



Laboratório de Instrumentação e Física Experimental de Partículas

Competence Center - Big Data

1st Informal Meeting - 16th June 2017

ATTENDEES (via Vidyo)

Ana Peixoto, Ana Sofia Nunes, Bernardo Tomé, Bruno Galhardo, Giles Strong, Guilherme Milhano, Helmut Wolters, Henrique Carvalho, Liliana Apolinário, Luís Silva, Mário David, Nuno Castro, Ruben Conceição, Tiago Vale

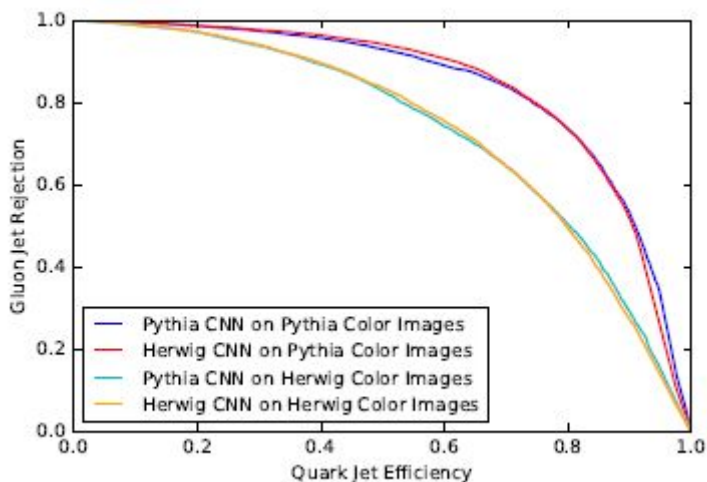
NOTES

- **Agenda:** <https://indico.lip.pt/indico/conferenceDisplay.py?confId=272>
- The goal of the Competence Center, namely the Big Data branch of it, is to provide a suitable forum to share and develop expertise on data analysis techniques (machine learning, but also efficient processing of large amounts of data).
 - **Short term plans:** foster our expertise in the field of Big Data (and Data Sciences in general); build a common background and share experience; identify common projects and services we could provide to the community (LIP and beyond)
 - **Medium term plans:** build a training program (internal and external to LIP); have dedicated teams engaged in different projects (synergies are expected)
 - **Long term plans:** aim at being a reference in this field, providing services to the academic and industrial communities in Portugal
- A key aspect is the mastering of advanced tools and the identification of the expertise and interests we have on that side. A survey will be organized. Please fill in your data [here](#).
 - Based on the information to be collected on this survey, a more concrete discussion on the needs for training and opportunities for synergies will be done at the next meeting.

- It is important to understand what other research groups are doing in this area. We want to understand where we can give a significant contribution.
- We need to develop a communication strategy, in close collaboration with the LIP ECO group, to ensure that our know-how is known and recognized by the scientific and industrial community. This is a key aspect to be able to provide services, within the scope of the Competence Center, to the community.
- Proposal to keep the meetings informal and divided in two parts:
 - a general discussion;
 - An oriented discussion on a specific topic, organized as a Journal Club, where a volunteer proposes and presents a relevant paper, project, tutorial, etc. The topic will be circulated in advance through the mailing list to allow a previous preparation of the topic by everyone interested.

JOURNAL CLUB

- **Guilherme Milhano**, *Deep learning in color: towards automated quark/gluon jet discrimination*, <https://arxiv.org/abs/1612.01551>
 - Nice pedagogical introduction to deep neural networks
 - Physics motivation: distinguish jets originating from quarks and gluons
 - Use of a convolutional NN using low level information with minimum physics input
 - Use of color to get jet images with more information
 - The method seems to be very robust: doing the training with Pythia or Herwig simulation leads to very similar results:



[plot from the discussed paper]

NEXT MEETING

14th July 2017, agenda:

<https://indico.lip.pt/indico/conferenceDisplay.py?confId=275>

N.B.: meetings will be held on a monthly basis. For the September meeting and onwards a doodle poll will be organized to define the suitable slot for our meetings.