Name: M Shan Iftikhar

Roll No: BSAI-4A-017

Lab task 1

Kaggle Competition: House Price Prediction

The goal of this challenge is to determine the sale prices of homes based on features such the size of the house, the neighborhood, the amount of rooms, etc. This is a regression problem, where the model must predict the continuous variable SalePrice.

Steps and Approach:

Data Loading:

We are using Pandas to load CSV file data into a DataFrame (train and test datasets).

Handling Missing Values:

For the numerical columns missing values are being filled up using mean of the respective column.

Most frequent category was used to fill the missing values for categorical columns.

Train-Test Split:

Split the dataset into X (features) and y (target, SalePrice) We will then train RandomForest model on this data.

Then model training, model evaluation, and prediction on test data are the last steps.

