ECE661 Fall 2024: Homework 10

TA: Rahul Deshmukh (deshmuk5@purdue.edu)
Due Date: 04 Dec 2024, 11.59 PM
Late submissions will be accepted only for special cases.

Turn in typed solutions via BrightSpace. Additional instructions can be found at BrightSpace. Since this is the last HW of the semester, please submit homework on time to ensure timely grading.

1 Theory Question

1. One of the biggest challenges in designing an AI/ML solution to a problem is making sure that your solution is not overfitting to the training data. The fact that the word "overfitting" occurs around 8 times in the second half of your instructor's Lecture 26 tutorial is a measure of its importance. Summarize in your own words your overall understanding of the notion "overfitting to the training data" as you see it mentioned at the different locations in the tutorial. To answer this question, you should read Section 5.2 of The Deep Learning Book by Goodfellow et. al. [1].

2. One of the things that makes a VAE difficult to understand is what happens between the Encoder that is used for creating the latent representation of the input and the Decoder that, at the least, should be able to reproduce the input at its output. What happens between the Encoder and the Decoder is referred to as the "Reparameterization Trick". It consists of the adding to the mean of the latent representation a random multiple of the learned standard-deviation. Explain in your own words how you understand the "Reparameterization Trick." To answer this question, you should understand the `Autoencoder` class in the provided script `autoencoder.py`.

2 Introduction

This homework consists of two parts:

1. **PCA, LDA and Autoencoder** for dimensionality reduction and nearest neighbor for face recognition.

2. **Cascaded AdaBoost Classifiers** for object detection.

3 Task 1: Face Recognition using PCA and LDA

The goal of this part of the homework is to classify an unknown face image given a database of labeled face images. Your overall approach will involve the following steps:

1. For this homework, you will use the provided subset of the FacePix database [2]. The images have been divided into a training set and a test set, each with 630 images. The name of each image is a string like `XX-YY.png` where `XX` is the identity of the human subject and `YY` is an integer index for the image for that human subject. Note that while the images in both the training set and the test set share the same names, they are different images.

2. Vectorize the images in the training set and compute the covariance matrix for the image vectors. Do not forget to subtract the mean and normalize each image vector to unit magnitude before you calculate the covariance matrix.

3. Use PCA and LDA to create a low-dimensional representation for the training face images. With each approach, you retain p eigenvectors:
   - For PCA, the p eigenvectors are directly the eigenvectors of the covariance matrix. Make sure you use the computational trick for calculating the eigenvectors of the covariance matrix as discussed in Section 2.1 of Prof. Kak's Optimal Subspaces Tutorial [3].
   - For LDA, the eigenvectors are of the matrix $S_W^{-1} S_B$ where S_W is the within-class scatter matrix and S_B is the between-class scatter matrix. Should it happen that your S_W is singular, then you must use Yu and Yang's algorithm for finding the LDA eigenvectors (refer to Section 3.5 of [3]).

4. For classification, first project all images of the training set into the p–dimensional subspace. Now project each image in the test set into the same p-dimensional subspace and use *your own implementation* of the Nearest Neighbor (NN) algorithm for classification. For this assignment, show the classification accuracy for both PCA and LDA as a function of the subspace dimensionality p. Subsequently, compare the results for PCA and LDA and how they change with p.

5. Finally, plot the p-dimensional embeddings along with labels for Train and Test using UMAP [4] embeddings as a scatter plot. Here are the steps to do so:
   - Install the UMAP library using the command: `pip install umap-learn`. Read the documentation to understand how to use it.
   - Use the UMAP library to reduce the dimensionality of the training and test data to 2 dimensions.
   - Plot the 2D UMAP embeddings of the training and test data as a scatter plot using different colors for different classes. For the training data, use the ground truth labels for coloring. For the test data, use the predicted labels from the nearest neighbor algorithm.
   - Plot the training and test data in two different plots. Include such plots for both PCA and LDA for a few values of p.
   - Discuss the results and compare the embeddings of PCA and LDA.

4 Task 2: Face Recognition using an Autoencoder

An autoencoder is a type of neural network that is commonly used for dimensionality reduction. It is comprised of two neural network based components: an encoder and a decoder. The encoder takes in an high-dimensional sample such as an image and outputs a p-dimensional vector representation just like PCA and LDA. The decoder then takes in the p-dimensional vector representation and outputs a sample (e.g. an image) in the original high-dimensional sample space. During training, the autoencoder refines itself by attempting to regenerate the input sample from the corresponding p-dimensional vector representation.

In this homework, you will be provided with an autoencoder that is pretrained on the same face dataset [2]. Your task is to implement the rest of the logic to perform the same face recognition task as described in Task 1. Here are the steps:

1. Familiarize yourself with the provided script `autoencoder.py`. It contains the autoencoder class, which is implemented in PyTorch, along with the necessary logic that handles: loading of the pre-trained weights, loading of the training and test data, and encoding images into the p-dimensional representation vectors. Additional guidelines can be found in the comments of the script. You can use the same `conda` environment that you created for previous homeworks for this task.

2. We provide you with three pretrained weights that corresponds to $p = 3, 8, 16$. In the script, you have everything you need to produce the representation vectors of the training set and the test set: `X_train` and `X_test`, respectively. Additionally, their labels are given in `y_train` and `y_test`.

3. Your task for this homework is to utilize your own Nearest Neighbor algorithm to perform the exact same face recognition task as in Task 1. You should repeat the task three times with the three provided model weights and plot the accuracies as a function of p in the same plot along with PCA and LDA.

4. Finally, plot the p-dimensional embeddings along with labels for Train and Test using UMAP embeddings as a scatter plot similar to Task 1. Make sure to include plots for the three pretrained weights. Discuss the results and compare the embeddings of Autoencoder with PCA and LDA.

5 Task 3: Object Detection using Cascaded AdaBoost Classifiers

The goal here is to use the Viola and Jones approach [5] to the design of an object detector with an arbitrarily low false positive rate[1]. The Viola and Jones algorithm carries out a cascaded implementation of AdaBoost classifiers, with each stage of the cascade trying to meet a target false-positive and true-detection rate. You should refer to Section 5 of Prof. Kak's AdaBoost Tutorial [6] for more details on the AdaBoost algorithm and Section 11 for the Viola and Jones approach on designing such a cascade. You

[1] In object detection, you have a false positive when you declare a non-object blob of pixels as an object.

will apply your implementation of cascaded AdaBoost classifiers to the proprietary car detection dataset that is provided to you. Please do not distribute the dataset.

6 Submission Instructions

Include a typed report explaining how did you solve the given programming tasks.

1. Turn in two files: (a) a typed self-contained pdf report with source code and results and (b) zipped source code files (only *.py/*.txt files are accepted). Name your .zip file as hw10_<First Name><Last Name>.zip and follow the same naming convention for your pdf report too.
2. Your pdf must include a description of
   - Task 1
     - A brief description of how you have implemented both PCA and LDA.
     - A plot showing the classification accuracy as a function of the subspace dimensionality p for both PCA and LDA. The classification accuracy is defined as:
     $$\text{accuracy} = \frac{\#\text{ of test images correctly classified}}{\text{total }\#\text{ of test images}} \tag{1}$$
     - A discussion on the results comparing PCA and LDA.
     - UMAP plots for both PCA and LDA for a few values of p and a discussion on the results.
   - Task 2
     - In the same plot, plot the classification accuracies as a function of p.
     - A discussion on the results comparing Autoencoder against PCA and LDA.
     - UMAP plots for Autoencoder for the three pretrained weights and a discussion on the results.
   - Task 3
     - A brief outline of how you implemented the cascaded classifiers, for both training and inference.
     - A plot showing the false positive (FP) and false negative (FN) rate after the first k stages of the cascade as a function of k:
     $$\text{FP} = \frac{\#\text{ of misclassified negative test images}}{\#\text{ of negative test images}}, \tag{2}$$
     $$\text{FN} = \frac{\#\text{ of misclassified positive test images}}{\#\text{ of positive test images}}. \tag{3}$$
   - Your source code. Make sure that your source code files are adequately commented and cleaned up.
3. The sample solutions from previous years are for reference only. **Your code and final report must be your own work.**

Homework File Size

For your homework submissions, please ensure that your reports (PDFs) are **under 10MB**.

Some ways to reduce your report sizes are:

- Downsample your input/output image when including in the reports.
- When including plots in your reports, save them as PDFs instead of PNG/JPEG. PDFs are vector formats which are lightweight and do not pixelate when zooming into the plot.

Strike a balance between the file size and image quality displayed in your report. This will help your TA in easy distribution of homeworks for grading without maxing out the disk space. Moreover, this will help you when you upload your homeworks to your individual git repos. You don't want to upload 40 MB PDF to your git repo!

Finally, do not include the images in your ZIP files for your homework submissions. Ideally, the ZIP file should be under 1MB because it only contains ASCII (*.py /*.txt) files.

References

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* MIT Press, 2016. http://www.deeplearningbook.org. 1

[2] "FacePix Database." https://cubic.asu.edu/content/facepix-database. 2, 3

[3] "PCA (Principal Components Analysis) and LDA (Linear Discriminant Analysis) for Image Recognition." https://engineering.purdue.edu/kak/Tutorials/OptimalSubspaces.pdf. 2

[4] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction." https://arxiv.org/abs/1802.03426, 2020. 2

[5] "Viola–Jones object detection framework." https://en.wikipedia.org/wiki/Viola-Jones_object_detection_framework. 3

[6] "The AdaBoost Algorithm for Designing Boosted Classifiers." https://engineering.purdue.edu/kak/Tutorials/AdaBoost.pdf. 3