Assignment : Intro to ML
Deadline : 10/oct/2021 (EOD)

Note : You can solve the problems below using scikit-learn libraries. Follow submission instructions posted by me previously. Many of you are not taking the report seriously, note that it constitutes 50% of total marks. All of your plots, results, analysis, conclusions, justifications must be in the report.

Problem 1:

Method 1:
(i)From Iris flower data set take only two features (sepal width and petal length). Project the labelled data on a 2D graph. Try to classify this data by drawing linear boundaries intuitively. (15 marks)
Method 2:
(ii) Now take all the features and apply MDA on this Iris flower data set and plot the labelled data on 2D graph(15 marks)
(iii)Show the plots of above two methods and write in your own words why method2 works better than method1 (10 marks)

Problem 2:

From MNIST Fashion data take 100 data points belonging to classes Sneaker, Pullover and Ankle boot.
(i) Project these data points on a 2D plane using data reduction techniques PCA, T-SNE and MDA.(25 marks)
(ii) Take 50 data points belonging to classes Sneaker, Pullover and Ankle boot from the test data set and compare the performance of these techniques(PCA, T-SNE, MDA) in terms of accuracy. (25 marks)
(iii) According to your view what is the best dimensionality reduction technique for solving this problem and justify your answer.(10 marks)