

✓ Task 13: PCA – Dimensionality Reduction

Dataset:

- Primary: MNIST dataset
- Alternative: Sklearn digits dataset

Tools:

- Python
- Scikit-learn
- Matplotlib
- Alternatives: OpenCV PCA

Hints / Mini Guide:

1. Load MNIST/digits dataset and flatten images into feature vectors.
2. Scale features using StandardScaler for correct PCA performance.
3. Apply PCA with multiple component counts like 2, 10, 30, 50.
4. Track explained variance ratio for each PCA setting.
5. Plot cumulative variance to choose best number of components.
6. Transform original dataset into reduced dimensional dataset.
7. Train Logistic Regression on reduced dataset.
8. Compare accuracy of original vs reduced dataset model.
9. Visualize PCA 2D scatter plot to observe separation.

Deliverables:

- Explained variance plot
- Reduced dataset
- Accuracy comparison report

Final Outcome:

Intern understands feature compression and variance trade-off.

Interview Questions Related To Above Task:

- What problem does PCA solve?
- What is explained variance?
- Why scaling required for PCA?
- PCA vs feature selection?
- What are PCA limitations?
-

Task Submission Guidelines

-  **Time Window:**

You can complete the task anytime between 10:00 AM to 10:00 PM on the given day. Submission link closes at 10:00 PM.

-  **Self-Research Allowed:**

You are free to explore, Google, or refer to tutorials to understand concepts and complete the task effectively.

-  **Debug Yourself:**

Try to resolve all errors by yourself. This helps you learn problem-solving and ensures you don't face the same issues in future tasks.

-  **No Paid Tools:**

If the task involves any paid software/tools, do not purchase anything. Just learn the process or find free alternatives.

-  **GitHub Submission:**

Create a new GitHub repository for each task.

Add everything you used for the task — code, datasets, screenshots (if any), and a short README.md explaining what you did.

Submit Here:

After completing the task, paste your GitHub repo link and submit it using the link below:

-  [\[Submission Link\]](#)

