Question 5

Dhruv Meena

October 24, 2024

Mechanism

I printed out the distances between the unseen test data and the original training data. The distances were quite large, to the order of 6-7, which is expected because the unseen images should not be anywhere close to the training data because they are pictures of completely different people.

I implemented a threshold mechanism where I manually change the threshold to see what changes, and how many false positives I can get. For a threshold of 0.99e7, I got 31 false positives, for example. Below I plot threshold vs false_positives.

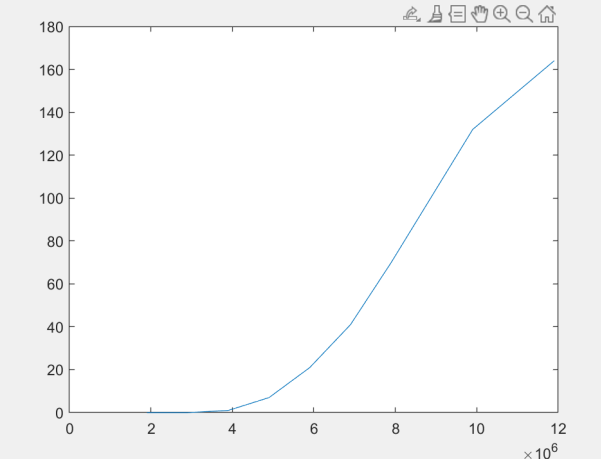

Figure 1: False Positives vs Threshold

We can see that the number of false positives keeps rising as we increase the threshold, which is expected because for higher thresholds it now classifies unseen images as matching with training data. Similarly for false negatives, increasing the threshold, increases the false negatives but it becomes constant after a while. For example, I got 128 false negatives for a threshold of 100, while for a very small value of 1e-7, I got only 8 false negatives. This makes sense as in seen image data, the test image has to match, only a very low threshold could give a false negative.

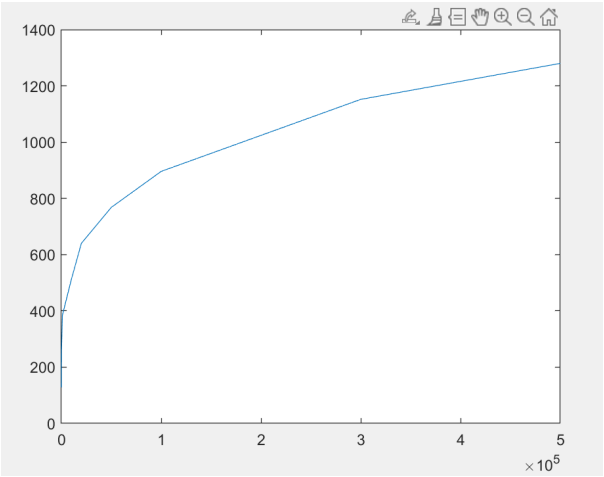

Figure 2: False Negatives vs Threshold