

**Big
ata
HEP**



Universidade do
Minho



TÉCNICO
LISBOA



LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia.

1ª reunião do projeto BigDataHEP

Coimbra, 11 de janeiro de 2019

Nuno Castro e Guilherme Milhano

BigDataHEP: Understanding Big Data in High Energy Physics

finding a needle in many haystacks

BDT kNN Octave SK-Learn
TMVA TensorFlow Numpy
Keras GlusterFS Pandas DNN CNNs
FPGAs RNNs ANN Distributed training Matlab
Pre-processing SVM RNNs K-fold GPUs CV
PCA NNs Theano XGBoost

BigDataHEP: Understanding Big Data in High Energy Physics

finding a needle in many haystacks

- Recursos computacionais
 - Minho:
 - servidor lipml.lip.pt: 1 GPU XP-Titan (NVIDIA GPU grant)
 - 2 servidores comprados pelo projeto (em instalação)
 - Lisboa:
 - 1 servidor comprado pelo projeto com 1 GPU Tesla P100 (será instalado logo que possível)
- Naturalmente que a ideia é partilhar todos os recursos disponíveis (e tentar obter mais...)



BigDataHEP: Understanding Big Data in High Energy Physics

finding a needle in many haystacks

- Recursos Humanos
 - pagos por verbas do projeto:
 - 1 investigador (18 meses): concurso aberto em Setembro
Rute começa em Março
 - 1 investigador (6+6 meses):
concurso abre dia 15 [<http://cern.ch/go/Gmx7>]
 - cada tarefa tem 12 meses de bolsa (lic. ou mestre) previstos
 - linha 5: Diogo (aluno de mestrado do DQ / Minho)

BigDataHEP: Understanding Big Data in High Energy Physics

finding a needle in many haystacks

- É fundamental não nos esquecermos de referenciar o projeto nos trabalhos desenvolvidos (apresentações, artigos, posters, teses,)
 - POCI-01-0145-FEDER-029147 - PTDC/FIS-PAR/29147/2017
- financiado por: OE/FCT, Lisboa 2020, Compete 2020 POCI, Portugal 2020 FEDER

Organizing Committee:

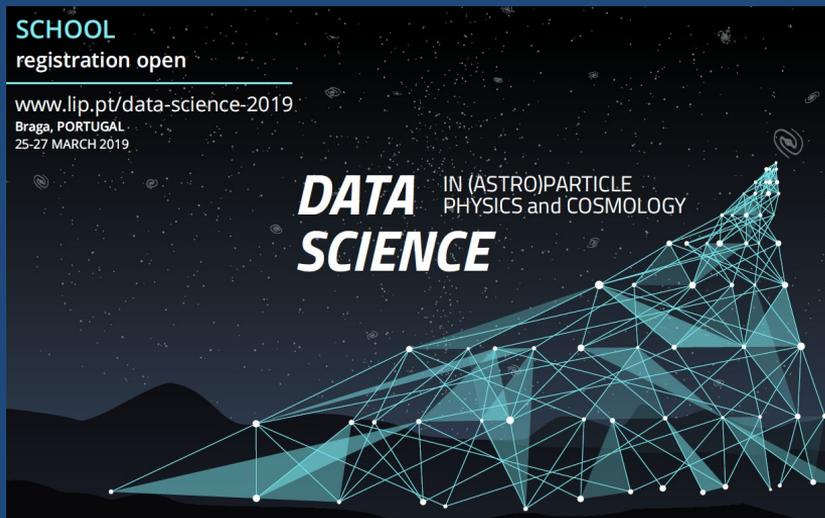
Liliana Apolinário
Annalisa Appice
Michal Bejger
Nuno Castro
Lorenzo Cazon
Ruben Conceição
Elena Cuoco
Ricardo Gonçalves
Nuno Leonardo
Valentina Lozza
Christopher Messenger
Guilherme Milhano
Guilherme Pereira
Mário Pimenta
Raul Sarmento
Bernardo Tomé
Agata Trovato
Filipe Vaz

Design and Development:

Henrique Carvalho
Carlos Manuel

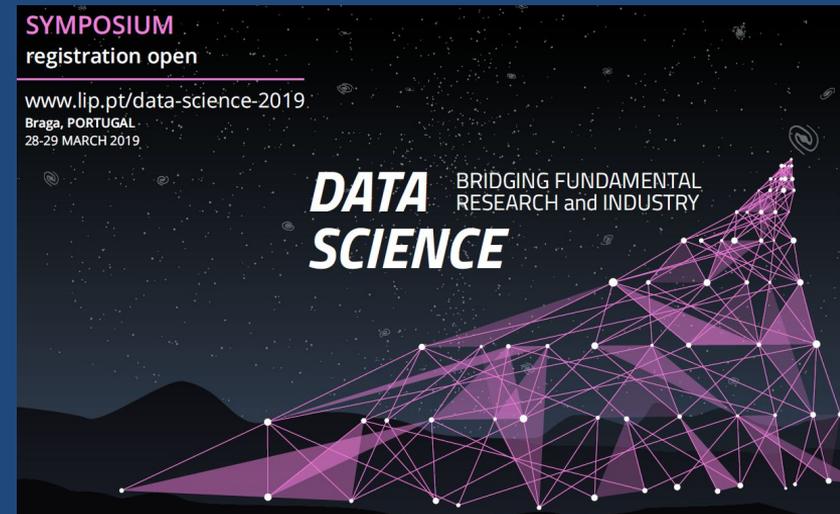
Secretariat:

Natália Antunes



<https://www.lip.pt/data-science-2019>

Universidade do Minho
Braga
25 a 29 de Março



BigDataHEP: Understanding Big Data in High Energy Physics

finding a needle in many haystacks

14:30 → 14:50 **Introdução**

Speakers: Guilherme Milhano (LIP), Nuno Castro (LIP, and University of Minho)

14:50 → 15:10 **Tarefa 1: Optimized Reuse in Repeated High-volume Data Analysis**

Speaker: António Pina (LIP Minho)

15:10 → 15:30 **Tarefa 2: Machine learning for LHC data: physics objects reconstruction and physics aware learning**

Speaker: Guilherme Milhano (LIP)

15:30 → 15:50 **Tarefa 3: Machine learning for LHC data: searching for new physics phenomena**

Speakers: Nuno Castro (LIP, and University of Minho), Tiago Vale (LIP)

15:50 → 16:20

coffee-break

16:20 → 16:40 **Tarefa 4: Machine learning for dark matter and rare event searches with the LZ detector**

Speakers: Paulo Brás (LIP e Universidade de Coimbra), Andrey Solovov (LIP e Universidade de Coimbra)

16:40 → 17:00 **Tarefa 5: Machine learning for the study of contamination in the electronic components industry**

Speakers: Barros Diogo (University of Minho), Pier Parpot (University of Minho)

17:00 → 17:20 **Outreach do projecto**

Speaker: Henrique Carvalho (LIP Minho)

17:20 → 17:50 **Discussão**