

Department of Electronics and Communication Engineering Visvesvaraya National Institute of Technology

Assignment 5: Machine Learning with Python Laboratory; ECL 443. Odd Semester, 2023-2024

Due Date: 20th October 2023.

1. You have to upload your digital form of solutions to the Google classroom page created for this purpose. You should prepare the laboratory report in the given format as mentioned in the introductory class for this course i.e. the laboratory report should contain the Abstract, Introduction, Method/Procedure, Results, Discussions, Conclusions and Appendices [you should paste your code here]. Alongwith the laboratory report, you should also upload the python file in such a manner that it can be ready executed in Spyder and all the required graphs in the correct format are generated. Please make sure that in the laboratory report, the plots are presented in clear and discernible manner.
2. You should name your python file the following. Your complete rollnumber_linreg. For example, if your roll number is BT20ECE001, then your function should be named as BT20ECE001_linreg.
3. Include the answer of all the questions in the Results Section of your report SEQUENTIALLY. The answers will be checked sequentially, so if any question is not according to the sequence, that will not be checked.
4. DO NOT COPY FROM EACH OTHER. TRY EACH QUESTION ON YOUR OWN. WRITE THE LABORATORY REPORT USING YOUR OWN LANGUAGE. REMEBRE, THIS IS A BASIC COURSE AND THE CONCEPTS WILL BE REQUIRED THROUGHOUT YOUR DEGREE PROGRAM AND POSSIBLEY, FOR YOUR JOB/HIGHER STUDY. IF YOU COPY, YOU ARE HARMING YOURSELF.

Implementation of Artificial Neural Network.

Download the ovarian cancer dataset [dataset provided by Mathworks]. Using the dataset, build a classifier that can distinguish between cancer and control/normal patients. However, before training the classifier, you have to compress the dataset using two different algorithms, PCA [Principle Component Analysis] and Autoencoder. Clearly mention all the relevant details like the architectural details of the ANN you used, optimisation function, error function, activation functions etc. Divide your dataset into training, validation and testing dataset using the function you created in the previous lab exercise.

1. Write your own code for performing PCA. For selecting the number of Principle components to be used, you can compute the ratio of the sum of the eigen values corresponding to the principle components to be used to the sum of all eigen values [You can use those Principle components for which the ratio will be 0.95]. Transform the original data in PC space and then perform the classification.
2. Build an Autoencoder with same dimension of latent space as that of the principle components. Train the Autoencoder and use the latent space representation for classification.
3. In all cases, provide confusion matrix, specificity, sensitivity values. Also provide ROC curve and AUROC values for every case. Compare the classification results with those you obtained for Assignment 2 [in which you performed the same classification task without any data compression.]