T 4: Deep Learning I

Zeit: Montag 16:00–18:30 Raum: H06

T 4.1 Mo 16:00 H06

Physics inspired feature engineering with Lorentz Boost Networks — •Yannik Rath, Martin Erdmann, Erik Geiser, and Marcel Rieger — III. Physikalisches Institut A, RWTH Aachen University

A large part of the success of deep learning in computer science can be attributed to the introduction of dedicated architectures exploiting the underlying structure of a given task. As deep learning methods are adopted for high energy physics, increasing attention is thus directed towards the development of new models incorporating physical knowledge.

In this talk, we present a network architecture that utilizes our knowledge of particle combinations and directly integrates Lorentz boosting to learn relevant physical features from basic four vectors. We explore two example applications, namely the discrimination of hadronic top-quark decays from light quark and gluon jets, and the separation of top-quark pair associated Higgs boson events from a $t\bar{t}$ background. We also investigate the learned combinations and boosts to gain insights into what the network is learning.

T 4.2 Mo 16:15 H06

Further development of the ATLAS Deep Learning flavour tagging algorithm — •Manuel Guth — Albert-Ludwigs Universität, Freiburg, DE

The development of machine learning techniques is making a lot of progress in the last few years. Already now, machine learning is deeply embedded in our daily life. Especially deep neural networks require a large amount of statistics for a robust training procedure in order to find yet unknown dependencies in data. The large amount of simulated data available in particle physics allows to use these new sophisticated techniques to improve the physics analyses. The identification of heavy flavour jets (tagging) plays an important role in almost all physics analyses at the ATLAS experiment. It is an essential tool for precision measurements as well as for searches for new physics phenomena. One of the frameworks within ATLAS for b-tagging is the Deep Learning tagger (DL1). It uses deep neural networks based on TensorFlow and Keras to distinguish b-, c- and light flavour jets using the information of several baseline b-taggers. A first introduction of the DL1 tagger is given, followed by detailed studies to improve the deep learning network architecture.

T 4.3 Mo 16:30 H06

Application of Deep Learning to Heavy Flavour Jet Identification with the CMS Experiment — Xavier Coubez[1,2], Luca Mastrolorenzo[1], •Spandan Mondal[1], Andrzej Novak[1], Andrey Pozdnyakov[1], and Alexander Schmidt[1] — [1]RWTH Aachen University, Germany — [2]Brown University, USA

Many physics analyses within the CMS experiment rely on the efficient identification of heavy flavour jets. Over the past few years, several algorithms have been developed to exploit the distinctive features of jets arising from heavy flavour quarks to distinguish them from those arising from light quarks. The CMS collaboration has recently shown that Deep Neural Networks (DNNs) can be used to achieve significantly higher efficiencies while tagging heavy flavour jets, compared to traditional Machine Learning approaches. In addition to standard b-tagging and c-tagging algorithms, Deep Learning has been implemented to develop flavour tagging algorithms specialized for boosted topologies, to aid physics analyses that focus on boosted regimes and heavy exotic particles. This talk focuses on new advances in the application of Deep Learning in heavy flavour jet identification at CMS as well as the performance measurements of tagging algorithms on CMS data.

T 4.4 Mo 16:45 H06

Validation of a Deep Neural Network Based Flavor Tagging Algorithm at Belle and Belle II — •Jochen Gemmler, Florian Bernlochner, Michael Feindt, and Pablo Goldenzweig for the Belle 2-Collaboration — ETP, KIT, Karlsruhe

Measurements of time dependent CP violation will be one of the key tasks of the Belle II experiment, which is located at the SuperKEKB collider in Tsukuba, Japan. Via electron-positron collisions, neutral B mesons are produced in an entangled state. For CP violation measurements, it is crucial to infer the flavor of the accompanying B_{tag} meson from its final states, exploiting the flavor specific decay topologies. This process is referred to as flavor tagging and the presented approach utilizes a Deep Neural Network (DNN) for this classification.

This talk shows current results of the validation of the DNN based approach on the full Belle dataset, which contains 772 Million B-Meson pairs, using the software framework of Belle II.

T 4.5 Mo 17:00 H06

Adversarial Neural Network-based data-simulation corrections for jet-tagging at CMS — Martin Erdmann, •Benjamin Fischer, Dennis Noll, Yannik Rath, Marcel Rieger, and David Schmidt — III. Physikalisches Institut A, RWTH Aachen University

Variable-dependent scale factors are commonly used in HEP to improve shape agreement of data and simulation. The choice of the underlying model is of great importance, but often requires a lot of manual tuning e.g. of bin sizes or fitted functions. This can be alleviated through the use of neural networks and their inherent powerful data modeling capabilities.

We present a novel and generalized method for producing scale factors using an adversarial neural network. This method is investigated in the context of the bottom-quark jet-tagging algorithms within the CMS experiment. The primary network uses the jet variables as inputs to derive the scale factor for a single jet. It is trained through the use of a second network, the adversary, which aims to differentiate between the data and rescaled simulation.

T 4.6 Mo 17:15 H06

Hyperparameter optimization of Adversarial Neural Networks in the tW dilepton channel using the ATLAS detector — •Christian Kirfel, Ian Brock, and Rui Zhang — Physikalisches Institut, Bonn, Deutschland

Neural networks are widely used for signal to background separation in high energy collider physics. Neural networks trained on Monte Carlo simulations can be highly sensitive to systematic uncertainties. A proposed technique to diminish this sensitivity is an adversarial neural network consisting of two networks that are trained against each other. In our case, the first network tries to separate between signal and background, while the second network tries to separate between a nominal signal sample and a signal sample with different settings. We are using a Minimax decision rule to achieve a good signal to background separation for the first network and a poor nominal to systematics separation for the second network. In this talk an adversarial neural network trained on tW dilepton channel Monte Carlo simulations with $t\bar{t}$ background using the ATLAS detector is introduced. Testing and tuning of the hyperparameters is presented for both networks as well as a comparison to a single neural network approach. Lastly the dependence for both approaches on systematic uncertainties is investigated.

T 4.7 Mo 17:30 H06

Adversarial Neural Networks zur Reduzierung des Einflusses von systematischen Unsicherheiten am Beispiel einer $\mathrm{t\bar{t}H}$ **-Analyse** — •Jörg Schindler, Karim El Morabit, Ulrich Husemann, Philip Keicher, Matthias Schröder, Jan van der Linden und Michael Wassmer — Institut für Experimentelle Teilchenphysik (ETP), Karlsruher Institut für Technologie (KIT)

Die Messung des Wirkungsquerschnitts für Higgs-Boson-Produktion in Assoziation mit einem Top-Quark-Antiquark-Paar ($\mathrm{t\bar{t}H}$) ermöglicht eine direkte Messung der Top-Higgs-Yukawa-Kopplung. Aufgrund des kleinen Wirkungsquerschnitts wird der Zerfall mit dem größten Verzweigungsverhältnis untersucht, der Zerfall in ein Bottom-Quark-Antiquark-Paar ($\mathrm{b\bar{b}}$). Dabei werden multivariate Analysemethoden verwendet, um Signal von Untergrund zu trennen.

Ein entscheidender Untergrund hierbei ist die $\mathrm{t\bar{t}}$-Produktion mit einem assozierten $\mathrm{b\bar{b}}$-Paar. Die verfügbaren Vorhersagen für diesen Prozess sind mit großen Unsicherheiten behaftet und weisen Unterschiede auf. Durch die Verwendung von Adversarial Neural Networks können die neuronalen Netze robust gegenüber diesen Unterschieden konstruiert werden.

In diesem Vortrag wird die Anwendung von Adversarial Neural Networks am Beispiel einer $\mathrm{t\bar{t}H}$ -Analyse im semileptonischen Kanal untersucht.

T 4.8 Mo 17:45 H06

Precise simulation of electromagnetic calorimeter showers using a Wasserstein Generative Adversarial Network — Martin Erdmann[1], Jonas Glombitza[1], and •Thorben Quast[1,2] — [1]Physikalisches Institut 3A, RWTH Aachen — [2]EP-LCD, CERN

The increased instantaneous luminosity at the High Luminosity LHC will raise the computing requirements for event reconstruction and analysis for current LHC-based experiments, hence limiting the available resources for the simulation of particles traversing matter. Developments on the performance of state-of-the-art simulation frameworks such as Geant4 are proceeding but are unlikely to fully compensate for this trend. Generative adversarial neural networks (GANs) have been shown to provide promising fast simulation models. Wasserstein GANs (WGANs) are a variant of this method. They employ a more robust metric for the adversarial training of the generator network. In this talk, we show our adaptation of the WGAN concept for the generation of electromagnetic showers inside a realistic setup of a multi-layer sampling calorimeter. In addition, conditioning on the energy of the incident particle and on its impact position is integrated through two auxiliary regression networks. Overall, the quality of these fast shower simulations with the WGAN reaches the level of showers generated with the GEANT4 program in most aspects. At the same time, the computational speed-up compared to traditional sequential simulations amounts to several orders of magnitudes.

T 4.9 Mo 18:00 H06

Parton showers with Generative Adversarial Networks — •Christof Sauer — Physikalisches Institut, Heidelberg, Deutschland

The prediction of physical processes are usually based on simulations – one example being parton showers. At present, the simulation of hadronic final states is done by dedicated software, such as Pythia, Herwig. This presentation intends to demonstrate a potential application of Generative Adversarial Networks (GANs) within the context of parton shower generation. Such machine learning techniques can be used to produce parton showers which are independent of any current shower model. It would allow to circumvent inherent problems in the simulation of parton showers. This method could be a applied in analyses that are too sensitive to parton shower effects in the modeling of the background and hence rely on an accurate background estimate.

As a first step, a network is trained on multijet events generated with Pythia, whereby the focus lies on training the network to produce realistic and consistent parton showers. The Monte Carlo samples serve as a surrogate to examine the applicability of this method under well controlled conditions before, subsequently, proceeding to use real data instead.

T 4.10 Mo 18:15 H06

Reinforced Sorting Networks for Particle Physics Analyses — Martin Erdmann, Benjamin Fischer, •Dennis Noll, Yannik Alexander Rath, Marcel Rieger, David Josef Schmidt, and Marcus Wirtz — III. Physikalisches Institut A, RWTH Aachen University

Deep learning architectures in particle physics are often strongly dependent on the order of their input variables. We present a two-stage deep learning architecture consisting of a network for sorting input objects and a subsequent network for data analysis. The sorting network (agent) is trained through reinforcement learning using feedback from the analysis network (environment). A tree search algorithm is used to examine the large space of different possible orders.

The optimal order depends on the environment and is learned by the agent in an unsupervised approach. Thus, the 2-stage system can choose an optimal solution which is not know to the physicist in advance.

We present the new approach and its application to various classification tasks.