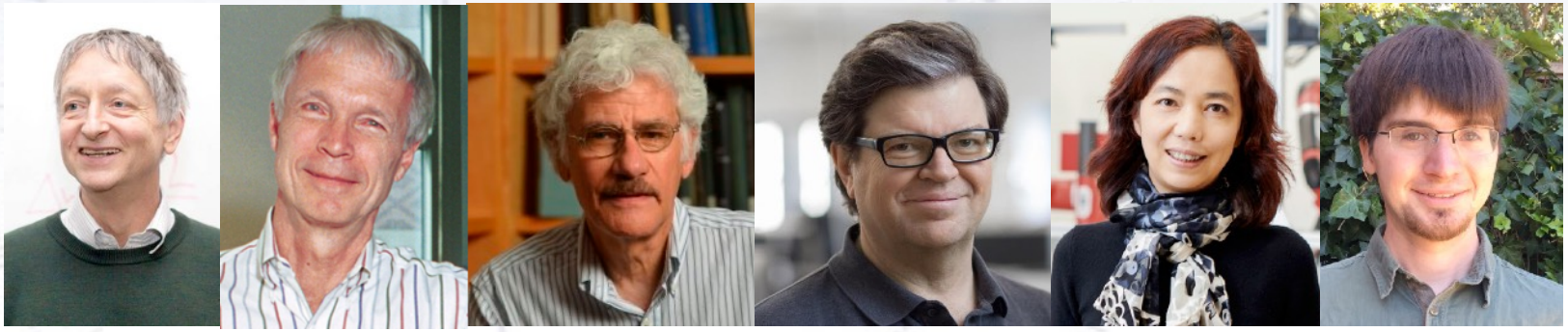
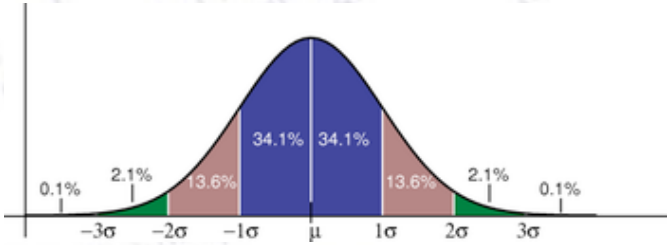
Machine Learning

Course information



Troels C. Petersen (NBI)



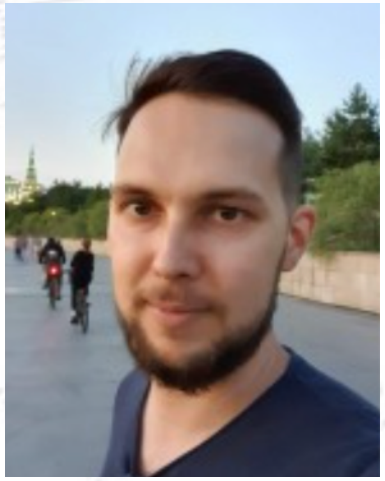
"Statistics is merely a quantisation of common sense - Machine Learning is a sharpening of it!"

Who are we?

Tilman Plehn

Thea Arrestad

Malte Algren

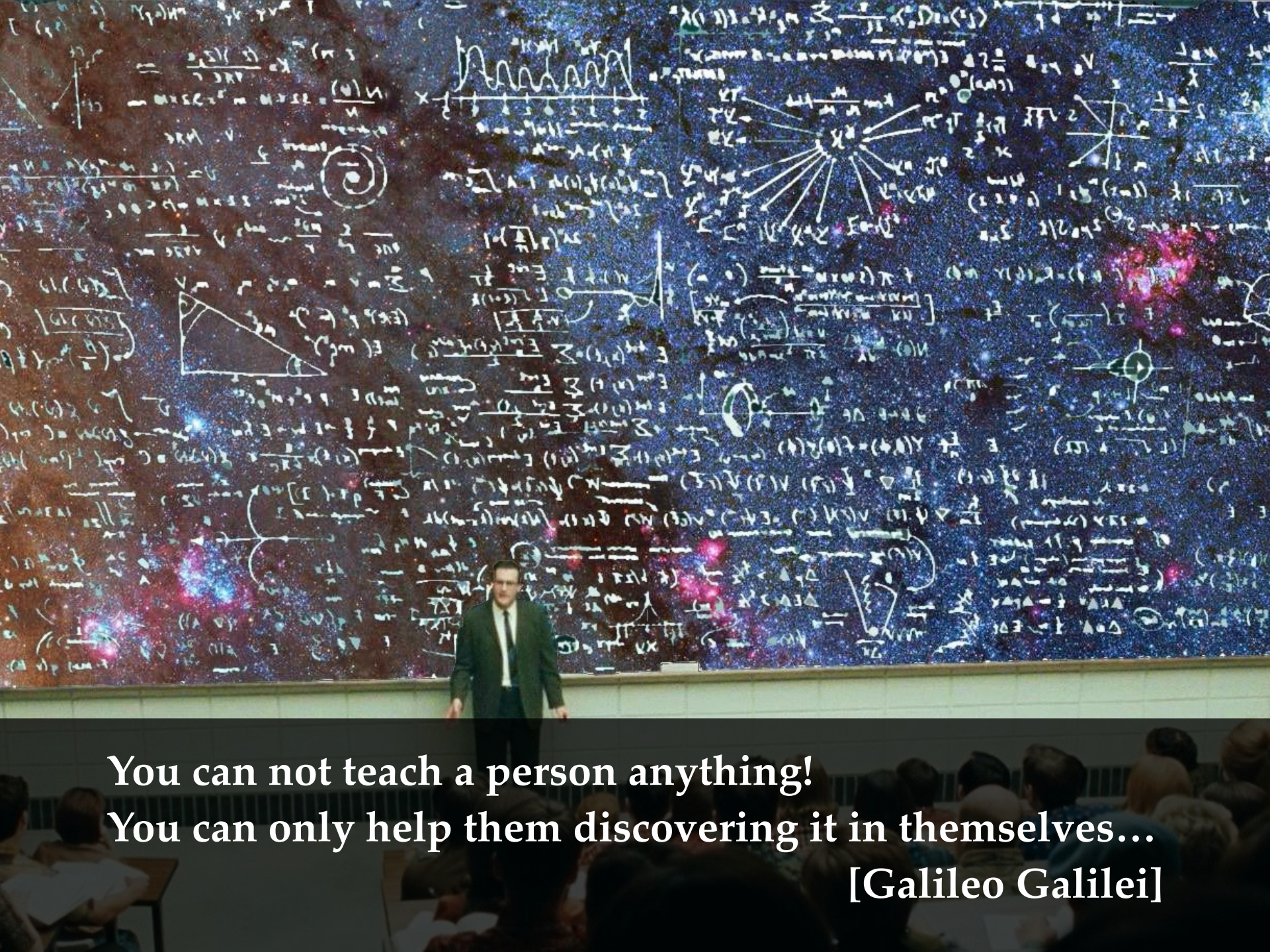


Inar Timiryasov

Daniel Murnane

Troels Petersen

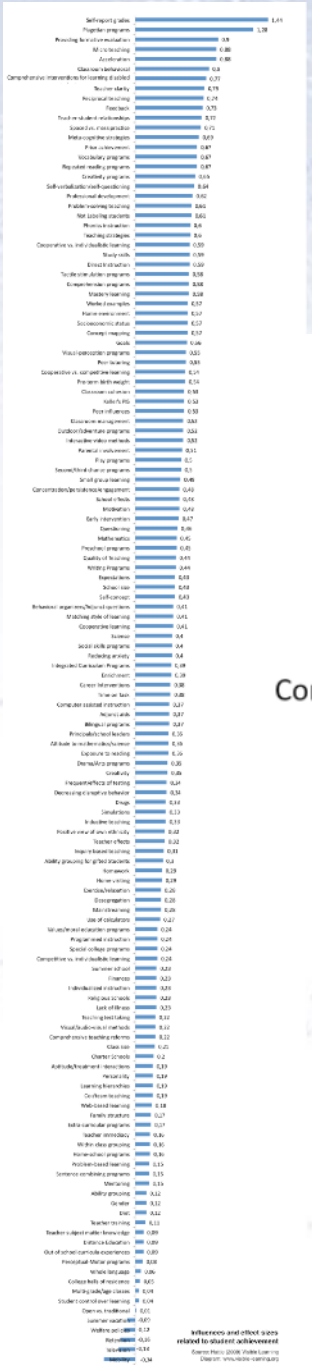
Our aim is for YOU to learn the most about ML and using it in physics research.
If you have any wishes or feedback, don't hesitate to write us.

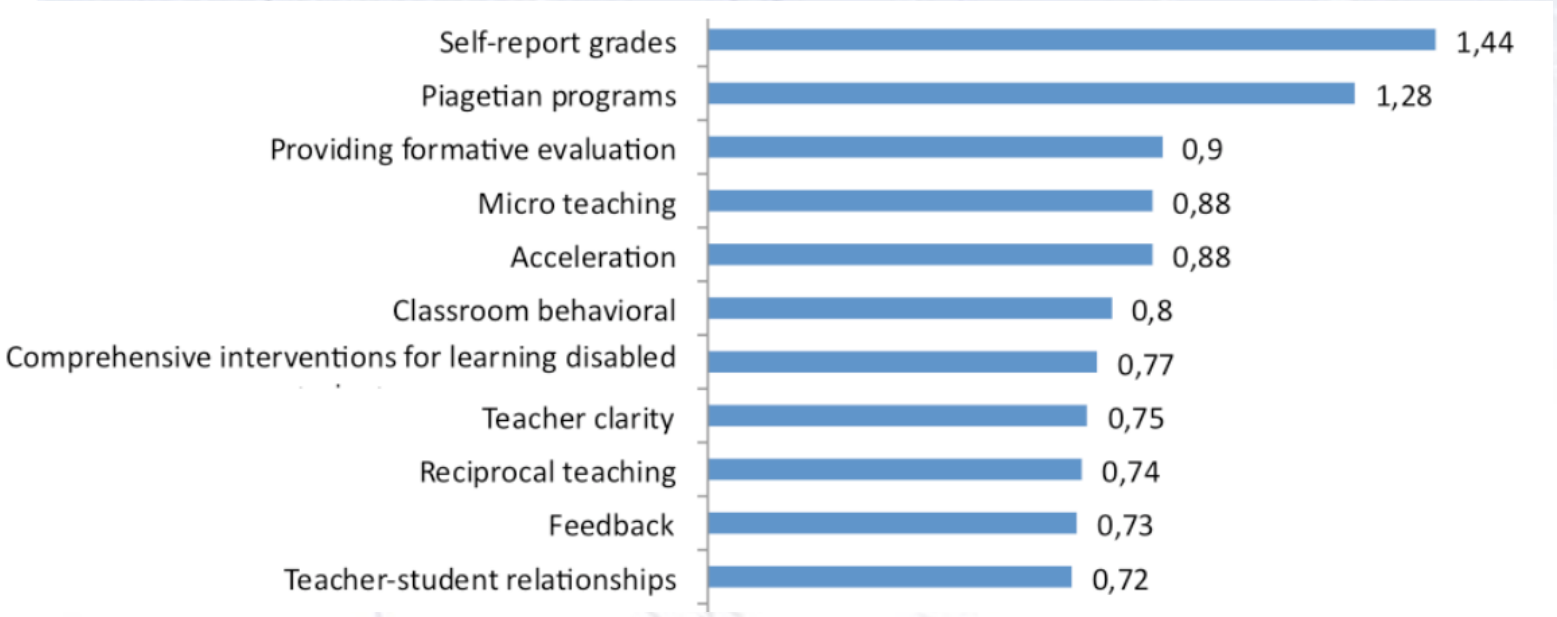


You can not teach a person anything!
You can only help them discovering it in themselves…
[Galileo Galilei]

What influences learning?

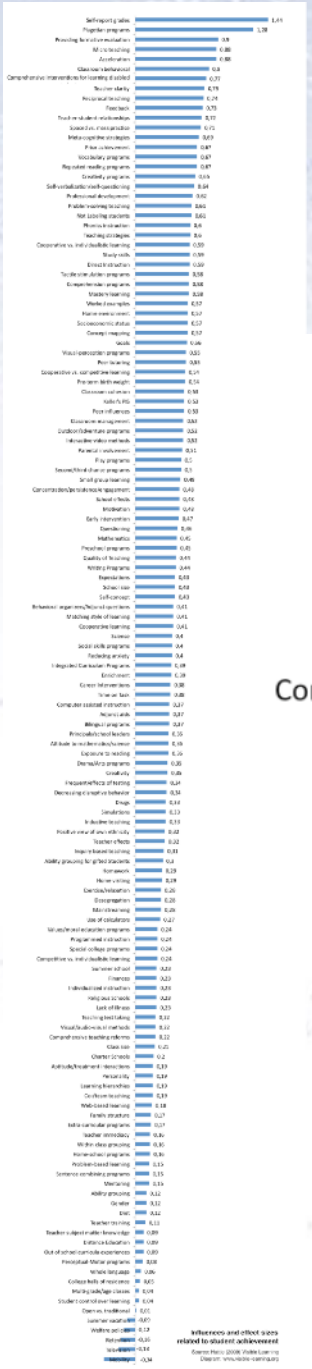
There are studies of this, one result shown below:

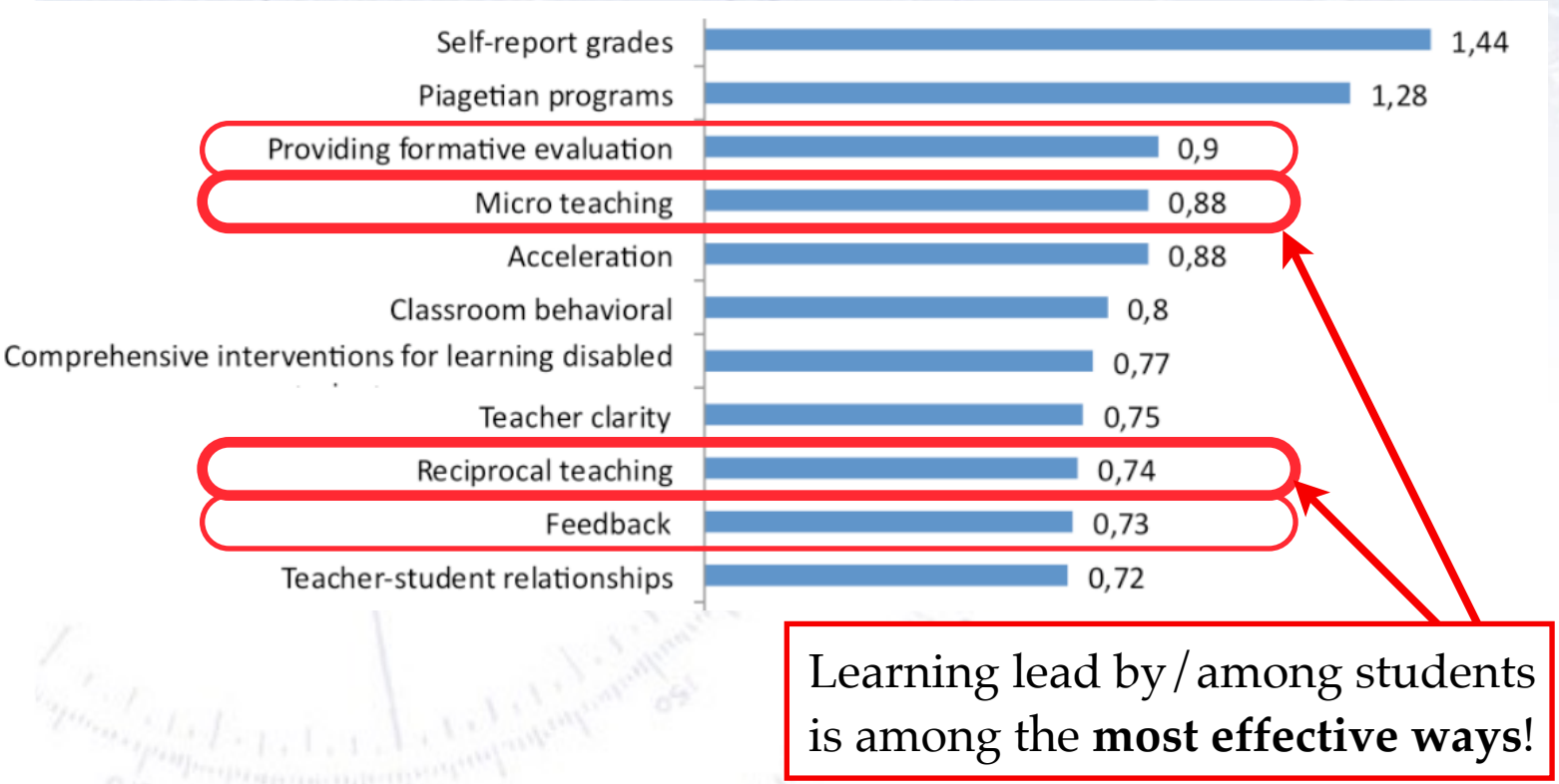




What influences learning?

There are studies of this, one result shown below:

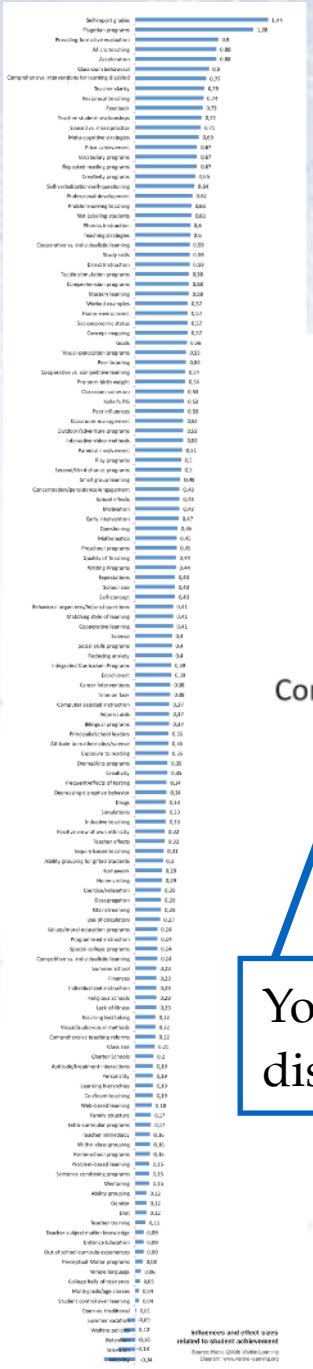


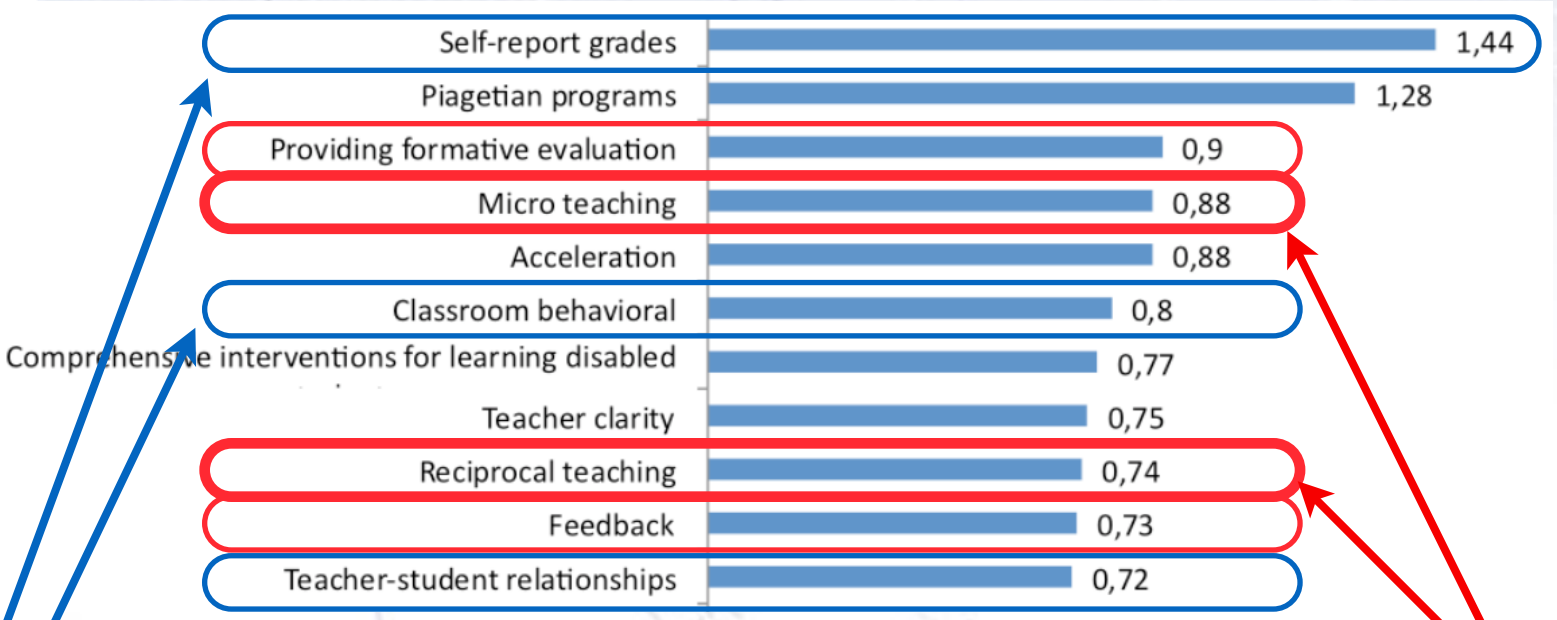


Learning lead by/among students is among the **most effective ways**!

What influences learning?

There are studies of this, one result shown below:



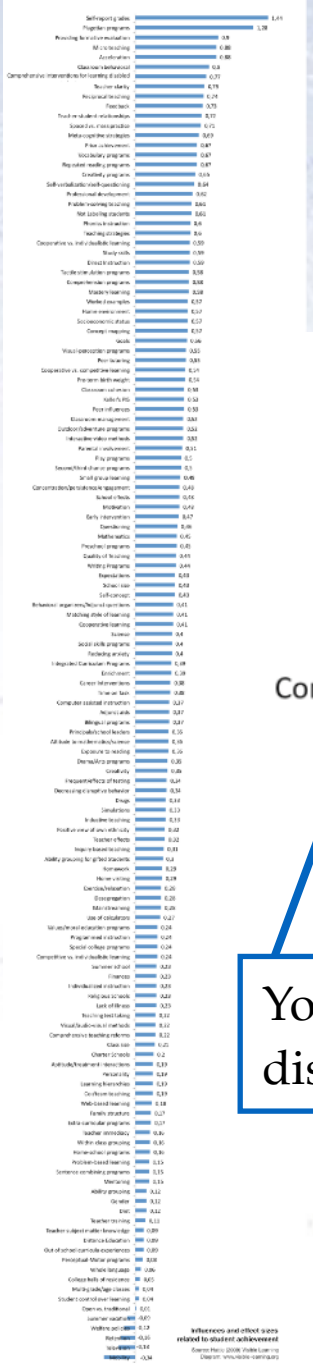


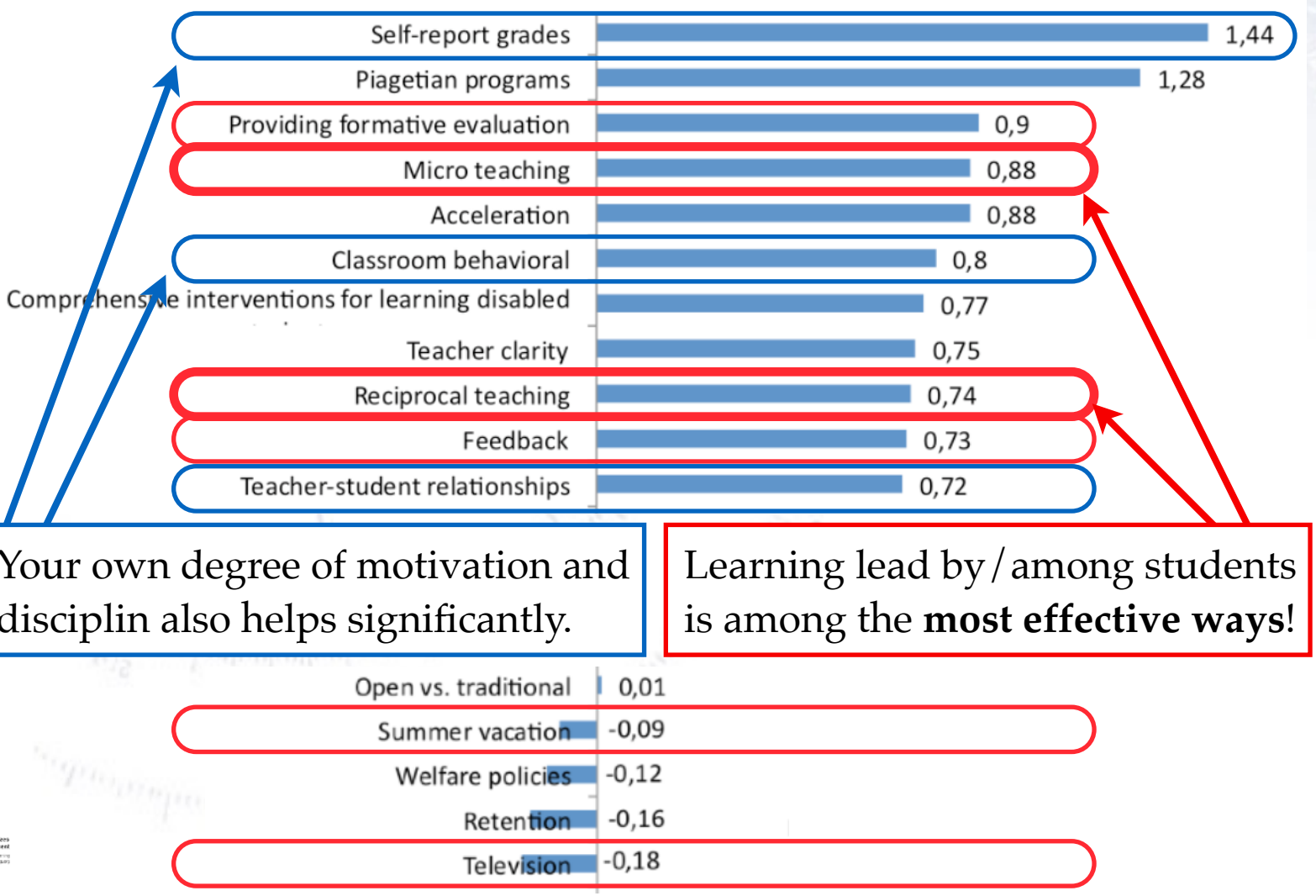
Your own degree of motivation and disciplin also helps significantly.

Learning lead by/among students is among the **most effective ways**!

What influences learning?

There are studies of this, one result shown below:



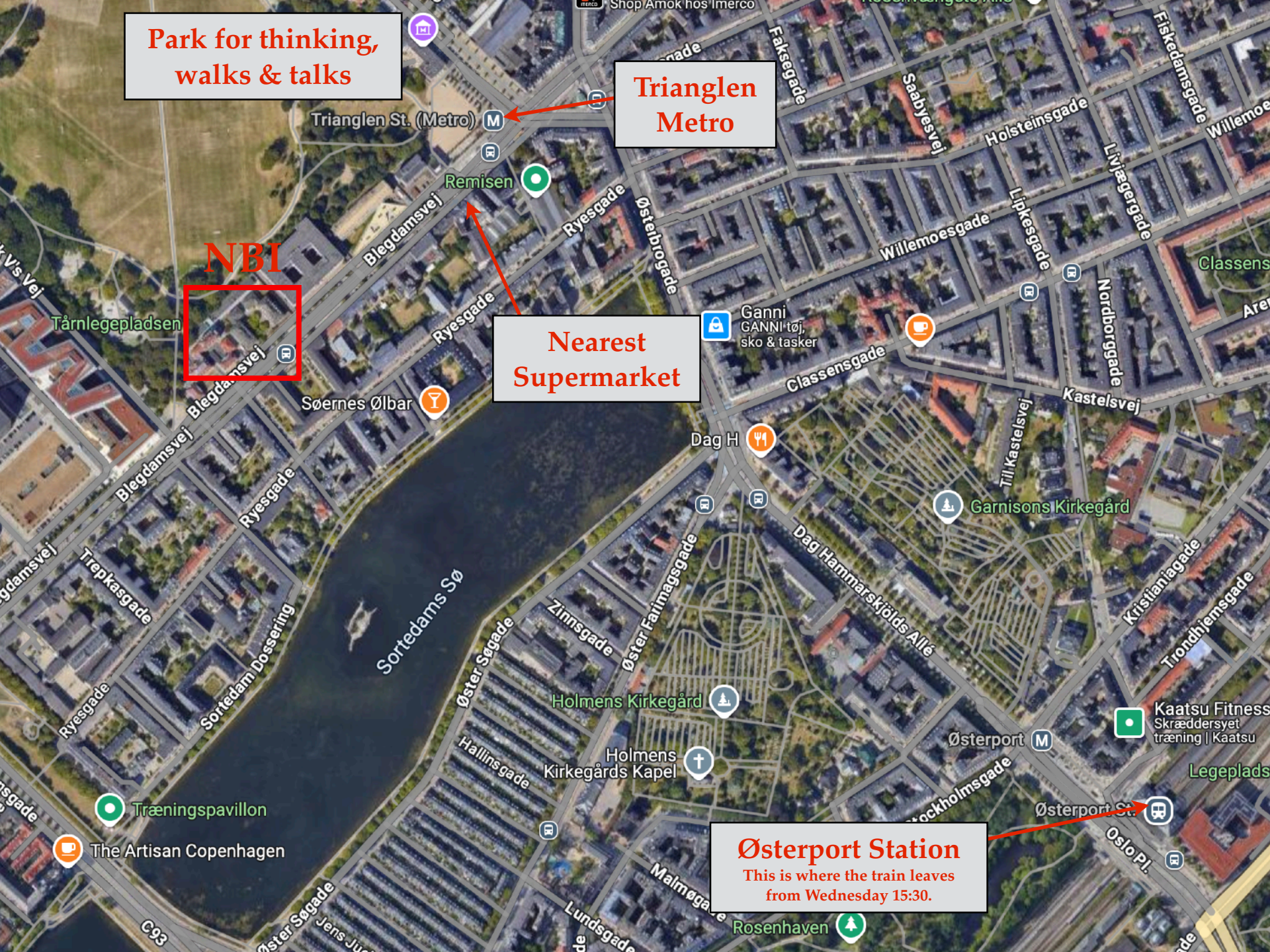


Your own degree of motivation and disciplin also helps significantly.

Learning lead by/among students is among the **most effective ways**!

Locations for Ph.D. School

Locations for Ph.D. School


Park for thinking, walks & talks
Trianglen Metro
Trianglen St. (Metro)
Remisen
NBI
Nearest Supermarket
Ganni
GANNI tøj, sko & tasker
Tårnlegepladsen
Søernes Ølbar
Dag H
Garnisons Kirkegård
Sortedams Sø
Holmens Kirkegård
Holmens Kirkegårds Kapel
Kaatsu Fitness
Skræddersyet træning | Kaatsu
Østerport
Træningspavillon
The Artisan Copenhagen
Østerport Station
This is where the train leaves from Wednesday 15:30.
Rosenhaven
Legeplads
Classens
Blegdamsvej
Ryesgade
Østerbrogade
Willemoesgade
Classensgade
Kastelsvej
Nordborggade
Lipkesgade
Livjægergade
Holsteinsgade
Saabyesvej
Faksegade
Fiskedamsgade
Trepkasgade
Sortedam Dossering
Zinnsgade
Øster Søgade
Øster Farimagsgade
Dag Hammarskjölds Allé
Kristianiagade
Trondhjemsgade
Hallinsgade
Oslo Pl.
Malmøgade
Lundsgade
C93

Copenhagen

All distances within red square are walkable.
For the rest of the city, a bicycle is very useful.
Alternatively, the metro is rather efficient.

A normal day in school

Essentially all days (except Wednesday) look alike:

A special day in school

Wednesday will be a bit different.

09:00 **12 - Lecture 5: Generative Adversarial Networks and introduction to Diffusion and Flow** — *Tilman Plehn*

Lecture (Aud. A)

10:00 *Auditorium A, Niels Bohr Institute* 09:00 - 10:15

Break

Auditorium A, Niels Bohr Institute 10:15 - 10:45

13 - Lecture 6: Diffusion and Flow techniques — *Malte Algren*

11:00

Lecture (Aud. A)

Auditorium A, Niels Bohr Institute 10:45 - 12:00

12:00 **Lunch break**

Auditorium A, Niels Bohr Institute 12:00 - 13:00

13:00 **14 - Discussion of Lectures and intro to exercises**

Discussion (Ke + Kc7)

Auditorium A, Niels Bohr Institute 13:00 - 13:30

15 - Work on exercises

14:00

Exercise (Ke + Kc7)

Auditorium A, Niels Bohr Institute 13:30 - 14:30

Packing up, walking to Østerport train station, and preparing for boarding train.

15:00

Excursion: Train, Castle, Drinks & Dinner

Auditorium A, Niels Bohr Institute 14:30 - 15:30

16 - Train ride to Helsingør, visit to Kronborg Castle, drinks on the ramparts, and dinner at Værftets Madmarket.

16:00

WebBlogs as literature

In ML, blogs/articles/tutorials are a very common (and great) source of literature on ML. For this reason, we've made a list of links that we find good:

Applied Machine Learning 2025 - Useful ML links

The field of Machine Learning (ML) is developing at a very fast pace and by an expanding number of practitioners. For this reason, text books on ML are typically few and slightly d describe the general concepts very well. Research papers have partially filled the gap, as these are more versatile and frequently updated. However, they typically only deal with a ver generel ML "phase space".
But whereas "classic" literature (i.e. books and papers) only covers partially, blogs and github repositories seem to fill out the rest, and typically in a much more accessible fashion wi example code. For this reason, we have tried to gather some of the more useful links to such blogs and repositories below. It is simply our (slightly random) selection of webpages tha found illustrative and useful. So use these at will, and also build your own list of reference sites.

Books:

- Applied Machine Learning by David Forsyth. Great book, at times even mildly entertaining, from one of the masters of ML vision.
- Elements of Statistical Learning II by Trevor Hastie et al. Second eddition of standard book on fundamentals.
- Machine Learning (from PDG) by Kyle Cranmer et al. (2021). A concise (60 page) overview
- Deep Learning by Ian Goodfellow et al. (2016). A short and good general introduction to ML can be found in Chapter 5 of Part 1.
- Pattern Recognition and Machine Learning by Christopher M. Bishop (2006).
- "Interpretable Machine Learning" by Christoph Molnar (2020). A Guide for Making Black Box Models Explainable.
- "Convolutional Neural Networks for Visual Recognition" by Andrej Karpathy (2017?). Used for teaching CNN in Stanford's cn231 class.
- "Deep Learning with Python" by author of Keras, Francois Chollet, now at Google AI (2017). Especially chapter 4 is a good overview of the fundamentals of Machine Learning.
- Book on neural networks and deep learning. Online book only!
- Pattern Recognition and Machine Learning (2006) by Christopher M. Bishop is a classic (if older) standard reference.

Papers:

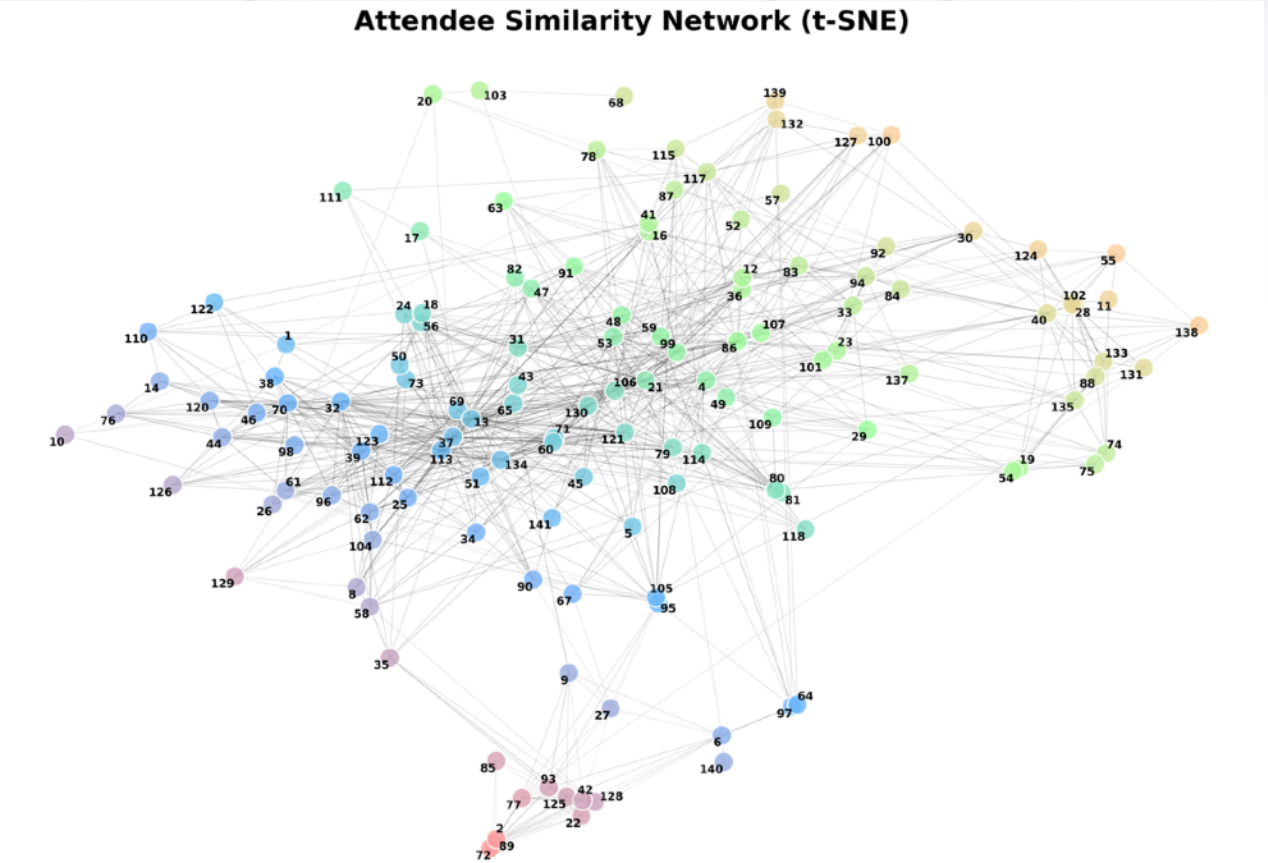
- XGBoost paper (2016). Highly readable paper showing the innovations of the XGBoost algorithm.
- LightGBM paper (2017). Explaining the great speedup and showing examples of execution times.
- More general paper (100 pages) on Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges (2021).
- ROCKET: Time series classification using random convolutional kernels
-

Blogs/Links/Tutorials - supervised learning:

- **Introduction to tree based learning**. Very good introduction to the basics of tree based learning.
- **Introduction to neural net based learning**. Very good introduction to the basics of Neural Net (NN) based learning.
- SciKit Learn tutorial. Gives a quick introduction to ML in general and has code examples for SciKit Learn.
- XGBoost vs. LightGBM. Discussion of differences, with code examples.
- XGBoost, LightGBM, and CatBoost. Discussion of differences and hyperparameters.
- Introduction to NGBoost, which is a tree based algorithm, which makes a probabilistic predictions (i.e. uncertainties).

Who are the others? Like me?

In order to let everyone get an idea of who is who at this school, we wanted to try to find out, who were alike/dislike. Using your registration information, one could imagine getting the following pattern came out:

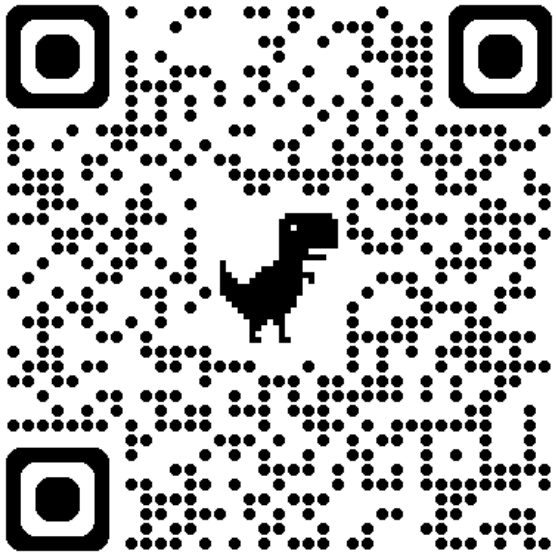


Who are the others? Like me?

Using an LLM, Daniel put together a short Physics-ML Bio of everyone, which can be found using the conference numbers here:

https://hamlet-conf.github.io/nbi-school-25/

Try to look up yourself…

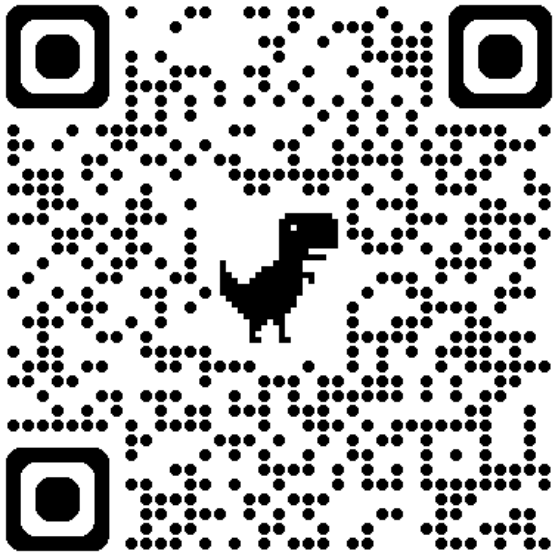


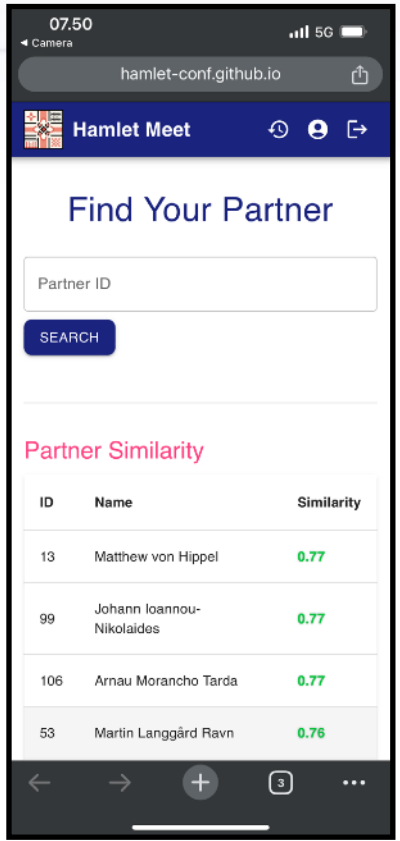
Who are the others? Like me?

Using an LLM, Daniel put together a short Physics-ML Bio of everyone, which can be found using the conference numbers here:

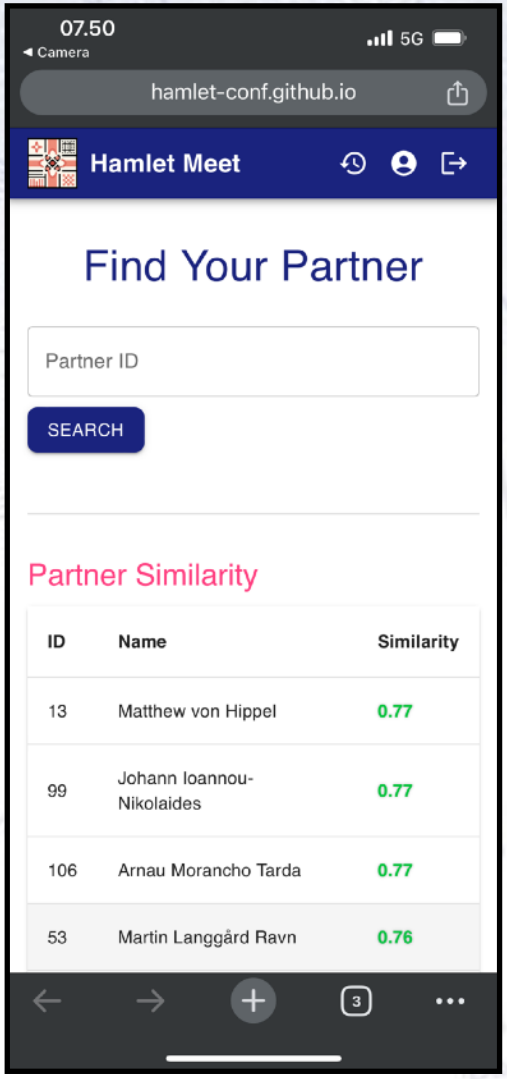
https://hamlet-conf.github.io/nbi-school-25/

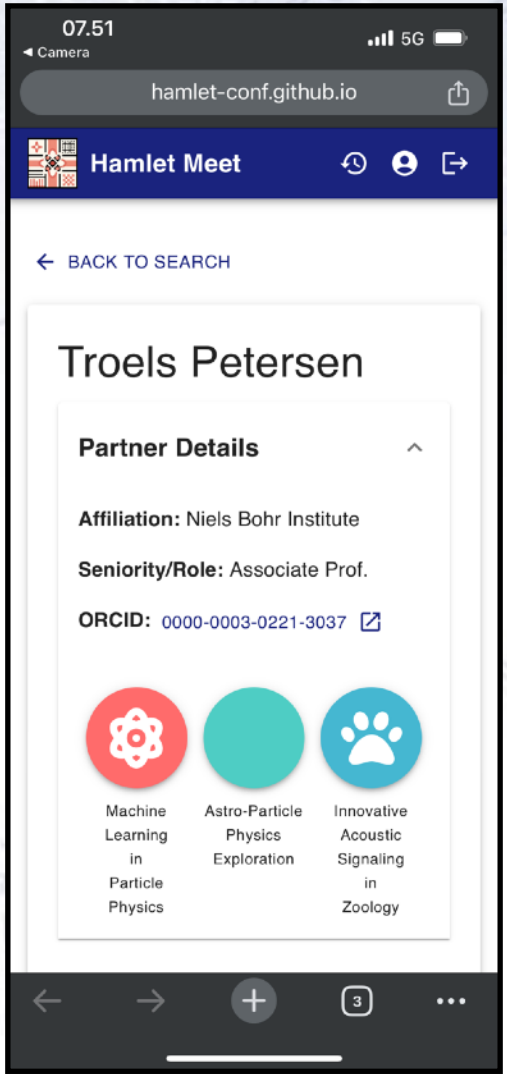
Try to look up yourself…

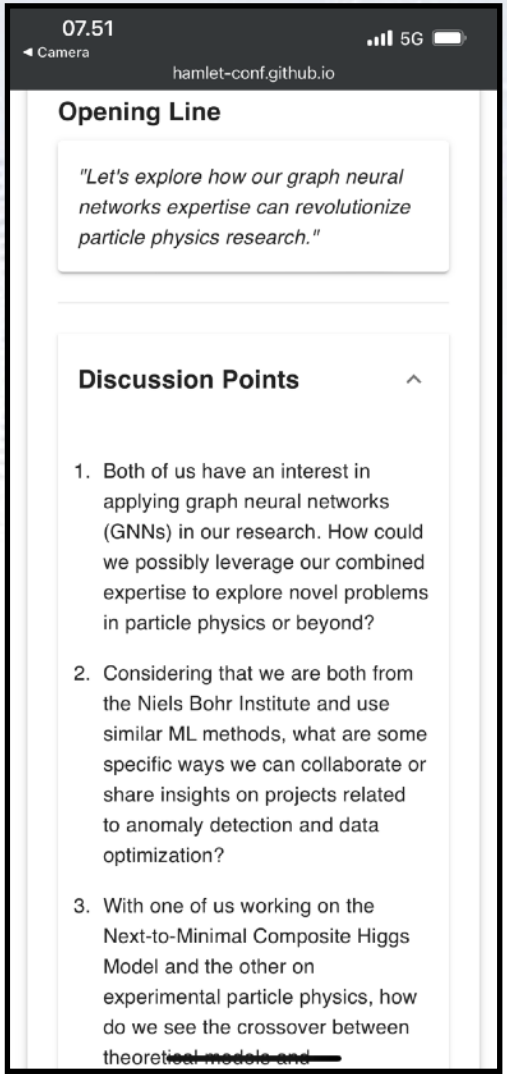




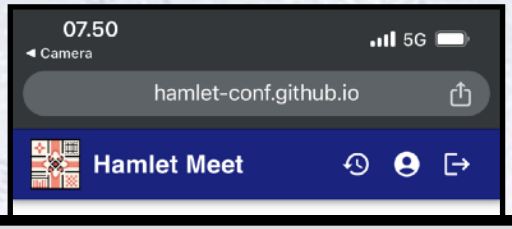
The HAMLET MeetUp App

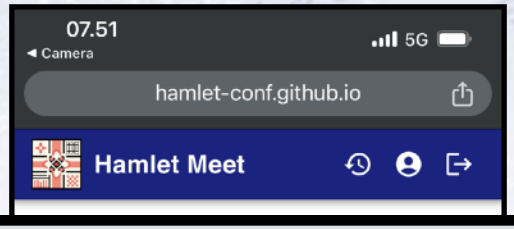


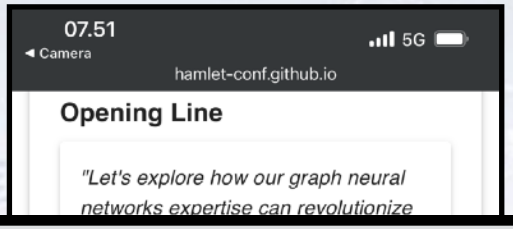




The HAMLET MeetUp App

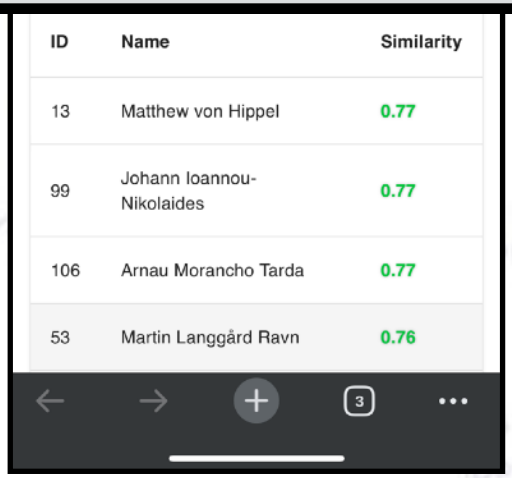


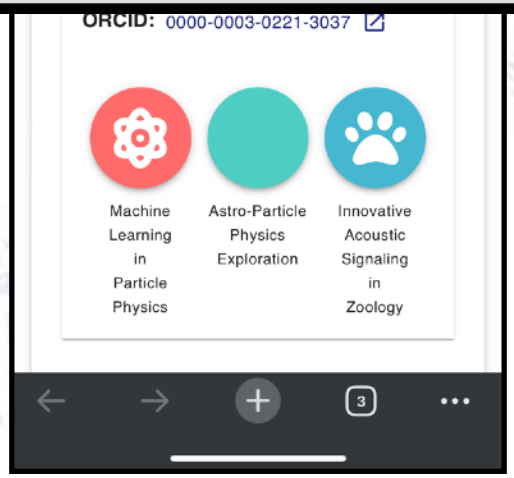


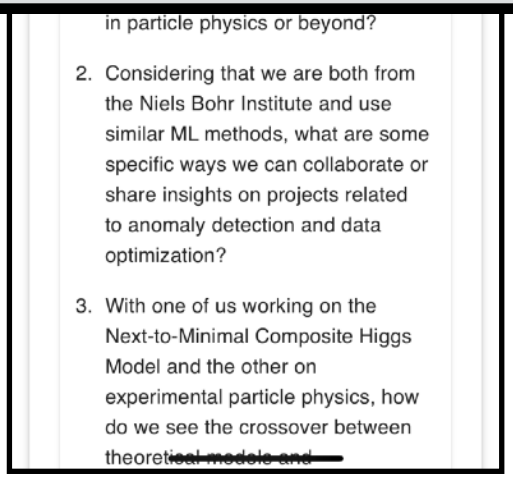


Play around with this tool, and don't hesitate to give us feedback.

Happy App'ing…







Code of Conduct

This Ph.D. school is a community event intended for presentations, networking and collaboration. We value a civil and respectful environment which encourages the free expression and exchange of scientific ideas.

All attendees are expected to adhere to the CERN Code of Conduct.

Should a lapse of professional decorum occur, attendees are encouraged to bring issues, in a confidential setting, to the advisors appointed by the conference organisers.
The advisors will suggest ways of redressing the matter and counsel the parties involved. The conference organisers may, after due consideration, take action as they deem appropriate, including, in severe cases, expulsion from the school.

If you believe someone is violating the code of conduct, we ask that you report it by emailing Daniel Murnane (daniel.murnane@nbi.ku.dk) and/or Troels C. Petersen (petersen@nbi.dk).

Essential Links

ML links:

https://www.nbi.dk/~petersen/Teaching/ML2025/MLlinks.html

GitHub for school exercises:

https://github.com/troelspetersen/PhDschool2025_AppML

Link to "Kaggle Challenge" setup/rules:

https://www.nbi.dk/~petersen/PhDschool2025/PhDschool2025_InitialProject.html

Google link for submitting "Kaggle Challenge" solutions:

https://docs.google.com/forms/d/1_IHeU9aBWtbQQza1c9thweWhc47qcZZBWxU-fKd0V9s/edit

Deadline is Thursday at noon - we will then start to evaluate the solutions.

The "School Proximity" MeetUp App:

https://hamlet-conf.github.io/nbi-school-25/

Computers & Internet

The Niels Bohr Institute have “eduroam” wifi network (thus the usual channel).

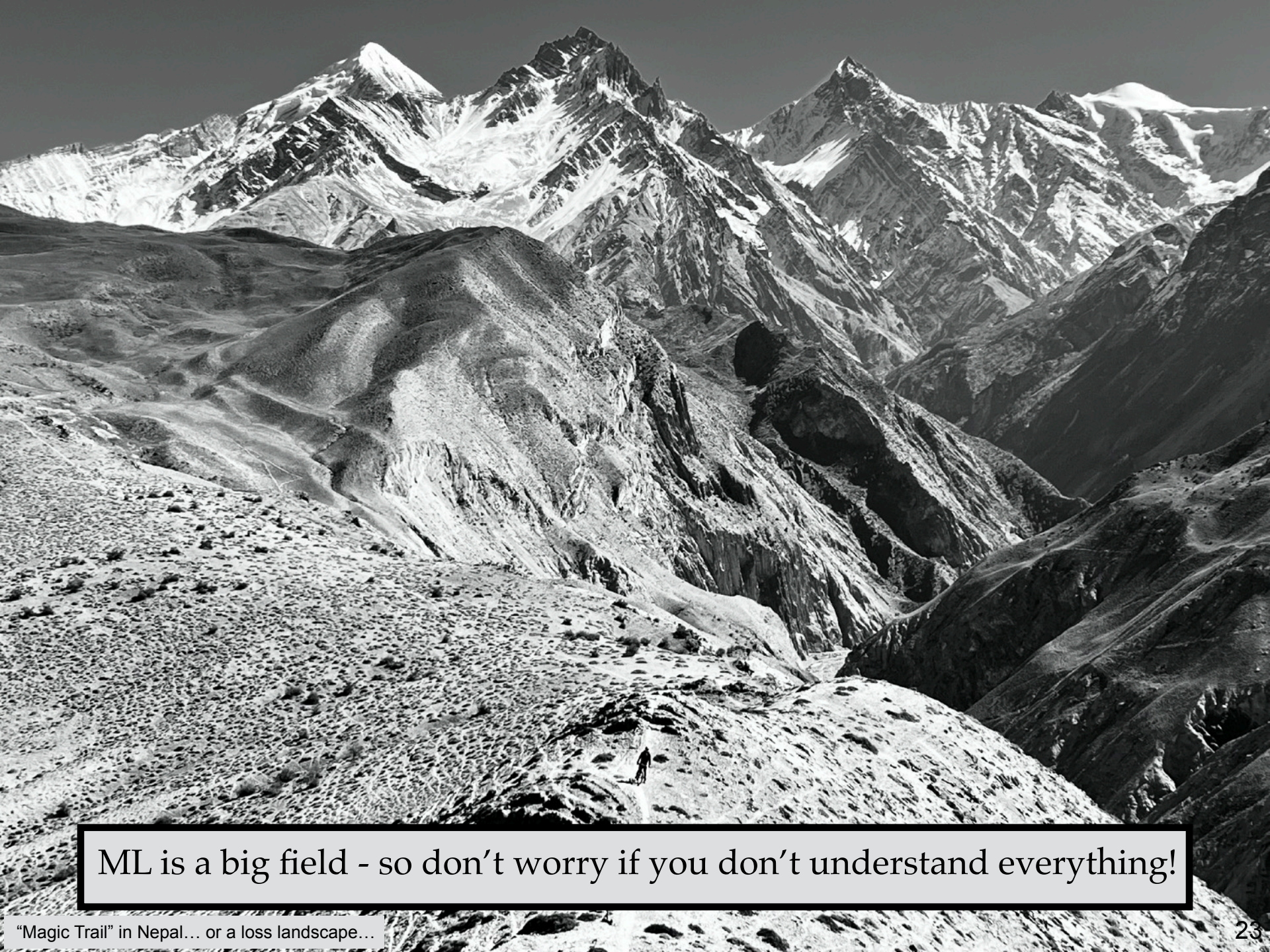
For the exercises (K-building), we have setup an addition WIFI source:

WiFi:

NC-SummerUniversity

Password:

Sommer2025!



ML is a big field - so don’t worry if you don’t understand everything!

“Magic Trail” in Nepal… or a loss landscape…

Between Models and Reality:

Ph.D. School on Machine Learning in Physics

Scientific advisory committee:
Anja Butter (LPNHE)
Tilman Plehn (Heidelberg)
Juergen Hesser (Heidelberg)
Bogdan Malaescu (LPNHE)
Bertrand Laforge (LPNHE)
Stefano Forte (Milan)
Stefano Carrazza (Milan)
Vincenzo Piuri (Milan)
Xavier Fresquet (Sorbo
Tobias Golling (Geneva)
Svyatoslav Voloshynovs va)

Speakers:
Tilman Plehn (Heidelberg)
Daniel Murnane (NBI)
Thea Aarrestad (ETH - Zuri
Malte Algren (Geneva)
Troels C. Petersen (NBI)

Happy School

Sponsors:

4eu+

Local organising committee:
Troels C. Petersen (NBI)
Johann Ioannou-Nikolaides (NBI)
Daniel Murnane (NBI)
Inar Timiryasov (NBI)
Jean-Loup Tastet (DIKU)
Rasmus Ørsøe (TUM)

All lectures will be held in the historic Auditorium A in central Copenhagen.

Niels Bohr Institute, Copenhagen, 2nd – 6th of June 2025

https://indico.nbi.ku.dk/event/2153/