Lab: Course project (7th sem ML)

Title: Building a custom project that address the real-world problem.

Objective: This lab assignment aims to provide a hands-on experience in tackling a real-world problem using custom model development. By implementing a solution using Python or C++ without relying on existing machine learning libraries, you will gain a deep understanding of the underlying concepts and processes involved in building models from scratch.

Task: Your task is to address a real-world problem by developing a custom model that provides a solution. You will work with a realistic dataset, implement the solution using Python or C++, and create a comprehensive project package, including a report, presentation, and code.

Problem Statement: Choose a real-world problem that interests you, classifying customer reviews, identifying disease outbreaks, or estimating energy consumption. Your solution should involve building a custom model that learns patterns from dataset.

Requirements:

1. **Dataset**:
   Choose or create a realistic dataset that is relevant to your chosen problem. The dataset should have enough features to support meaningful analysis and model development.
2. **Model Development**:
   a. Develop a custom model from scratch to solve the chosen problem. You can explore algorithms like linear regression, k-nearest neighbors, decision trees, or neural networks.
   b. Implement data preprocessing steps, such as handling missing values, normalizing data, or encoding categorical variables.
3. **Visualization of Dataset**:
   a. Generate visualizations to explore the dataset's characteristics
   b. Use plots, graphs, and charts to provide insights into the data's patterns and distributions.
4. **Results and Evaluation**:
   a. Train your custom model using the prepared dataset.
   b. Evaluate the model's performance using appropriate metrics, such as accuracy, precision, recall, or mean squared error, depending on the problem type.
5. **Handling Overfitting or Underfitting**:
   a. Analyze if your model is suffering from overfitting or underfitting.
   b. If needed, propose strategies to mitigate these issues, such as adjusting model complexity, incorporating regularization, or modifying training data.
6. **Preparing document**:
   A concise report detailing the problem, dataset, methodology, results, insights, and challenges faced. A presentation summarizing the project's key aspects, solution approach, results, and visualizations. Well-commented code files showcasing your custom model implementation, data preprocessing, and visualization code.

Submission:
Submit the final project containing the report, presentation, and code files. Ensure that the code is well-structured and easy to understand.

Grading:
Evaluation will consider:
The relevance and complexity of the chosen problem. Effectiveness of the custom model in addressing the problem. Quality and clarity of the report and presentation. Demonstrated understanding of model evaluation and handling of overfitting/underfitting.