**CS5691 –** Pattern Recognition in Machine Learning

Programming Assignment 1

Lab Report

Name: - Jay Jitendra Prajapati

Roll No.: - ME21B143

**TABLE OF CONTENTS**

Tasks Page No.

**1) PCA Algorithm on MNIST dataset 3**

1.1 PCA algorithm . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3

1.1.1 Code for algorithm . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.1.2 Images of principal components . . . . . . . . . . . . . . . . . . . . . .

1.1.3 Variance in the dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.2 Reconstructing the dataset using different representations . . . .

1.2.1 Dimension k for downstream task . . . . . . . . . . . . . . . . . . . . .

1.2.2 Images of new dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.3 Kernel PCA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.3.1 Polynomial Kernel . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.3.2 Gaussian Kernel . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.4 Best suited Kernel for this dataset . . . . . . . . . . . . . . . . . . . . . . . . . .

**2) Clustering Algorithm on given dataset**

2.1 K-Means Algorithm with different random initialization . . . . . . . .

2.2 K-Means Algorithm for K = {2,3,4,5} . . . . . . . . . . . . . . . . . . . . . . . . .

2.3 Spectral Clustering . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

* 1. Variation in Spectral Clustering . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1. **PCA Algorithm on MNIST Dataset**

* Loading the MNIST Dataset from Hugging face using load\_dataset library.
* Extracting 1000 images from dataset by maintaining a count array
* Convert the images into NumPy arrays and flatten them.
  1. **PCA Algorithm**
     1. **Code for Algorithm**
* Step 1: - Center the dataset by calculating the mean and subtracting it from every data point.
* Step 2: - Calculate the Covariance matrix using the formula

C =