

Assignment 3: Multi Layer Perceptron classifier

1 Instructions

- Use python programming language for your implementation.
- Use appropriate approach if you find some attribute is missing in your data.
- Report must contain step-wise description of your implementation and analysis of results. Since data analysis is a crucial task for any machine learning algorithm, report should demonstrate detailed analysis of results and conclusion. It should also clearly mention the steps to run your code.
- Learn the projection matrix for any dimension reduction technique using the train split only. Once the projection matrix has been trained using the train split, use that matrix to reduce the dimension of validation and test splits.
- You can use any python library function to complete the assignment. Use packages of pytorch only.

2 Dataset:

Download Optical Recognition of Handwritten Digits Data Set:

<https://archive.ics.uci.edu/ml/datasets/optical+recognition+of+handwritten+digits>

3 Problem statement: Multi Layer Perceptron classifier

Build a MLP classifier for the given dataset. If required, any data cleaning or pre-processing can be done. Mention them in the report. Use stochastic gradient decent for each of the models.

1. Find the number of nodes in input and output layer according to the dataset and justify it in the report. Specify and justify any other hyper parameter that is/are needed. **5 Marks**

2. Vary the number of hidden layers and number of nodes in each hidden layer. Consider the following architectures.

- (a) 0 hidden layer
- (b) 1 hidden layer with 2 nodes
- (c) 1 hidden layer with 6 nodes
- (d) 2 hidden layers with 2 and 3 nodes respectively
- (e) 2 hidden layers with 3 and 2 nodes respectively

For each of the architectures, vary the learning rates as 0.1, 0.01, 0.001, 0.0001, 0.00001. Any other learning rate may be explored as well. **25 Marks**

3. Plot graph for the results in the previous parts with respect to accuracy. (Learning rate vs accuracy for each model (in one plot) and model vs accuracy for each learning rate (in one plot).) **15 Marks**
4. Mention the architecture and hyper parameters of the best found model in the report. Try to justify it. **5 Marks**
5. Reduce the feature dimension of the data into a two dimensional feature space using Principle Component Analysis (PCA). Plot the reduced dimensional data in a 2d plane. In the plot, all data points of a single class should have same color and data points from different classes should have different colors. **15 Marks**
6. Apply MLP of step 2 in the reduced feature space. Compare with classification output generated from step 2. you may use bet learning rate obtained from step 3 **15 Marks**
7. Prepare a detailed report. **20 Marks**