# Advanced Machine Learning Techniques Task Report

## Objective

The objective of this task was to apply advanced machine learning techniques, including ensemble methods and dimensionality reduction, on the Iris dataset. The techniques used include Random Forest, Gradient Boosting, and Principal Component Analysis (PCA).

## Dataset

**Iris Dataset**: The dataset contains 150 samples of iris flowers, each described by four features: sepal length, sepal width, petal length, and petal width. The objective is to classify the flowers into three species: setosa, versicolor, and virginica.

## Steps Performed

### 1. Loading and Preprocessing the Data

- Loaded the Iris dataset.

- Split the data into training and testing sets with a 70-30 split.

### 2. Implementing and Evaluating Advanced ML Models

**Random Forest Classifier**:

- Built a Random Forest classifier, which constructs multiple decision trees and aggregates their outputs.

- Trained the model on the training set and evaluated it on the test set.

- Achieved perfect accuracy, indicating very high performance on this dataset.

**Gradient Boosting Classifier**

- Built a Gradient Boosting classifier, which iteratively builds models that correct the errors of previous models.

- Trained the model on the training set and evaluated it on the test set.

- Also achieved perfect accuracy, demonstrating the effectiveness of gradient boosting even on small datasets.

**Principal Component Analysis (PCA):**

- Applied PCA to reduce the dimensionality of the dataset.

- Visualized the data in a two-dimensional space using the first two principal components, effectively separating the three iris species.

### 3. Fine-tuning and Hyperparameter Optimization

- Conducted hyperparameter tuning for the Gradient Boosting model using GridSearchCV to find the optimal parameters.

- Evaluated the optimized model on the test set, which continued to achieve perfect accuracy, validating the tuning process.

## Model Evaluation and Comparison

**Random Forest Classifier**:

-**Accuracy**: 1.0

- **Precision, Recall, F1-score**: All metrics were 1.0 for each species (setosa, versicolor, virginica).

**Gradient Boosting Classifier**:

- **Accuracy**: 1.0

- **Precision, Recall, F1-score**: All metrics were 1.0 for each species.

**Best Gradient Boosting Classifier (after hyperparameter tuning)**:

-**Accuracy**: 1.0

- **Precision, Recall, F1-score**: All metrics were 1.0 for each species.

## Summary of Results

All three models (Random Forest, Gradient Boosting, and the optimized Gradient Boosting) achieved perfect accuracy on the Iris dataset. This indicates that the models were highly effective in classifying the iris species based on the given features. The application of PCA provided additional insights by visualizing the separation between different species in a lower-dimensional space.

## Conclusion

The task successfully demonstrated the application of advanced machine learning techniques on the Iris dataset. Both ensemble methods (Random Forest and Gradient Boosting) and dimensionality reduction (PCA) were effectively implemented, evaluated, and fine-tuned, leading to perfect classification accuracy. This task highlights the robustness and effectiveness of these advanced techniques in handling classification problems with well-structured datasets.