1. Introduction

The purpose of this assignment was to implement, compare, and evaluate different machine learning classifiers on two benchmark datasets: the **Wine dataset** and the **Handwritten Digits dataset**. The classifiers explored were Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), and Random Forests (RF).

The evaluation considered both original and PCA-reduced feature sets. Performance was reported across multiple train-test splits (50:50, 60:40, 70:30, 80:20).

---

2. Methodology

2.1 Datasets

- **Wine Dataset:** 178 samples, 13 features, 3 classes.
- **Digits Dataset:** 1797 samples, 64 features (8×8 images), 10 classes.

2.2 Classifiers Implemented

1. **SVM Classifier**
   - Kernels: Linear, Polynomial, RBF (Gaussian), Sigmoid
   - Parameter tuning via GridSearchCV (`C`, `gamma`, `degree`).
2. **MLP Classifier**
   - Tuned momentum values (0.7, 0.9), learning rates (0.001, 0.01), epoch sizes (200–400).
   - Momentum considered using `solver='sgd'`.
   - Loss curves recorded for convergence study.
3. **Random Forest Classifier**
   - Number of trees (`n_estimators`) and maximum depth (`max_depth`) tuned.
   - Robust to feature scaling but kept consistent preprocessing.

2.3 Evaluation Metrics

- Accuracy, Precision, Recall, F1-score
- Confusion matrix (heatmap)
- ROC curves and AUC
- Learning curves (training vs validation accuracy)
- Dimensionality reduction with PCA (retain 95% variance)

2.4 Experimental Setup

- Train-test splits: **50:50, 60:40, 70:30, 80:20**
- All classifiers tested both with **default parameters** and with **hyperparameter tuning**.
- Outputs automatically saved as plots and CSVs for reproducibility.

---

3. Results

3.1 SVM

- RBF kernel achieved the highest performance among SVMs.
- With tuning, RBF reached **>90% accuracy** on both Wine and Digits datasets.
- Polynomial and Sigmoid kernels performed worse in comparison.

3.2 MLP

- Training and validation curves demonstrated convergence trends.
- Proper tuning of momentum and learning rate improved accuracy significantly.
- Loss curves clearly showed overfitting risk with small datasets (Wine).

3.3 Random Forest

- Achieved stable high accuracy (often >90%) with sufficient trees (100+).
- Robust results across all splits, less sensitive to hyperparameters.

3.4 PCA Impact

- PCA reduced feature dimensions (Wine: 13→~6, Digits: 64→~40).
- In most cases, PCA reduced training time while retaining ~same accuracy.
- Slight performance drop in some SVM kernels, but overall efficiency improved.

---

4. Discussion

- **Best performing classifier:** SVM with RBF kernel and tuned parameters.
- **Most robust classifier:** Random Forest, consistent across datasets and splits.
- **Most sensitive classifier:** MLP, requiring careful tuning of learning rate and momentum.
- **Effect of PCA:** Dimensionality reduction preserved performance while speeding computation.

The assignment illustrates the importance of both algorithm selection and hyperparameter tuning.

---

5. Conclusion

- Implemented SVM (Linear, Poly, RBF, Sigmoid), MLP, and Random Forest classifiers.
- Compared performance across different train-test splits on Wine and Digits datasets.
- Evaluated using multiple metrics (Accuracy, Precision, Recall, F1, AUC, Confusion matrices).
- Added PCA for dimensionality reduction and observed trade-offs.
- Achieved **≥90% accuracy** with tuned models (SVM RBF, Random Forest).

This assignment provided hands-on experience with classical machine learning models, evaluation techniques, and result interpretation.

---

6. References

- Scikit-learn Documentation: https://scikit-learn.org/stable/
- UCI Machine Learning Repository: Wine Dataset
- Scikit-learn load_digits dataset

Github repo link- https://github.com/immu729/Machine-Learning-Lab.git