Machine Learning Laboratory

Session 1 – Basics all at once

Theoretical Preparation

In the first session, focus is on understanding the basics of programming in Python. You will understand Python data structures, some important Python packages, and also learn image manipulations. Some common Machine Learning terminology will also be introduced, and using the publicly available Iris dataset, we will learn Principal Component Analysis (PCA), including a quick recap on Singular Value Decomposition (SVD).

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 Machine (SVM) and k-nearest neighbours (KNN). In addition, we also have a short look at a couple of Unsupervised Classification methods including K-Means Clustering and Gaussian Mixture Model (GMM).

Prepare yourself on the mentioned topics using the following resources. These are important for you to have a successful session (hints for the test):

- A crash course on Python and NumPy: Link
- SVD: Link (You might be asked to calculate SVD of a very simple matrix.)
- PCA: Link (Understand the concept and geometric interpretation.)
- LDA: Link (Understand the concept.)
- SVM: Link (Understand the concept and geometric interpretation.)
- KNN Link (Understand the concept.)
- Clustering algorithms (e.g., K-Means, GMM, Mean-Shift): Link (Understand the concept.)

Further Reading

Wikipedia pages of these algorithms are generally very good. Besides, YouTubers "StatQuest with Josh Starmer" and "3Blue1Brown" have posted many nice tutorials. For details, you can easily find books of SVM or PCA.