2.1. For each one of the three Biometric systems, what score threshold (a.k.a. operating point) should you use? Please explain your answer and describe how you have obtained each one of the respective system thresholds. (1 point)

The following is the threshold I would use for each system:

Name	Threshold
s1.csv	62.958
s2.csv	0.5849
s3.csv	2.3469

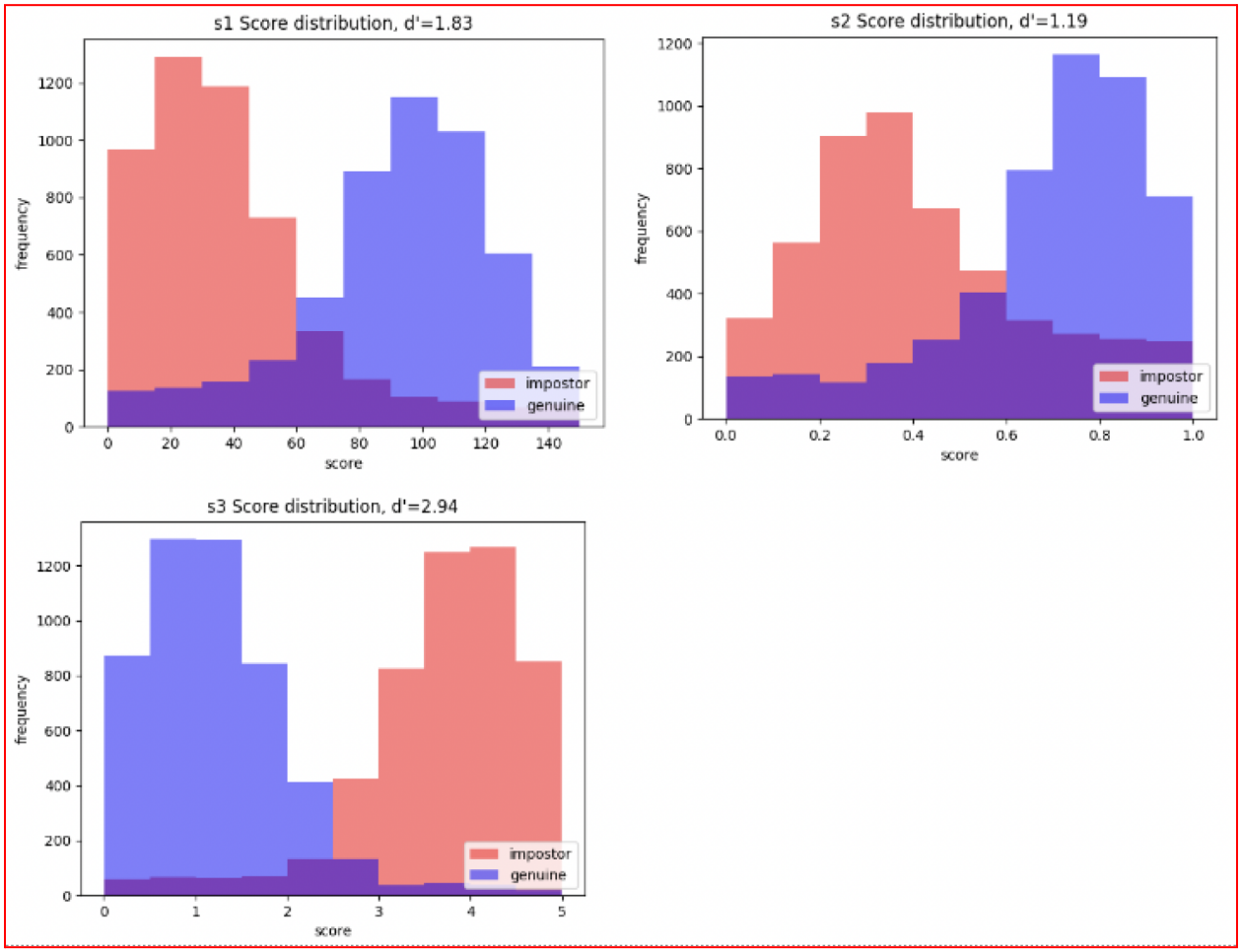
I arrived at these numbers for the threshold by computing the value at which the FMR and FNMR values were equal. In other words, I computed the threshold that would be needed to achieve EER (equal error rate).

For models s1 and s2, these values could be computed by utilizing the given function, `compute_sim_fmr_fnmr_eer provided` in the metrics python file. The value for s3 was computed by adapting the functions for computing FMR and FNMR at a given threshold for distances instead of similarities. The process for computing the threshold at EER is nearly identical to s1 and s2, with the only difference being the utilization of the adapted functions.

I decided to use the threshold at EER simply because not much is known about the nature of these models, so it would be safe to choose a value at which the errors are equivalent. However, as discussed in class, with more information it might be preferable for one type of error to be smaller than the other.

2.2. For each system, plot and provide a graph with the distributions of their respective scores. (1 point)

Graphs:

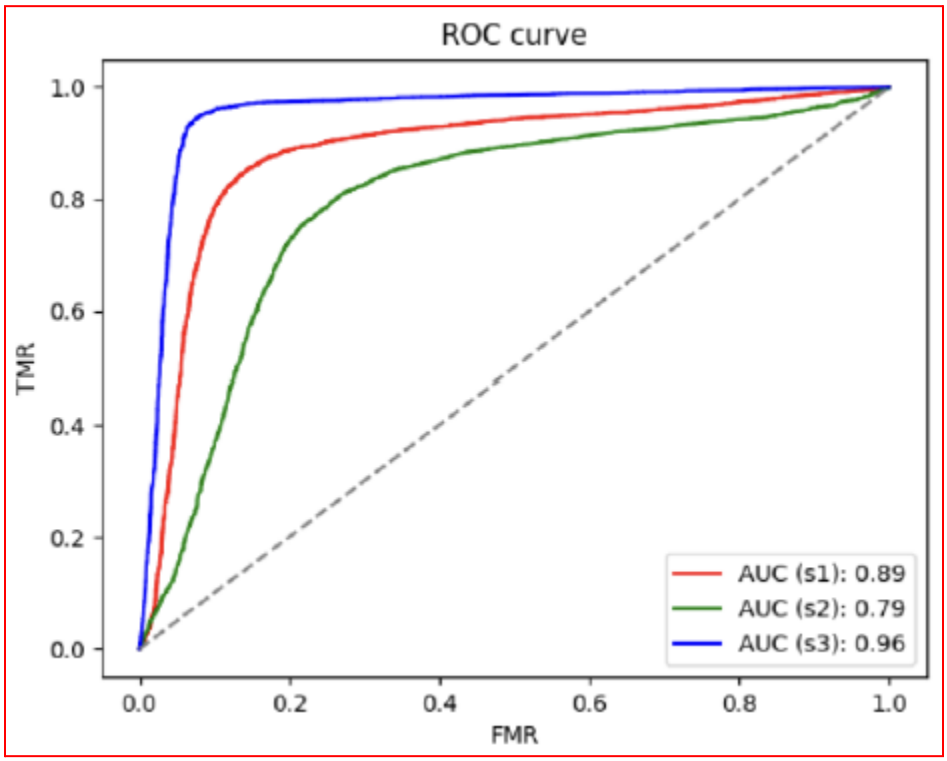


2.3. According to the d' (d-prime) values that one might compute for each system, which of the three should you use if you had to select only one for identification? Please justify your answer. (1 point)

According to the d' values of the three systems, System 3 is the most ideal because it has the highest d' value, indicating that there is a greater separation between the impostor and genuine distributions, meaning that this system is better at distinguishing between the two groups.

2.4. Plot and provide a single graph with the ROC curves and AUCs of all three systems together. A reference to help: https://tinyurl.com/23fkm3jf. (1 point)

Graph:



2.5. According to the ROC curves and AUC values, which one of the three systems should you use if you had to select only one for identification? Please justify your answer. (1 point)

I would choose s3 because it has the highest AUC value (as seen in the graph above). A higher AUC value indicates a lower FMR and higher TMR (or lower FNMR), and this can be seen with the actual curves as well, since the blue curve, representing s3, is closest to the edges.

2.6. Re-do question 2.4 but this time leveraging the **optimized** *scitkit-learn* (https://en.wikipedia.org/wiki/Scikit-learn) implementation of AUC calculation (see https://tinyurl.com/nsyu2k9b). Which implementation is faster; the naive one presented in class or the one using scikit-learn? Please justify your answer with collected runtimes. (5 points)

The implementation with the optimized sci-kit learn function is significantly faster, as shown below:

```
import time

s1 = load_data('s1.csv')

start_time = time.time()
compute_fmr_tmr_auc(s1)
end_time = time.time()

delta_time = end_time - start_time
print('Execution time:', delta_time, 'seconds.')
```

```
Execution time: 37.076 seconds.
```

```
import numpy as np
from sklearn import metrics

s1 = load_data('s1.csv')
y = [o[0] for o in s1]
pred = [o[1] for o in s1]

start_time = time.time()
fpr, tpr, _ = metrics.roc_curve(y, pred, pos_label=1)
metrics.auc(fpr, tpr)
end_time = time.time()

delta_time = end_time - start_time
print('Execution time:', delta_time, 'seconds.')
```

```
Execution time: 0.015 seconds.
```