low Risk)
- o Rejected (High Risk)

OR

- o Likely to Repay
- o Likely to Default

## Data Science Tasks

- Data cleaning and preprocessing
- Handling missing values
- Feature selection (income, credit score, loan amount, etc.)
- Handling class imbalance
- Model training using:
  - o Logistic Regression
  - o Decision Tree
  - o Random Forest
  - o Support Vector Machine (SVM)
  - o Gradient Boosting (optional)

## Data Science Success Criteria

- High prediction accuracy
- High Recall for defaulters (to reduce risky approvals)
- Balanced Precision and Recall
- Low false approval rate
- Stable performance on new data

## 4. Project Plan: Loan Approval Analysis

### Project Plan Overview

This project aims to analyse banking loan data to predict customer repayment behaviour and support data-driven loan approval decisions using machine learning techniques.

### Project Stages and Activities

Stage Activity		Duration
1	Business Understanding	1 week
2	Data Collection & Understanding	2 weeks
3	Data Cleaning & Preparation	2 weeks
4	Model Building	2 weeks
5	Model Evaluation	1 week
6	Deployment	1 week
7	Monitoring & Maintenance	Continuous

## **Resources Needed**

### **Human Resources**

- o Data scientists
- o Data analysts
- o Loan officers
- o Banking domain experts

### **Technical Resources**

- o Secure database systems
- o Cloud computing infrastructure
- o Risk assessment software

### **Tools & Techniques**

- Python – Data analysis and model building
- SQL – Data extraction and management
- Machine Learning Algorithms – Classification models
- Power BI / Tableau – Reporting and visualization

### **Final Outcome**

- A deployed loan prediction system
- Accurate identification of risky applicants
- Faster and automated loan approval process
- Reduced loan defaults
- Improved profitability and customer satisfaction