

**Department of Artificial Intelligence**

**College of Computer Science and Information Technology**

**Due Date: Tuesday November 5, 2024 @ 11:59 PM**

**Late Submissions:**

* Q: Can I skip the lab and submit the solution?
  + You will receive a mark of **zero** if you do not attend the lab, even if you complete the exercise. Attending the labs is compulsory for evaluation. If you have a justified excuse, you may receive a partial mark depending on the circumstances. See the next question for information on late submissions.
* **Q:** If I submit it at 12:00am, you’ll still mark it, right?
  + **A:** 11:59pm and earlier is on time. Anything after 11:59pm is late. Anything late will **NOT** be probably marked. If I find you have a legitimate cause, you will be graded according to the following rules (24 hours after deadline 🡪 assignment is marked out of 75% only, 48 hours after deadline 🡪 assignment is marked out of 50% only, 72 hours after deadline 🡪 assignment is marked out of 25% only)

The objective of this task is to explore the application of Principal Component Analysis (PCA) for image compression. We will be using the MNIST dataset, which consists of hand-written digits, to investigate how PCA can be utilized to reduce the dimensionality of the images.

**Step 1: Dataset Import:**

1. Import the MNIST dataset using a library like scikit-learn or use

**from** sklearn.datasets **import** fetch\_openml

mnist **=** fetch\_openml('mnist\_784')

1. Display sample images from the dataset to give an overview of the data.

**Step 2: Data Preprocessing:**

1. Reshape the images from 28x28 pixels to a 1D array of 784 elements.
2. Normalize the pixel values to a range between 0 and 1.

**Step 3: Implement PCA:**

1. Implement PCA using a library like scikit-learn.
2. Choose an appropriate number of principal components for compression based on experimentation.

**Step 4: Compression:**

1. Transform the original images using the computed principal components.
2. Reconstruct the images from the compressed data.

**Step 5: Visualization:**

1. Display original and compressed images side by side.

Provide an explanation of your code in your own words. This is to ensure that you have a deep understanding of the code you've written and its underlying concepts. You are expected to comments on the main parts and functions of the code.

Guidelines:

* Your explanation should be original and in your own words. Do not copy explanations from textbooks, online resources, or peers.
* Go beyond just describing what the code does. Explain why you chose certain methods or approaches and how they benefit the solution.
* Document your usage of generative AI if any as per the announced policy (see syllabus).

**Assessment**

1. Each student will show all the above parts running as demo to the Lab Instructor **before leaving the lab.** Total marks for the lab is as follows

|  |  |
| --- | --- |
| Task 1 | Marks (demo + report) |
| 1 | 10 |
| Total | 10 |

1. Students will prepare a report in which they will submit the snapshots taken while they worked on each part. They will explain the figures to make sure that they understood what they did.