

# Lab 7 - Dimensionality Reduction using PCA

*Lab 7 on Principal Component Analysis for DS3010 - Machine Learning*

## OVERVIEW & PURPOSE

In this lab, you will experiment with Dimensionality Reduction Algorithm of Machine Learning

## Instructions

1. Please submit the assignment through Moodle in .ipynb format (python notebook)
2. The submission should contain a single notebook containing all the solutions, including the requested documentation, observations, and findings .
3. The naming convention for the notebook  
**<firstname>\_<lastname>\_<rollnumber>.ipynb** .
4. You must adequately comment on the code to improve its readability.
5. The lab is worth **5 points**.
6. This graded lab is due on **10th November at 6 pm** .
7. **Your code and answers will be checked for plagiarism and if found plagiarised, zero marks will be provided for assignment 7. So make sure you actually code and solve the questions rather than noting down the answers .**
8. **Download Data -**  
<https://drive.google.com/file/d/1n2xbgfA9y8RFoYYovFe2OWX3ZpQ8UQML/view>

## Lab

### 1. PCA on MNIST Images Dataset (5 Points)

- a. Read the provided 'mnist.csv' file using Pandas Dataframe. (0.25 Points)
- b. Extract the data of any single label of your choice and same dataset will be used for all the questions mentioned below.
- c. Plot the visualization of data.
- d. Write a PCA algorithm from scratch. (3 Points)
- e. Plot the eigen values.(0.5 Points)
- f. Plot the reconstruction error for 1-dimension to 784-dimension.(0.5 Points)
- g. Plot the visualization with transformed features.(0.25 Points)
- h. Write your observations.(0.5 Points)