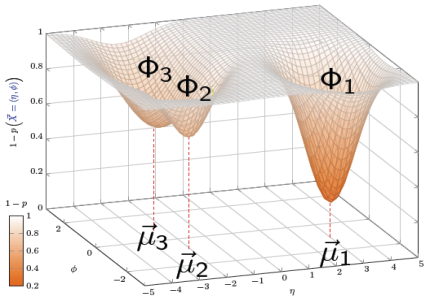
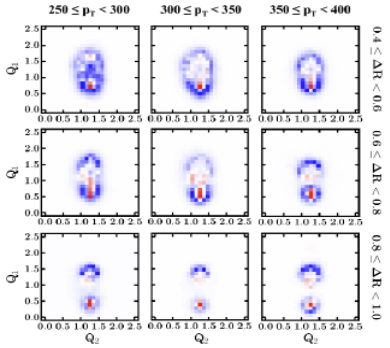
Machine Learning Physics Jamboree





Ariel Schwartzman
SLAC

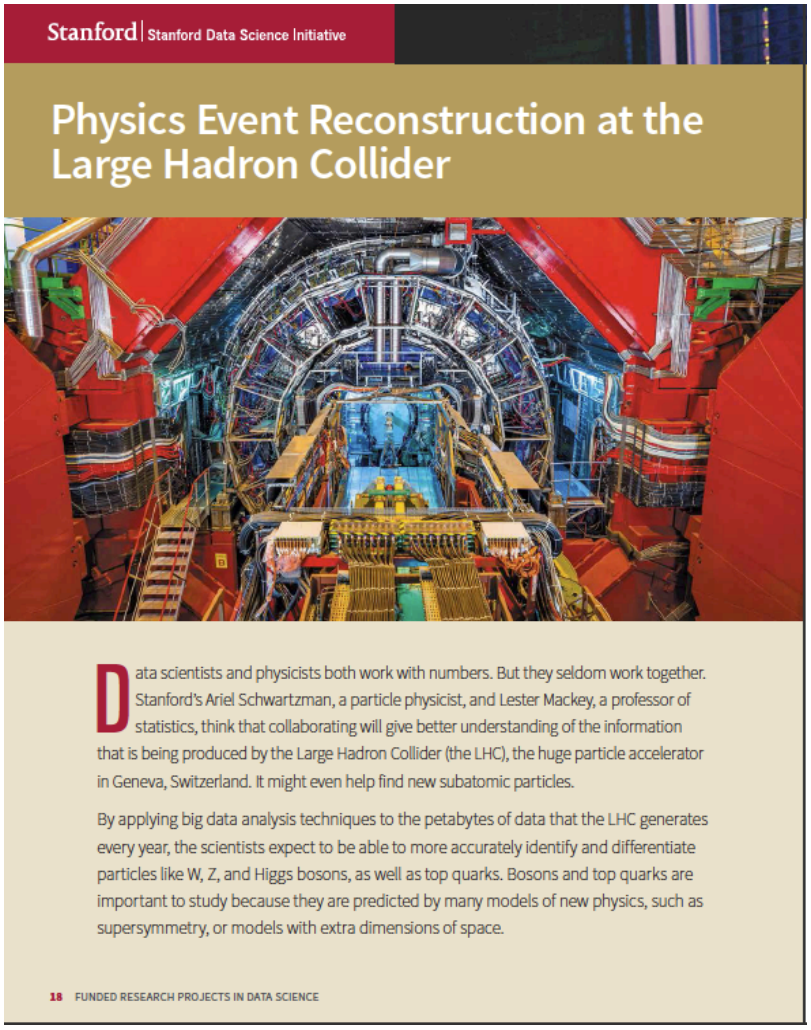
Machine learning in jet physics

- **Machine learning is widely used at hadron colliders**
  - **Use physical intuition to reduce information down to a small set of input variables into a multivariate discriminant**
    - Do not make use of all information available
    - Require the invention of variables that capture the main features of signals and backgrounds
    - No guidance when the optimal performance achieved

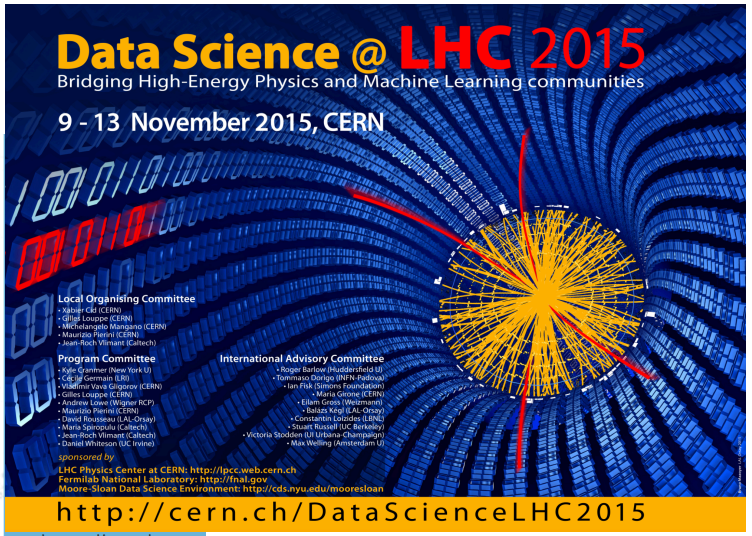
- **A new approach:**
  - Connect the fields of jets and jet substructure with machine learning, computer vision, statistics, and information theory
  - Apply advanced machine learning and pattern recognition methods to to **analyze LHC data in new ways**
  - Use machine learning to **understand the physics within jets**, to help design more powerful analysis methods

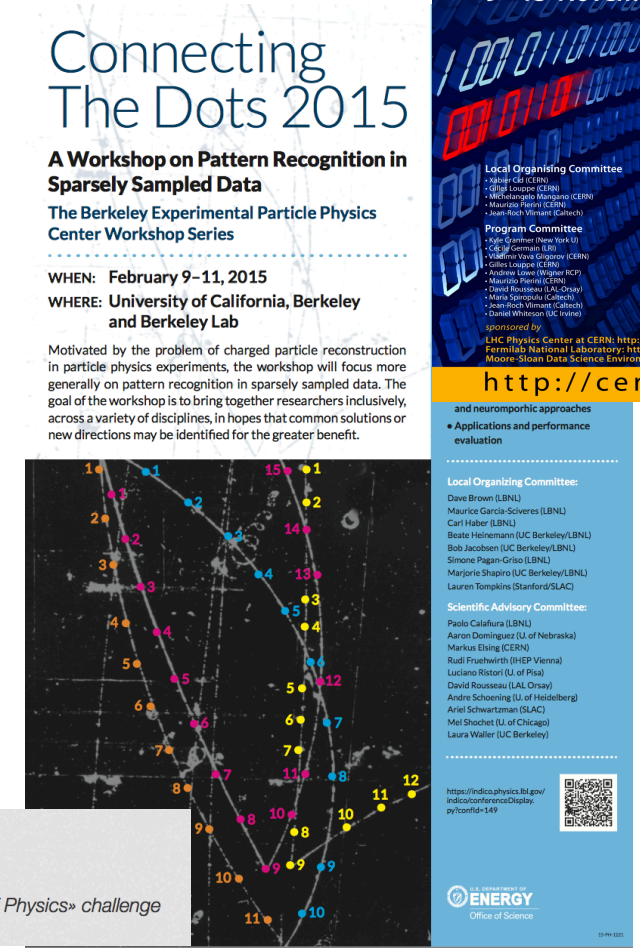
Stanford Data Science Initiative (SDSI)

https://sdsi.stanford.edu



ML in HEP





ALEPH Workshop @ NIPS 2015

Applying (machine) Learning to Experimental Physics (ALEPH) and «Flavours of Physics» challenge

SDSI LHC projects

- Three main frontiers:
  - **Learning discriminative, interpretable representations for jet tagging:**
    - Can we learn more discriminative representations of jets?
    - What do these representations teach us about physics?
    - Goal: improve classification and physical understanding with learned jet representations
    - Jet images (computer vision), deep learning, probabilistic models (fuzzy jets)
  - **Quantifying sub-optimality**
    - For a given representation of a jet and a given family of classifiers, what is the best achievable classification performance?
    - Information-theoretic lower bounds
  - **Suppressing pileup:**
    - Can we learn to identify pileup particles and remove their influence?
    - Probabilistic models to capture correlations structure among particles, regression algorithms for pileup corrections, and image processing noise reduction methods

CS229 and CS221 Projects

- **Color flow tagging**
  - Bonner Lenny, Guy Amdur, Anton Apostolatos
- **Jet charge**
  - Dylan Rueter, Krishna Soni
- **b-quark jet tagging**
  - Qi Zeng, Zihao Jiang
- **Pileup suppression**
  - Aviv Cukierman,
- **Pileup subtraction**
  - Vein Kong, Jiakun Li, Yujia Zhang

Jamboree goals

- **Informal discussions about current projects, insights gained, lessons learned, future ideas and directions**
- **Identify connections with other areas of physics**
  - Particle physics theory, neutrinos, accelerator physics, …
- **Identify possible new "grand challenges"**
- **Informal workshop: please participate!**

13:00 - 13:20 **Introduction** *20'*
Speaker: Ariel Gustavo Schwartzman (SLAC National Accelerator Laboratory (US))
MLPJ_ariel.pdf

13:30 - 13:50 **Machine learning with probabilistic models for jet physics** *20'*
Speaker: Ben Nachman (SLAC National Accelerator Laboratory (US))

14:00 - 14:20 **Machine learning with computer vision for jet physics** *20'*
Speakers: Ben Nachman (SLAC National Accelerator Laboratory (US)), Michael Aaron Kagan (SLAC National Accelerator Laboratory (US)), Luke Percival De Oliveira (SLAC National Accelerator Laboratory (US))

14:30 - 14:50 **Machine learning with recurrent neural networks for b-quark identification** *20'*
Speakers: Qi Zeng (SLAC National Accelerator Laboratory (US)), Michael Aaron Kagan (SLAC National Accelerator Laboratory (US)), Zihao Jiang (Stanford University (US))

15:00 - 15:20 **Measuring color flow with machine learning** *20'*
Speakers: Bronner Lenny (Computer Science Department, Stanford University), Guy Amdur (Computer Science Department, Stanford University), Anton Apostolatos (Computer Science Department, Stanford University)

15:30 - 16:00 Coffee break

16:00 - 16:20 **Exploring the Limits of Classification Accuracy** *20'*
Speakers: Carolyn Kim (Computer Science Department, Stanford University), Lester Mackey (Statistics Department, Stanford University)

16:30 - 16:50 **Using artificial intelligence to suppress pileup** *20'*
Speaker: Aviv Ruben Cukierman (SLAC National Accelerator Laboratory (US))

17:00 - 17:20 **Pileup subtraction using regression methods** *20'*
Speakers: Vein Kong (Electrical Engineering Department, Stanford University), Jiakun Li (Electrical Engineering Department, Stanford University), Yujia Zhang (Electrical Engineering Department, Stanford University)
pileup_presentatio...

17:30 - 17:50 **Concluding remarks** *20'*
Speakers: Ben Nachman (SLAC National Accelerator Laboratory (US)), Michael Aaron Kagan (SLAC National Accelerator Laboratory (US))