Machine Learning Lab Course

Machine learning is playing an increasingly important role in our society and everyday life. Autonomous driving, AI-generated images, and chat bots that seem almost human are no longer science fiction. Of course, this also affects scientific research, where a lot of data often has to be analyzed as effectively as possible.
An extreme example of this is the ATLAS experiment at CERN. There, 10.000 TB of data from high-energy proton-proton collisions are recorded annually.
Such gigantic amounts of data are a great challenge for further processing, but also absolutely necessary to discover new particles or processes that have never been measured before. In 2012, scientists at CERN discovered a particle that had previously been based only on theoretical considerations - the Higgs boson. This particle plays a key role in the Standard Model of particle physics, as the associated Higgs field gives particles their mass.
This lab course combines these two interesting fields to search for the Higgs boson using neural networks. Using official simulated data from the ATLAS experiment, neural networks will be trained to distinguish events with Higgs bosons from the rest of the measured background events. The lab course will focus primarily on machine learning and its application to simulated data.
The first part of the lab deals with a minimal example to find the right choice for the training setup. The choice of the right setup is the key to successful Machine Learning.
Once the right setup is found, the next step is to design your own neural network. An already trained network is loaded as a "competitor" whose performance must be surpassed. Finally, the network is evaluated to see what contribution it could make to the Higgs search.
Since this lab course deals solely with the creation, training, and evaluation of neural networks, no additional protocol has to be written at the end. The report will be handed in as a detailed documented Jupyter notebook together with your saved final neural networks.

Machine Learning Introduction

To understand what machine learning is and how a neural network works check out the first two videos in the following playlist by 3Blue1Brown:
https://youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi

Introduction to H→ZZ

One of the possible decay channels for the Higgs boson is H→ZZ→4ℓ. This channel is also called the “golden channel” due to its clear and clean signature. A short visualization can be found at https://youtu.be/wZpfHzlhuzg.
Use the PDG website to check the fundamental properties of the Higgs boson, Z boson, electron, and muon (like charge, mass,...):
https://pdglive.lbl.gov/Viewer.action

Setup the environment

We use Anaconda to get the correct environment and all necessary packages. Follow these steps to get the correct environment and test it:

1. Install Anaconda:
   https://docs.conda.io/projects/conda/en/latest/user-guide/install/index.html
2. Import the environment `environment.yml`
   Within the shell use:
   `conda env create -n ml_labcourse --file environment.yml`
3. Activate the conda environment
   `conda activate ml_labcourse`
4. Open a jupyter notebook
   Within the shell use:
   `jupyter notebook`
5. Open `test_setup.ipynb` and execute the cells

If no error occurred you are done :)