

3rd IML Machine Learning Workshop

Report of Contributions

Contribution ID: 1

Type: **not specified**

Gauge Fields in Deep Learning

Tuesday, 16 April 2019 16:30 (1h 30m)

Invited colloquium by Prof. Dr. Max Welling

Part of the IML workshop, but listed under the CERN colloquium category.

Joint work with Taco Cohen, Maurice Weiler and Berkay Kicanaoglu

ABSTRACT:

Gauge field theory is the foundation of modern physics, including general relativity and the standard model of physics. It describes how a theory of physics should transform under symmetry transformations. For instance, in electrodynamics, electric forces may transform into magnetic forces if we transform a static observer to one that moves at constant speed. Similarly, in general relativity acceleration and gravity are equated to each other under symmetry transformations. Gauge fields also play a crucial role in modern quantum field theory and the standard model of physics, where they describe the forces between particles that transform into each other under (abstract) symmetry transformations.

In this work we describe how the mathematics of gauge groups becomes inevitable when you are interested in deep learning on manifolds. Defining a convolution on a manifold involves transporting geometric objects such as feature vectors and kernels across the manifold, which due to curvature become path dependent. As such it becomes impossible to represent these objects in a global reference frame and one is forced to consider local frames. These reference frames are arbitrary and changing between them is called a (local) gauge transformation. Since we do not want our computations to depend on the specific choice of frames we are forced to consider equivariance of our convolutions under gauge transformations. These considerations result in the first fully general theory of deep learning on manifolds, with gauge equivariant convolutions as the necessary key ingredient.

Preferred contribution length

Session Classification: CERN Colloquium

Contribution ID: 2

Type: **not specified**

Coffee break

Session Classification: Industry talks and panel

Contribution ID: 3

Type: **not specified**

Coffee break

Session Classification: Submitted contributions

Contribution ID: 4

Type: **not specified**

Introduction to the basics of deep learning

Thursday, 18 April 2019 09:00 (1h 30m)

Preferred contribution length

Primary author: RATH, Yannik Alexander (RWTH Aachen University (DE))

Presenter: RATH, Yannik Alexander (RWTH Aachen University (DE))

Session Classification: Tutorials

Contribution ID: 5

Type: **not specified**

Advanced Generative Adversarial Network Techniques

Thursday, 18 April 2019 16:45 (1 hour)

Preferred contribution length

Primary author: GLOMBITZA, Jonas (Rheinisch-Westfaelische Tech. Hoch. (DE))

Presenter: GLOMBITZA, Jonas (Rheinisch-Westfaelische Tech. Hoch. (DE))

Session Classification: Tutorials

Contribution ID: 6

Type: **not specified**

Physics inspired Autonomous Feature Engineering

Thursday, 18 April 2019 11:00 (1 hour)

Preferred contribution length

Primary author: RIEGER, Marcel (RWTH Aachen University (DE))

Presenter: RIEGER, Marcel (RWTH Aachen University (DE))

Session Classification: Tutorials

Contribution ID: 7

Type: **not specified**

Tutorial, part 1

Session Classification: Tutorials

Contribution ID: 8

Type: **not specified**

Tutorial, part 2

Session Classification: Tutorials

Contribution ID: 9

Type: **not specified**

Amazon AI: Tensor and Higher-Order Generalizations of the GSVD with Applications to Personalized Cancer Medicine

Monday, 15 April 2019 14:20 (45 minutes)

The number of high-dimensional datasets recording multiple aspects of interrelated phenomena is increasing in many areas, from medicine to finance. This drives the need for mathematical frameworks that can simultaneously identify the similar and dissimilar among multiple matrices and tensors, and thus create a single coherent model from the multiple datasets. The generalized singular value decomposition (GSVD) was formulated as such a comparative spectral decomposition of two column-matched but row-independent matrices. I will, first, define a higher-order GSVD (HO GSVD) and a tensor GSVD and prove that they extend almost all of the mathematical properties of the GSVD to multiple matrices and two tensors. Second, I will describe the development of a tensor HO GSVD for multiple tensors. Third, I will describe the use of these decompositions in the comparisons of cancer and normal genomes, where they uncover patterns that predict survival and response to treatment. The data had been publicly available for years, but the patterns remained unknown until the data were modeled by using the decompositions, illustrating their ability to find what other methods miss.

Preferred contribution length

Primary author: Dr PONNAPALLI, Priya (Amazon AI)

Presenter: Dr PONNAPALLI, Priya (Amazon AI)

Session Classification: Industry talks and panel

Contribution ID: **10**

Type: **not specified**

Introduction

Monday, 15 April 2019 14:00 (20 minutes)

Preferred contribution length

Presenter: STOYE, Markus (CERN)

Session Classification: Industry talks and panel

Contribution ID: 11

Type: **not specified**

Google DeepMind: Compressing neural networks

Monday, 15 April 2019 15:05 (45 minutes)

Empirically it has been observed numerous times that trained neural networks often have high degrees of parameter-redundancy. It remains an open theoretical question why this parameter redundancy cannot be reduced before training by using smaller neural networks. On the other hand, the recent scientific literature reports a plethora of practical methods to “compress” neural networks during or after training with (almost) arbitrarily small sacrifices to task performance while significantly reducing the computational demands of a neural network model. This talk gives an overview over the field of neural network compression methods and then introduces one family of approaches based on Bayesian neural networks. Some of the appealing theoretical properties of Bayesian approaches to neural network compression are discussed and practical implementations for modern neural network training are sketched. The talk concludes with discussing difficulties with neural network compression in practice and an outlook towards exploiting noise and redundancy in the data for more compute-efficient neural networks.

Preferred contribution length

Primary author: Dr GENEWEIN, Tim (DeepMind)

Presenter: Dr GENEWEIN, Tim (DeepMind)

Session Classification: Industry talks and panel

Contribution ID: 12

Type: **not specified**

B12 Consulting: Big value out of small data

Monday, 15 April 2019 16:30 (30 minutes)

Machine Learning (and especially Deep Learning) algorithms often require large amounts of data to accomplish their tasks. However, a common problem when such approaches are applied in business contexts is that only relatively small datasets are initially accessible, leading to a fundamental question: how to apply ML tools when there is apparently not enough data available? In this talk, I will discuss why those “small data” projects are actually of high relevance and discuss three concrete strategies data scientists can utilise to alleviate data deficiency and their strengths/weaknesses. For each strategy, I will present a business case scenario from the portfolio B12 Consulting, illustrating how to partially overcome analysis issues with insufficient data.

Preferred contribution length

Primary author: Dr HERQUET, Michel (B12)

Presenter: Dr HERQUET, Michel (B12)

Session Classification: Industry talks and panel

Contribution ID: 13

Type: **not specified**

Industry panel

Monday, 15 April 2019 17:05 (55 minutes)

Preferred contribution length

Primary authors: Dr HERQUET, Michel (B12); Dr GENEWEIN, Tim (DeepMind)

Presenters: Dr HERQUET, Michel (B12); Dr GENEWEIN, Tim (DeepMind)

Session Classification: Industry talks and panel

Contribution ID: 14

Type: **not specified**

Welcome

Monday, 15 April 2019 09:00 (20 minutes)

Preferred contribution length

Presenter: SCHRAMM, Steven Randolph (Universite de Geneve (CH))

Session Classification: Invited keynote talks

Contribution ID: 15

Type: **not specified**

Conceptual overview of ML in HEP

Monday, 15 April 2019 09:20 (45 minutes)

This talk is intended to provide a conceptual overview of how ML is used in particle physics, in order to provide a basic level of understanding for all attendees such that people can follow the subsequent talks. Detailed tutorials on the application of ML techniques will instead take place on Friday.

Presenter: Dr GLEYZER, Sergei (University of Florida (US))

Session Classification: Invited keynote talks

Contribution ID: 16

Type: **not specified**

Future areas of focus for ML in particle physics

Monday, 15 April 2019 10:05 (45 minutes)

Preferred contribution length

Presenter: CRANMER, Kyle Stuart (New York University (US))

Session Classification: Invited keynote talks

Contribution ID: 17

Type: **not specified**

On the Statistical Mechanics and Information Theory of Deep Learning for Particle Physicists

Monday, 15 April 2019 11:15 (1 hour)

This is an introduction to Deep Learning, for Particle Physicists. These techniques can provide excellent performance for separating different categories of events, pattern recognition, etc. In order to achieve this it is highly desirable for users to understand the basic concepts underlying their operation. The Statistical Mechanics and Information Theory aspects of this will be addressed in this talk.

Preferred contribution length

Primary author: Prof. TISHBY, Naftali (Hebrew University of Jerusalem)

Presenter: Prof. TISHBY, Naftali (Hebrew University of Jerusalem)

Session Classification: Joint lecture/seminar

Contribution ID: **18**

Type: **not specified**

Daily announcements

Tuesday, 16 April 2019 09:00 (5 minutes)

Session Classification: Submitted contributions

Contribution ID: **19**

Type: **not specified**

Daily announcements

Wednesday, 17 April 2019 09:00 (5 minutes)

Session Classification: Submitted contributions

Contribution ID: 20

Type: **not specified**

Close-out

Wednesday, 17 April 2019 17:30 (10 minutes)

Preferred contribution length

Primary author: SCHRAMM, Steven Randolph (Universite de Geneve (CH))

Presenter: SCHRAMM, Steven Randolph (Universite de Geneve (CH))

Session Classification: Submitted contributions

Contribution ID: 21

Type: **not specified**

Novelty Detection Meets Collider Physics

Tuesday, 16 April 2019 14:00 (30 minutes)

Novelty detection is the machine learning task to recognize data, which belong to an unknown pattern. Complementary to supervised learning, it allows to analyze data model-independently. We demonstrate the potential role of novelty detection in collider physics, using autoencoder-based deep neural network. Explicitly, we develop a set of density-based novelty evaluators, which are sensitive to the clustering of unknown-pattern testing data or new-physics signal events, for the design of detection algorithms. We also explore the influence of the known-pattern data fluctuations, arising from non-signal regions, on detection sensitivity. Strategies to address it are proposed. The algorithms are applied to detecting fermionic di-top partner and resonant di-top productions at LHC, and exotic Higgs decays of two specific modes at a future e^+e^- collider. With parton-level analysis, we conclude that potentially the new-physics benchmarks can be recognized with high efficiency.

Preferred contribution length

30 minutes

Primary authors: Dr HAJER, Jan; Ms LI, Ying-Ying (HKUST); Prof. LIU, Tao; Mr WANG, He

Presenter: Ms LI, Ying-Ying (HKUST)

Session Classification: Submitted contributions

Contribution ID: 22

Type: **not specified**

Model-Assisted GANs for the optimisation of simulation parameters and as an algorithm for fast Monte Carlo production

Wednesday, 17 April 2019 11:25 (20 minutes)

We propose and demonstrate the use of a Model-Assisted Generative Adversarial Network to produce simulated images that accurately match true images through the variation of underlying model parameters that describe the image generation process. The generator learns the parameter values that give images that best match the true images. The best match parameter values that produce the most accurate simulated images can be extracted and used to re-tune the default simulation to minimise any bias when applying image recognition techniques to simulated and true images. In the case of a real-world experiment, the true data is replaced by experimental data with unknown true parameter values. The Model-Assisted Generative Adversarial Network uses a convolutional neural network to emulate the simulation for all parameter values that, when trained, can be used as a conditional generator for fast image production.

Preferred contribution length

20 minutes

Primary authors: Mr ALONSO MONSALVE, Saul (CERN); WHITEHEAD, Leigh Howard (University of Cambridge (GB))

Presenter: Mr ALONSO MONSALVE, Saul (CERN)

Session Classification: Submitted contributions

Contribution ID: 23

Type: **not specified**

Learning representations of irregular particle-detector geometry with distance-weighted graph networks

Wednesday, 17 April 2019 15:30 (20 minutes)

We explore the possibility of using graph networks to deal with irregular-geometry detectors when reconstructing particles. Thanks to their representation-learning capabilities, graph networks can exploit the detector granularity, while dealing with the event sparsity and the irregular detector geometry. In this context, we introduce two distance-weighted graph network architectures, the GarNet and the GravNet layers and we apply them to a typical particle reconstruction task. As an example, we consider a high granularity calorimeter, loosely inspired by the endcap calorimeter to be installed in the CMS detector for the High-Luminosity LHC phase. We focus the study on the basis for calorimeter reconstruction, clustering, and provide a quantitative comparison to alternative approaches. The proposed methods outperform previous methods or reach competitive performance while keeping favourable computing-resource consumption. Being geometry agnostic, they can be easily generalized to other use cases and to other detectors, e.g., tracking in silicon detectors.

Preferred contribution length

20 minutes

Primary authors: KIESELER, Jan (CERN); IYAMA, Yutaro (CERN); QASIM, Shah Rukh (SEECS - School of Electrical Engineering and Computer Science); PIERINI, Maurizio (CERN)

Presenter: KIESELER, Jan (CERN)

Session Classification: Submitted contributions

Contribution ID: 24

Type: **not specified**

The Tracking Machine Learning challenge

Tuesday, 16 April 2019 15:30 (30 minutes)

The HL-LHC will see ATLAS and CMS see proton bunch collisions reaching track multiplicity up to 10.000 charged tracks per event. Algorithms need to be developed to harness the increased combinatorial complexity. To engage the Computer Science community to contribute new ideas, we have organized a Tracking Machine Learning challenge (TrackML). Participants were provided events with 100k 3D points, and are asked to group the points into tracks; they are also given a 100GB training dataset including the ground truth. The challenge is run in two phases. The first “Accuracy” phase has run on Kaggle platform from May to August 2018; algorithms were judged only on a score related to the fraction of correctly assigned hits. The second “Throughput” phase ran Sep 2018 to March 2019 on Codalab, required code submission; algorithms were then ranked by combining accuracy and speed. The first phase has seen 653 participants, with top performers with innovative approaches. The second phase has just finished and featured some astonishingly fast solution. The talk will report on the first lessons from the challenge.

Preferred contribution length

30 minutes

Primary authors: ROUSSEAU, David (LAL-Orsay, FR); AMROUCHE, Sabrina (Université de Genève (CH)); BASARA, Laurent Roger Igor (Universita degli Studi di Trento è INFN (IT)); CALAFIURA, Paolo (Lawrence Berkeley National Lab. (US)); Mr ESTRADE, Victor (LRI); FARRELL, Steven (Lawrence Berkeley National Lab (US)); GERMAIN, Cecile (Universite Paris Sud); GUYON, Isabelle; GLIGOROV, Vladimir (Centre National de la Recherche Scientifique (FR)); GOLLING, Tobias (Universite de Geneve (CH)); GRAY, Heather (LBNL); HUSHCHYN, Mikhail (Yandex School of Data Analysis (RU)); INNOCENTE, Vincenzo (CERN); KIEHN, Moritz (Universite de Geneve (CH)); MOYSE, Edward (University of Massachusetts (US)); SALZBURGER, Andreas (CERN); USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU)); Dr VLIMANT, Jean-Roch (California Institute of Technology (US)); YILMAZ, Yetkin (LAL-Orsay (FR))

Presenter: ROUSSEAU, David (LAL-Orsay, FR)

Session Classification: Submitted contributions

Contribution ID: 25

Type: **not specified**

GroomRL: jet grooming through reinforcement learning

Wednesday, 17 April 2019 16:30 (20 minutes)

We introduce a novel implementation of a reinforcement learning algorithm which is adapted to the problem of jet grooming, a crucial component of jet physics at hadron colliders. We show that the grooming policies trained using a Deep Q-Network model outperform state-of-the-art tools used at the LHC such as Recursive Soft Drop, allowing for improved resolution of the mass of boosted objects. The algorithm learns how to optimally remove soft wide-angle radiation, allowing for a modular jet grooming tool that can be applied in a wide range of contexts.

Preferred contribution length

20 minutes

Primary author: DREYER, Frederic Alexandre (Oxford)**Presenter:** DREYER, Frederic Alexandre (Oxford)**Session Classification:** Submitted contributions

Contribution ID: 26

Type: **not specified**

Fast Deep Learning on FPGAs for the Phase-II L0 Muon Barrel Trigger of the ATLAS Experiment

Wednesday, 17 April 2019 17:10 (20 minutes)

The Level-0 Muon Trigger system of the ATLAS experiment will undergo a full upgrade for HL-LHC to stand the challenging performances requested with the increasing instantaneous luminosity. The upgraded trigger system foresees to send RPC raw hit data to the off-detector trigger processors, where the trigger algorithms run on new generation of Field-Programmable Gate Arrays (FPGAs). The FPGA represents an optimal solution in this context, because of its flexibility, wide availability of logical resources and high processing speed. Studies and simulations of different trigger algorithms have been performed, and novel low precision deep neural network architectures (based on ternary dense and convnet networks) optimized to run on FPGAs and to cope with sparse data are presented. Both physics performances in terms of efficiency and fake rates, and FPGA logic resource occupancy and timing obtained with the developed algorithms are presented.

Preferred contribution length

20 minutes

Primary authors: GIAGU, Stefano (Sapienza Universita e INFN, Roma I (IT)); FRANCESCATO, Simone (Sapienza Universita e INFN, Roma I (IT)); SEBASTIANI, Cristiano (INFN Roma and Sapienza Universita' di Roma (IT)); SABETTA, Luigi (Sapienza Universita e INFN, Roma I (IT))

Presenter: SABETTA, Luigi (Sapienza Universita e INFN, Roma I (IT))

Session Classification: Submitted contributions

Contribution ID: 27

Type: **not specified**

A hybrid deep learning approach to vertexing

Wednesday, 17 April 2019 14:00 (20 minutes)

In the transition to Run 3 in 2021, LHCb will undergo a major luminosity upgrade, going from 1.1 to 5.6 expected visible Primary Vertices (PVs) per event, and will adopt a purely software trigger. This has fueled increased interest in alternative highly-parallel and GPU friendly algorithms for tracking and reconstruction. We will present a novel prototype algorithm for vertexing in the LHCb upgrade conditions.

We use a custom kernel to transform the sparse 3D space of hits and tracks into a dense 1D dataset, and then apply Deep Learning techniques to find PV locations. By training networks on our kernels using several Convolutional Neural Network layers, we have achieved better than 90% efficiency with no more than 0.2 False Positives (FPs) per event. Beyond its physics performance, this algorithm also provides a rich collection of possibilities for visualization and study of 1D convolutional networks. We will discuss the design, performance, and future potential areas of improvement and study, such as possible ways to recover the full 3D vertex information.

Preferred contribution length

20 minutes

Primary authors: FANG, Rui; SCHREINER, Henry Fredrick (University of Cincinnati (US)); SOKOLOFF, Michael David (University of Cincinnati (US)); Mr WEISSER, Constantin Niko (Massachusetts Inst. of Technology (US)); WILLIAMS, J Michael (Massachusetts Inst. of Technology (US))

Presenter: SCHREINER, Henry Fredrick (University of Cincinnati (US))

Session Classification: Submitted contributions

Contribution ID: 28

Type: **not specified**

Exploring SMEFT in VH channel with Machine Learning

Tuesday, 16 April 2019 15:00 (30 minutes)

We use Machine Learning (ML) techniques to exploit kinematic information in VH, the production of a Higgs in association with a massive vector boson. We parametrize the effect of new physics in terms of the SMEFT framework. We find that the use of a shallow neural network allows us to dramatically increase the sensitivity to deviations in VH respect to previous estimates. We also discuss the relation between the usual measures of performance in Machine Learning, such as AUC or accuracy, with the more adept measure of Asimov significance. This relation is particularly relevant when parametrizing systematic uncertainties. Our results show the potential of incorporating Machine Learning techniques to the SMEFT studies using the current datasets.

Preferred contribution length

30 minutes

Primary author: KHOSA, Charanjit Kaur**Co-authors:** FERREIRA DE FREITAS, Felipe; Prof. SANZ GONZALEZ, Veronica**Presenter:** KHOSA, Charanjit Kaur**Session Classification:** Submitted contributions

Contribution ID: 29

Type: **not specified**

LUMIN - a deep learning and data science ecosystem for high-energy physics

Wednesday, 17 April 2019 12:05 (20 minutes)

LUMIN aims to become a deep-learning and data-analysis ecosystem for High-Energy Physics, and perhaps other scientific domains in the future. Similar to Keras and fastai it is a wrapper framework for a graph computation library (PyTorch), but includes many useful functions to handle domain-specific requirements and problems. It also intends to provide easy access to to state-of-the-art methods, but still be flexible enough for users to inherit from base classes and override methods to meet their own demands.

In this talk I will be introducing the library, discussing some of its distinguishing characteristics, and going through an example workflow. There will also be a general invitation for people to test out the library and provide feedback, suggestions, or contributions.

Preferred contribution length

20 minutes

Primary author: STRONG, Giles Chatham (LIP Laboratorio de Instrumentacao e Fisica Experimental de Part)

Presenter: STRONG, Giles Chatham (LIP Laboratorio de Instrumentacao e Fisica Experimental de Part)

Session Classification: Submitted contributions

Contribution ID: **30**Type: **not specified**

ML Techniques for heavy flavour identification in CMS

Wednesday, 17 April 2019 14:40 (30 minutes)

Jet flavour identification is a fundamental component for the physics program of the LHC-based experiments. The presence of multiple flavours to be identified leads to a multiclass classification problem. Moreover, the classification of boosted jets has acquired an increasing importance in the physics program of CMS. In this presentation we will present the performance on both simulated and real data of our latest resolved and boosted heavy flavour taggers as well as the future prospects for the evolution of these techniques and the technical strategies adopted to deploy them in the harsh computing environment of a large-scale HEP computing software stack.

Preferred contribution length

30 minutes

Primary author: BOLS, Emil Sorensen (Vrije Universiteit Brussel (BE))**Presenter:** BOLS, Emil Sorensen (Vrije Universiteit Brussel (BE))**Session Classification:** Submitted contributions

Contribution ID: 31

Type: **not specified**

Uncertain Networks

Tuesday, 16 April 2019 14:30 (30 minutes)

Machine learning methods are being increasingly and successfully applied to many different physics problems. However, currently uncertainties in machine learning methods are not modelled well, if at all. In this talk I will discuss how using Bayesian neural networks can give us a handle on uncertainties in machine learning. I will use tagging tops vs. QCD as an example of how these networks are competitive with other neural network taggers with the advantage of providing an event-by-event uncertainty on the classification. I will then further discuss how this uncertainty changes with experimental systematic effects, using pile-up and jet energy scale as examples.

Preferred contribution length

30 minutes

Primary authors: KASIECZKA, Gregor (Hamburg University (DE)); THOMPSON, Jennifer (ITP Heidelberg); PLEHN, Tilman

Presenter: THOMPSON, Jennifer (ITP Heidelberg)

Session Classification: Submitted contributions

Contribution ID: 32

Type: **not specified**

Towards Fast Displaced Vertex Finding

Many standard model extensions predict long-lived massive particles that can be detected by looking for displaced decay vertices in the inner detector volume. Current approaches to seek for these events in high-energy particle collisions rely on the presence of additional energetic signatures to make an online selection during data-taking. Enabling trigger-level reconstruction of displaced vertices would strongly enhance the reach of such searches.

A proposed strategy is to exploit machine learning techniques to reconstruct approximate vertex positions from raw detector hits. In particular, a two-step procedure would first reconstruct the primary interaction position and exploit trigger-level objects, such as high energy muons, to define a region of interest. The second step would also use the raw detector hits to search for a displaced vertex.

This work focuses on the first step of approximating the location of the primary vertex in an idealised detector geometry using dense neural networks for regression. Properties of vertex reconstruction, such as position resolution and dependency on pile-up interactions will be discussed. Finally, future work will be tackled, including promising approaches for displaced vertex reconstruction given a limited search area.

Preferred contribution length

20 minutes

Primary authors: ALBERTSSON, Kim (Lulea University of Technology (SE)); MELONI, Federico (Deutsches Elektronen-Synchrotron (DE))

Presenter: ALBERTSSON, Kim (Lulea University of Technology (SE))

Session Classification: Submitted contributions

Contribution ID: 33

Type: **not specified**

Traditional approach to network architecture optimisation

Thursday, 18 April 2019 12:00 (30 minutes)

Preferred contribution length

Primary author: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Presenter: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Session Classification: Tutorials

Contribution ID: 34

Type: **not specified**

Differentiable architecture search

Thursday, 18 April 2019 12:30 (40 minutes)

Preferred contribution length

Primary author: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Presenter: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Session Classification: Tutorials

Contribution ID: 35

Type: **not specified**

Bayesian approach to network design

Thursday, 18 April 2019 15:00 (30 minutes)

Preferred contribution length

Primary author: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Presenter: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Session Classification: Tutorials

Contribution ID: 36

Type: **not specified**

Bayesian dropout explanation and examples

Thursday, 18 April 2019 15:30 (40 minutes)

Preferred contribution length

Primary author: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Presenter: USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU))

Session Classification: Tutorials

Contribution ID: 37

Type: **not specified**

Machine Learning Uncertainties with Adversarial Neural Networks

Tuesday, 16 April 2019 09:05 (20 minutes)

Machine learning is a powerful tool to reveal and exploit correlations in a multi-dimensional parameter space. Making predictions from such correlations is a highly non-trivial task, in particular when the details of the underlying dynamics of a theoretical model are not fully understood. Uncertainties in the training data add towards the complexity of performing machine learning tasks such as event classification. However, it has been shown that adversarial neural networks can be used to decorrelate the trained model from systematic uncertainties that affect the kinematics on an event-by-event basis. Here we show that this approach can be extended to theoretical uncertainties (e.g. renormalization and factorization scale uncertainties) that affect the event sample as a whole. The result of the adversarial training is a classifier that is insensitive to these uncertainties by having learned to avoid regions of phase space (or feature space) that are affected by the uncertainties. This paves the way to a more reliable event classification, as well as novel approaches to perform parameter fits of particle physics data. We demonstrate the benefits of the method explicitly in an example considering effective field theory extensions of Higgs boson production in association with jets.

Preferred contribution length

20 minutes

Primary authors: ENGLERT, Christoph Peter (University of Glasgow (GB)); SPANNOWSKY, Michael (University of Durham (GB)); Dr GALLER, Peter (University of Glasgow); HARRIS, Philip Coleman (Massachusetts Inst. of Technology (US))

Presenter: Dr GALLER, Peter (University of Glasgow)

Session Classification: Submitted contributions

Contribution ID: 38

Type: **not specified**

Event Generation and Statistical Sampling with Deep Generative Models

Wednesday, 17 April 2019 11:45 (20 minutes)

We present a study for the generation of events from a physical process with generative deep learning. To simulate physical processes it is not only important to produce physical events, but also to produce the events with the right frequency of occurrence (density). We investigate the feasibility to learn the event generation and the frequency of occurrence with Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) to produce events like Monte Carlo generators. We study three toy models from high energy physics, i.e. a simple two-body decay, the processes $e^+e^- \rightarrow Z \rightarrow l^+l^-$ and $pp \rightarrow t\bar{t}$ including the decay of the top quarks and a simulation of the detector response. We show that GANs and the standard VAE do not produce the right distributions. By buffering density information of Monte Carlo events in latent space given the encoder of a VAE we are able to construct a prior for the sampling of new events from the decoder that yields distributions that are in very good agreement with real Monte Carlo events and are generated $\mathcal{O}(10^8)$ times faster. Applications of this work include generic density estimation and sampling, targeted event generation via a principal component analysis of encoded events in the latent space and the possibility to generate better random numbers for importance sampling, e.g. for the phase space integration of matrix elements in quantum perturbation theories. The method also allows to build event generators directly from real data events.

Preferred contribution length

20 minutes

Primary authors: OTTEN, Sydney (Radboud Universiteit Nijmegen); CARON, Sascha (Nikhef National institute for subatomic physics (NL)); Ms DE SWART, Wieske (Radboud University Nijmegen); VAN BEEKVELD, Melissa (R); HENDRIKS, Luc (Nikhef); Mr PODAREANU, Damian (SURFsara); Mr VAN LEEUWEN, Caspar (SURFsara); RUIZ DE AUSTRI, Roberto; VERHEYEN, Rob (Radboud University Nijmegen)

Presenter: OTTEN, Sydney (Radboud Universiteit Nijmegen)

Session Classification: Submitted contributions

Contribution ID: 39

Type: **not specified**

The information theory of Deep Learning

Tuesday, 16 April 2019 11:15 (1 hour)

Preferred contribution length

Primary author: Prof. TISHBY, Naftali (Hebrew University of Jerusalem)

Presenter: Prof. TISHBY, Naftali (Hebrew University of Jerusalem)

Session Classification: Invited plenary talk

Contribution ID: 40

Type: **not specified**

ParticleNet: Jet Tagging via Particle Clouds

Wednesday, 17 April 2019 15:10 (20 minutes)

How to represent a jet is at the core of machine learning on jet physics. Inspired by the notion of point cloud, we propose a new approach that considers a jet as an unordered set of its constituent particles, effectively a “particle cloud”. Such particle cloud representation of jets is efficient in incorporating raw information of jets and also explicitly respects the permutation symmetry. Based on the particle cloud representation, we propose ParticleNet, a customized neural network architecture using Dynamic Graph CNN for jet tagging problems. The ParticleNet architecture achieves state-of-the-art performance on two representative jet tagging benchmarks and improves significantly over existing methods.

Preferred contribution length

20 minutes

Primary authors: QU, Huilin (Univ. of California Santa Barbara (US)); GOUSKOS, Loukas (Univ. of California Santa Barbara (US))

Presenter: QU, Huilin (Univ. of California Santa Barbara (US))

Session Classification: Submitted contributions

Contribution ID: 41

Type: **not specified**

Deep generative models for fast shower simulation in ATLAS

Wednesday, 17 April 2019 10:15 (20 minutes)

The extensive physics program of the ATLAS experiment at the Large Hadron Collider (LHC) relies on large scale and high fidelity simulation of the detector response to particle interactions. Current full simulation techniques using Geant4 provide accurate modeling of the underlying physics processes, but are inherently resource intensive. In light of the high-luminosity upgrade of the LHC and the need for ever larger simulated datasets to support physics analysis, the development of new faster simulation techniques is crucial. Building on the recent success of deep learning algorithms, Variational Auto-Encoders and Generative Adversarial Networks are investigated for modeling the response of the ATLAS electromagnetic calorimeter for photons in a central calorimeter region over a range of energies. The properties of synthesized showers using deep neural networks are compared to showers from a full detector simulation using Geant4. With this feasibility study we demonstrate the potential of using such algorithms for fast calorimeter simulation for the ATLAS experiment in the future, complementing current simulation techniques.

Preferred contribution length

20 minutes

Primary author: GHOSH, Aishik (Centre National de la Recherche Scientifique (FR))**Presenter:** GHOSH, Aishik (Centre National de la Recherche Scientifique (FR))**Session Classification:** Submitted contributions

Contribution ID: 42

Type: **not specified**

NeuralRinger: An Ensemble of Neural Networks Fed from Calorimeter Ring Sums for Triggering on Electrons

Wednesday, 17 April 2019 16:50 (20 minutes)

In 2017, the ATLAS experiment implemented an ensemble of neural networks (NeuralRinger algorithm) dedicated to reduce the latency of the first, fast, online software (HLT) selection stage for electrons with transverse energy above 15 GeV. In order to minimize detector response and shower development fluctuations, and being inspired in the ensemble of likelihood models currently operating in the offline and final HLT selections, the ensemble comprises Multi-Layer Perceptron (MLP) models tuned for pseudo-rapidity and transverse energy bins. The MLPs are fed from calorimetry information formatted into concentric ring energy sums, which are built around the particle axis and normalized by its total transverse energy. Although triggering algorithms are typically developed from offline models adapted to operate in stringent online conditions, the NeuralRinger development was starting from the online perspective. We describe the analysis performed during the NeuralRinger development in the Run 2, the trigger commissioning and its online performance during the 2017 and 2018 data-taking. It is estimated that the Neural Ringer allowed a reduction of 25% in the e/γ slice processing demands when considering its operation on top of all other improvements done in the electron chains. Statistical tests performed on crucial offline calorimeter response discriminant features show negligible impact ($<1\sigma$) in offline reconstruction.

Preferred contribution length

20 minutes

Primary authors: SPOLIDORO FREUND, Werner (Federal University of of Rio de Janeiro (BR)); ATLAS COLLABORATION

Presenter: SPOLIDORO FREUND, Werner (Federal University of of Rio de Janeiro (BR))

Session Classification: Submitted contributions

Contribution ID: 43

Type: **not specified**

Learning Invariant Representations using Mutual Information Regularization

Tuesday, 16 April 2019 09:45 (20 minutes)

Invariance of learned representations of neural networks against certain sensitive attributes of the input data is a desirable trait in many modern-day applications of machine learning, such as precision measurements in experimental high-energy physics and enforcing algorithmic fairness in the social and financial domain. We present a method for enforcing this invariance through regularization of the mutual information between the target variable and the classifier output. Applications of the proposed technique to rare decay searches in experimental high-energy physics are presented, and demonstrate improvement in statistical significance over conventionally trained neural networks and classical machine learning techniques.

Preferred contribution length

20 minutes

Primary author: Mr TAN, Justin (University of Melbourne)**Presenter:** Mr TAN, Justin (University of Melbourne)**Session Classification:** Submitted contributions

Contribution ID: 44

Type: **not specified**

Containers for Machine Learning in HEP

Tuesday, 16 April 2019 10:25 (20 minutes)

Physicists want to use modern open source machine learning tools developed by industry for machine learning projects and analyses in high energy physics. The software environment that a physicist prototypes, tests, and runs these projects in is ideally the same regardless of compute site (be it their laptop or on the GRID). However, historically it has been difficult to find compute sites that have both the desired hardware resources for machine learning (i.e. GPUs) and a compatible software environment for the project, resulting in suboptimal use of resources and wasted researcher time tuning their software requirements to the imposed constraints. Container technologies, such as Docker and Singularity, provide a scalable and robust solution to this problem.

We present work by *Heinrich* demonstrating the use of containers to run analysis jobs in reproducible compute environments at GRID endpoints with GPU resources that support Singularity. We additionally present complimentary work by *Feickert* that provides publicly available “base” Docker images of a HEP orientated machine learning environment: the CentOS 7 file system with the ATLAS “standalone” analysis release AnalysisBase, HDF5 support and utilities, and modern Python 3 with pip with libraries such as NumPy, TensorFlow and uproot installed. We further present ongoing synergetic work to expand both these efforts.

Preferred contribution length

20 minutes

Primary authors: FEICKERT, Matthew (Southern Methodist University (US)); HEINRICH, Lukas Alexander (CERN); GUEST, Dan (University of California Irvine (US))

Presenter: FEICKERT, Matthew (Southern Methodist University (US))

Session Classification: Submitted contributions

Contribution ID: 45

Type: **not specified**

High Granularity Calorimeter Simulation using Generative Adversarial Networks

Wednesday, 17 April 2019 09:55 (20 minutes)

High Energy Physics simulation typically involves Monte Carlo method. Today >50% of WLCG resources are used for simulation that will increase further as detector granularity and luminosity increase. Machine learning has been very successful in the field of image recognition and generation. We have explored image generation techniques for speeding up HEP detector simulation. Calorimeter responses can be treated as images with energy deposition interpreted as “pixel intensities”. One important difference is that pixel luminosity usually cover a range of 0-255 while energy depositions can vary over many orders of magnitude. We have implemented a three dimensional detector simulation tool using Generative Adversarial Networks. Our initial implementation could generate detector response for different energies of the incoming particles at fixed angles in a 25x25x25 cell grid. We present an upgraded version able to simulate electron showers for variable angles of impact in addition to variable primary energies. The inclusion of angles required increasing the sample size in transverse direction to 51x51x25 cells, multiplying by four the number of outputs. Due to the complexity of the task, the range of primary energies has been initially limited to 100-200 GeV. Training was has been improved raising cell energies to power less than one. A check for correct angle is added to the cost function together with comparisons of cell energy distribution. Currently, the accuracy of the result is a bit lower than the fixed angle version but still within 10% for relevant shower parameters.

Preferred contribution length

30 minutes

Primary authors: CARMINATI, Federico; KHATTAK, Gul Rukh (University of Peshawar (PK)); VALLECORSIA, Sofia

Presenter: KHATTAK, Gul Rukh (University of Peshawar (PK))

Session Classification: Submitted contributions

Contribution ID: 46

Type: **not specified**

Fast Simulation Using Generative Adversarial Network in LHCb

Wednesday, 17 April 2019 11:05 (20 minutes)

LHCb is one of the major experiments operating at the Large Hadron Collider at CERN. The richness of the physics program and the increasing precision of the measurements in LHCb lead to the need of ever larger simulated samples. This need will increase further when the upgraded LHCb detector will start collecting data in the LHC Run 3. Given the computing resources pledged for the production of Monte Carlo simulated events in the next years, the use of fast simulation techniques will be mandatory to cope with the expected dataset size. In LHCb generative models, which are nowadays widely used for computer vision and image processing are being investigated in order to accelerate the generation of showers in the calorimeter and high-level responses of Cherenkov detector. We demonstrate that this approach provides high-fidelity results along with a significant speed increase and discuss possible implication of these results. We also present an implementation of this algorithm into LHCb simulation software and validation tests.

Preferred contribution length

30 minutes

Primary authors: RATNIKOV, Fedor (Yandex School of Data Analysis (RU)); DERKACH, Denis (National Research University Higher School of Economics (RU)); MAEVSKIY, Artem (National Research University Higher School of Economics (RU)); Mr ARTEMIEV, Maxim (National Research University - Higher School of Economics); KHAIUROV, Ruslan (Yandex School of Data Analysis (RU)); Mr KAZEEV, Nikita (Yandex School of Data Analysis (RU)); USTYUZHANIN, Andrey (Yandex School of Data Analysis (RU)); ANDERLINI, Lucio (Universita e INFN, Firenze (IT)); ZAKHAROV, Egor (Skolkovo Institute of Science and Technology)

Presenter: MAEVSKIY, Artem (National Research University Higher School of Economics (RU))

Session Classification: Submitted contributions

Contribution ID: 47

Type: **not specified**

DijetGAN: A Generative-Adversarial Network Approach for the Simulation of QCD Dijet Events at the LHC

Wednesday, 17 April 2019 09:35 (20 minutes)

In this talk, I will present a Generative-Adversarial Network (GAN) based on convolutional neural networks that is used to simulate the production of pairs of jets at the LHC. The GAN is trained on events generated using MadGraph5 + Pythia8, and Delphes3 fast detector simulation. A number of kinematic distributions both at Monte Carlo truth level and after the detector simulation can be reproduced by the generator network with a very good level of agreement.

Preferred contribution length

20 minutes

Primary authors: FAUCCI GIANNELLI, Michele (University of Edinburgh); DI SIPIO, Riccardo (University of Toronto (CA)); KETABCHI, Sana (University of Toronto (CA)); PALAZZO, Serena (The University of Edinburgh (GB))

Presenter: PALAZZO, Serena (The University of Edinburgh (GB))

Session Classification: Submitted contributions

Contribution ID: 48

Type: **not specified**

Feature ranking based on subtraction methods

Wednesday, 17 April 2019 14:20 (20 minutes)

The input variables of ML methods in physics analysis are often highly correlated and figuring out which ones are the most important ones for the classification turns out to be a non-trivial task. We compare the standard method of TMVA to rank variables with several newly developed methods based on iterative removal for the use case of a search for top pair associated Higgs production (ttH) in the Higgs to b-pair decay channel.

Preferred contribution length

20 minutes

Primary authors: GLAYSHER, Paul (DESY); AN, Sitong (CERN, Carnegie Mellon University (US)); KATZY, Judith (Deutsches Elektronen-Synchrotron (DE))

Presenter: GLAYSHER, Paul (DESY)

Session Classification: Submitted contributions

Contribution ID: 49

Type: **not specified**

Decoding Physics Information in DNNs

Tuesday, 16 April 2019 09:25 (20 minutes)

A more dedicated study on the information flow in DNNs will help us understand their behaviour and the deep connection between DNN models and the corresponding tasks. Taking into account our well-established physics analysis framework (observable-based), we present a novel way to interpret DNNs results for HEP, which not only gives a clear physics picture but also inspires interfaces with the theoretical foundation. Information captured by DNNs can thus be used as a fine-tailored general-purpose encoder. As a concrete example, we showcase using encoded information to help with physics searches at the LHC.

Preferred contribution length

20 minutes

Primary author: CHENG, Taoli (University of Montreal)**Presenter:** CHENG, Taoli (University of Montreal)**Session Classification:** Submitted contributions

Contribution ID: 50

Type: **not specified**

Applying Generative Models to Scientific Research

Wednesday, 17 April 2019 09:05 (30 minutes)

Surrogate generative models demonstrate extraordinary progress in current years. Although most applications are dedicated to image generation and similar commercial goals, this approach is also very promising for natural sciences, especially for tasks like fast event simulation in HEP experiments. However, application of such generative models to scientific research implies specific requirements and expectations from these models. In the presentation, I'll discuss specific points which need attention when using generative models for scientific research. This includes ensuring that models satisfy different boundary conditions and match scientifically important but marginal statistics. We also need to establish procedures to evaluate the quality of the particular model, propagate model imperfection into systematic uncertainties of the final scientific result, and so on.

Preferred contribution length

30 minutes

Primary author: RATNIKOV, Fedor (Yandex School of Data Analysis (RU))**Presenter:** RATNIKOV, Fedor (Yandex School of Data Analysis (RU))**Session Classification:** Submitted contributions

Contribution ID: 51

Type: **not specified**

Neural networks for the abstraction of the physical symmetries in the nature

Tuesday, 16 April 2019 10:05 (20 minutes)

Neural networks are so powerful universal approximator of complicated patterns in large-scale data, leading the explosive developments of AI in terms of deep learning. However, in many cases, usual neural networks are trained to possess poor level of abstraction, so that the model's predictability and generalizability can be quite unstable, depending on the quality and amount of the data used for training. In this presentation, we introduce a new neural network architecture which has improved capability of capturing the key features and the physical laws hidden in data, in a mathematically more robust and simpler way. We demonstrate the performance of the new architecture, with an application for high energy particle scattering processes at the LHC.

Preferred contribution length

20 minutes

Primary author: CHO, Wonsang (Seoul National University)**Presenter:** CHO, Wonsang (Seoul National University)**Session Classification:** Submitted contributions

Contribution ID: 52

Type: **not specified**

Talk 1

Session Classification: Submitted contributions

Contribution ID: 53

Type: **not specified**

Talk 2

Session Classification: Submitted contributions

Contribution ID: 54

Type: **not specified**

Talk 3

Session Classification: Submitted contributions

Contribution ID: 55

Type: **not specified**

Talk 1

Session Classification: Submitted contributions

Contribution ID: 56

Type: **not specified**

Talk 2

Session Classification: Submitted contributions

Contribution ID: 57

Type: **not specified**

Talk 3

Session Classification: Submitted contributions

Contribution ID: **58**

Type: **not specified**

Talk 4

Session Classification: Submitted contributions

Contribution ID: 59

Type: **not specified**

Talk 5

Session Classification: Submitted contributions

Contribution ID: **60**

Type: **not specified**

Talk 1

Session Classification: Submitted contributions

Contribution ID: **61**

Type: **not specified**

Talk 2

Session Classification: Submitted contributions

Contribution ID: **62**

Type: **not specified**

Talk 3

Session Classification: Submitted contributions

Contribution ID: **63**

Type: **not specified**

Talk 5

Session Classification: Submitted contributions

Contribution ID: **64**

Type: **not specified**

Talk 6

Session Classification: Submitted contributions

Contribution ID: 65

Type: **not specified**

Talk 7

Session Classification: Submitted contributions

Contribution ID: **66**

Type: **not specified**

Talk 1

Session Classification: Submitted contributions

Contribution ID: 67

Type: **not specified**

Talk 2

Session Classification: Submitted contributions

Contribution ID: **68**

Type: **not specified**

Talk 3

Session Classification: Submitted contributions

Contribution ID: **69**

Type: **not specified**

Talk 4

Session Classification: Submitted contributions

Contribution ID: 70

Type: **not specified**

Talk 5

Session Classification: Submitted contributions

Contribution ID: 71

Type: **not specified**

Talk 6

Session Classification: Submitted contributions

Contribution ID: 72

Type: **not specified**

Talk 7

Session Classification: Submitted contributions

Contribution ID: 73

Type: **not specified**

Talk 8

Session Classification: Submitted contributions

Contribution ID: 74

Type: **not specified**

Talk 4

Session Classification: Submitted contributions

Contribution ID: 75

Type: **not specified**

Talk 4

Session Classification: Submitted contributions

Contribution ID: 76

Type: **not specified**

Talk 5

Session Classification: Submitted contributions

Contribution ID: 77

Type: **not specified**

Talk 8

Session Classification: Submitted contributions

Contribution ID: 78

Type: **not specified**

Talk 9

Session Classification: Submitted contributions

Contribution ID: **80**

Type: **not specified**

Q&A with Prof. Naftali Tishby

Tuesday, 16 April 2019 14:15 (45 minutes)

Preferred contribution length

Primary author: Prof. TISHBY, Naftali (Hebrew University of Jerusalem)

Presenter: Prof. TISHBY, Naftali (Hebrew University of Jerusalem)

Session Classification: Q&A with Prof. Naftali Tishby