

I Workshop de Computing y Software de la Red Española de LHC

Wednesday 28 April 2021 - Thursday 29 April 2021

Book of Abstracts

Contents

Hardware for Machine Learning and ML as a service	1
Anomalies detection with Machine Learning at the LHC	1
Application of Machine Learning for LHC experiment operations	1
Topical presentations	1
Future of COMCHA	1
Strategy and perspectives	1
Short presentations	2
Resources and exploitation	2
Data Management	2
Analysis Facilities	2
Welcome	2
Welcome and Introduction	3
Test	3
Lightweight site federation for CMS support	3
Exploitation of network-segregated CPU resources at BSC for CMS	3
A Deep Learning approach to LHCb Calorimeter reconstruction using a Cellular Automaton	3
ATLAS EventIndex: A Catalogue of all ATLAS Events	4
Hunting dark matter signals with deep learning at the LHC	4
High-dimensional sampling algorithms comparison for particle and astrophysics applications	4
Jet classification in t-tbar decays of heavy BSM resonances in ATLAS using ML techniques	5
A Spanish data cache service for the CMS experiment	6
DNN used in the search for scalar top quark pair production in the top corridor region	6

COMCHA / 1

Hardware for Machine Learning and ML as a service

Corresponding Author: mdelamor@us.es

COMCHA / 2

Anomalies detection with Machine Learning at the LHC

Corresponding Author: judita.mamuzic@cern.ch

COMCHA / 3

Application of Machine Learning for LHC experiment operations

Author: Pablo Martinez Ruiz Del Arbol¹

¹ *Universidad de Cantabria and CSIC (ES)*

Corresponding Author: pablo.martinez@cern.ch

COMCHA / 4

Topical presentations

COMCHA / 5

Future of COMCHA

Author: Arantza De Oyanguren Campos¹

¹ *Univ. of Valencia and CSIC (ES)*

Corresponding Author: arantza.de.oyanguren.campos@cern.ch

LHC Computing / 6

Strategy and perspectives

Authors: Andreu Pacheco Pages¹; Jose Hernandez²; Jose Salt³

¹ *Institut de Física d'Altes Energies - Barcelona (ES)*

² *CIEMAT*

³ *Univ. of Valencia and CSIC (ES)*

Corresponding Authors: jose.salt@cern.ch, chema.hernandez.calama@gmail.com, pacheco@ifae.es

LHC Computing / 7

Short presentations

LHC Computing / 8

Resources and exploitation

Author: Santiago Gonzalez De La Hoz¹

¹ *Univ. of Valencia and CSIC (ES)*

Corresponding Author: santiago.gonzalezdelahoz@cern.ch

LHC Computing / 9

Data Management

Corresponding Authors: chema.hernandez.calama@gmail.com, jose.flix.molina@cern.ch

LHC Computing / 10

Analysis Facilities

Author: Antonio Perez-Calero Yzquierdo¹

Co-author: Andreu Pacheco Pages²

¹ *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

² *Institut de Física d'Altes Energies - Barcelona (ES)*

Corresponding Authors: antonio.perez.calero.yzquierdo@cern.ch, pacheco@ifae.es

COMCHA / 11

Welcome

Author: Carmen Garcia¹

¹ *Univ. of Valencia and CSIC (ES)*

Corresponding Author: carmen.garcia@cern.ch

LHC Computing / 12

Welcome and Introduction

Author: Jose Salt¹

¹ *Univ. of Valencia and CSIC (ES)*

Corresponding Author: jose.salt@cern.ch

LHC Computing / 14

Test

This is a test

LHC Computing / 18

Lightweight site federation for CMS support

Authors: Carlos Acosta Silva¹; Antonio Delgado Peris²; Jose Flix Molina²; Jaime Gómez del Pulgar Ruano³; Jose Miguel Guerrero³; Jose Hernandez³; Antonio Perez-Calero Yzquierdo²; Francisco Javier Rodriguez-Calonge²

¹ *PIC*

² *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

³ *CIEMAT*

Corresponding Author: antonio.delgado.peris@cern.ch

LHC Computing / 19

Exploitation of network-segregated CPU resources at BSC for CMS

Authors: Jose Hernandez¹; Antonio Delgado Peris²; Antonio Perez-Calero Yzquierdo²; Jose Flix Molina²

¹ *CIEMAT*

² *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

Corresponding Author: chema.hernandez.calama@gmail.com

COMCHA / 20

A Deep Learning approach to LHCb Calorimeter reconstruction using a Cellular Automaton

Author: Nuria Valls Canudas¹

Co-authors: Xavier Vilasis Cardona¹; Miriam Calvo Gomez¹; Elisabet Golobardes Ribe²

¹ *La Salle, Ramon Llull University (ES)*

² *University of Barcelona (ES)*

The optimization of reconstruction algorithms has become a key aspect in LHCb as it is currently undergoing a major upgrade that will considerably increase the data processing rate. Aiming to accelerate the second most time consuming reconstruction process of the trigger, we propose an alternative reconstruction algorithm for the Electromagnetic Calorimeter of LHCb. Together with the use of deep learning techniques and the understanding of the current algorithm, our proposal decomposes the reconstruction process into small parts that benefit the generalized learning of small neural network architectures and simplifies the training dataset. This approach takes as input the full simulation data of the calorimeter and outputs a list of reconstructed clusters in a nearly constant time without any dependency in the event complexity.

LHC Computing / 21

ATLAS EventIndex: A Catalogue of all ATLAS Events

Authors: Carlos Garcia Montoro¹; Alvaro Fernandez Casani¹; Santiago Gonzalez De La Hoz¹; Jose Salt¹; Javier Sanchez²; Miguel Villaplana¹

¹ *Univ. of Valencia and CSIC (ES)*

² *Universidad de Valencia (ES)*

Corresponding Author: carlos.garcia.montoro@cern.ch

COMCHA / 22

Hunting dark matter signals with deep learning at the LHC

Authors: Andres Daniel Perez¹; Ernesto Arganda²; Anibal Medina¹; Alejandro Szykman¹

¹ *IFLP - UNLP/CONICET*

² *Instituto de Física Teórica UAM/CSIC*

Corresponding Author: perez.andres.daniel@gmail.com

We study several simplified dark matter models and their signatures at the LHC using Neural Networks. We focus on the usual monojet plus missing transverse energy channel, but to train the algorithms we organize the data in 2D histograms instead of event-by-event arrays. This results in a huge performance boost to distinguish between standard model (SM) only and SM plus new physics signals. We found that Neural Network results do not change with the number of background events if they are shown as a function of S/\sqrt{B} , where S and B are the number of signal and background events per histogram, respectively. This provides flexibility to the method, since testing a particular model is straightforward, only the new physics monojet cross-section is needed. Furthermore, we also discuss the network performance under incorrect assumptions. Finally, we propose multimodel classifiers to search and identify new signals in a more general way, for the next LHC run.

COMCHA / 23

High-dimensional sampling algorithms comparison for particle and astrophysics applications

Authors: Roberto Ruiz De Austri¹; Judita Mamuzic²

¹ *Instituto de Fisica Corpuscular (ES)*

² *IFIC Valencia*

High energy physics analyses often encounter optimization problems in high-dimensional parameter space, and finding an optimal algorithm represents an important initial step in each analysis, where typically an optimum of a very complicated function needs to be found. We performed an optimization algorithm comparison using a number of test functions, for increasing dimensionality, and compared them to a random sampling algorithm. Additionally we made a performance comparison for a realistic high dimensional example, derived from a recent global fit of weak-scale Supersymmetry. We find that for different test functions different algorithms show to be optimal, and we present comparison of performance for random sampling, Differential Evolution, Particle Swarm Optimization, the Covariance Matrix Adaptation Evolution Strategy, Bayesian Optimization, Grey Wolf Optimization, and the PyGMO Artificial Bee Colony, Gaussian Particle Filter and Adaptive Memory Programming for Global Optimization algorithms in [1].

[1] <https://arxiv.org/abs/2101.04525>

COMCHA / 24

Jet classification in t-tbar decays of heavy BSM resonances in ATLAS using ML techniques

Authors: Jorge Juan Martinez De Lejarza Samper¹; Julio Lozano Bahilo¹; Jose Salt²

¹ *Univ. of Valencia and CSIC (ES)*

² *Instituto de Fisica Corpuscular (IFIC) - Universidad de Valencia*

Corresponding Author: jormard2@alumni.uv.es

The search for heavy resonances with diverse mass and width values is the ultimate goal of the ATLAS physics group where we are developing this study. The analysis selects events where one of the top quarks decays hadronically and the other one semileptonically. The reconstruction of the different components of the decay allows us to build an invariant mass distribution which would show some excess around the mass of the resonance if it exists.

The improvement in the reconstruction of the ttbar decay is the main purpose of our study. Therefore we are replacing a traditional Chi2 technique based on intermediate invariant mass values and transverse momentum balance with a novel approach making use of Machine Learning algorithms with any variable that might contribute to a classification of the jets of the event.

After a selection of appropriate events we store a set of variables for each jet in the event including kinematic variables, angular distances, invariant masses and tagging variables. Our classification involves 4 different classes: b jet from hadronic decay; b jet from semileptonic decay; jets coming from the desintegration of the W boson of the hadronic top quark decay and any other jet present in the event but not related to the resonance decay process. Given the fact that we prefer to characterise the jets in a non-binary way,

our problem would be termed as multiclass multilabel. We have used several algorithms capable of providing classification for that kind of problem, namely Deep Neural Networks, Random Forest and eXtreme Gradient Boosting in order to train our dataset of jets.

As a first step we study the relevance of the variables with the help of the Permutation Importance method (DNN) and the Boruta method (RF, XGB) and discard those that do not contribute to any algorithm and resonance mass. Then we proceed with the optimization of the hyperparameters for each kind of algorithm. The training is followed by an assignment of jet roles within each event that gives our final reconstruction efficiency.

Preliminary results show a reconstruction efficiency at $M(Z')=1\text{TeV}$ of around 77% of the events where all jets were matched (reference jets) at Monte Carlo level for XGB and slightly smaller efficiencies for RF and DNN. The traditional Chi2 method gives values just above 70% for the same mass value. Current work involves the addition of the Chi2 assignment to the set of variables and the use of a clustering algorithm to allow further optimization of the ML classifying algorithms.

LHC Computing / 25

A Spanish data cache service for the CMS experiment

Author: Jose Flix Molina¹

Co-authors: Jose Hernandez²; Antonio Perez-Calero Yzquierdo¹; Carlos Acosta Silva³; Antonio Delgado⁴; Carlos Perez Dengra⁵

¹ *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

² *CIEMAT*

³ *PIC*

⁴ *University of Notre Dame (US)*

⁵ *PIC-CIEMAT*

Corresponding Author: jose.flix.molina@cern.ch

COMCHA / 26

DNN used in the search for scalar top quark pair production in the top corridor region

Author: Andrea Trapote Fernandez¹

¹ *Universidad de Oviedo (ES)*

Corresponding Author: andrea.trapote.fernandez@cern.ch

A parametric DNN algorithm is used to separate signal from background in the search for scalar top quark pair production at the LHC. This search targets a region of parameter space where the kinematics of top squark pair production and top quark pair production are very similar, because of the mass difference between the top squark and the neutralino being close to the top quark mass. The search is performed with the full run 2 data set of proton-proton collisions at a centre-of-mass energy of 13 TeV, collected by the CMS detector.

COMCHA / 27

Information

Corresponding Authors: jose.salt@cern.ch, chema.hernandez.calama@gmail.com, pacheco@ifae.es, luca.fiorini@cern.ch