Wine Classifier Dashboard Project Report

This report presents the implementation of a Machine Learning based interactive dashboard developed using Streamlit and Scikit-learn. The project focuses on the classification of wines using the UCI Wine dataset, which contains chemical features of different wine samples and their corresponding classes. The purpose of this project is to provide an interactive tool where users can train models, adjust hyperparameters, input feature values, and predict wine classes in real-time.

Objectives

- To load and preprocess the UCI Wine dataset.
- To implement machine learning models including Random Forest and Support Vector Classifier (SVC).
- To provide an interactive dashboard using Streamlit for visualization and prediction.
- To evaluate model performance using accuracy, confusion matrix, and classification reports.

Methodology

1. **Dataset Loading:** The UCI Wine dataset was loaded using Scikit-learn's built-in dataset loader. The dataset contains 178 samples with 13 chemical features.
2. **Model Selection:** Users are given an option to choose between Random Forest and Support Vector Classifier (RBF kernel).
3. **Interactive Hyperparameters:** The dashboard allows the adjustment of hyperparameters such as test size, random state, number of estimators, maximum depth (for Random Forest), and C & gamma (for SVC).
4. **Training and Testing:** The dataset is split into training and testing subsets, and the chosen model is trained on the training set.
5. **Performance Metrics:** Accuracy, confusion matrix, classification reports, and feature importances (for Random Forest) are generated.
6. **User Prediction Interface:** A dynamic form allows users to input feature values or use a random sample from the test set to make predictions with probability scores.

Results and Features

- The dashboard displays the test accuracy of the trained model in real-time.
- Quick dataset facts such as number of samples, features, and classes are shown.
- A feature input section allows users to experiment with custom wine feature values and obtain predictions.
- Confusion matrix and classification report provide insights into model performance.
- Random Forest additionally provides feature importance rankings to interpret the model.

Conclusion

The Wine Classifier Dashboard successfully demonstrates how machine learning models can be integrated into an interactive web application using Streamlit. The project allows both exploration of dataset characteristics and real-time predictions, making it an effective learning and demonstration tool for machine learning applications. This project highlights

the practical use of Python libraries such as Scikit-learn, Pandas, and Streamlit in building end-to-end ML applications.