**Assignment Description:**

This assignment focuses on predicting loan default using the **LendingClub dataset**. Participants will begin by exploring LendingClub’s business model to understand the importance of default prediction in the peer-to-peer lending ecosystem. Using the dataset, participants will perform bivariate analysis of predictors, build a classification model to predict defaults, identify the top predictors, and evaluate model performance using standard classification metrics.

**Assignment Objectives:**

By completing this assignment, participants will be able to:

* **Analyse** LendingClub’s business model and its relevance to credit risk prediction.
* **Explore** and perform bivariate analysis between predictors and loan default outcomes.
* **Build** and implement a classification model to predict loan defaults.
* **Evaluate** model performance using appropriate metrics and **interpret** key predictors.

**Assignment Guidelines:**

* Start with a brief introduction (3–5 sentences) about **LendingClub** and explain its business model.
* Perform **data preprocessing** steps (handling missing values, encoding, scaling if necessary).
* Conduct **bivariate analysis**: For each predictor, analyse its individual relationship with loan default (e.g., income vs. default, loan amount vs. default). Visualizations are encouraged.
* Build a **classification model** (Logistic Regression, Decision Tree, Random Forest, or any suitable algorithm).
* Report the **top 5 features** influencing loan default. Justify findings with visualizations/interpretations.
* Evaluate model performance using metrics such as **Accuracy, Precision, Recall, F1-score, ROC-AUC**.
* Summarize insights and recommendations based on model outputs.

**Deliverables:**

* A **report document** (PDF/DOCX) containing:
* Introduction to LendingClub and business model
* Data preprocessing steps
* Bivariate analysis results with interpretation and visuals
* Model building process and evaluation results
* Insights on top 5 predictive features
* A **presentation deck** (PPT/PDF) summarizing key findings, visuals, and insights.

**Submission Details:**

Deadline: — 10th October,2025

Submit via the assignment link created in the LMS as a zipped folder with the report and presentation deck.

**Evaluation Criteria:**

* Understanding of LendingClub & business context – **10%**
* Data preprocessing & bivariate analysis – **25%**
* Model development & top 5 features – **30%**
* Model evaluation & interpretation – **20%**
* Report quality, presentation clarity, and originality – **15%**  
  **Total: 10 Marks**

**Plagiarism Policy:**

* Ensure all submissions are original.
* Cite all sources used for research.