Machine Intelligence

inter-disciplinary introduction

Advanced Computing and Machine Learning Course
25-27 May 2020, IGFAE

Andrey Ustyuzhanin [1,2]

[1] NRU HSE

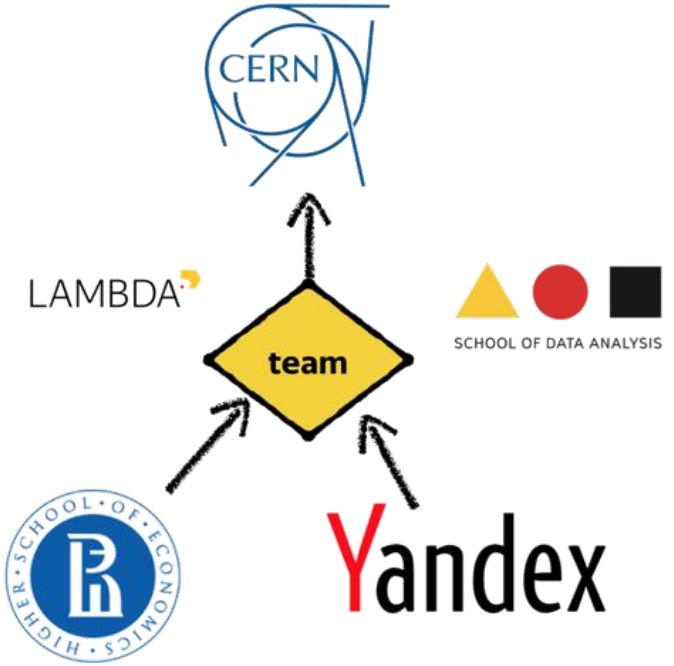
[2] NUST MISiS

LAMBDA • HSE

May 25, 2020

A few words about your lecturer

- Head of Laboratory of methods for Big Data Analysis at Higher School of Economics (HSE),
  - Applications of Machine Learning to natural science challenges
  - HSE has joined LHCb in 2018
- Head of LHCb Yandex School of Data Analysis (YSDA) team
- Co-organizer of Flavours of Physics Kaggle competitions, 2015
- Co-organizer of TrackML challenge (2018)
- Education activities (5 MLHEP schools, ML at ICL, Clermont-Ferrand, URL Barcelona, Coursera)



Outlook

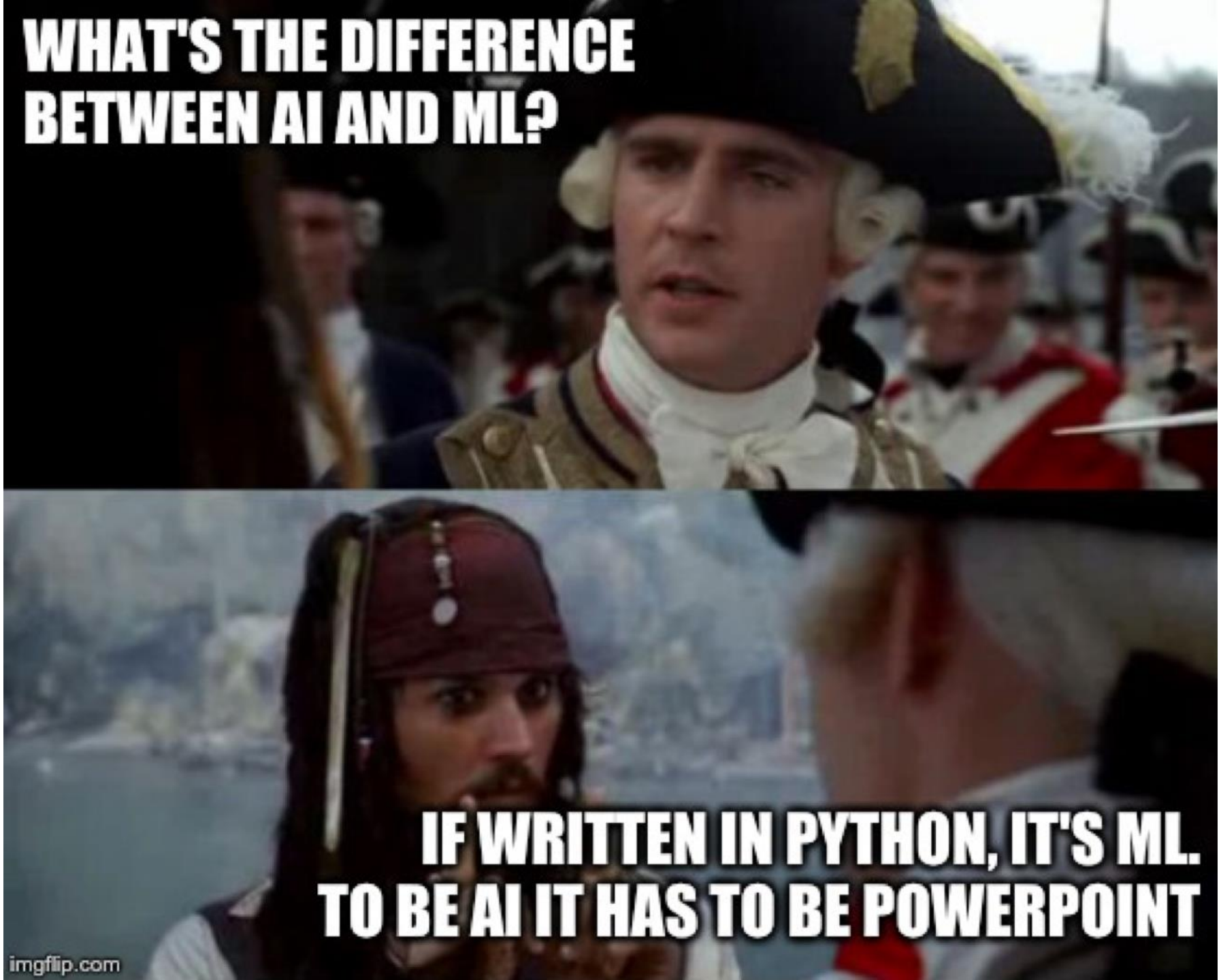
- **Introduction**
- Magic of Machine Intelligence
- For natural sciences
- Linear models
- Figures of merit
- Overfitting: how to fool a linear regression
- Regularization
- Conclusion

Machine Intelligence Magic

Remarkable examples of AI technologies progress

- Human-level playing in computer games (Go, StarCraft, Dota 2) and winning world's Go champion Lee Sedol by Google AlphaGo;
- Understanding and generation of human-readable and understandable texts;
- Recognition and generation of images indistinguishable from photos by the naked eye;
- Simulation of complicated physics processes;
- Controlling complicated real-time systems like quantum qubits;
- Controlling autonomous vehicles in populated regions;
- and many others.

"Any sufficiently advanced technology is indistinguishable from magic."
Arthur Clarke


WHAT'S THE DIFFERENCE
BETWEEN AI AND ML?
IF WRITTEN IN PYTHON, IT'S ML.
TO BE AI IT HAS TO BE POWERPOINT
imgflip.com

Human-level text understanding and generation

- Human-written input:

– "A train carriage containing controlled nuclear materials was stolen in Cincinnati today. Its whereabouts are unknown."

- GPT2 models by OpenAI

- "The incident occurred on the downtown train line, which runs from Covington and Ashland stations."
- "In an email to Ohio news outlets, the U.S. Department of Energy said it is working with the Federal Railroad Administration to find the thief."
- "The theft of this nuclear material will have significant negative consequences on public and environmental health, our workforce and the economy of our nation," said Tom Hicks, the U.S. Energy Secretary, in a statement. "Our top priority is to secure the theft and ensure it doesn't happen again."
- "The stolen material was taken from the University of Cincinnati's Research Triangle Park nuclear research site, according to a news release from Department officials." …

https://openai.com/blog/better-language-models/#sample2

Leaderboard Version: **2.0**

Rank	Name	Model	URL	Score	BoolQ	CB	COPA	MultiRC	ReCoRD	RTE	WiC	WSC	AX-b	AX-g
1	SuperGLUE Human Baselines	SuperGLUE Human Baselines		89.8	89.0	95.8/98.9	100.0	81.8/51.9	91.7/91.3	93.6	80.0	100.0	76.6	99.3/99.7
2	T5 Team - Google	T5		89.3	91.2	93.9/96.8	94.8	88.1/63.3	94.1/93.4	92.5	76.9	93.8	65.6	92.7/91.9
3	Zhuiyi Technology	RoBERTa-mtl-adv		85.7	87.1	92.4/95.6	91.2	85.1/54.3	91.7/91.3	88.1	72.1	91.8	58.5	91.0/78.1
4	Facebook AI	RoBERTa		84.6	87.1	90.5/95.2	90.6	84.4/52.5	90.6/90.0	88.2	69.9	89.0	57.9	91.0/78.1
5	Huawei Noah's Ark Lab	NEZHA-Large		80.8	82.3	89.7/93.6	91.2	74.2/32.1	77.1/76.4	88.5	72.7	90.4	48.1	86.5/78.7
6	IBM Research AI	BERT-mtl		73.5	84.8	89.6/94.0	73.8	73.2/30.5	74.6/74.0	84.1	66.2	61.0	29.6	97.8/57.3

Comprehension of written text. SuperGLUE metric

City street view simulation

This time lap shows the original scene (left), segmentation map (bottom right) and neural-network produced scene (right) by NVIDIA.

https://www.youtube.com/watch?v=ayPqjPekn7g

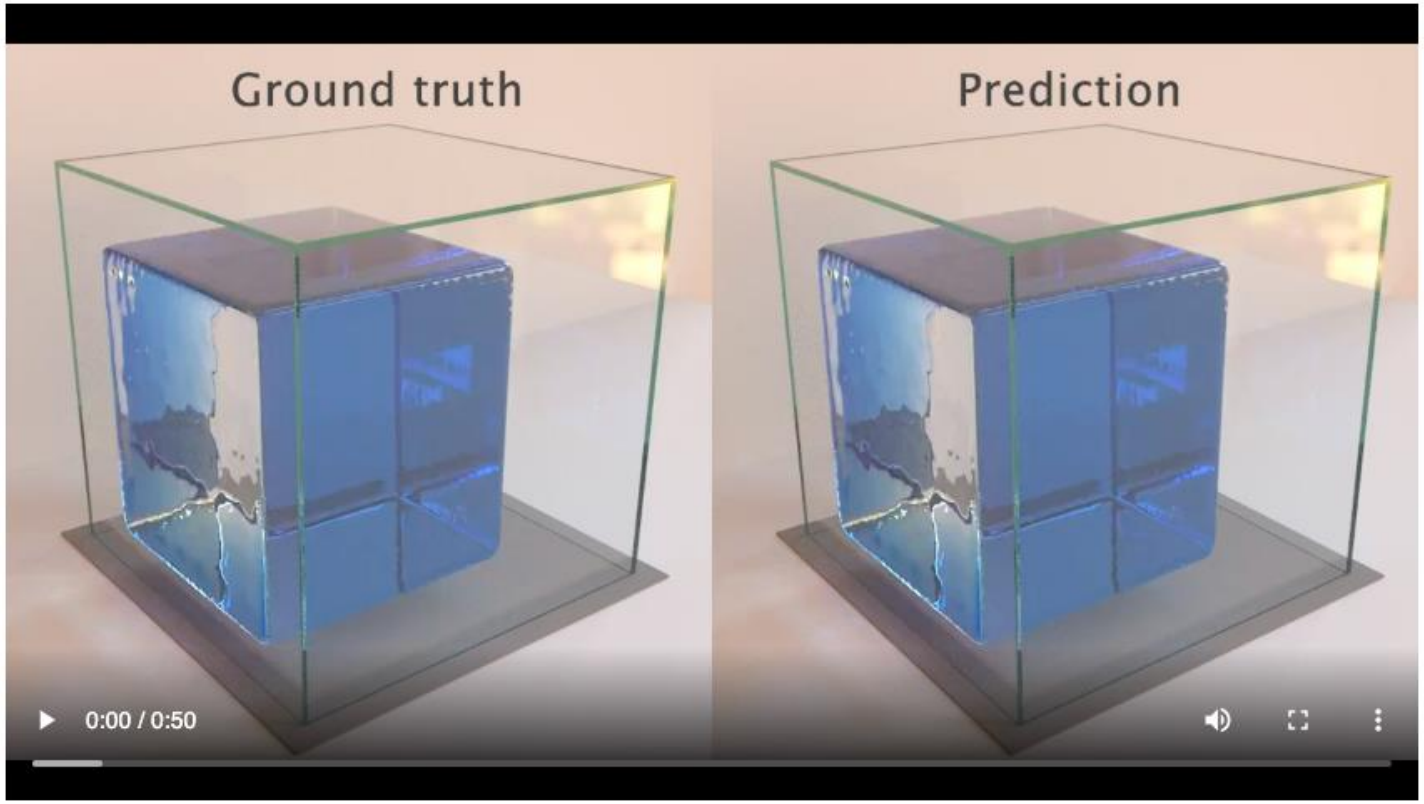
https://arxiv.org/abs/1808.06601

Quantum supremacy

- experimental confirmation that computations on quantum processors could be practically performed exponentially faster than on regular supercomputers
  https://www.nature.com/articles/s41586-019-1666-5
- keep 53 qubits strictly in sync during computation is not an easy task
- How to control qubits with Reinforcement Learning:
  https://www.nature.com/articles/s41534-019-0141-3

Fluid dynamics computation

- Paper by Alvaro Sanchez-Gonzalez et al. from Google (link)



- The time lap shows properly simulated water volume evolution (left) and simulated evolution by the trained neural network (right) with "Graph Network-based Simulators" .

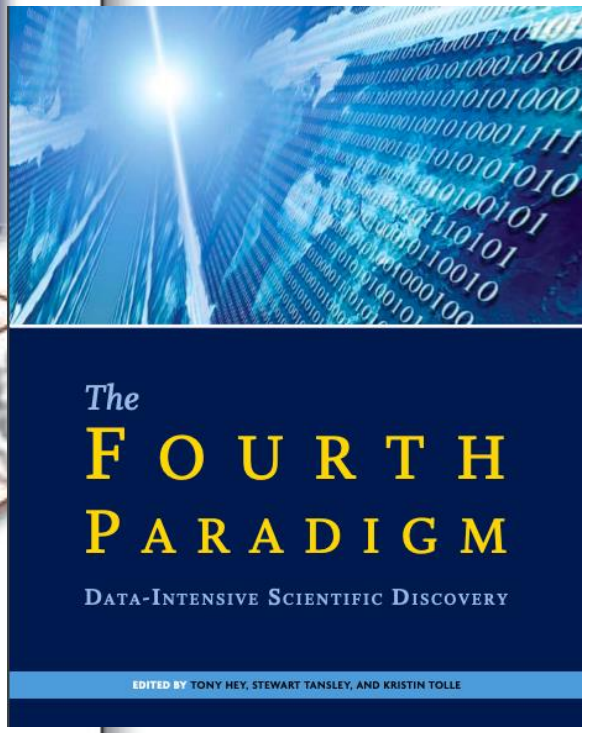
Abridged History of science

Jim Gray's vision of science, 2009

Science Paradigms

- Thousand years ago:
  science was **empirical**
  describing natural phenomena
- Last few hundred years:
  theoretical branch
  using models, generalizations
- Last few decades:
  a **computational** branch
  simulating complex phenomena
- Today: **data exploration** (eScience)
  unify theory, experiment, and simulation
  - Data captured by instruments or generated by simulator
  - Processed by software
  - Information/knowledge stored in computer
  - Scientist analyzes database/files using data management and statistics

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G p}{3} - K\frac{c^2}{a^2}$$



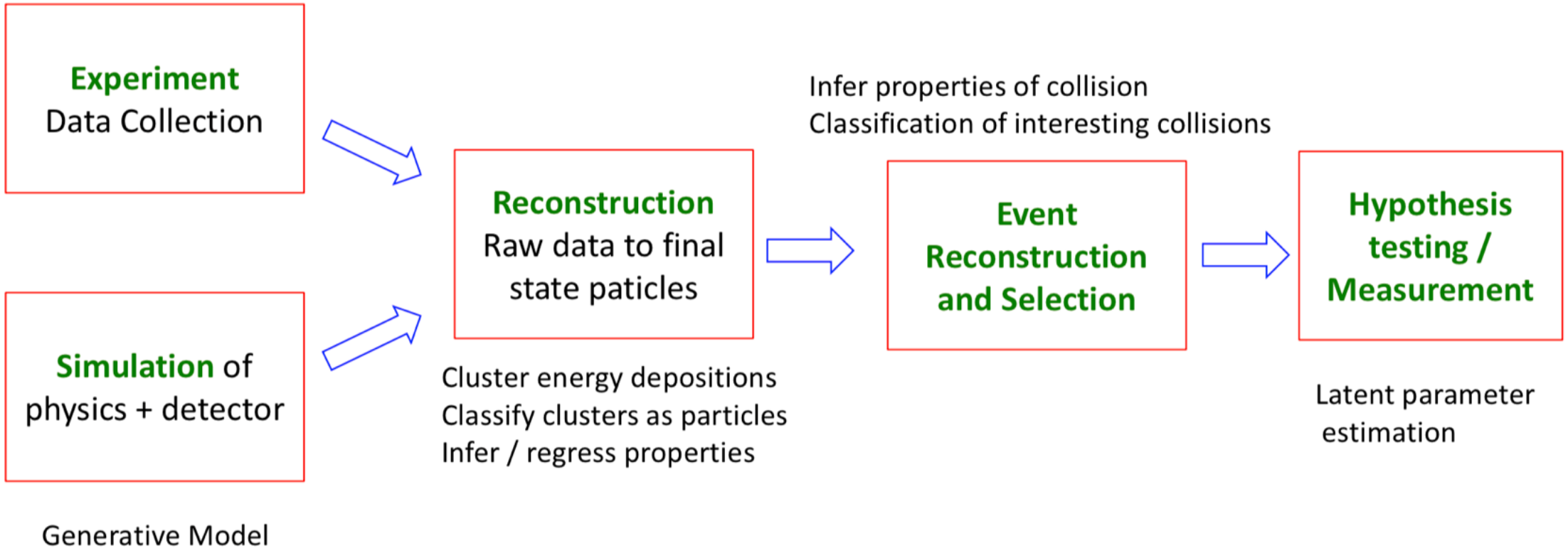
link

Physics-inspired approaches in ML

- Simulated Annealing
- MCMC techniques
- Gibbs sampling
- Gaussian process
- Gradient descent
- Boltzmann Machine
- Energy-based GANs

Arxiv:1903.10563

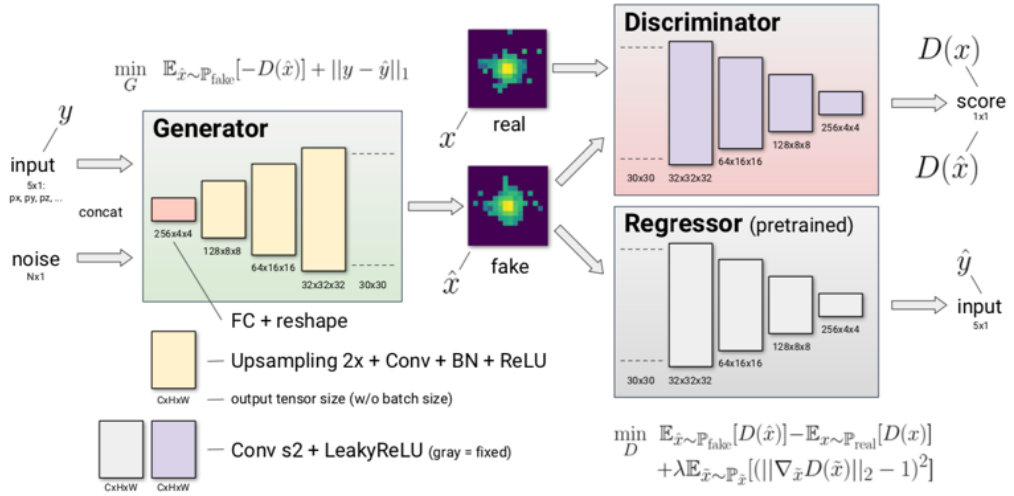
Important ML application areas

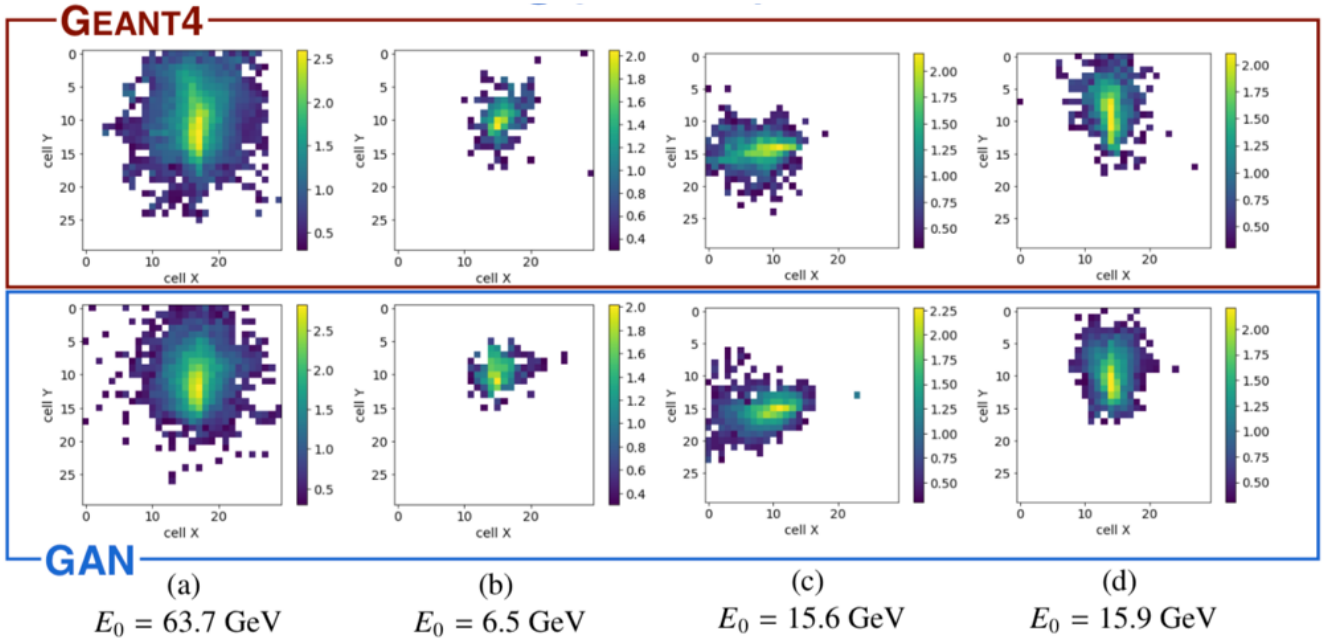


http://bit.ly/2FHWTZ4

Example: Fast Calorimetry Simulation

- LHCb-like calorimeter 30x30
- 5 conditional parameters per particle (3D momentum, 2D coordinate)
- Electrons from particle gun shot at 1x1 cm square at the center of the calorimeter face
- Approach: use GANs
- 10^5 x speed-up!





(a) $E_0 = 63.7$ GeV (b) $E_0 = 6.5$ GeV (c) $E_0 = 15.6$ GeV (d) $E_0 = 15.9$ GeV

https://doi.org/10.1051/epjconf/201921402034

Modern AI Challenges in Physics

- Better understanding how to use computer vision and natural language processing techniques, e.g., dealing with data structures, like trees and graphs, that can be analyzed with Deep Learning
- Can ML help with our most computationally costly problems, like simulation or the combinatorial challenge of tracking?
- Can fast O(ns-μs) NN inference be done with FPGAs to put ML early in the trigger / data acquisition process?
- Can we design better architectures and training algorithms to tackle our HEP challenges?
- How can we make best use of our simulation for inference without the PDF, e.g., Likelihood Free Inference?

https://nature.com/articles/s41586-018-0361-2