**Experiential Learning (EC -1) – Assignment 1**

**15 Marks**

**Group Size – 1 to 5 students**

**Problem Statement:**

Dataset credit.csv contains information about credit card debt for hundreds of customers. The response is balance (average credit card debt for each individual) and there are several quantitative predictors: age, cards (number of credit cards), education (years of education), income (in thousands of dollars), limit (credit limit), and rating (credit rating). **You need carry out detailed Exploratory Data Analysis, create linear regression models using OLS, Ridge and Lasso methods and carry out detailed outlier and residual analysis. Proper hyperparameter tuning has to be carried out. Also, comment on the bias variance trade off, that is, overfitting and under fitting aspect of the models created.**

*Income* - income in thousands of dollars

*Limit* - credit limit

*Rating* - credit rating

*Cards* - number of credit cards

*Age* - Age of the customer

*Education* - years of education

*Own* - house ownership

*Student* - student status

*Married* - marital status

*Region* - East, West or South

*Balance* - average credit card debt for each individual

Note that this is an open ended assignment, you are expected to carry out data exploration, extensive model selection process, model diagnostics and testing etc. A Machine Learning project is generally not linear but will consist broadly of the following major steps.

Define Problem

Prepare Data

Data understanding (EDA)

Model building

Evaluate

Improve Model

Report Results

Evaluation will be relative based upon your results and efforts put in.

**What will you have to submit?**

You need to submit a jupyter notebook (.ipynb) file with all the work done via the elearn(Taxila) portal. Naming convention of the file should be <BITS-ID-number>.ipynb.

First cell of the Jupyter notebook should have the list of Group members and their BITS ID.

**Deadline:**

15th September, 2025

Note that this is a hard deadline and there is NO scope of any extension. Late submissions will NOT be entertained. Copying efforts will be dealt very seriously, students involved, if caught, will be given zero marks.