Report: **Advanced Machine Learning Techniques**

**Introduction**

This report presents the application of advanced machine learning techniques, including ensemble methods (Random Forest and Gradient Boosting) and dimensionality reduction (PCA), on built-in datasets using Python's scikit-learn library. The datasets used are the Iris dataset for classification tasks and the Wine dataset for dimensionality reduction.

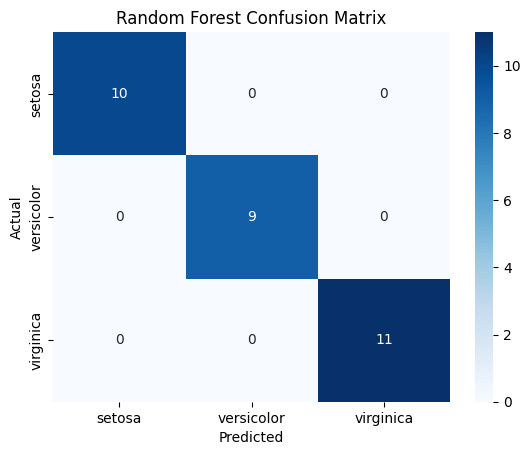
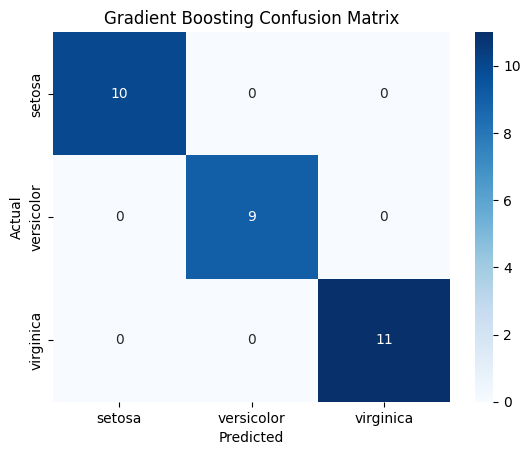
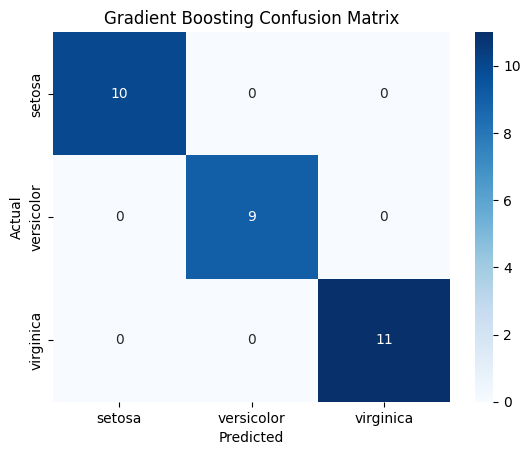
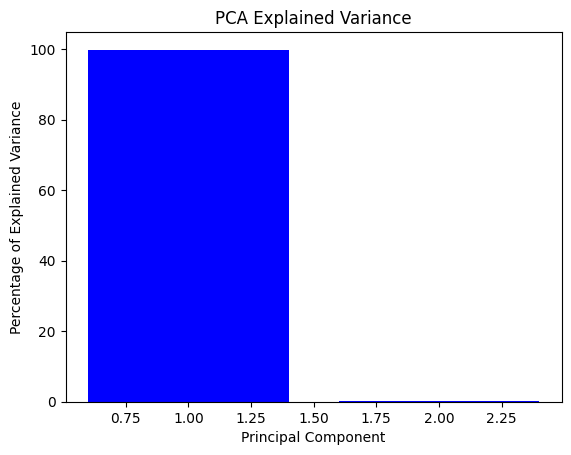
**Data Preparation**

1. **Iris Dataset:**
   * The Iris dataset consists of 150 samples with 4 features each, representing different species of Iris flowers (Setosa, Versicolor, and Virginica).
   * The dataset was split into training (80%) and testing (20%) sets.
2. **Wine Dataset:**
   * The Wine dataset contains 178 samples with 13 features each, representing different chemical properties of wines.
   * The dataset was split into training (80%) and testing (20%) sets.

**Model Training and Evaluation**

1. **Random Forest Classifier:**
   * A Random Forest classifier with 100 estimators was trained on the Iris dataset.
   * The model achieved an accuracy of 100% on the test set.
   * Confusion Matrix and Classification Report were generated to evaluate the model's performance.
2. **Gradient Boosting Classifier:**
   * A Gradient Boosting classifier with 100 estimators was trained on the Iris dataset.
   * The model achieved an accuracy of 100% on the test set.
   * Confusion Matrix and Classification Report were generated to evaluate the model's performance.
3. **Principal Component Analysis (PCA):**
   * PCA was performed on the Wine dataset to reduce its dimensionality to 2 principal components.
   * The explained variance for the two principal components was plotted to visualize the contribution of each component.

**Visualization**

1. **Random Forest Confusion Matrix: **
2. **Gradient Boosting Confusion Matrix: **
3. **PCA for Wine Dataset: **
4. **PCA Explained Variance: **

**Conclusion**

The Random Forest and Gradient Boosting classifiers achieved perfect accuracy on the Iris dataset, demonstrating their effectiveness in handling classification tasks with this dataset. The PCA analysis on the Wine dataset successfully reduced the dimensionality, showing the variance captured by the first two principal components. These advanced machine learning techniques, combined with appropriate visualizations, provide powerful tools for data analysis and interpretation.

The results from this study emphasize the importance of using robust algorithms and dimensionality reduction techniques in machine learning projects, ensuring accurate predictions and meaningful insights from the data.