PAUL SCHERRER INSTITUT

PSI

Nicole Hiller :: Operation :: Paul Scherrer Institut

Machine Learning at PSI

10th Workshop on Longitudinal Electron Bunch Diagnostics

Machine Learning

wikipedia.org: **Machine learning** (**ML**) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence.

Samuel, Arthur L. (1959). "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*. **44**: 206–226. CiteSeerX 10.1.1.368.2254. doi:10.1147/rd.441.0206

- **Learning**: (un-)supervised, reinforcement, self-learning, feature learning, anomaly detection, clustering…
- **Models**: neural networks, decision trees, support vector machines, …
- **Strong ties to:** Optimisation, data mining, statistics -> data science

ML at PSI - What for?

- Tuning / Optimisation / Control
  (e.g. Bayesian optimisation at SwissFEL & HIPA)

- Prognostics / alarm handling / anomaly-breakout detection
  (e.g. interlock prediction at HIPA, virtual diagnostics -> Photon diagnostics looking into this)

- Data analysis
  (e.g. clustering, pattern recognition -> beamline people are looking into this)

- Simulations / Modeling
  (e.g. surrogate models when simulations are very time-consuming or there is no clear model -> predicting detector calibration factors; training a fast “online” model of an accelerator)

- Other projects in the planning
  (e.g. efficient patient scheduling for proton therapy, …)

Bayesian Optimisation at SwissFEL

Goal: Optimise the FEL pulse intensity
-> Many Photons = Happy Users

Challenges:

- Exponential dependence of FEL output on many, coupled machine parameters
- Manual tuning is time consuming & inefficient
- Use many knobs (~40)
- Don’t drive the machine into the wall…

J. Kirschner[1], M. Nonnenmacher[1], M. Mutny[1], A. Krause[1], N. Hiller[2], R. Ischebeck[2], A. Adelmann[2]

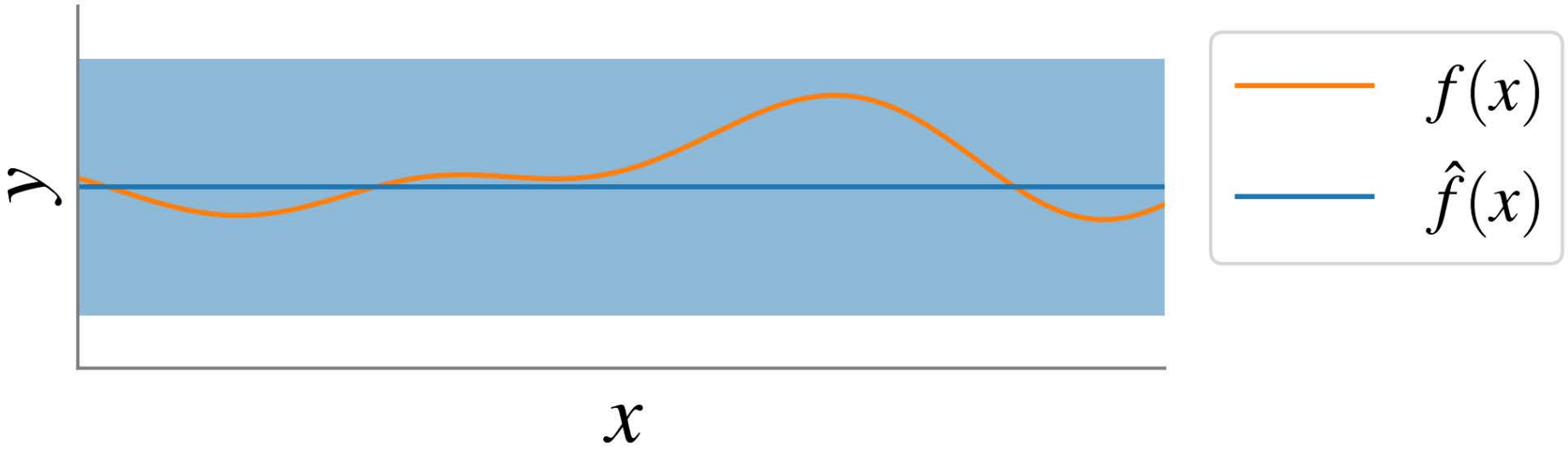
[1] ETH Zürich, Department of Computer Science, Zürich, Switzerland

[2] Paul Scherrer Institut (PSI), CH-5232 Villigen PSI, Switzerland

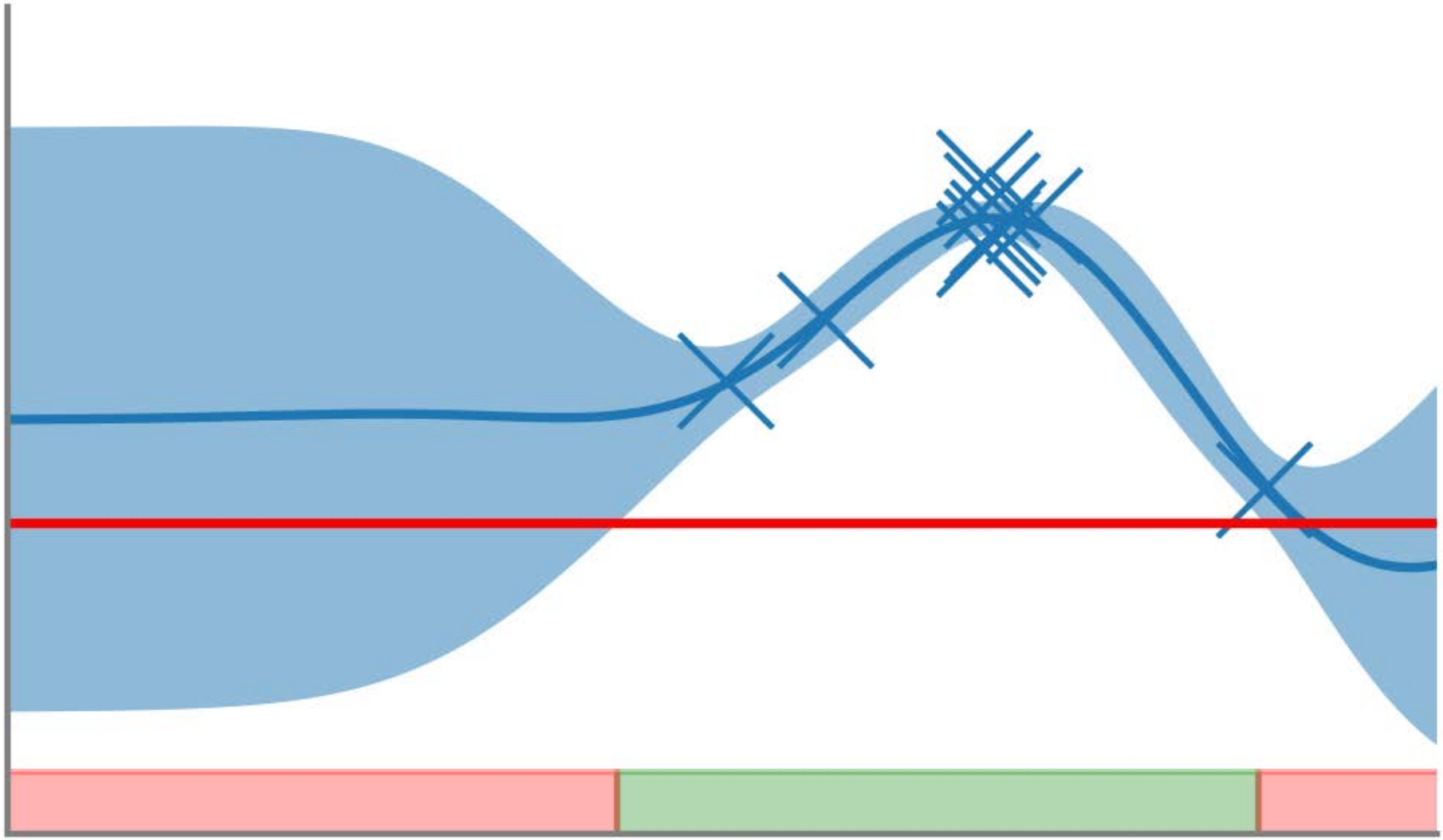
– The work is supported by the Swiss National Science Foundation under grant 200020_159557

Bayesian Optimisation at SwissFEL

- The ML Experts from ETHZ suggested Bayesian Optimisation!
  -> How does this work?

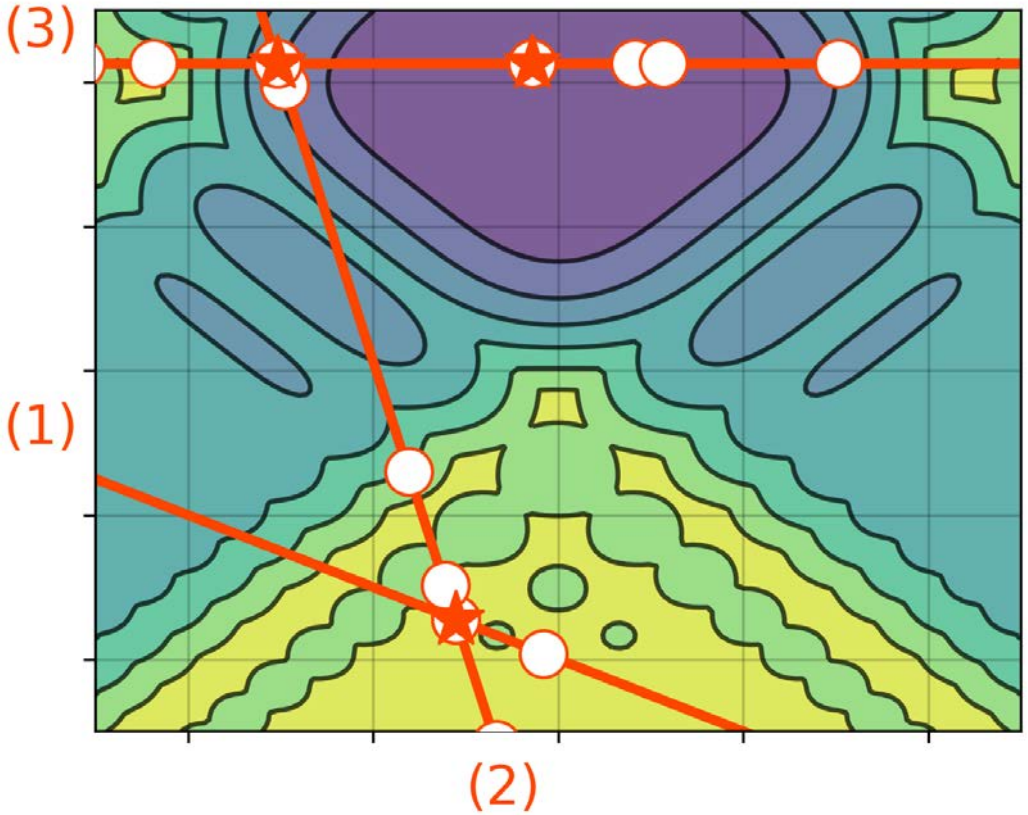


Bayesian Optimisation at SwissFEL

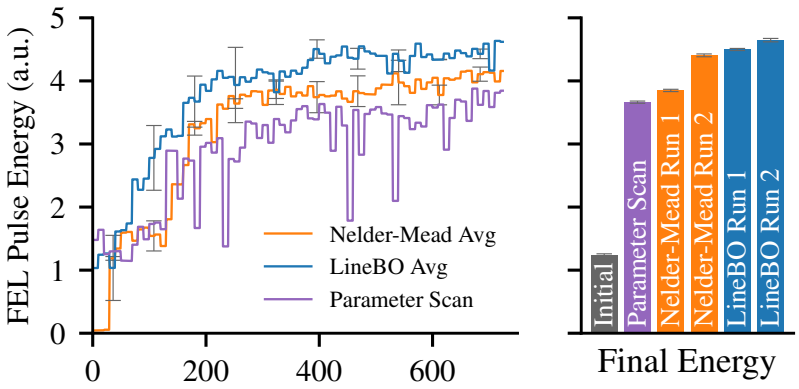


Bayesian Optimisation at SwissFEL

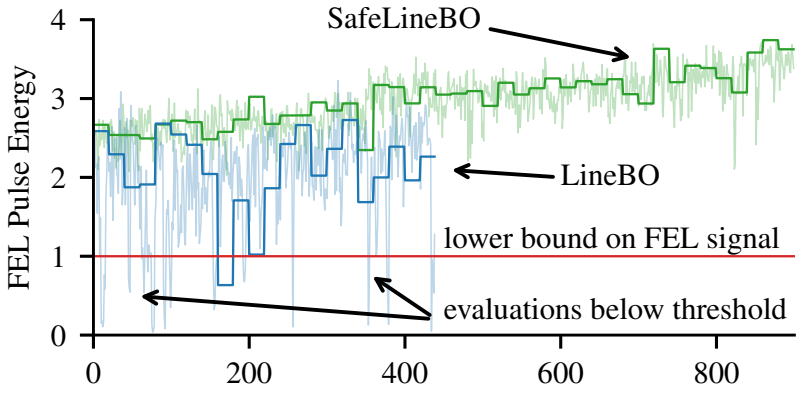
- How does it scale for many parameters? -> Poorly…
- What do we do? -> Combination of gradient method (local sampling) and 1D line approach over several lines



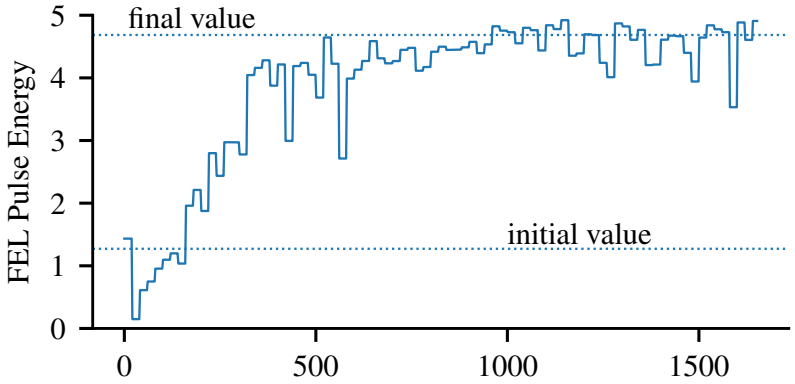
Bayesian Optimisation at SwissFEL



Benchmarking against other algorithms!
(24 Parameters)



Adding Safety Constraints
(24 Parameters)



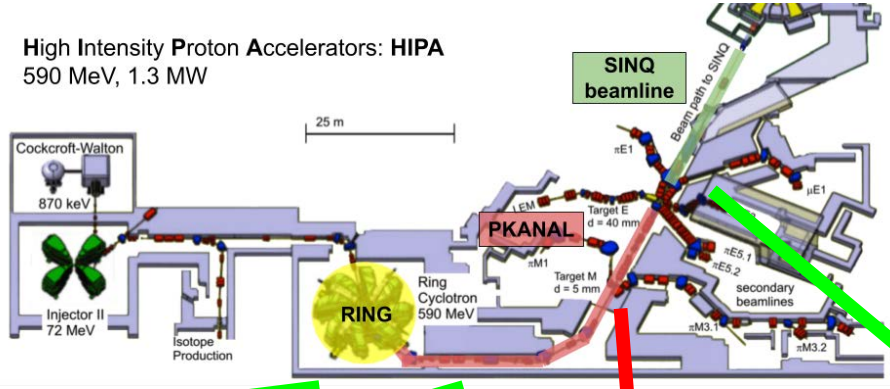
Tuning at about 1 Hz: 500 steps = 8 min

Still works fast with 40 Parameters!

Bayesian Optimisation at SwissFEL

- Where does the journey go to?
  - Versatile Framework with a GUI that
    - can be used by operators
    - can be used at other accelerators -> we are also tuning HIPA with it now!

Interlock Prediction at HIPA



Magnets settings

Transition monitors

Beam position monitors

Temperatures etc.

Interlock system

Shut off the beam

~20% of all beam losses

Quite some could be avoided

Roughly 80 different types, but 5-6 common types

Andreas Adelmann, Jaime Coello, Davide Reggiani, Jochem Snuverink, Melissa Zacharias (PSI); Anastasia Pentina (Swiss Data Science Center)

Interlock Prediction at HIPA

- What would be your approach?

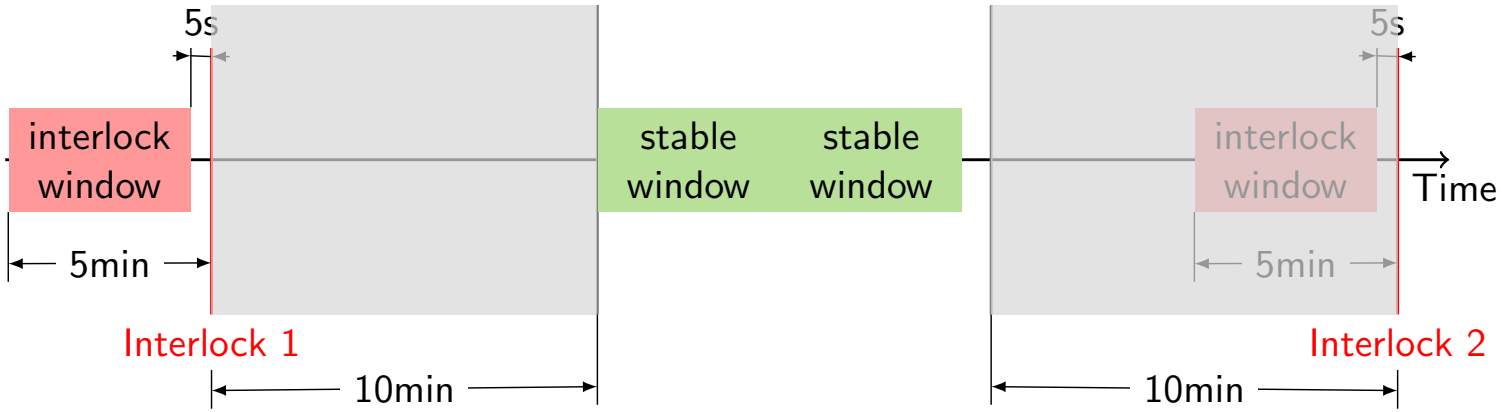
Interlock Prediction at HIPA

Goal: Train a model (supervised learning) that can then predict interlocks and prevent them by lowering the beam current temporarily

Approach:

- Training on archived data (450 channels) -> lots of data preparation needed
- Finding the best model

- **Interlock-window**: interlock happens in 5 seconds
- **Stable-window**: interlocks are far from happening

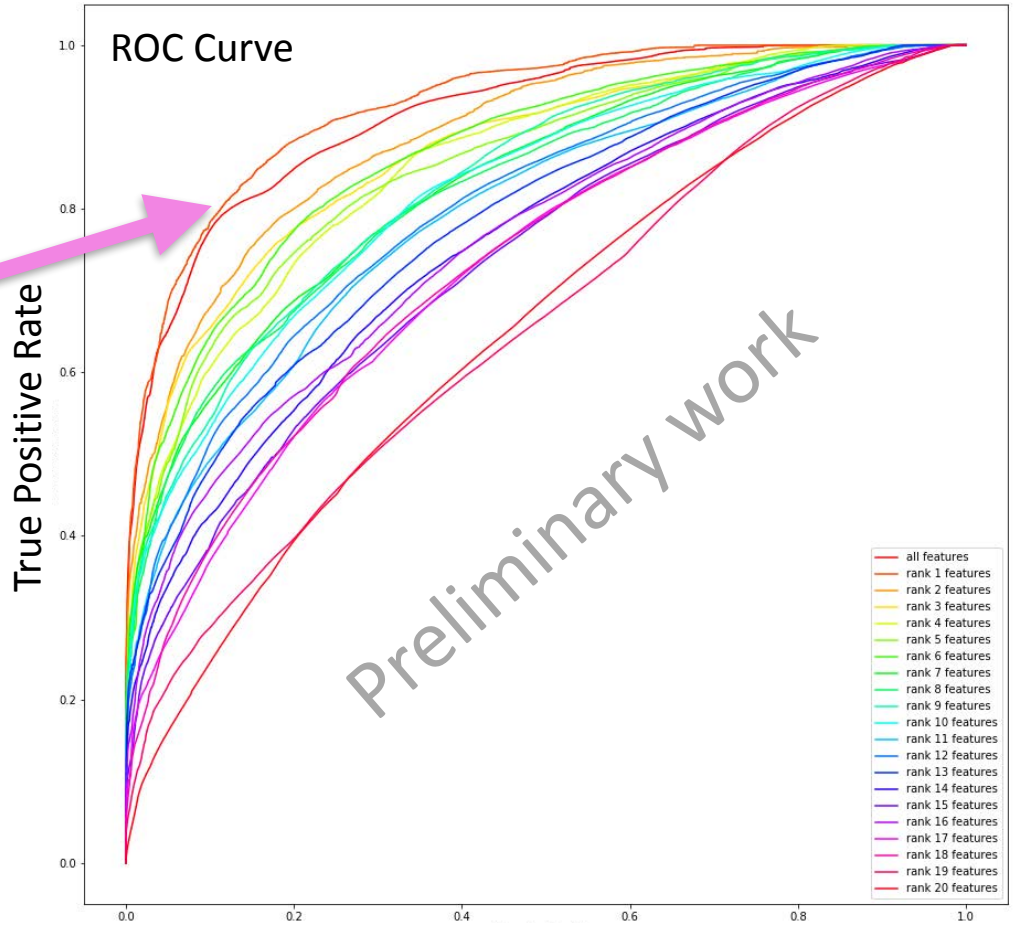


Interlock Prediction at HIPA

In an interlock happens: 25 s no beam
If the charge is temporarily lowered: corresponds to 6 s no beam
-> We want a high positive true prediction, but a very low positive false one!

So far best results with:
Random Forest

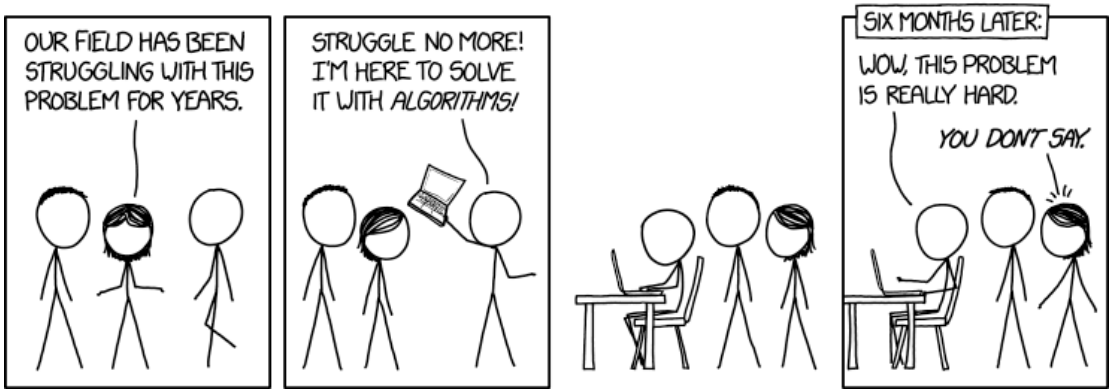
Best results with 50/450 features!
-> Maybe we can also learn some physics from that



ROC curve = receiver operating characteristic curve

Lessons learned so far

- **Machine learning techniques can be great tools especially when**
  - Your problem is hard to formulate analytically (pattern recognition,…)
  - You have a lot of data
  - You would have to frequently use very time-consuming simulations
  - Your “standard tools” do not work very well
  - You need a high degree of automation
- **but**
  - You still need to understand your problem! -> it does not do that for you
  - You still need to prepare your data (data science skills are crucial) -> do not underestimate this part
  - Quite often a lot of “hyper-parameters” are involved -> you will need to learn the “tricks of the trade”



www.xkcd.com

ML @ PSI - Networking

- Regular (every 6 weeks) discussion rounds: ML Luncheons (11:30-13:00 with free sandwiches) -> open discussions, informal presentations
- ML Awareness event -> Half day event with invited speakers
- ML Mailing List & SLACK channel for announcements & discussions
- Collaboration with the Adaptive Systems Group at ETHZ (Prof. Andreas Krause):
  - PhD & Master Students
- Project grants for the Swiss Data Science Center (SDSC)
  - www.datascience.ch -> professional support for projects (e.g. data scientist is at PSI once per week)
  - 3rd ICFA Mini-Workshop on Machine Learning Applications for Charged Particle Accelerators 18-21 Feb 2020 at PAL, Korea https://www.indico.kr/event/5/

Gartner Hype Cycle for Artificial Intelligence, 2019

gartner.com/SmarterWithGartner

Source: Gartner

© 2019 Gartner, Inc. and/or its affiliates. All rights reserved.

Gartner

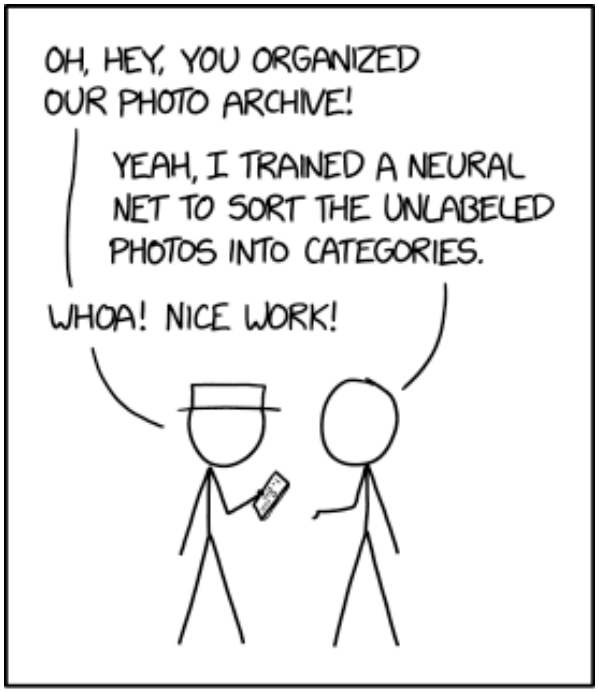
PAUL SCHERRER INSTITUT

PSI

What are your opinions/ experiences?

"The machine learning algorithm wants to know if we'd like a dozen wireless mice to feed the Python book we just bought."

How are things at your institute?



ENGINEERING TIP:
WHEN YOU DO A TASK BY HAND, YOU CAN TECHNICALLY SAY YOU TRAINED A NEURAL NET TO DO IT.

www.xkcd.com