LAB10 REPORT

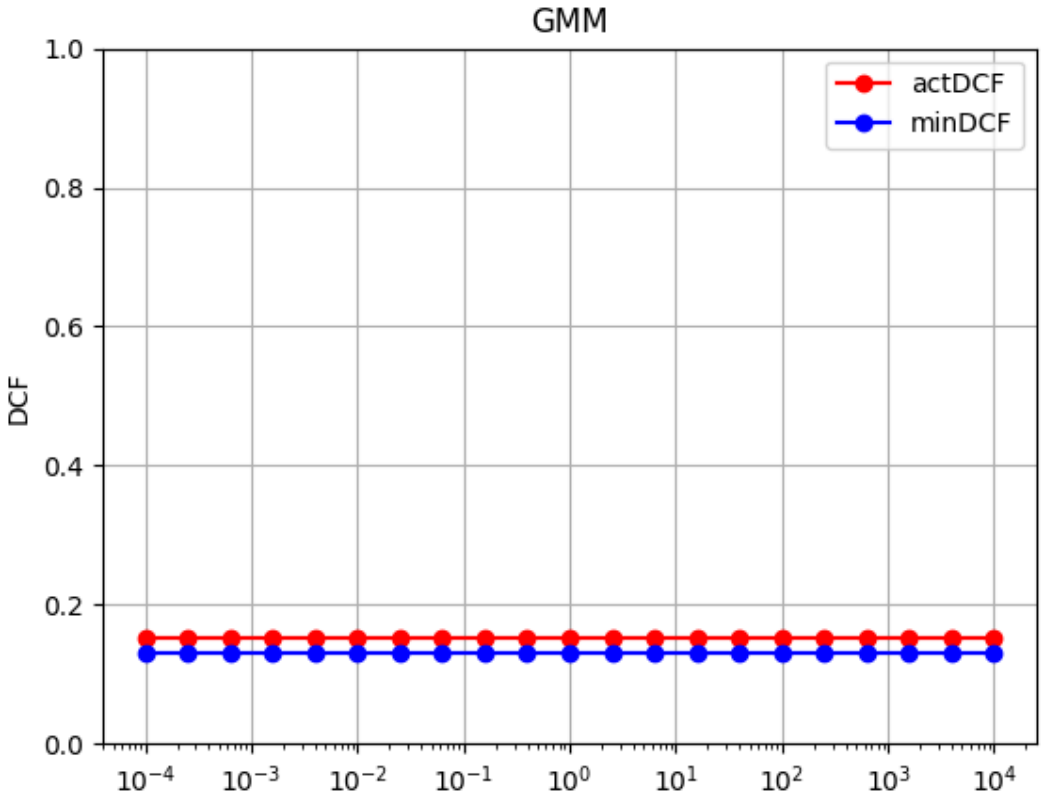
We now try to use GMM for classification.A a GMM can be interpreted as the sum of n gaussian, each with its relative weight. This allows us to exit the constraint GMM imposed on our dataset distribution, which is it to be modeled as a gaussian. with a GMM our class posterior probability can represent way better our data distribution.
I tried to model the GMM using 1,2,4,7,16,32 components for all combinations for the true and false class.
The results are very good, with low minDCF and actDCF, the combinations and number of components impacts those metrics as well. The best result is obtained for:

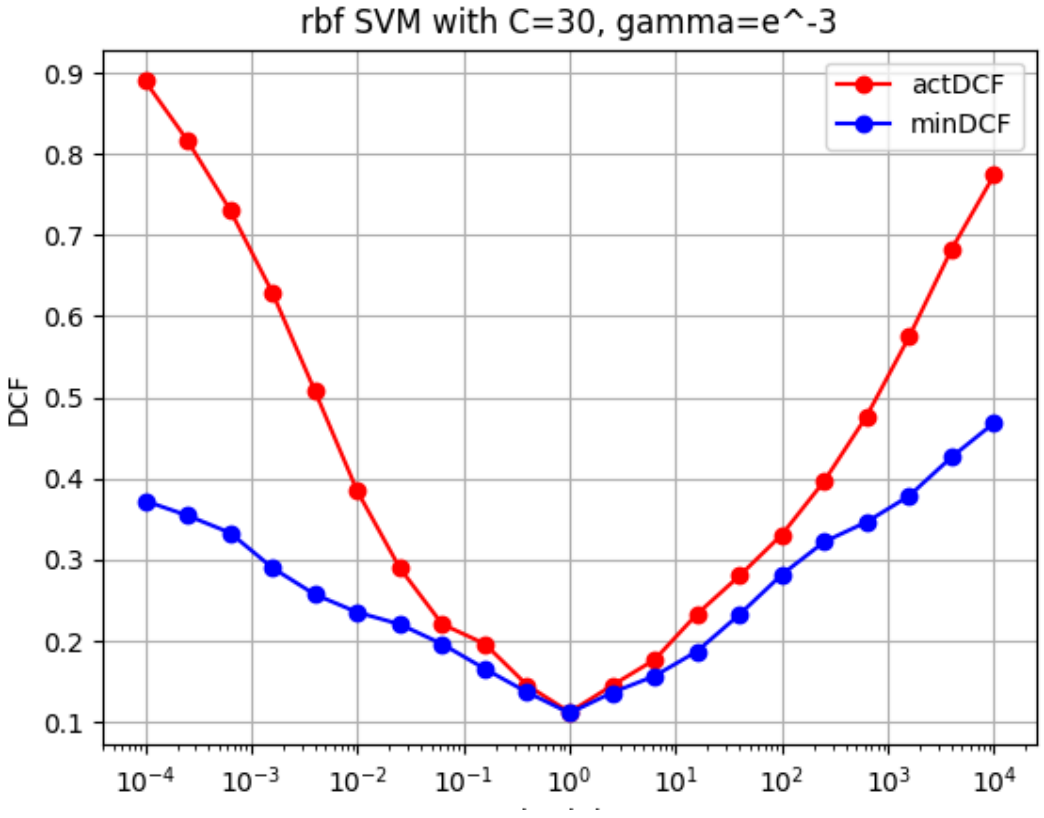
```
----------------------------------------
GMM diagonal covariance matrix components true class:32 components fake class:8
actDCF=0.15165770609318996
minDCF=0.13124039938556067
----------------------------------------
```

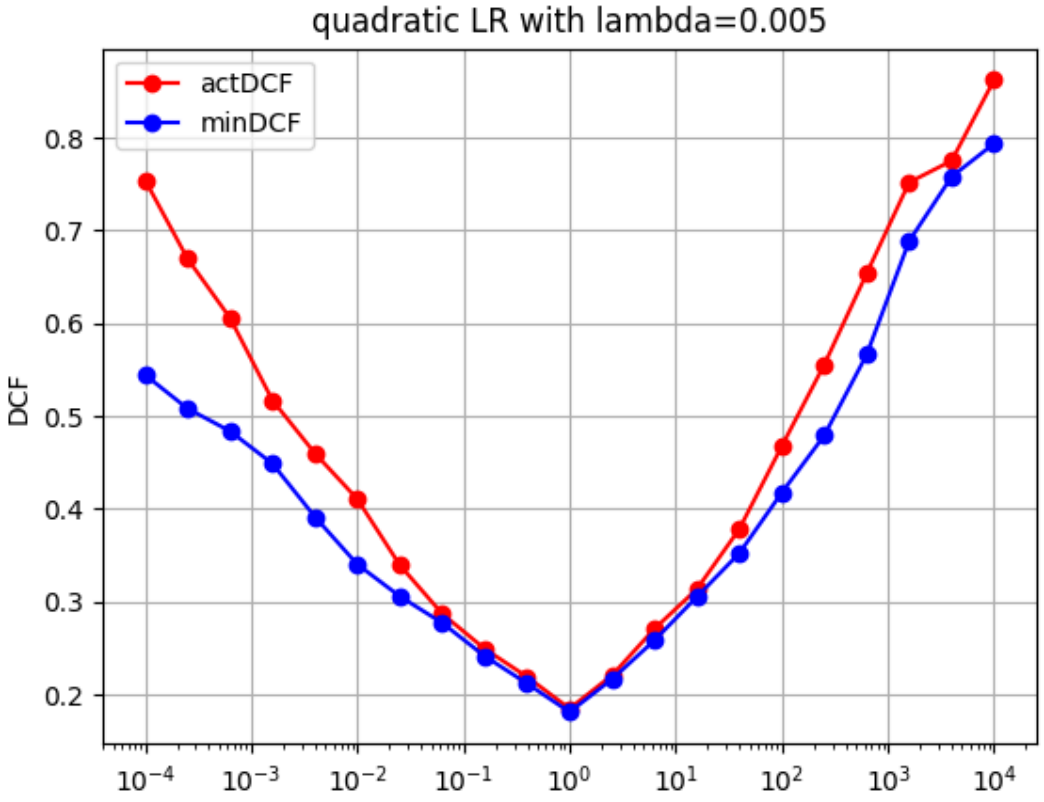
The calibration error is not big, the classifier works nicely on our dataset! The results are in line with my expectations since not all features of our sample were represented accurately by a gaussian. This model allows us to better approximate the dataset distribution.

So far, we have that the best classifier is GMM with those combinations of features. For LR the best is the quadratic one (although it would benefit a lot by calibration), while for SVM the best is the one with the rbf kernel (also here, calibration would help this classifier because it would regain a probabilistic interpretation for the scores).

Now we analyze the best model (quadratic LR, SVM with rbf kernel,GMM with 8 components for the fake class and 32 components for the true class)


GMM
actDCF
minDCF
DCF
1.0
0.8
0.6
0.4
0.2
0.0
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}


rbf SVM with C=30, gamma=e^-3
actDCF
minDCF
DCF
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}



I'm actually very surprised by the GMM model, it has very good actDCF and minDCF for all the applications priors, there still is a little miss calibration but overall it has very good dcf for all various applications.

The SVM with rbf kernel has very good DCF and calibration for small priors , but there is significant miss calibration for other ranges and can actually be harmful to our application.

The quadratic LR on the other hand has better actDCF, good DCF for small priors and overall it has less calibration error. It still isn't suitable for some applications