ML in Fundamental Physics

Ludwig-Maximilians-Universität München – Dr. Sven Krippendorf

---

Exercise Sheet 7

Submission by Saturday 27th 17:00 via Moodle.
Corrections due by Wednesday 1st 16:00 via Moodle.
Format for all questions: individual Jupyter notebooks .ipynb.

1 Clustering (3 points)

The aim of this exercise is to familiarise yourself with the analysis of clustering hyperparameters. For this purpose, let us consider the AgglomerativeClustering algorithm in sklearn. Using `make_circles`, `make_moons`, `make_blobs` visualise the different performance of different linkages (ward, complete, average, single) (1 point). Perform a little bit of hyperparameter tuning and show for the different linkage models the best model you have found. (1 point)

In the second part of this exercise, compare these results with the performance of t-SNE and k-means.[1] (1point)

2 AutoEncoder (4 + 2 points)

The aim of this exercise is to implement simple autoencoders.

- (1 point) Generate 2-d images (1-channel) showing polynomials up to a maximum degree (e.g. 40x40) in two variables.
- (2 points) Build two autoencoder architectures (e.g. similar to the ones presented in the lecture) which involve a single hidden dense layer and several hidden dense layers.
- (1 point) For quadratic polynomials and two latent dimensions, visualise the results of your latent dimensions. Is an interpretation of your latent parameters easily visible?
- (2 points) *Add a custom loss function which de-correlates the latent parameters. Is it possible to find an interpretation for the latent parameters in this case?

3 KL – divergence (1 point)

Show by example that $D_{KL}(Q||P) \neq D_{KL}(P||Q)$ in general.

[1] You are not required to perform extensive parameter scanning but do not just use default parameters.