hugomcp@di.ubi.pt, 2025/26

Machine Learning

Practical Sheet 6: Experimental Setup

Consider the well-known "AR.zip" dataset, available at the course web page. It contains 3,315 images. It is a well-known benchmark for face recognition and expression analysis.

Created by Aleix M. Martinez and Robert Benavente at CVC, Universitat Autònoma de Barcelona, the purpose is to evaluate face recognition systems under controlled variations in facial expression, illumination, and occlusion.

- Subjects: 126 people (70 men and 56 women).
- Images per person: 26 images.
- Image type: Color frontal face images.
- Resolution: 768×576 pixels.
- Conditions captured:
    - Neutral, smiling, angry, and screaming expressions.
    - Different lighting conditions (left, right, and all lights on).
    - Occlusions with sunglasses and scarves (common for robustness tests).

Suppose that we are interested in discriminating between the "male" and "female" images in this set, which in this case is denoted by the first symbol of each filename ("m" for males and "w" for females).

1. Create a script that divides the available data into "K" folders, keeping the original proportions of images per class in each folder.

2. Using the K-fold cross-validation paradigm (for k=5), divide the existing data into "train" (3 folds), "validation" (1 fold) and "test" (1 fold).

3. Obtain the “PCA” representation of the original data, keeping 95% of the original amount of information in the dataset.

4. Plot the resulting “top-k” eigenvectors and show the corresponding eigenvalues (sorted).

5. Using Python + Keras, create a 3 layered neural network that discriminates between males and females. Use the “patience” as stopping criterium.

6. Plot the Precision/Recall (P/R) plot.

7. Plot the Receiver Operating Characteristic (ROC) curve and obtain the corresponding “AUC” value.

8. Repeat the 5-7 steps, according to the 5-fold cross validation paradigm. Obtain the mean and standard deviation accuracy and AUC values. Also, for the P/R and ROC curves, provide the shaded confidence intervals.