2. Questions

2.1. As explained in class, the provided third-party face recognition library is able to extract a 512-dimensional float feature vector from a given face image, as well as calculate the angular distance between two feature vectors, using ArcFace (https://bit.ly/3J8Tgtc). The expected behavior for the software is to generate small distances for two face images that depict the same individual (genuine pair), and large distances for two images that depict different individuals (impostor pair).

Leveraging the content of **only** the "dataset" folder within the provided data, the third-party face recognition library, and metrics learned in class, please determine **what is a good angular distance threshold to separate genuine from impostor pairs**. While providing your answer for the distance threshold, please explain in detail how you computed it. (1 point)

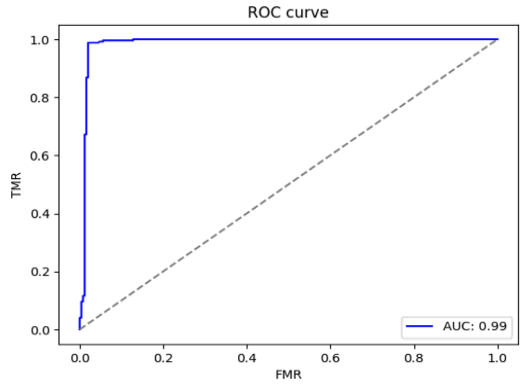
To determine the optimal angular distance threshold for separating genuine from impostor pairs, I used the "dataset" partition to calculate angular distances between image pairs with the third-party face recognition library. I used 250 genuine pairs and 250 impostor pairs from the given dataset. For each distance, I computed the False Match Rate (FMR) and False Non-Match Rate (FNMR) and identified the Equal Error Rate (EER) threshold – where FMR equals FNMR – as the optimal threshold. Using the computed data, the EER threshold was determined to be 0.9392, with both FMR and FNMR values at 0.02. This threshold provides a balanced trade-off between false matches and false non-matches, effectively distinguishing genuine from impostor pairs. This was calculated by using the function *compute_dist_fmr_fnmr_eer()*.

```
[63] print(compute_dist_fmr_fnmr_eer(load_data('/content/output3.csv')))

     (0.02, 0.02, 0.9392)
```

2.2. **What is the AUC** of the face recognition system you are using? In addition, please provide a graph with the system's **ROC curve**. (1 point)

The AUC of the face recognition system that I am using is 0.99.



2.3. By leveraging the face recognition system and the distance threshold previously computed, and by either capturing your face with your webcam or providing an image with your face, find within the “dataset” folder what is the individual that is the most similar to you. **Please provide the subject ID and the angular distance between your face and theirs**. In your opinion, do you have anything in common with this subject (e.g., gender, ethnicity, age, etc.)? If yes, what is it? (2 points)

The subject ID that the system found my image to be most similar to was subject 10. The angular distance was found to be 1.4331976. The subject that returned as the most similar to my picture does share some characteristics similar to mine. For one, we seem to share the same ethnicity, as I believe the subject in the image is Indian. In addition, we share somewhat similar facial features, especially around the eyebrow region.

2.4. By leveraging the face recognition system and the distance threshold previously computed, **please provide the subject ID** (or “**UNKNOWN**”, if the individual does not have a face within the “datasets” folder), as well as **the respective angular distances** that supported your decision, for each one of the 15 images provided within the “queries” folder. (6 points)

Query 4168: closest match: Subject ID = subject01, Angular Distance = 0.1335.
Query 4387: Closest match: Subject ID = subject02, Angular Distance = 0.1584.
Query 4507: Closest match: Subject ID = subject07, Angular Distance = 0.0968.
Query 4535: Closest match: Subject ID = subject03, Angular Distance = 0.2440.
Query 5314: Closest match: Subject ID = subject08, Angular Distance = 0.2157.
Query 6012: No match found within the threshold. Unknown.
Query 6510: No match found within the threshold. Unknown.
Query 6653: Closest match: Subject ID = subject05, Angular Distance = 0.1023.
Query 6706: No match found within the threshold. Unknown.
Query 7076: Closest match: Subject ID = subject06, Angular Distance = 0.1153.
Query 7549: Closest match: Subject ID = subject09, Angular Distance = 0.2223.
Query 7633: Closest match: Subject ID = subject04, Angular Distance = 0.0749.
Query 7745: No match found within the threshold. Unknown.
Query 9395: Closest match: Subject ID = subject10, Angular Distance = 0.0776.
Query 9708: No match found within the threshold. Unknown.