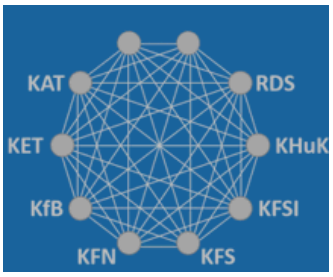

Contribution ID: **37** Type: **not specified**

Dealing with Systematic Uncertainties in Machine Learning

Wednesday 1 January 2025 10:30 (5 minutes)

Data analysis in particle physics today is characterized by large datasets of independent "events". ML methods are ubiquitous in all analysis steps from physics object selection to classification of events into signal and background categories. The parameters of the underlying physics models are extracted using maximum-likelihood methods, using detailed simulated (synthetic) datasets.

With even larger datasets expected for the next decade, the accuracy of many physics results will be limited by systematic uncertainties. Known systematic uncertainties are initially estimated by auxiliary measurements or variations of simulation parameters. They are then translated as nuisance parameters of the physics model and further constrained in the fit to data. However, the modeling of data in simulation may be imperfect, possibly leading to unwanted biases in the physics result. On the other hand, differences between data and simulation may also be due to new-physics effects.
Our group has explored several approaches to reduce and/or quantify the impact of systematic uncertainties on ML methods in the context of Higgs-boson studies at the CERN Large Hadron Collider. The studies, documented in several undergraduate theses, include adversarial networks, Bayesian neural networks, and Domain Adaptation,

List of Committees:

Please describe your expertise/areas in which you would like to contribute / advise.

Application of deep learning to science problems.

Do you consent to the data usage and public abstract data posting in the ErUM-Data Community Information Exchange?

Yes

In ErUM-Data, what kind of data are you dealing with?

Large datasets of statistically independent events, typically O(few 100) input features for each event. Simulated datasets used for supervised learning.

What is your expertise in computing and / or software development?

Parallelized data processing pipelines, standard ML tools, likelihood methods, multivariate data analysis.

What is your field and role?

University professor in experimental particle physics, leading teams in ML-based data analysis and silicon detector development.

Your ErUM - Committee is

KET - Komitee für Elementarteilchenphysik

Please describe areas in which you can contribute to “data handling”teaching.

Simulated data in particle physics.

My current most burning research question, I like to find partners for, is:

How can we make ML techniques aware of, and resilient against, systematic uncertainties of the data?

Please describe areas in which you would like to improve your knowledge / skills.

State-of-the-art ML algorithms.

Primary author: HUSEMANN, Ulrich (Institute of Experimental Particle Physics - Karlsruhe Institute of Technology - KIT)

Presenter: HUSEMANN, Ulrich (Institute of Experimental Particle Physics - Karlsruhe Institute of Technology - KIT)