Assignment 04
Principal Component Analysis (PCA)

Due Date: 4th June 2022

Guidelines

- Submit all of your code, saved models, and results in a single zip file with the name FirstName_RollNumber_04.zip
- Submit a single zip file containing:
  a) Code (task01.ipynb, task02.ipynb)
  b) Report
- There should be a report.pdf detailing your experience and highlighting any interesting results.
- Email the instructor or TAs if there are any questions.
- **Don't resubmit the dataset provided.**

Overview

In this assignment, you will be implementing a speaker verification system using PCA. The goals of this assignment are:

- Understand how to perform PCA on the given dataset.
- How to extract features from audio data.
- Understand how PCA can be used for speaker verification.
- Understand how PCA can be used for speaker classification.

Dataset

For this assignment, you will be using "**Free Spoken Digit Dataset (FSDD)"**. FSDD is a simple audio/speech dataset consisting of recordings of spoken digits in wav files at 8kHz. The recordings are trimmed so that they have near minimal silence at the beginnings and ends. The dataset consists of:

- 6 speakers
- 3,000 recordings (50 of each digit per speaker)
- English pronunciations

Dataset folder consists of three files **"features.npy"** (contains audio file features where each row represents a single audio file feature)**, "labels.npy"** (contains corresponding labels of audio file)**,** and **"metadata.py"** (contains labels to speaker attributes mapping)**.**

Task 01: Speaker Verification using PCA (45 Marks)

In this task, you need to perform speaker verification using PCA i.e. given two audio files you need to tell whether they belong to the same person or not. For this task, you will be using FSDD dataset.

1) Load the dataset and perform an 80/20 train/test split. **Hint: You can use np.load() to load the .npy files.**

Training:

2) Compute the mean of train data using python.
3) Mean subtract the data using python.
4) Find covariance matrix using python.
5) Compute eigenvalues and eigenvectors of the covariance matrix using python.
6) Compute the principal components and find how much information is captured by each principal component.
7) From the test data randomly select any two audio files and tell whether they belong to the same speaker or not using the nearest neighbor in a low dimensional space. **You will not use any built-in library for KNN or PCA.**
8) Report precision, recall, accuracy, F1 score, confusion matrix, ROC Curve for each class and for all classes combined ROC curve. Also, report Area-Under-the-Curve.

Task 02: Speaker Identification using PCA (45 Marks)

In this task, you need to perform speaker identification using PCA i.e. given an audio file you need to predict the speaker using FSDD dataset.

1) Load the dataset and perform an 80/20 train/test split.

Training:

2) Compute the principal components class-wise.
3) Perform speaker classification using PCA. **(Hint: see the discussion of face classification in the class using PCA.)**

Testing:

4) Report precision, recall, accuracy, F1 score, confusion matrix, ROC Curve for each class and for all classes combined ROC curve. Also, report Area-Under-the-Curve.

Note: For this task, you are not allowed to use any classification algorithm or neural network.

Report (10 Marks)

Task 01:

- Show classification performance with the dimensionality reduction. Which dimension is best for classification using the nearest neighbor.
- Analyze your results and report precision, recall, accuracy, F1 score, confusion matrix, ROC Curve for each class and for all classes combined ROC curve. Also, report Area-Under-the-Curve.

Task 02:

- Write your algorithm for speaker identification
- Analyze your results and report precision, recall, accuracy, F1 score, confusion matrix, ROC Curve for each class and for all classes combined ROC curve. Also, report Area-Under-the-Curve.