Department of Artificial Intelligence, SVNIT,SURAT
B.Tech-III ,SEM-V
Subject- Machine Learning(AI301)

LAB ASSIGNMENT-1

1. The goal of this assignment is to **develop and implement a simple linear regression model** using Python to understand the linear relationship between two continuous variables. Simple linear regression is a fundamental supervised machine learning technique that models the relationship between a single independent variable XXX and a dependent variable YYY.

This aims to:

- Load and preprocess a dataset containing two numeric variables.
- Visualize the data distribution and the correlation between variables.
- Fit a linear regression model to the data.
- Evaluate the model's performance using appropriate metrics (e.g., Mean Squared Error, R^2 score).
- Make predictions on new data using the learned regression model.

The implementation is conducted using Python with key libraries such as `pandas`, `numpy`, `matplotlib`, `seaborn`, and `scikit-learn`.

2. **Problem Statement:**

Consider that a real estate company has the data of real estate prices in Delhi. The company wants to optimise the selling price of the properties, based on important factors such as area, bedrooms, parking, etc.

Essentially, the company wants:

- To identify the variables affecting house prices, e.g., area, number of rooms, bathrooms, etc.
- To create a linear model that quantitatively relates house prices with variables, such as the number of rooms, area, number of bathrooms, etc.
- To know the accuracy of the model, i.e. how well do these variables predict the house prices