

ASSIGNMENT #2

Construct a machine learning based model for classification using Python for the following UCI datasets:

UCI datasets (can be loaded from the package itself):

- a. Wine dataset: <https://archive.ics.uci.edu/ml/machine-learning-databases/wine/wine.data>
- b. Handwritten Digit Dataset:
<https://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits>

Implement and compare the following ML classifiers for all the 2 datasets and show the classification results (Accuracy, Precision, Recall, F-score, confusion matrix) with and without parameter tuning:

1. SVM classifier (Linear, Polynomial, Gaussian, & Sigmoid)
 2. MLP classifier (Momentum term, Epoch size and learning rate)
 3. Random Forest classifier
- Apply different values of train-test set splits (Eg.: 50:50, 60:40, 70:30 and 80:20) and report the corresponding results for both the classifiers.
 - Generate the image (heat map) of the confusion matrix for every experimentation. Generate the images of training & loss generation curves for each classifier and for every dataset.
 - For each dataset, generate an image illustrating Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) for every classifier.
 - Use Principal Component Analysis (PCA) for feature dimensionality reduction and again apply the above 3 ML classifiers on the reduced feature set. Show the classification results (Accuracy, Precision, Recall, F-score, confusion matrix).
 - Try to achieve accuracy $\geq 90\%$. Show the performance comparison among classifiers in a table.

Save the assignment in a single pdf file with the naming convention “Full Class Roll No_Full Name.pdf” and upload the report by using the Google form link:

<https://forms.gle/j4wfKa45AmKt9Bsj9>

Submission Deadline: 29th August 2025 Friday (11:59 pm) EOD