Group 7

Braun Fabian, Marija Nikolic , Tiago de Freitas Pereira

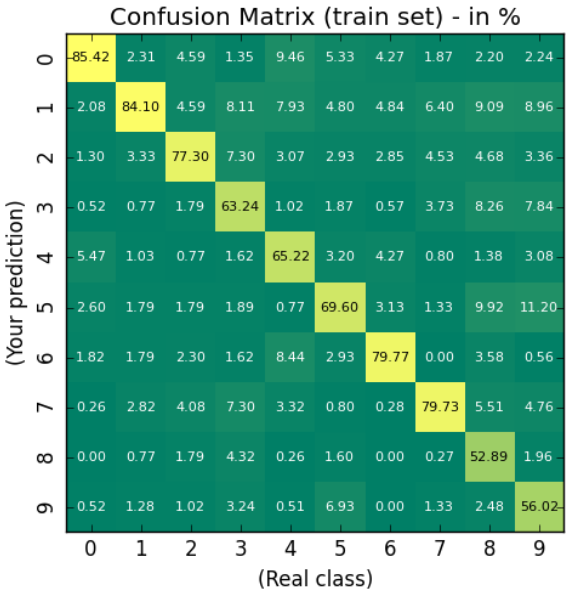
KNN+PCA

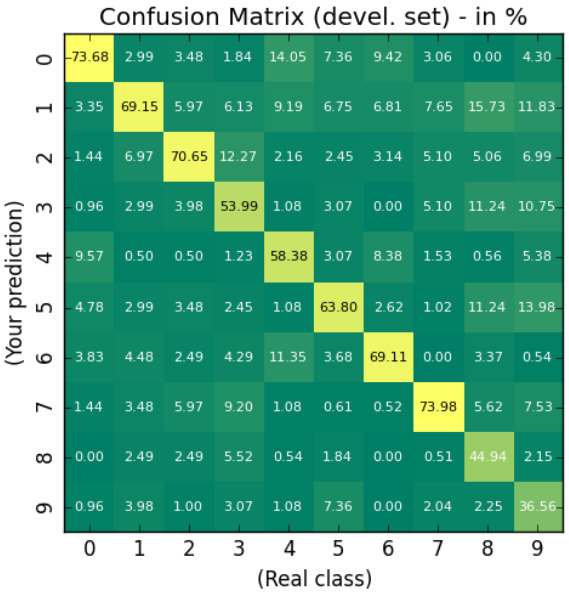
- Simple strategy (capacity=0)
- PCA for dimensionality reduction
- Selected configuration: K=9 and PCA=10 (25.91% of the energy)
- Good tradeoff between dimensionality and CER with 10 PC

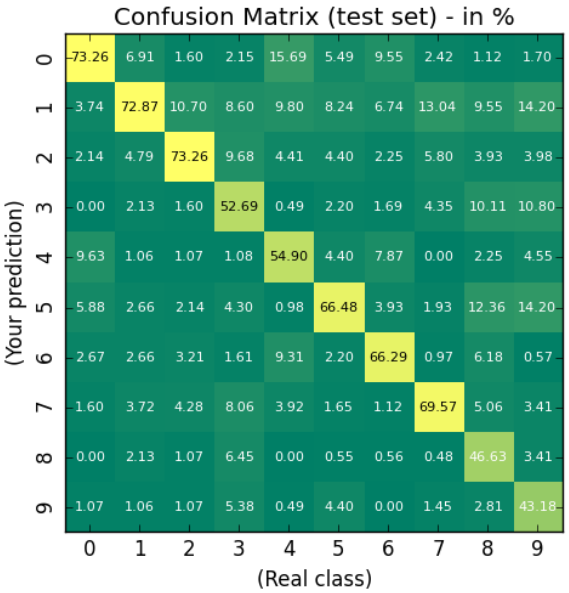
```
python mini_project.py -PCA -c 10 -kNN -nn 9
CER Train: 28.50%
CER Dev: 38.07%
CER Eval: 37.91%
```

KNN+PCA

- Confusion matrix







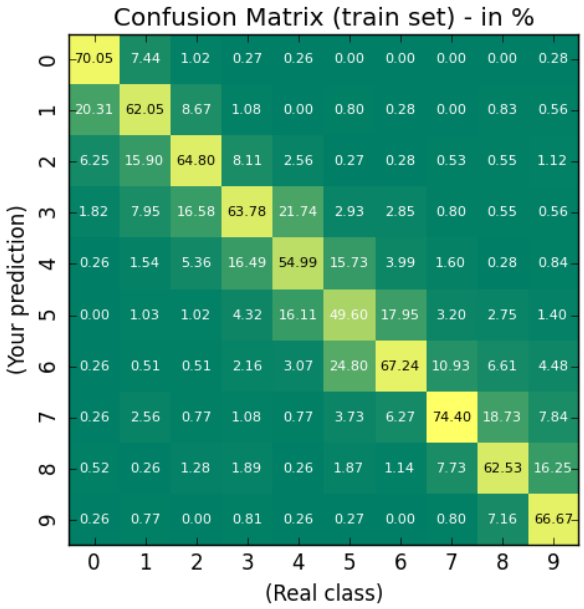
GMM+PCA

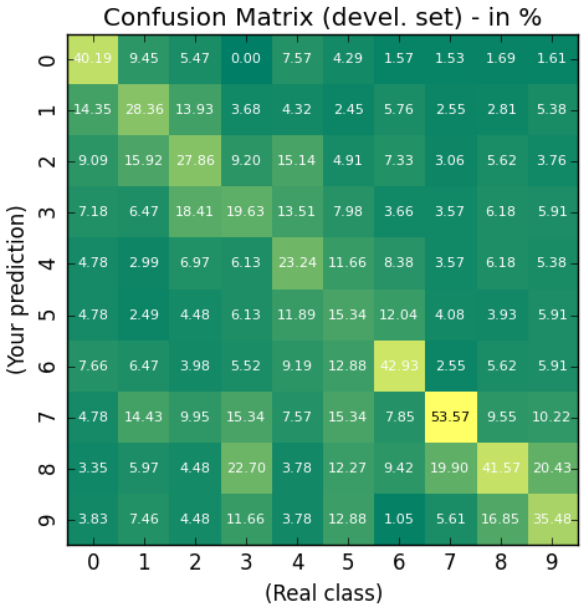
- Generative approach to model the digits
- PCA for dimensionality reduction
- Number of gaussian components 150 (capacity=450)
- Clear overfitting

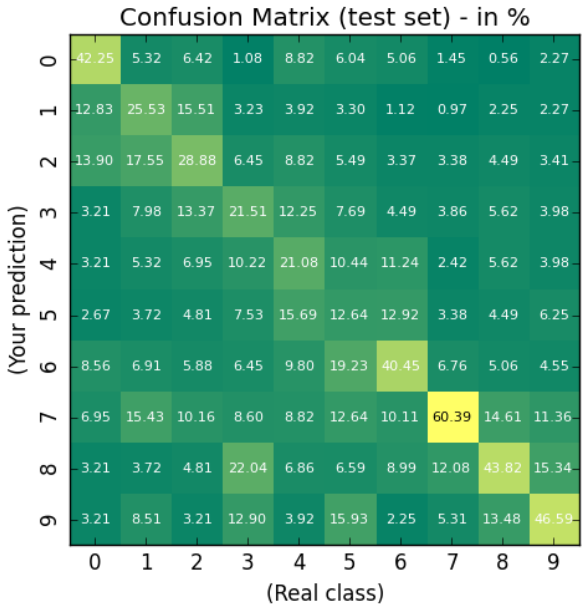
```
 python mini_project.py -PCA -c 10 –GMM –nb_gaus
150
    CER Train: 36.45%
    CER Dev: 66.68%
    CER Eval: 65.62%
```

GMM+PCA

- Confusion matrix







GMM+PCA

Modelling with less gaussian components (capacity=30):

```
  python mini_project.py -PCA -c 10 –GMM –nb_gaus 16
    CER Train: 62.54%
    CER Dev: 67.70%
    CER Eval: 65.46%
```

- We have a question regarding the strategy adopted

Conclusion

- Simple solution was the best (KNN+PCA)
- Possible to reproduce the results
- Suggestion: Model one GMM per digit