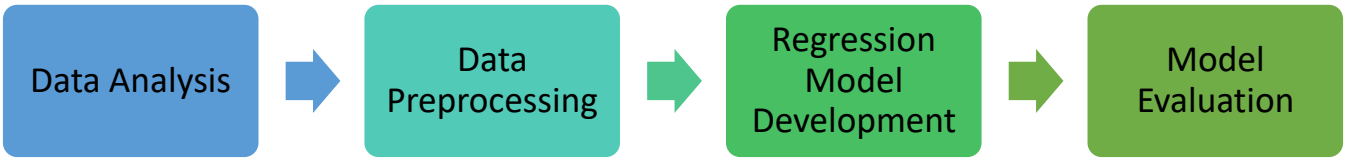
LAB WORK 1 (20 MARKS)

HOUSE PRICING PREDICTION USING REGRESSION

In this lab work assignment, your task is to develop a program that utilizes your knowledge of regression techniques to predict house prices based on the given features. You will need to analyze the dataset, preprocess the data, train, and evaluate different regression models, and compare their performance. The dataset for this task is provided, named 'House Pricing.csv'.

The following is the process flow as your reference:



1. **Data Analysis** - Load and analyze the given dataset of house prices. Explore the dataset to gain insights into the data. Identify the features (e.g., number of bedrooms, area, location, etc.) and the target variable (house prices) in the dataset.
2. **Data Preprocessing** - Perform data preprocessing tasks such as handling missing values, handling categorical variables, and scaling/normalizing the numerical features. Split the dataset into training and testing sets for model evaluation.
3. **Regression Model Development** - Implement a regression model (e.g., linear regression, polynomial regression, etc.) to predict house prices based on the features. Train the regression model using the training dataset and evaluate the model with testing data using appropriate evaluation metrics (e.g., mean squared error, root mean squared error, etc.).
4. **Model Evaluation** - Analyze the performance of the regression model using the evaluation metrics obtained in step 3. Visualize the results, such as predicted vs. actual house prices, and analyze the accuracy of the predictions. Experiment with different hyperparameters (e.g., regularization, learning rate, etc.) to optimize the model's performance.

Submission:

Python notebook/code that implements the regression model for house pricing prediction.

<u>Deadlines for submission: 4th May 2023</u>