MACHINE LEARNING IN HIGH ENERGY PHYSICS

This is a snapshot of current practice based on my experience in the field over the past 20 years.

Adrian Bevan

a.j.bevan@qmul.ac.uk

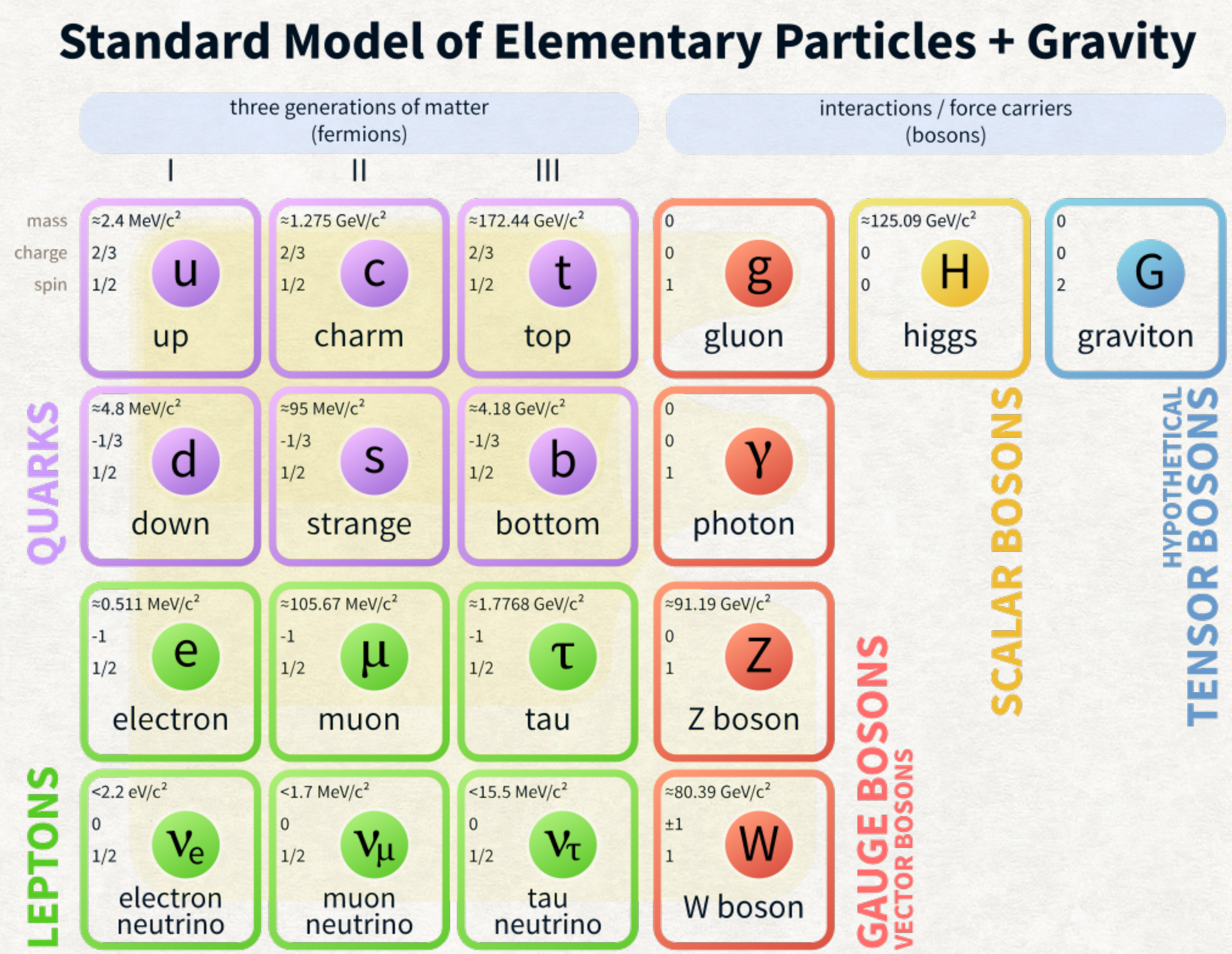
Queen Mary University of London

HORSE2017

CONTENTS

INTRODUCTION

- Physicists aspire to develop a Grand Unified Theory of everything (GUT).
    - One theory to explain the particles and all their interactions

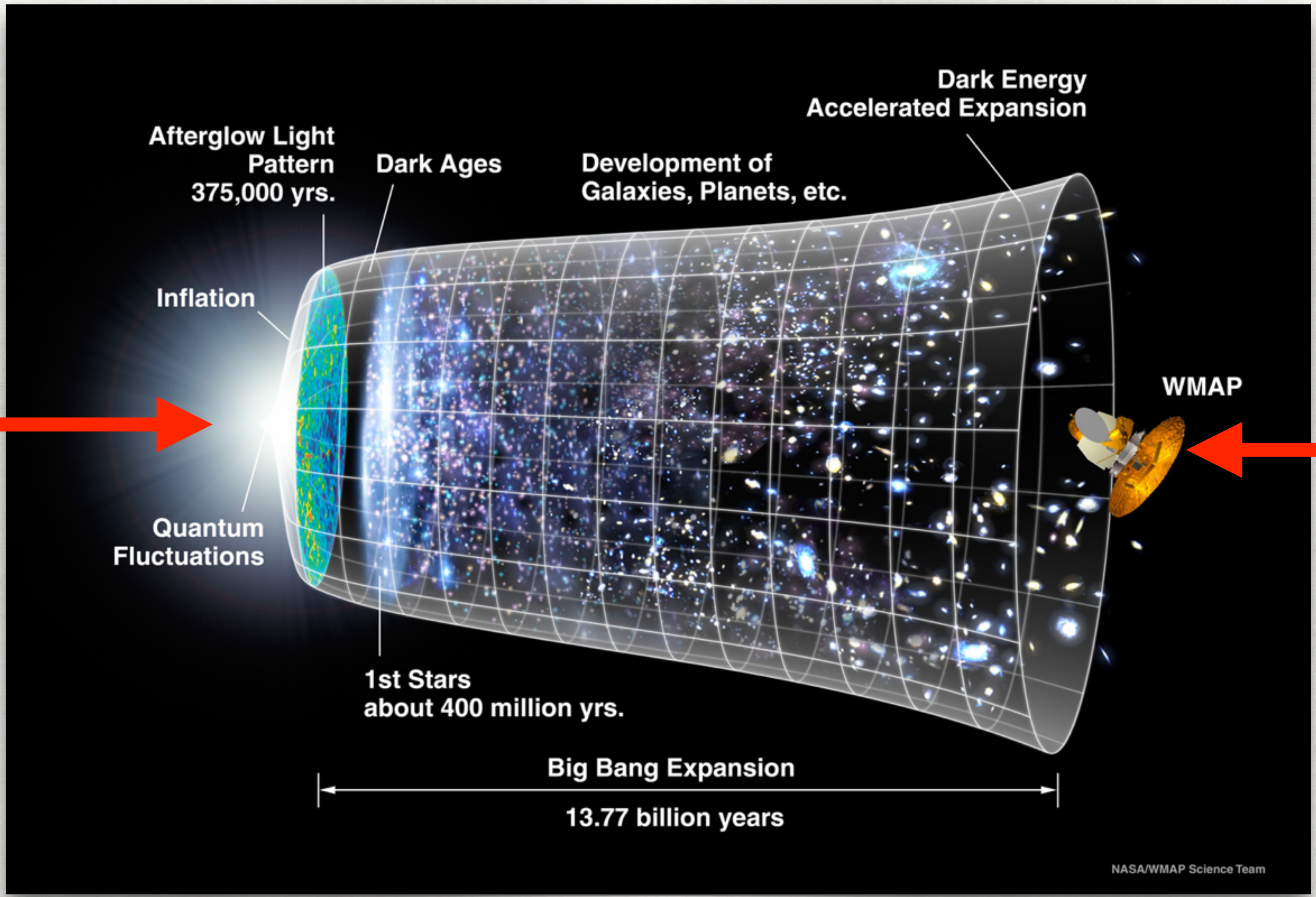


- There are many missing pieces to the puzzle.

INTRODUCTION

- Physicists aspire to develop a Grand Unified Theory of everything (GUT).
  - One theory to explain the particles and all their interactions

The Big Bang



Now

- We want to understand nature over a huge range of conditions.

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- Experiments all take the following form

SOMETHING HAPPENS → DETECT SIGNATURE (PASSIVE/ACTIVE DETECTOR) → PRE-PROCESS DATA → ANALYSE DATA

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- Experiments all take the following form

SOMETHING HAPPENS → DETECT SIGNATURE (PASSIVE/ACTIVE DETECTOR) → PRE-PROCESS DATA → ANALYSE DATA

Collisions in an accelerator

Fixed target interactions (accelerator/ cosmic rays/ neutrino)

→ Signatures found in sub-detectors → Throw away stuff known to be junk → Try and understand what is left

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- Experiments all take the following form

SOMETHING HAPPENS → DETECT SIGNATURE (PASSIVE/ACTIVE DETECTOR) → PRE-PROCESS DATA → ANALYSE DATA

Collisions in an accelerator

Fixed target interactions (accelerator/ cosmic rays/ neutrino)

→ Signatures found in sub-detectors → Throw away stuff known to be junk → Try and understand what is left

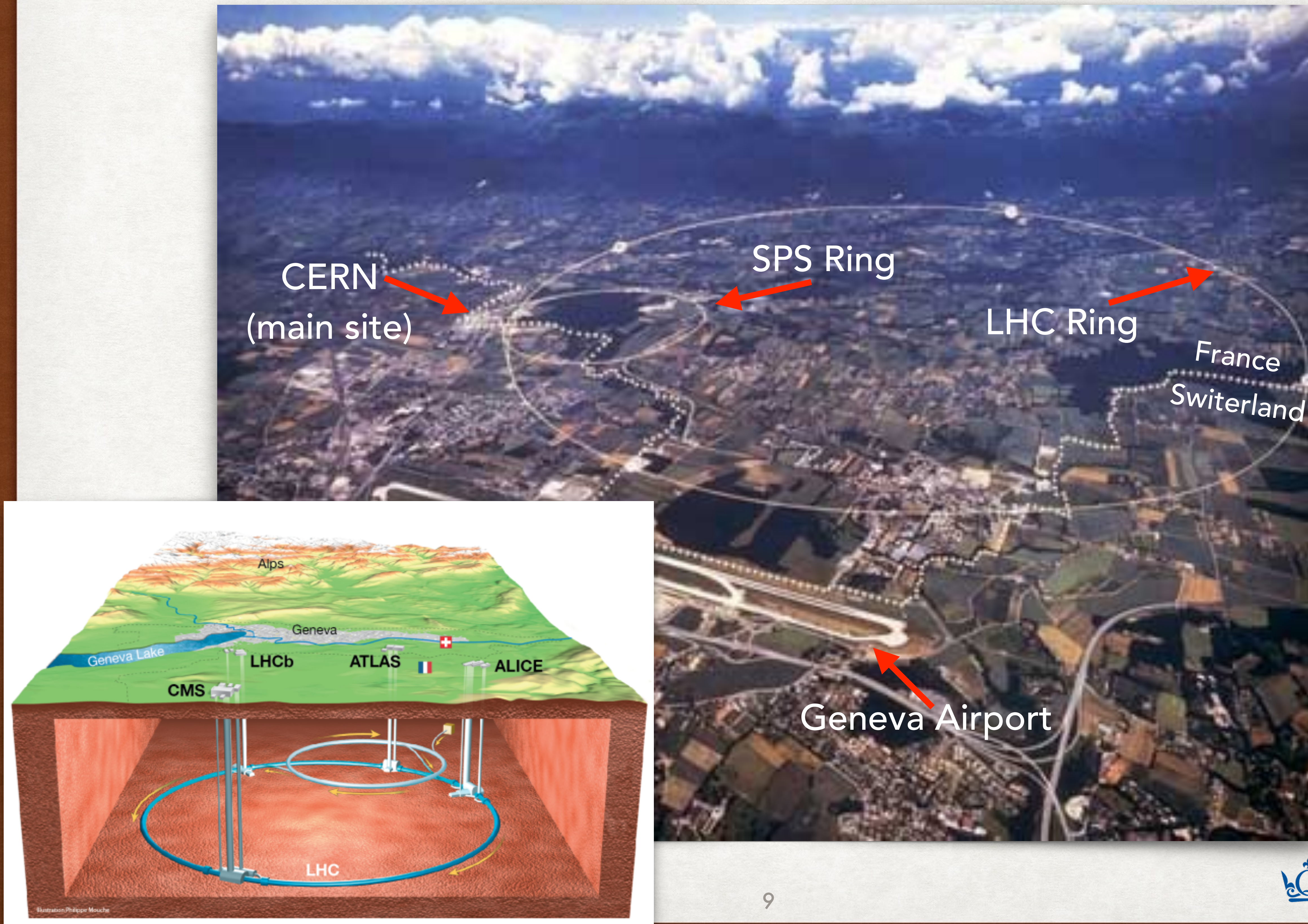
Some of us use terms like feature space and examples … but usual jargon is input variable and event; occasionally a little translation is required.

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- e.g. The Large Hadron Collider at CERN (Geneva, Switzerland)

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- e.g. The Large Hadron Collider at CERN (Geneva, Switzerland)



Queen Mary
University of London

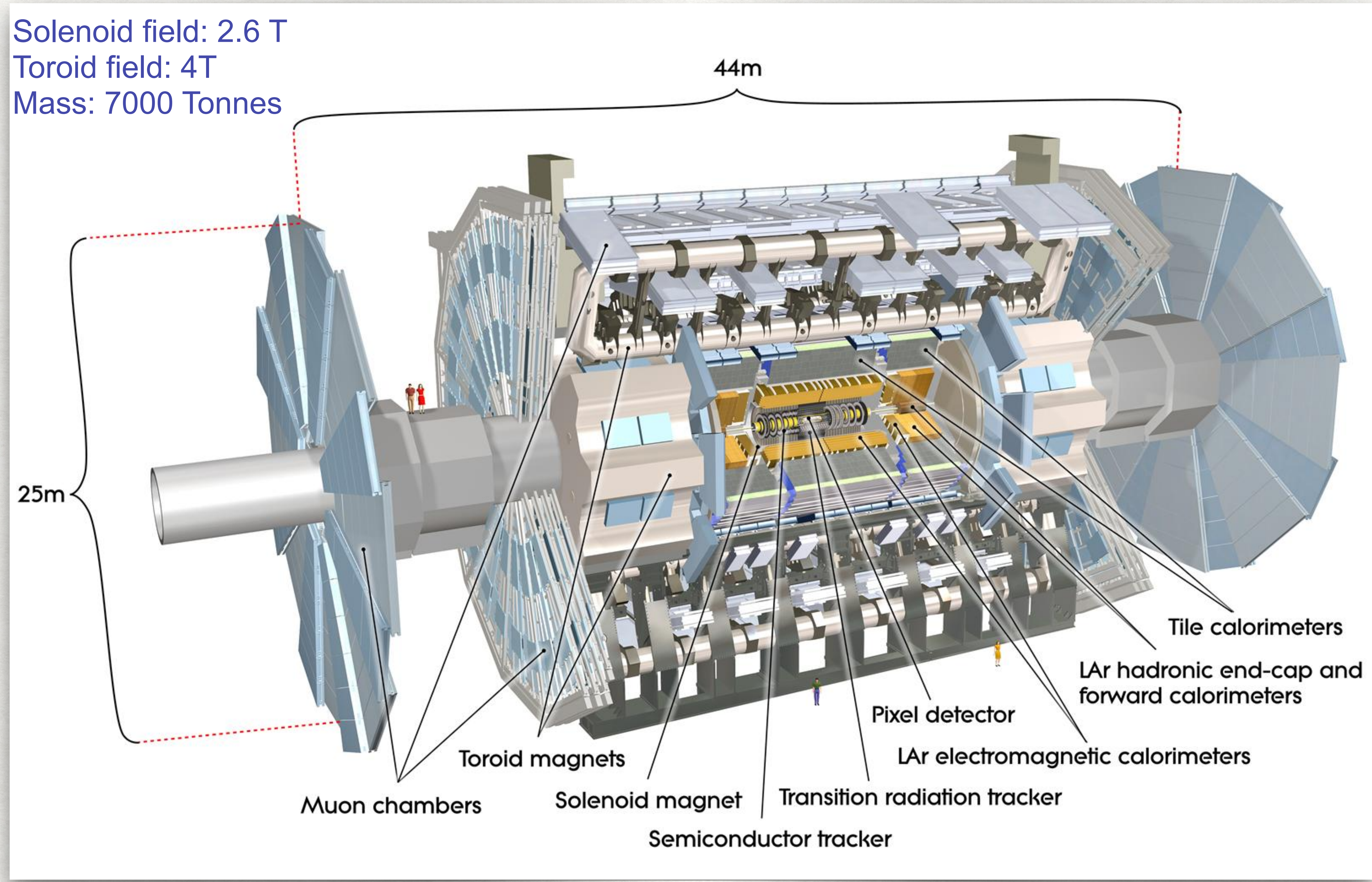
TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- e.g. The Large Hadron Collider at CERN (Geneva, Switzerland)

Queen Mary
University of London

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

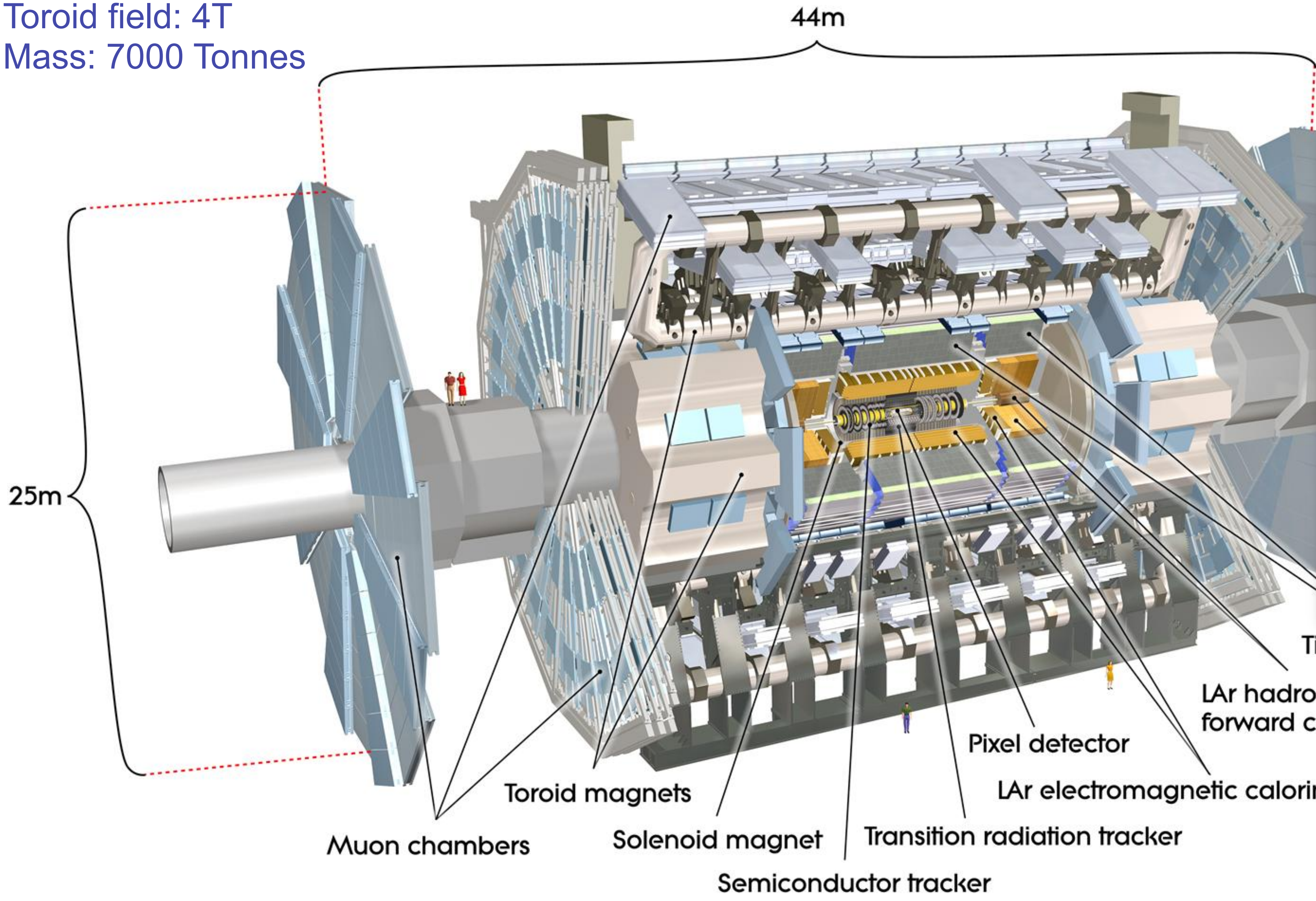
- e.g. The ATLAS experiment at CERN: basically a very large camera



TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- e.g. The ATLAS experiment at CERN: basically a very large camera

Solenoid field: 2.6 T
Toroid field: 4T
Mass: 7000 Tonnes



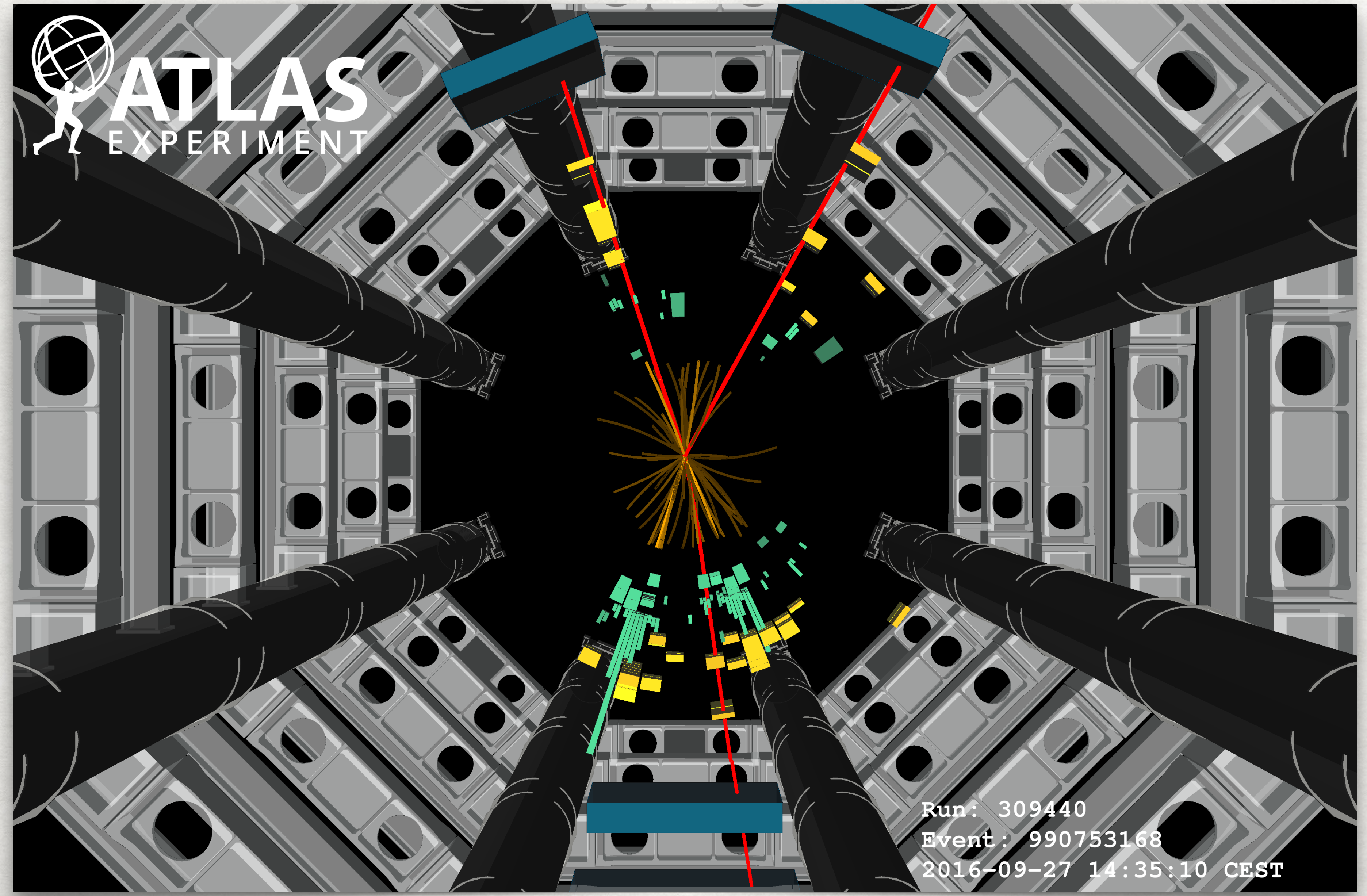
G. O. Jones building (Physics) for comparison.

TYPICAL PROBLEMS: DETECTOR PARTICLE SEARCH

- Look in a 25ns time window:
  - Hundreds of thousands of electronic readout channels to digitize.
    - Amplify, shape, discriminate thresholds to leave just the important channels that have electric charge deposited.
  - Turn that vast set of raw data into kinematic quantities that describe particles and objects (like jets = collections of charged particles).
  - Figure out if it is interesting or not.
  - If interesting - store the event for offline analysis (pre-processing and analysis).
- Repeat

WHAT DOES A 3D PICTURE LOOK LIKE?

- This is a candidate event for a Higgs boson decay into two b-quarks. It could equally be background (two top quarks decaying).
- Impossible for us to say for sure!



Analysis method uses the following workflow:

1. Pre-selection
2. Final selection
3. BDT *(usually)*
4. Likelihood fit

WHERE DOES ML FIT IN?

- Feature spaces are often low dimensionality
  - Typically 10-20 dimensions are used to discriminate between signal and background event signatures.. can have O(600) features.

- People nowadays normally use:
  - Boosted Decision Trees.

- Sometimes use Multilayer perceptrons (older people tend to use these).

- Recent years - exploring use of / starting to use
  - Support Vector Machines.
  - Deep Learning (Deep MLP's, CNNs, RNNs).

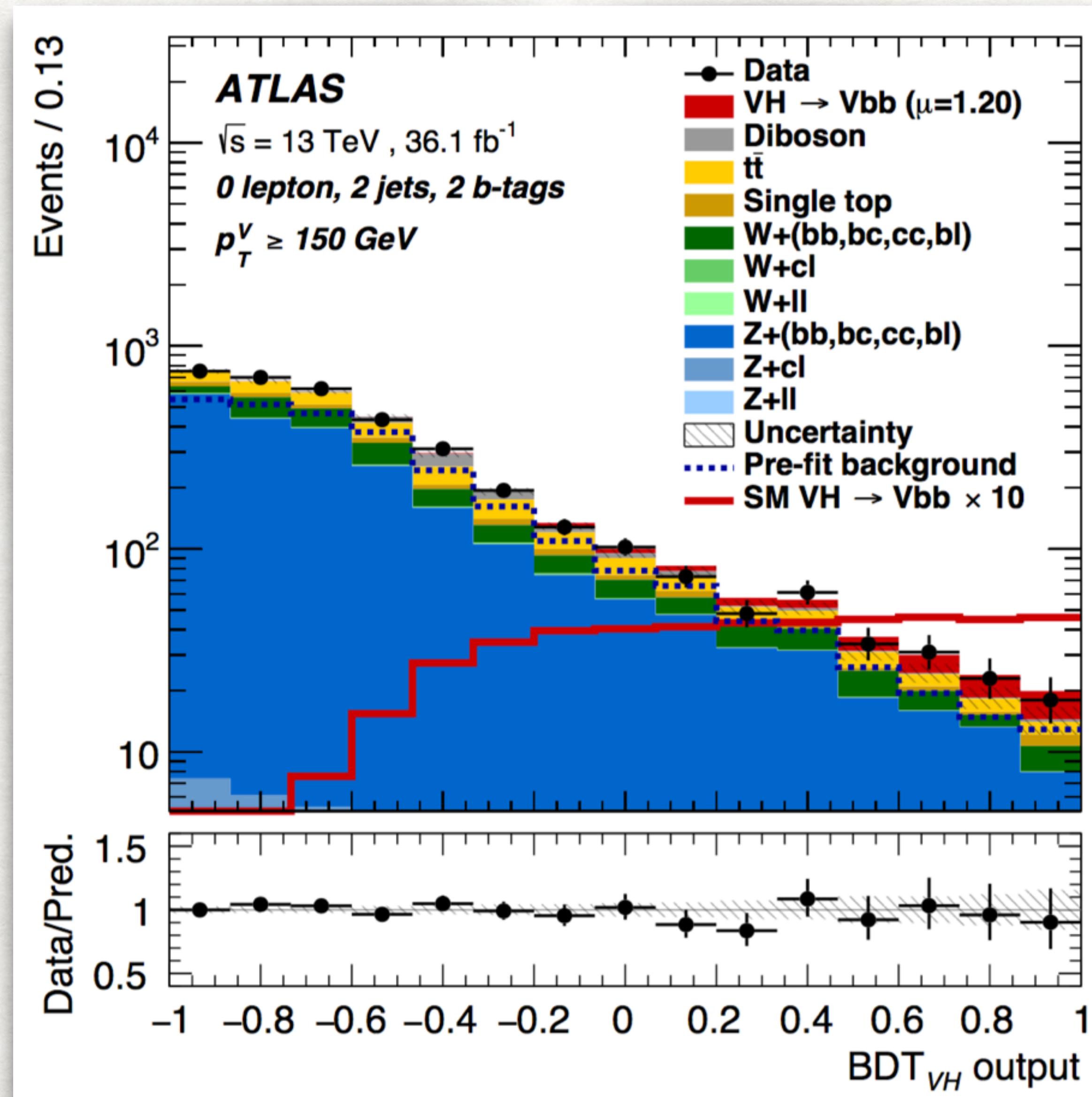
https://arxiv.org/pdf/1708.03299.pdf

HIGGS DECAYING INTO TWO B QUARKS

- O(10) dimensional feature space
- Different BDT configurations for different physical signatures

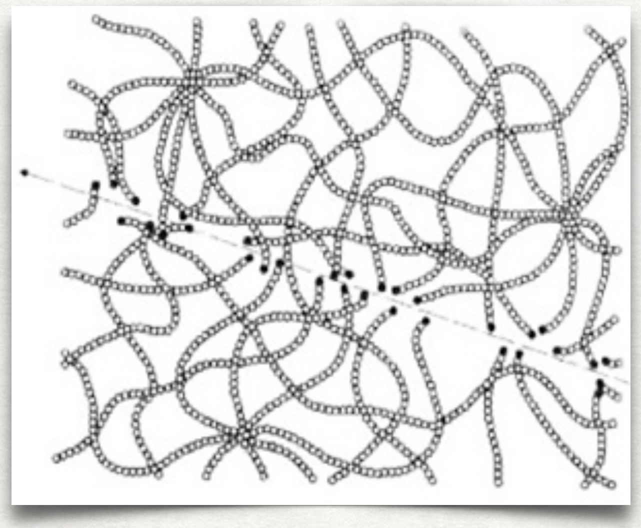
Variable	0-lepton	1-lepton	2-lepton
p_T^V	$\equiv E_T^{miss}$	$\times$	$\times$
E_T^{miss}	$\times$	$\times$	$\times$
$p_T^{b_1}$	$\times$	$\times$	$\times$
$p_T^{b_2}$	$\times$	$\times$	$\times$
m_{bb}	$\times$	$\times$	$\times$
$\Delta R(\boldsymbol{b}_1, \boldsymbol{b}_2)$	$\times$	$\times$	$\times$
$\lvert\Delta\eta(\boldsymbol{b}_1, \boldsymbol{b}_2)\rvert$	$\times$		
$\Delta\phi(\boldsymbol{V}, \boldsymbol{bb})$	$\times$	$\times$	$\times$
$\lvert\Delta\eta(\boldsymbol{V}, \boldsymbol{bb})\rvert$			$\times$
m_{eff}	$\times$		
$\min[\Delta\phi(\boldsymbol{\ell}, \boldsymbol{b})]$		$\times$	
m_T^W		$\times$	
$m_{\ell\ell}$			$\times$
m_{top}		$\times$	
$\lvert\Delta Y(\boldsymbol{V}, \boldsymbol{bb})\rvert$		$\times$	
	Only in 3-jet events		
$p_T^{\mathrm{jet}_3}$	$\times$	$\times$	$\times$
m_{bbj}	$\times$	$\times$	$\times$

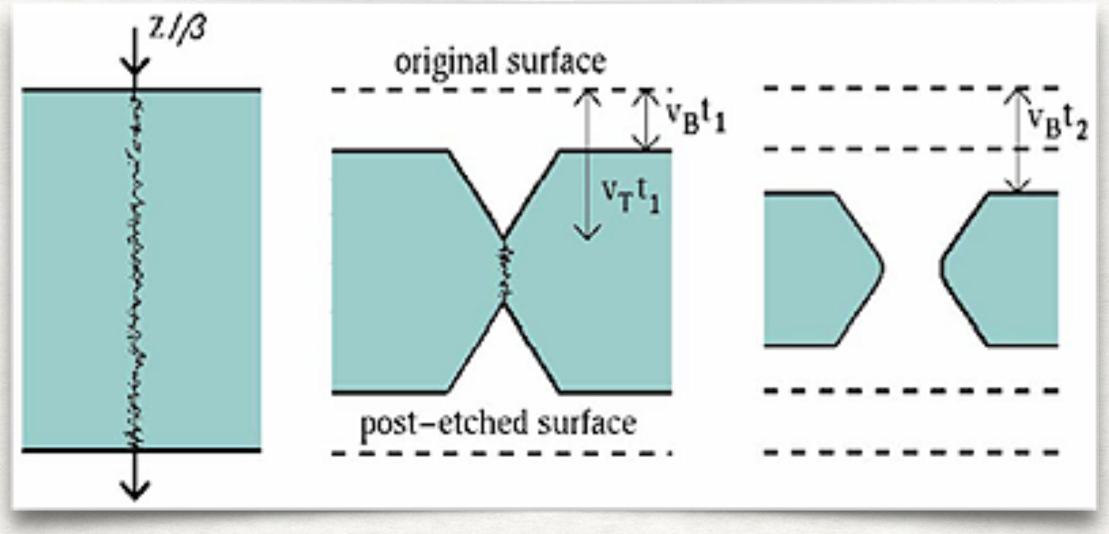
$$Z \to \nu\bar{\nu}, H \to b\bar{b}$$

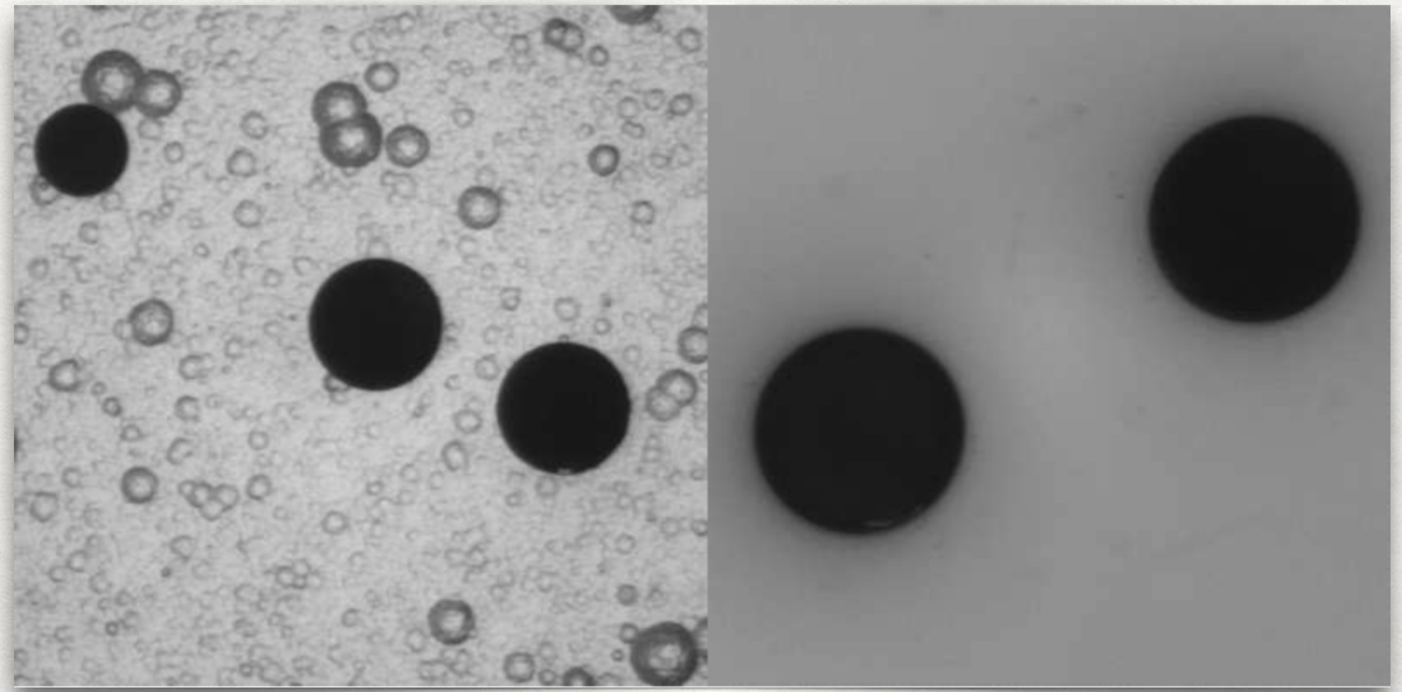


TYPICAL PROBLEMS: IMAGE SEARCH

- e.g. MoEDAL:
  - Search for magnetic monopoles in nuclear track emulsion (think of looking for a dot in a 35mm photographic film negative).







Holes from Pb ions traversing the plastic detecting elements.

Stacks of plastic allow us to search for heavily ionising particles that, if found, would be new and could explain some deep concepts in modern physics.

TYPICAL PROBLEMS: IMAGE SEARCH

- Have been using Citizen Science to search for regions of interest (ROIs).
- Also using traditional Vision Science techniques to identify ROIs.
- Some investigation has been done using KNN and SVM methods to identify ROIs.
- At QMUL we are using CNNs for this problem; training examples are an issue and we have been "simulating" images by oversampling a limited number of reference samples.

TOOLS

- Before circa 2015:
  - linear discriminants
  - Multilayer Perceptrons (not deep)
  - Boosted Decision Trees
  - (rarely) Support Vector Machines
  - Generally home grown or libsvm solutions.

★ 2014 Kaggle Higgs data challenge. This had a huge impact on the HEP community.

- Since 2015:
  - Adoping a broader base of possible tools (interfaced to the HEP analysis framework ROOT).
  - TensorFlow, keras, caffe, SciKit Learn, R, …
    - We use or have dabbled with some of these toolkits.
    - Developing a TensorFlow-based level 4 module for next summer.

GENERALISATION

- Typical solution: use the hold out method: split data in two, test and validate.
  - Comparison is based on similarity metrics between two histograms output(*).
- We've introduced k-fold cross validation to the community toolkit.
  - Not widely adopted (yet).
- Deep learning practitioners generally use drop out (some have used maxout) to promote hyper-parameter generalisation and mitigate the risk of overtraining.

(*)Some caveats are required here as some metrics we use are not really valid; and effective we do a $\mathbf{X}^2$ by eye to see if things look reasonable.

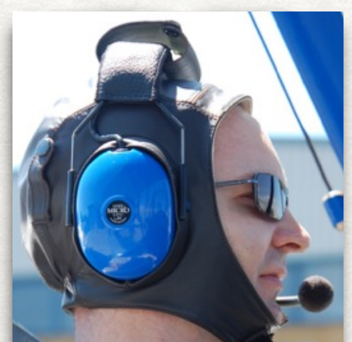
HOW MUCH DO YOU NEED TO UNDERSTAND?

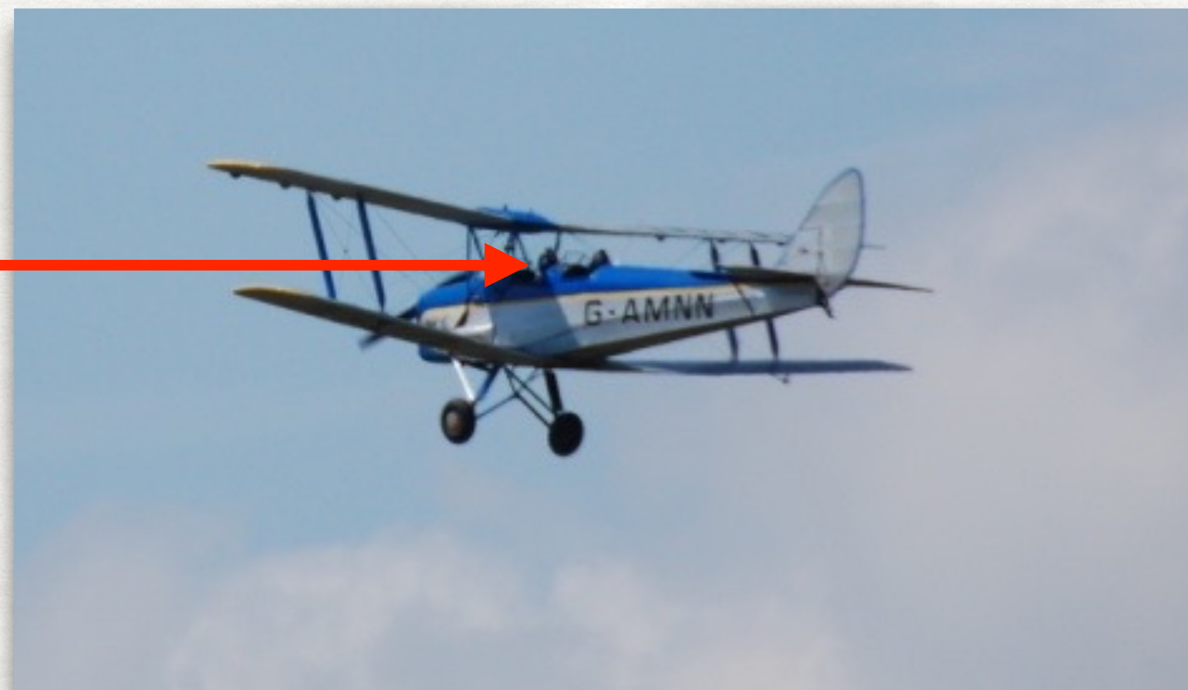
- ML algorithms and AI are scientific subjects in their own right.
  - These necessitate careful understanding of the techniques.
  - Once a plane is in the air it is easy to fly… (just don't hit the ground)

HOW MUCH DO YOU NEED TO UNDERSTAND?

- ML algorithms and AI are scientific subjects in their own right.
  - These necessitate careful understanding of the techniques.
  - Once a plane is in the air it is easy to fly… (just don't hit the ground)
  - Having flown it would you trust yourself to land safely?

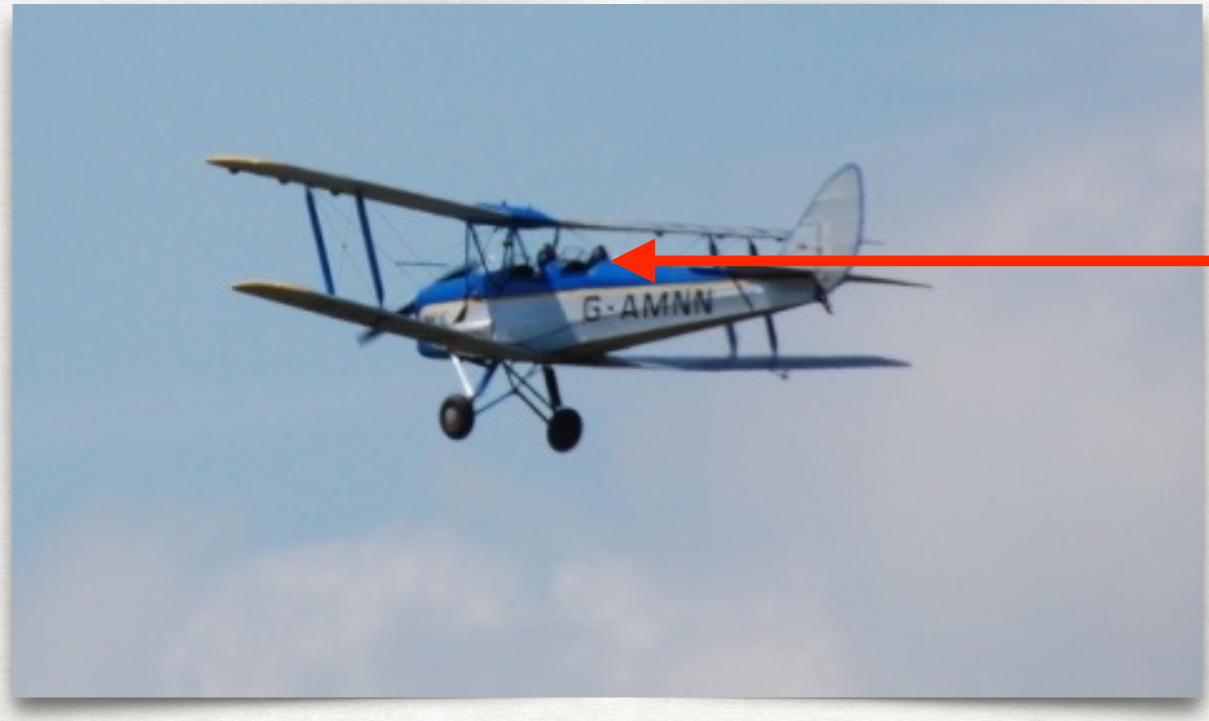
Me





HOW MUCH DO YOU NEED TO UNDERSTAND?

- ML algorithms and AI are scientific subjects in their own right.
  - These necessitate careful understanding of the techniques.
  - Once a plane is in the air it is easy to fly… (don't hit the ground)
  - Having flown it would you trust yourself to land safely?



The expert

- It takes much more skill to land a plane safely than it does to fly it once it is in the area.
- Likewise — It takes much more skill to use an ML algorithm properly than to train it using a black box of tools.

LIKE NEO, WE ALL HAVE A CHOICE

Take the blue pill: Use the black box and never understand the true potential of it.

Take the red pill: Jump down the rabbit hole; open your eyes to the algorithm and realise the potential of machine learning.

Queen Mary University of London

SOME RECENT PAPERS

- A selection of physics results:
  - Evidence for the $H \rightarrow b\bar{b}$ decay with the ATLAS detector: https://arxiv.org/pdf/1708.03299.pdf
  - Physics of the B Factories (Chapter 4; also see Ch 5, 8 and index): https://arxiv.org/abs/1406.6311
  - A Hierarchical NeuroBayes-based Algorithm for Full Reconstruction of B Mesons at B Factories: https://arxiv.org/abs/1102.3876
  - A Study of B0 to rho+rho- Decays and Constraints on the CKM Angle alpha: https://arxiv.org/abs/0705.2157

- Machine learning experts having a go at particle physics:
  - Enhanced Higgs to $\tau^+\tau^-$ Searches with Deep Learning: https://arxiv.org/abs/1410.3469
  - Gabor Melis (winner of the Kaggle Higgs Challenge): https://www.kaggle.com/melisgl

- Reconstructing events/objects:
  - Jet-parton assignment using deep learning: https://arxiv.org/abs/1706.01117
  - Event classification: https://arxiv.org/pdf/1708.07034.pdf
  - Machine Learning on CMS: https://indico.cern.ch/event/665947/
  - Pileup rejection with a CNN: https://arxiv.org/abs/1707.08600

- Community pages/meetings:
  - Inter Experiment Machine Learning LHC working group: http://iml.web.cern.ch. This group holds regular meetings about topics. They welcome input from ML experts as well as talks about experiences and developments from experiments. The initial remit was LHC because of the people involved in starting this up, but that remit has grown.
  - Machine Learning in Physics: https://indico.ph.qmul.ac.uk/indico/conferenceDisplay.py?confId=130
  - Computing in High Energy Physics: CHEP 2016: http://chep2016.org

CONCLUSION

- Particle physics relies on modern machine learning in many areas of operation.
- Variety of techniques are used; including deep learning methods.
- Unique data sets that may be of broader interest:
  - e.g. A Higgs signal event is topologically very similar to a background event.
  - Training data is generally more limited than we would like.
  - Computing resources are also limited.
- Our problems could have dual issues in the real world.
  - Where there are synergies we would like to learn from CS experts and/or collaborate on solving problems.