

Hands-on Machine Learning Training

Session 2 – Basic Principles

Theoretical Preparation

In this session, focus is on understanding principles for performing supervised classification. You will find out how training data can be used to fit models which can be finally utilized to classify novel data samples. Here, we start work with designing our own classifier by using simple linear functions to fit the given dataset. We also focus on some important classification algorithms like Support Vector Machines (SVMs) and k-nearest neighbours (knn). We also learn to evaluate the various models and look at some basic accuracy measures. In addition, we also have a short look at a couple of Unsupervised Classification methods including K-means Clustering and Gaussian Mixture Models (GMMs). For preparation we provide you with various links which are suitable for learning the concepts required for the practical session. You are also strongly encouraged to read further on your own in addition to the links we suggest.

- Read and understand the following:
 - Refer to the enclosed document for the links on the various topics
- Based on this literature you should have an understanding of
 - SVM ¹
 - knn ²
 - K-means clustering ³
 - GMM ⁴
 - Linear Regression ⁵
 - PCA ⁶

¹<https://machinelearningmastery.com/support-vector-machines-for-machine-learning/>

²<https://machinelearningmastery.com/k-nearest-neighbors-for-machine-learning/>

³<https://brilliant.org/wiki/k-means-clustering/>

⁴<https://brilliant.org/wiki/gaussian-mixture-model/>

⁵<https://machinelearningmastery.com/linear-regression-for-machine-learning/>

⁶<https://medium.com/@aptrishu/understanding-principle-component-analysis-e32be0253ef0>

Further Reading

The following (advanced) papers on the concept of classification might be interesting for further understanding and more profound knowledge:

- Jason Brownlee et al., "Naive Bayes Classifier From Scratch in Python" ⁷
- Philipp Hagenlocher et al., "Decision Tree Learning " ⁸
- Gilles Louppe., "Understanding Random Forests: From Theory to Practice" ⁹

⁷<https://machinelearningmastery.com/naive-bayes-classifier-scratch-python/>

⁸https://www5.in.tum.de/lehre/seminare/datamining/ss17/paper_pres/08_decision_tree/paper.pdf

⁹<https://arxiv.org/pdf/1407.7502.pdf>