

5MIN RESPONSE TO  
DEEP LEARNING APPLICATIONS IN THE NATURAL SCIENCES  
BY P. BALDI

**BALÁZS KÉGL**  
CNRS & University Paris Saclay

# TWO POINTS

- Automated scientist? We need to fix the epistemology.
- Deep learning in HEP: the rawer the better.

# AUTOMATED SCIENTIST

# AUTOMATED SCIENTIST

- Yes, the scientific method can/will be automated
  - come up with model, build experiments/detectors, observe data, reject model, iterate

# AUTOMATED SCIENTIST

- Yes, the scientific method can/will be automated
  - come up with model, build experiments/detectors, observe data, reject model, iterate
- The epistemology of **automated hypothesis generation** is a mess: the **“come up with a model”** part.
  - for a lot of people, this is the big data revolution: we will find stuff that we could/would have not thought of

# DEEP LEARNING IN HEP

- There is no real success story yet
  - deep learning was important to win the HiggsML challenge, but the **improvement was marginal**
  - simulated data sets are **too small**
  - digested/engineered **features are too uncorrelated**
  - the **systematics** is killing you

# DEEP LEARNING IN HEP

- Go back to **raw data**
  - tracker, pixel calorimeter
- Work on the **convolutional layers**
  - type of input, detector geometry
- Figure out how to include the **systematics** into the training
  - see Kyle's stuff