Assignment 1 Question:

Title: Predicting Concrete Compressive Strength

Description: In this assignment, you will explore the Concrete Compressive Strength Dataset, which contains various ingredients used in the production of concrete and their respective compressive strength. Your task is to build a machine learning model that predicts the compressive strength of concrete based on these features.

Tasks:

1. **Data Exploration:**
    - Load the Concrete Compressive Strength Dataset into your programming environment.
    - Explore the dataset to understand its structure, features, and target variable (compressive strength).
    - Visualize the distribution of the target variable and explore relationships between features and compressive strength.
2. **Data Preprocessing:**
    - Handle any missing values or outliers in the dataset.
    - Preprocess the data as needed (e.g., feature scaling, encoding categorical variables).
3. **Model Building:**
    - Choose an appropriate machine learning algorithm for regression based on the nature of the problem and the dataset.
    - Split the dataset into training and testing sets.
    - Train the machine learning model on the training data.
4. **Model Evaluation:**
    - Evaluate the performance of the trained model on the testing data using appropriate evaluation metrics for regression (e.g., Mean Absolute Error, Root Mean Squared Error).
    - Interpret the results and analyze the model's performance.
5. **Model Optimization:**
    - Experiment with different machine learning algorithms and hyperparameters to improve the model's performance.
    - Perform cross-validation and hyperparameter tuning to optimize the model.
6. **Conclusion:**
    - Summarize your findings and conclusions from the analysis.
    - Discuss the potential applications and implications of the developed model for predicting concrete compressive strength.

Submission:

- Submit your assignment report, including code snippets, visualizations, model evaluation results, and interpretations using Jupyter Notebook within 7 days after it is assigned.
- Provide explanations and insights into your approach, choices, and findings throughout the assignment.

Note: You need to use these regression algorithms: Linear Regression, Decision Tree Regression, Random Forest Regression for building the predictive model. Experiment with focusing on data preprocessing techniques such as data manipulation, creation of new features etc. to achieve the best performance.