Regression

Regression Assignment Objective:
The objective of this assignment is to evaluate your understanding of regression techniques in supervised learning by applying them to a real-world dataset.

Dataset:
Use the California Housing dataset available in the sklearn library. This dataset contains information about various features of houses in California and their respective median prices.

Key Components to be Fulfilled:

1. Loading and Preprocessing (2 marks):
    - Load the California Housing dataset using the fetch_california_housing function from sklearn.
    - Convert the dataset into a pandas DataFrame for easier handling. Handle missing values (if any) and perform necessary feature scaling (e.g., standardization).
    - Explain the preprocessing steps you performed and justify why they are necessary for this dataset.

2. Regression Algorithm Implementation (5 marks):
    - Implement the following regression algorithms:
        - Linear Regression
        - Decision Tree Regressor
        - Random Forest Regressor
        - Gradient Boosting Regressor
        - Support Vector Regressor (SVR)
    - For each algorithm: Provide a brief explanation of how it works. Explain why it might be suitable for this dataset.

3. Model Evaluation and Comparison (2 marks):
    - Evaluate the performance of each algorithm using the following metrics:
        - Mean Squared Error (MSE)
        - Mean Absolute Error (MAE)
        - R-squared Score (R^2)
    - Compare the results of all models and identify: The best-performing algorithm with justification and the worst-performing algorithm with reasoning.

4. Timely Submission (1 mark):
    - Submit your code in a Jupyter Notebook format via a GitHub link.
    - Ensure that your code is well-documented and explanations are clear and concise.
    - Submission Guidelines: Use Python and Jupyter Notebook for implementation. Submit the GitHub repository link containing the code and all necessary files. Ensure proper documentation, including code comments and markdown explanations.

Total Score: 10