LAB WORK 1 (20 MARKS)

Crab Age Prediction Using Regression

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

The following is the process flow as your reference:

1. Data Analysis - Load and analyze the given dataset of crab age. Explore the dataset to gain insights into the data. Identify the features (e.g., sex, length, diameter, height, etc.) and the target variable (age) 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 crab ages based on the features. Train the regression model using the training dataset obtained from step 2. Experiment with different hyperparameters (e.g. regularization, learning rate, etc.) and compare their performance in step 4.

4. Model Evaluation - Analyze the best performance of the regression model on testing set. Please report the evaluation result using minimum MSE, RMSE and R^2.

Submission:
Please submit a link to your google colab/Kaggle notebook or a link to the drive that host your .ipynb file

Deadline for submission: 11:59pm, 5th April 2024