**Project B: Gender Identification from Face Images**

**CMPSC 441**

**Artificial Intelligence**

**Spring 2024**

------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

# Problem Description

You are going to identify the gender of a given person from his/her face images. You will be given a face dataset. You will have to extract features from the data, deploy the classifiers (using Scikit-learn/Matlab/Octave/R/your preferred programming language), and finally measure the performances (in terms of accuracy, precision, and recall rate) of the classifier. You need to experiment with the following classifiers:

1. K-Nearest Neighbors
2. Artificial Neural Network
3. Naïve Bayes
4. Logistic Regression

# Definition: Identification

In gender identification you have to identify the gender (male/female) of a person using training dataset. Initially template males and females should be created. When a claim comes from a person p, the test pattern is matched with the stored templates. If the test pattern is found similar to male template, the test person is identified as male. Otherwise, the system will say she is a female.

# Data Sets and Experiment Design

**Face Database**

We are using AR face database which is public and access is free. To enable detailed testing and model building the AR face images have been manually labelled with 22 facial features on each face. The 22 points chosen are consistent across all images (Please read next paragraph: Data Format). This labelled database contains face images of 136 persons (76 men and 60 women). Images feature frontal view faces with different facial expressions and illumination conditions.

* Male images are stored as: m-xx-yy.pts
* Females as: w-xx-yy.pts
* 'xx' is a unique person identifier (from "00" to "76" for men and from "00" to "60" for women). 'yy' specifies expression or lighting condition. Its meanings are described as follows:

1: Neutral expression

2: Smile

3: Anger

5: left light on

**Data Format**

Each face image contains 22 markup points as shown in Figure 1. For each point, we have the (X, Y) coordinates (*i.e.,* abscissa and mantissa) which shows the spatial location of the markup point in the image. Therefore, each data file contains 22 coordinates corresponding 22 points in Figure 1.

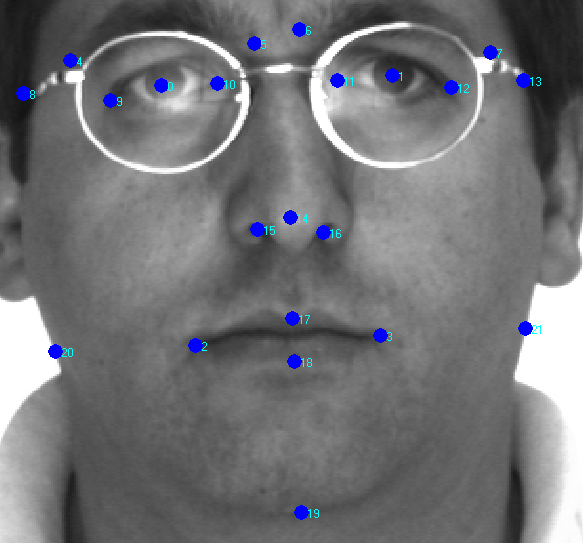


Figure 1: A face sample with 22 markup points

# Feature Definition and Extraction

**Defining Features**

Following are seven features-

1. Eye length ratio: length of eye (maximum of two) over distance between points 8 and 13.
2. Eye distance ratio: distance between center of two eyes over distance between points 8 and 13.
3. Nose ratio: Distance between points 15 and 16 over distance between 20 and 21.
4. Lip size ratio: Distance between points 2 and 3 over distance between 17 and 18.
5. Lip length ratio: Distance between points 2 and 3 over distance between 20 and 21.
6. Eye-brow length ratio: Distance between points 4 and 5 (or distance between points 6 and 7 whichever is larger) over distance between 8 and 13.
7. Aggressive ratio: Distance between points 10 and 19 over distance between 20 and 21.

**Feature Extraction**

You will have to extract the above mention features. Compute the feature values as per the definition of the features given above. Use the Euclidian distance when you compute the distance between two points.

# Experimental Results and Analysis

1. You will have to prepare a confusion matrix and calculate precision, recall rate, and accuracy for each of the classifiers.
2. You need to analyze (discuss) the results.

# Appendices

1. Put the values of seven features.
2. Put the source code.

# Deliverables

You should upload the following items:

1. Report having: introduction, problem description, data sets and experiment design, defining features and feature extraction, result and analysis, conclusion, references, and appendices.
2. Source code.

# Extra Points

You are welcome to experiment with more classifiers and more features. Use of more classifiers and features will give you extra points.