



Machine learning in research: dealing with the non-ideal.

Workshop Machine-Learning for Research 2020
Florian Huber, eScience Center

Outline

1. Expectation management
2. In science it's different
3. Expectation management

If this experiment fails
it will tear a hole in
the fabric of space
and Time!



Wow!!
Really?



No... if it actually fails
the room will stink like eggs
and we'll have to leave
for a bit!



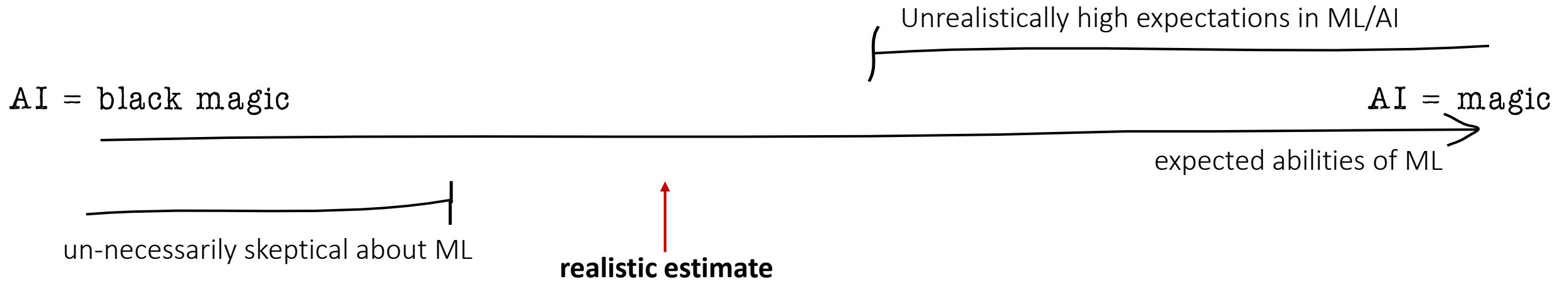
Oh...



@twisteddoodles

1. Expectation management

Key part of our job at eScience Center:
Expectation management.

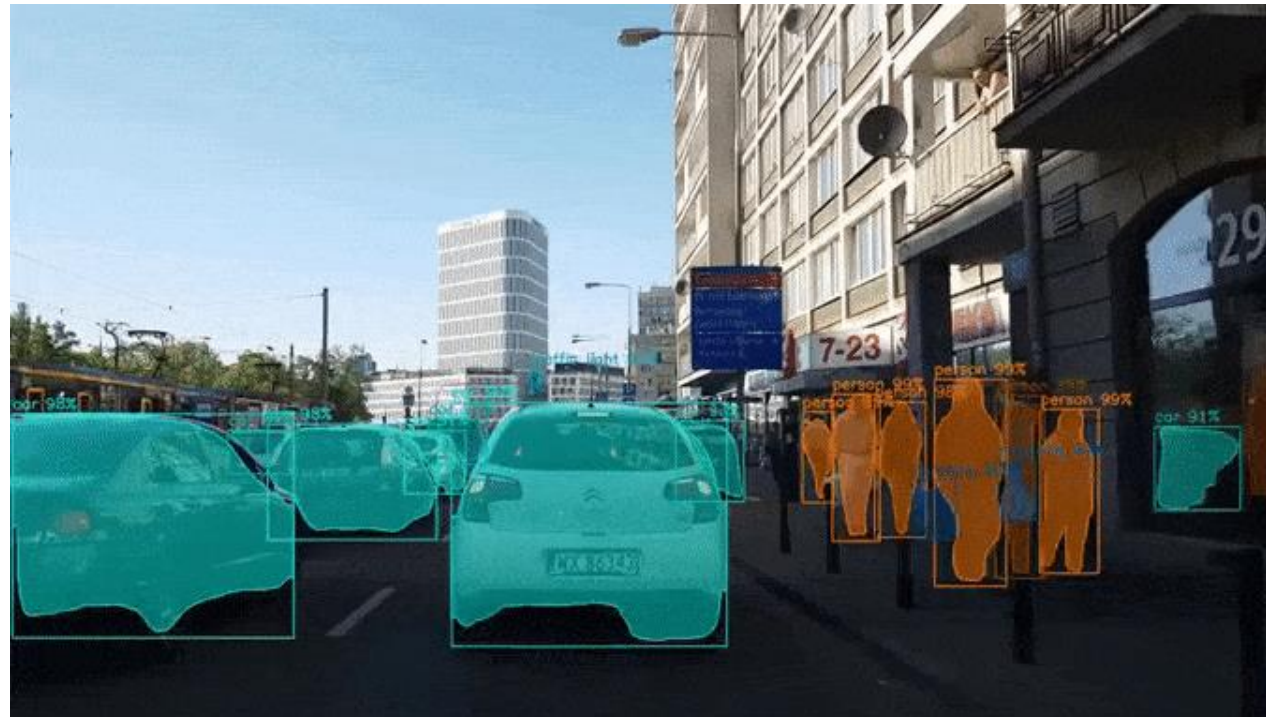


What triggers unrealistic expectations towards ML?

1. Most popular are the greatest success stories.

Usually not representative for most research problems!

- ImageNet-type image classification. (CNN boom boosted by AlexNet, 2012)
- Alpha Go, Alpha Zero and Co.
- Rise of NLP (text translation, search engines, Watson, recommender systems)
- Autonomous driving (Tesla, google, ...)



1. Most popular are the greatest success stories.

Usually not representative for most research problems!

Most shiny cases	Typical case (in science/research)
Huge dataset (e.g. ImageNet, google news)	Often: medium to small sized datasets
Mostly homogeneous data (same size/resolution etc.)	Can be very heterogeneous, from different sources, of different sizes etc.
Prober labels (often for all instances), clear class/category distinctions	Incomplete labels, false labels, many borderline cases
Virtually unlimited compute power	Restricted by available/accessibile infrastructure
Final accuracy is all that matters.	Strong focus on 'understandable' machine-learning, robustness, reproducibility.

2. Humanized interpretation of results.

Agreement in prediction does not imply same way of reasoning!



2. Humanized interpretation of results.

Agreement in prediction does not imply same way of reasoning!



(a) Texture image

81.4%	Indian elephant
10.3%	indri
8.2%	black swan



(b) Content image

71.1%	tabby cat
17.3%	grey fox
3.3%	Siamese cat



(c) Texture-shape cue conflict

63.9%	Indian elephant
26.4%	indri
9.6%	black swan

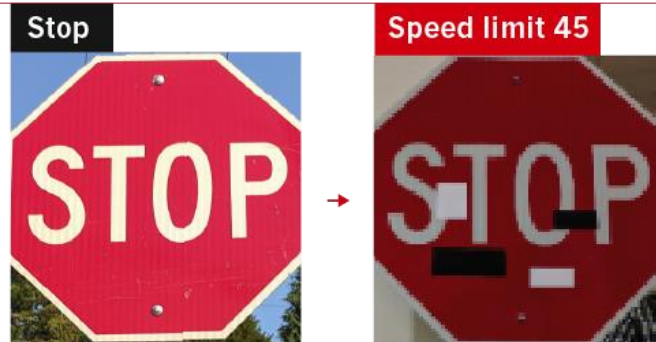
2. Humanized interpretation of results.

Agreement in prediction does not imply same way of reasoning!

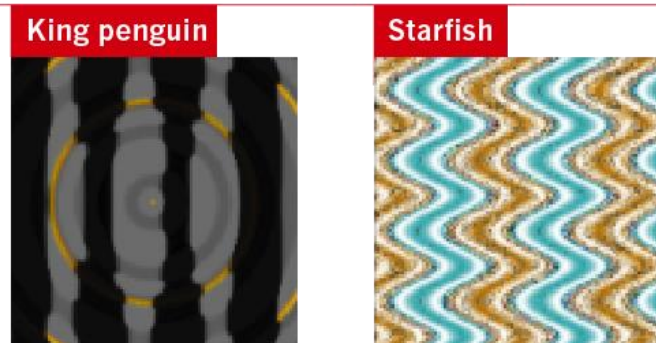
FOOLING THE AI

Deep neural networks (DNNs) are brilliant at image recognition — but they can be easily hacked.

These stickers made an artificial-intelligence system read this stop sign as 'speed limit 45'.

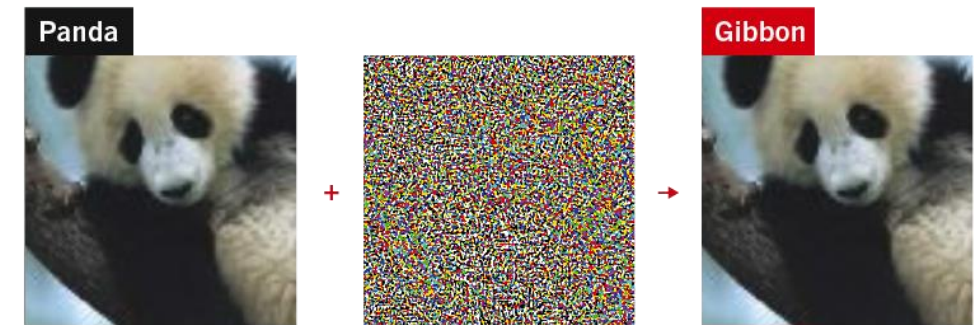


Scientists have evolved images that look like abstract patterns — but which DNNs see as familiar objects.



PERCEPTION PROBLEMS

Adding carefully crafted noise to a picture can create a new image that people would see as identical, but which a DNN sees as utterly different.



In this way, any starting image can be tweaked so a DNN misclassifies it as any target image a researcher chooses.

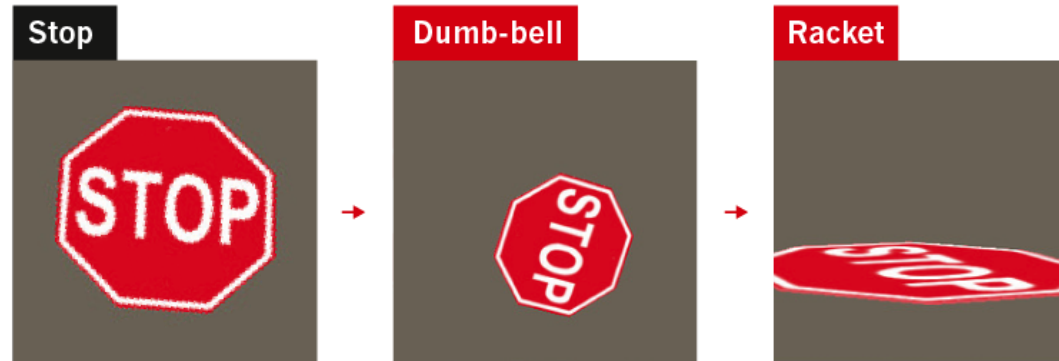


2. Humanized interpretation of results.

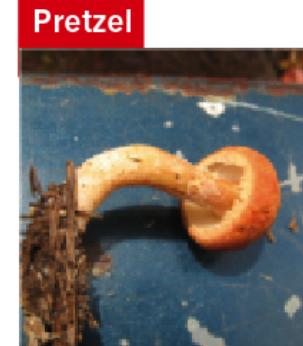
Agreement in prediction does not imply same way of reasoning!

LATEST TRICKS

Rotating objects in an image confuses DNNs, probably because they are too different from the types of image used to train the network.



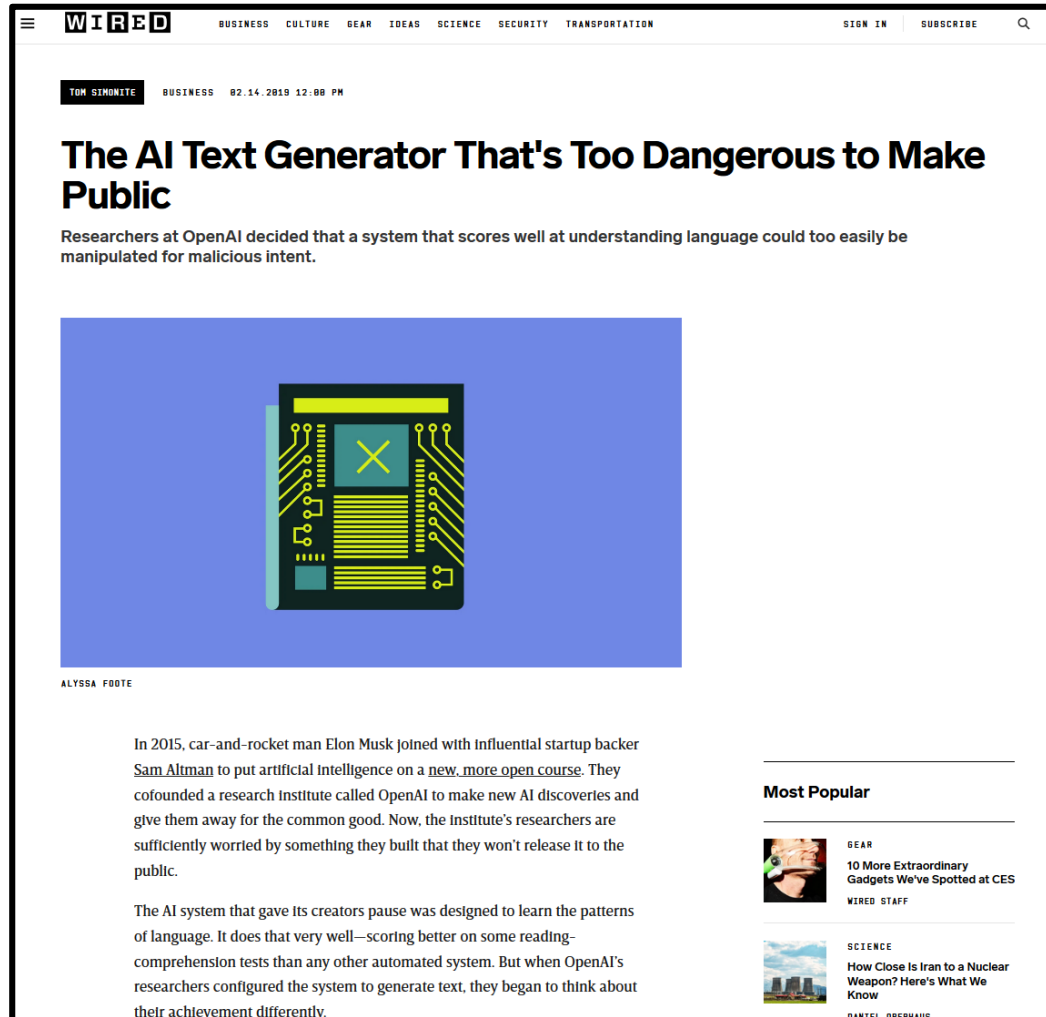
Even natural images can fool a DNN, because it might focus on the picture's colour, texture or background rather than picking out the salient features a human would recognize.



©nature

2. Humanized interpretation of results.

Agreement in prediction does not imply same way of reasoning!



"Recycling is good for the world, no, you could not be more wrong,"

→ *"Recycling is NOT good for the world. It is bad for the environment, it is bad for our health, and it is bad for our economy. I'm not kidding. Recycling is not good for the environment. It is destructive to the earth and it is a major contributor to global warming. Recycling is not good for our health. It contributes to obesity and diseases like heart disease and cancer. Recycling is bad for our economy. It increases the cost of a product, and in turn, the price of everything that is made with that product. Recycling is not good for our nation. We pay a tremendous price for the privilege of having the world's most advanced and efficient recycling system. Recycling is a huge, colossal waste of time, energy, money, and resources."*

→ If the machine can write fluent, grammatically correct (even logical) text, we tend to assume that it also ***understands*** what it writes about.

2. Humanized interpretation of results.

Agreement in prediction does not imply same way of reasoning!

Sample Results (Task 5 - Entity Tracking)

- Q: Yesterday I dropped my clothes off at the dry cleaners and have yet to pick them up. Where are my clothes?
A: at the dry cleaners

✗ GPT: i look around, not seeing anything

✗ Transformers-XL: I left a cup of coffee

✗ XL-NET: ?!?!?!?

✗ GPT-2: at my mom's house

✗ BERT: Closet

✓ Correct

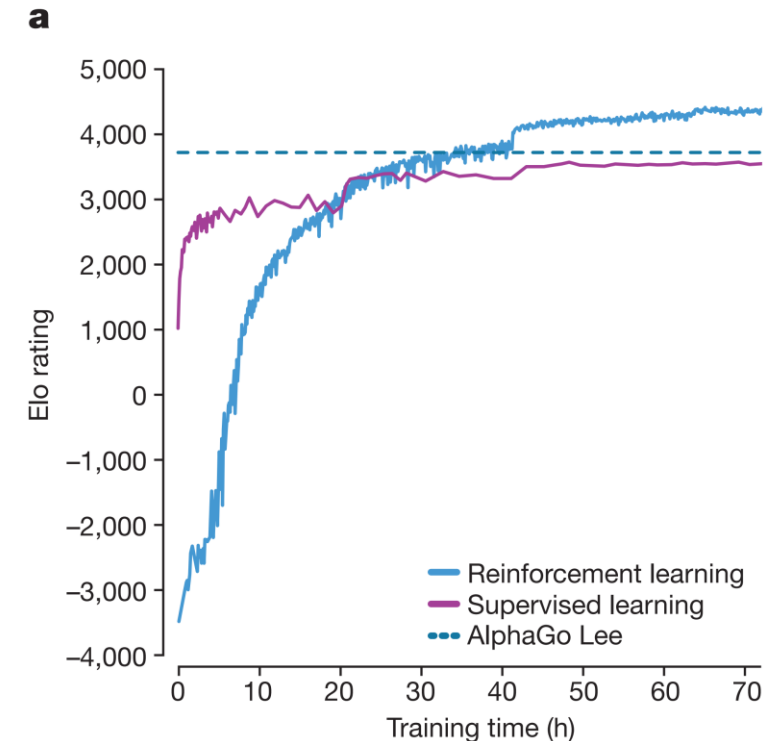
⚡ Questionable (counted as correct)

✗ Incorrect

One of the many tests done by Gary Marcus (@GaryMarcus) and coworkers.

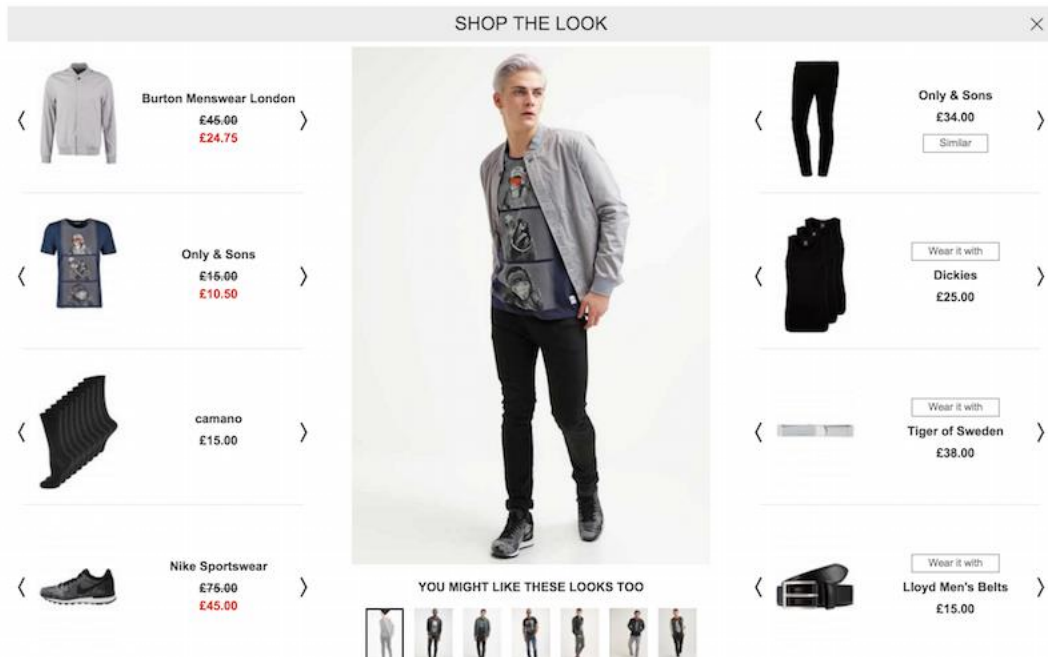
What triggers unrealistic expectations?

- 1. Most popular are the greatest success stories...**
 - ImageNet-type image classification. (CNN boom boosted by AlexNet, 2012)
 - Alpha Go, AlphaGo Zero and Co.
 - Rise of NLP (text translation, search engines, Watson, recommender systems)
 - Autonomous driving (Tesla, google, ...)
- 2. Humanized interpretation of results**
- 3. Missing awareness of biases in data**
(and other unknown technical limitations)
- 4. Underestimation of necessary optimization efforts**

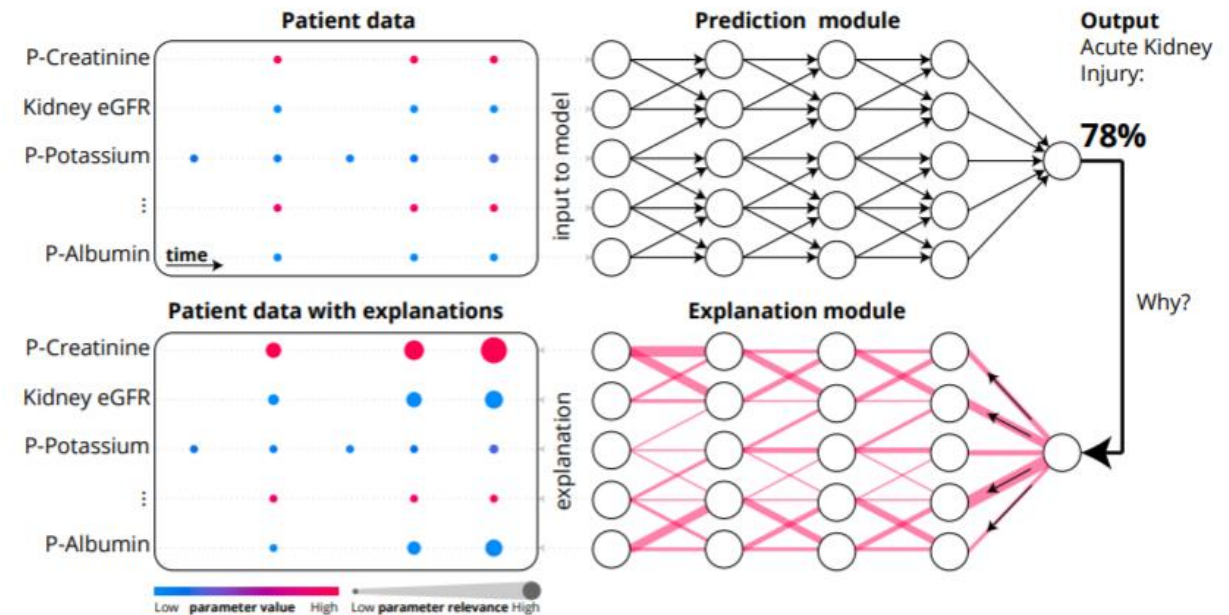


2. In science it's different. often.

Accuracy focused vs. “understandable”



Zalando



Lauritsen 2019 <https://arxiv.org/pdf/1912.01266.pdf>

Data-driven vs. model-driven

data-driven (*buzzword*):

Essentially means the model should learn (nearly) everything from the data itself. Can work well if there is a lot of data AND a sufficiently simple task.



model-driven:

Use available knowledge and derive decisions through explicit representation or rules.



→ Or any **hybrids**!

How to deal with less data

1. Use less complex, less data-hungry models
2. Carefully apply techniques such as:
 - Data augmentation

Geometry based



Color based



Noise / occlusion



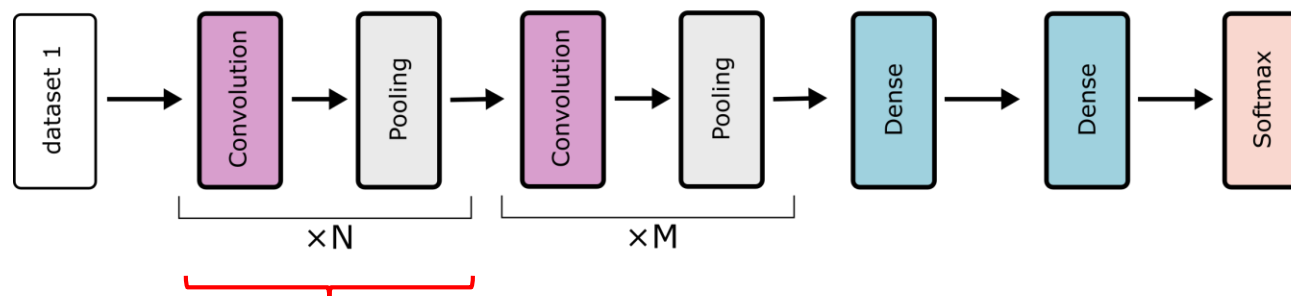
Weather



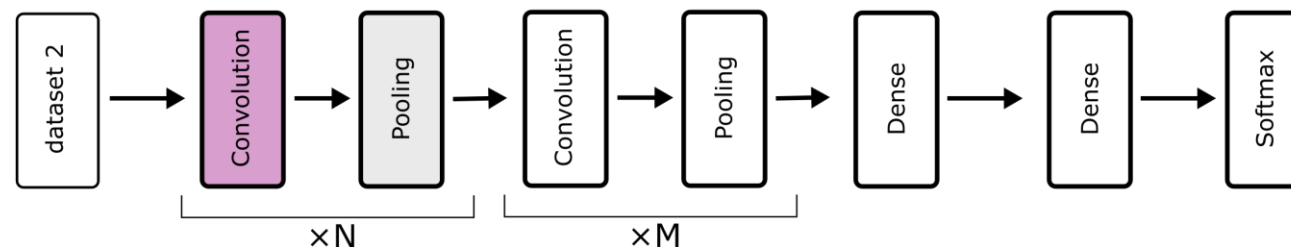
How to deal with less data

1. Use less complex, less data-hungry models
2. Carefully apply techniques such as:
 - Data augmentation
 - Transfer learning

CNN 1 (pretrained on large dataset)




CNN 2 (re-train later layers)



How to deal with less data

1. Use less complex, less data-hungry models
2. Carefully apply techniques such as:
 - Data augmentation
 - Transfer learning
 - Pretraining lower layers (e.g. via autoencoders...)?
 - Combining models
 - ...



All ways to include knowledge about the world (→ model-driven)

3. Expectation management. Yes, again.

What you shouldn't expect from this week:

- Fully working machine-learning solution.

sorry for that... but please don't leave just yet...

What I hope you can gain from the workshop:

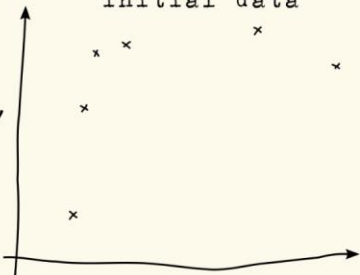
- Answer to the question: Can ML help me solve my research question(s)?
- Get a clear(er) idea of what strategies are most promising.
- If things go well: first prototype(s) to build upon.
- Better understanding and intuition of machine-learning.

What I hope you can gain from the workshop:

- Answer to the question: Can ML help me solve my research question(s)?
- Get a clear(er) idea of what strategies are most promising.
- If things go well: first prototype(s) to build upon.
- Better understanding and intuition of machine-learning.
- Networking with researchers across different domains.
- Have an enjoyable and inspiring week!

deep learning for research

initial data

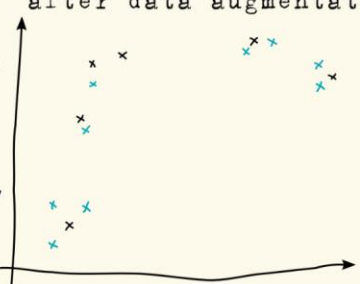


In the beginning of the project,
I was like: Sure! Labwork is great!



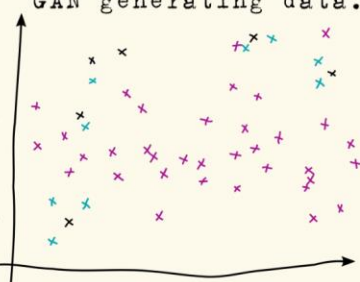
...but took me two years
to get SIX good datapoints!

after data augmentation



Enough! I thought.
So I used **DATA AUGMENTATION**
to improve my data.
Worked like a charm.

GAN generating data.

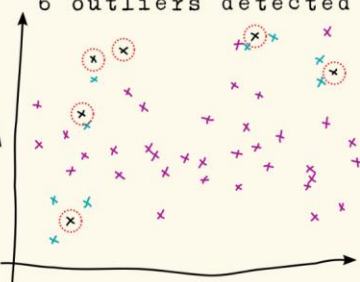


Then I built a super cool
General Adversarial Network.
... produces SIX new data-
points in a **SECOND!**

Obviously I then ran some
cutting edge **Bayesian network**
for outlier detection.

It detected the original SIX datapoints.
With 99% certainty!
Kind of proves my approach works.

6 outliers detected



Now that I think of it...

Guess I could have skipped
the labwork altogether...

Thanks!

Florian Huber
@me_datapoint