Towards Artificial Intelligence

Mirco Schönfeld
University of Bayreuth

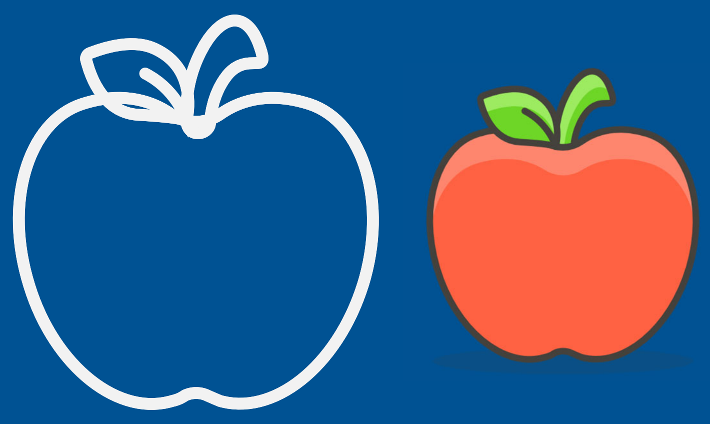
mirco.schoenfeld@uni-bayreuth.de
@TWlyY29

Types of Machine Learning

Unsupervised Learning

Data:
Just data, no labels

Goal:
Learn underlying structure



This thing is like the other thing

Supervised Learning

Data:
Labeled data

Goal:
Learn function mapping data to labels

This thing is an apple

Reinforcement Learning

Data:
State-action pairs

Goal:
Maximize future rewards over many time steps

Eat this thing because it keeps you healthy

Reinforcement Learning

UNIVERSITÄT BAYREUTH

General-purpose framework for artificial intelligence

- Based on the concept of *agents* having the capacity to *act*
- Each *action* influences the agent‘s future *state*
- Success is measurable by some *reward*

RL-based system have a *goal* or an *objective*

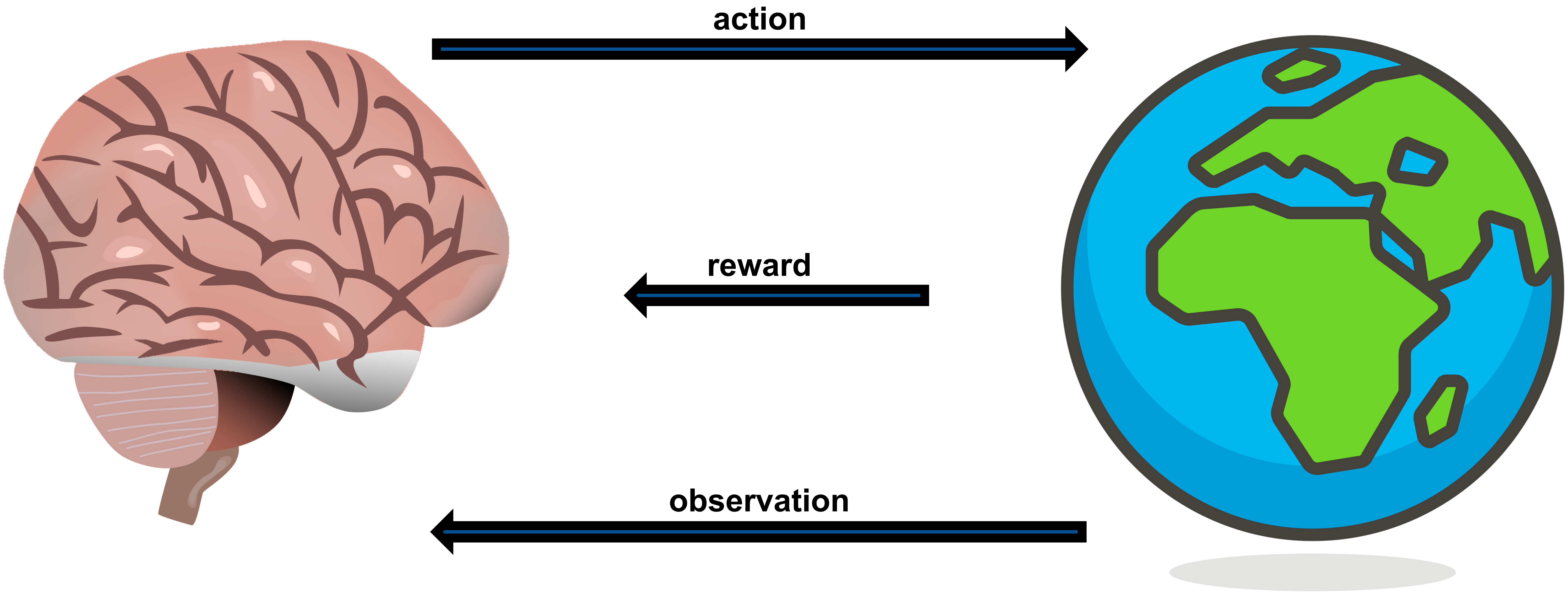
Ultimately, the aim is to learn a sequence of actions to maximise future reward

Reinforcement Learning vs. Un/Supervised Learning

UNIVERSITÄT BAYREUTH

- No human supervision, only a reward signal
- Feedback is delayed, not instantaneous
- Time matters (sequential data)
- Agent‘s actions affect subsequent actions

Key Concepts of (Reinforcement) Learning



https://commons.wikimedia.org/wiki/File:201704_brain.svg
https://commons.wikimedia.org/wiki/File:Globe-Africa-and-Europe%26Asiaparts.svg

Examples of (Reinforcement) Learning

- Fly stunt manoeuvres in a helicopter
- Manage an investment portfolio
- Make a humanoid robot walk
- Place many different computer games better than humans

How would you measure *success*?

Measuring Success in (Reinforcement) Learning

UNIVERSITÄT BAYREUTH

- Fly stunt manoeuvres in a helicopter
  + (positive reward) for following a desired trajectory
  - (negative reward) for crashing
- Manage an investment portfolio
  + for every € earned
- Make a humanoid robot walk
  + for forward motion
  - for falling over
- Place many different computer games better than humans
  + for increasing a score
  - for decreasing the score

Reward

A **reward** is a scalar feedback signal
indicating how well an agent is doing at a certain point in time

Immediate *feedback*

Agent‘s goal: maximise cumulative reward

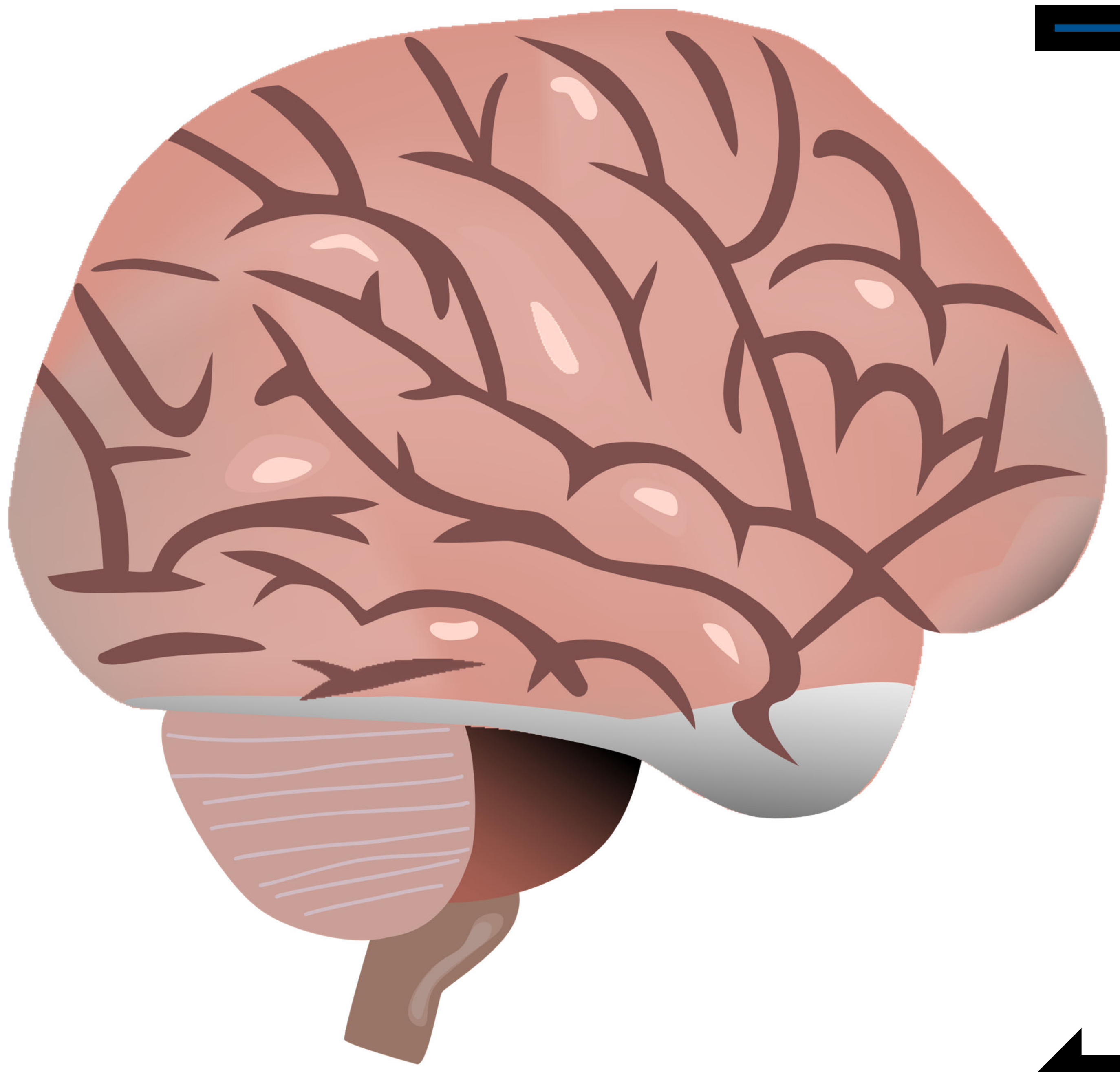
Learning to Succeed

Goal:

Select a sequence of actions to maximise *total future reward*

- Actions may have long term consequences
  e.g. a financial investment may take months to mature
- Reward may be delayed
  e.g. repairing a helicopter might prevent a crash in the future
- Sacrificing immediate reward might be better to gain more long-term reward
  e.g. blocking opponent moves might increase winning chances many moves later

Key Concepts of Reinforcement Learning



1. action

Action:
A move the agent can make in the environment. Chosen from **Action space *A***, i.e. the set of possible actions.

3. reward

Reward:
Feedback that measures success or failure of the agent‘s actions

2. observation

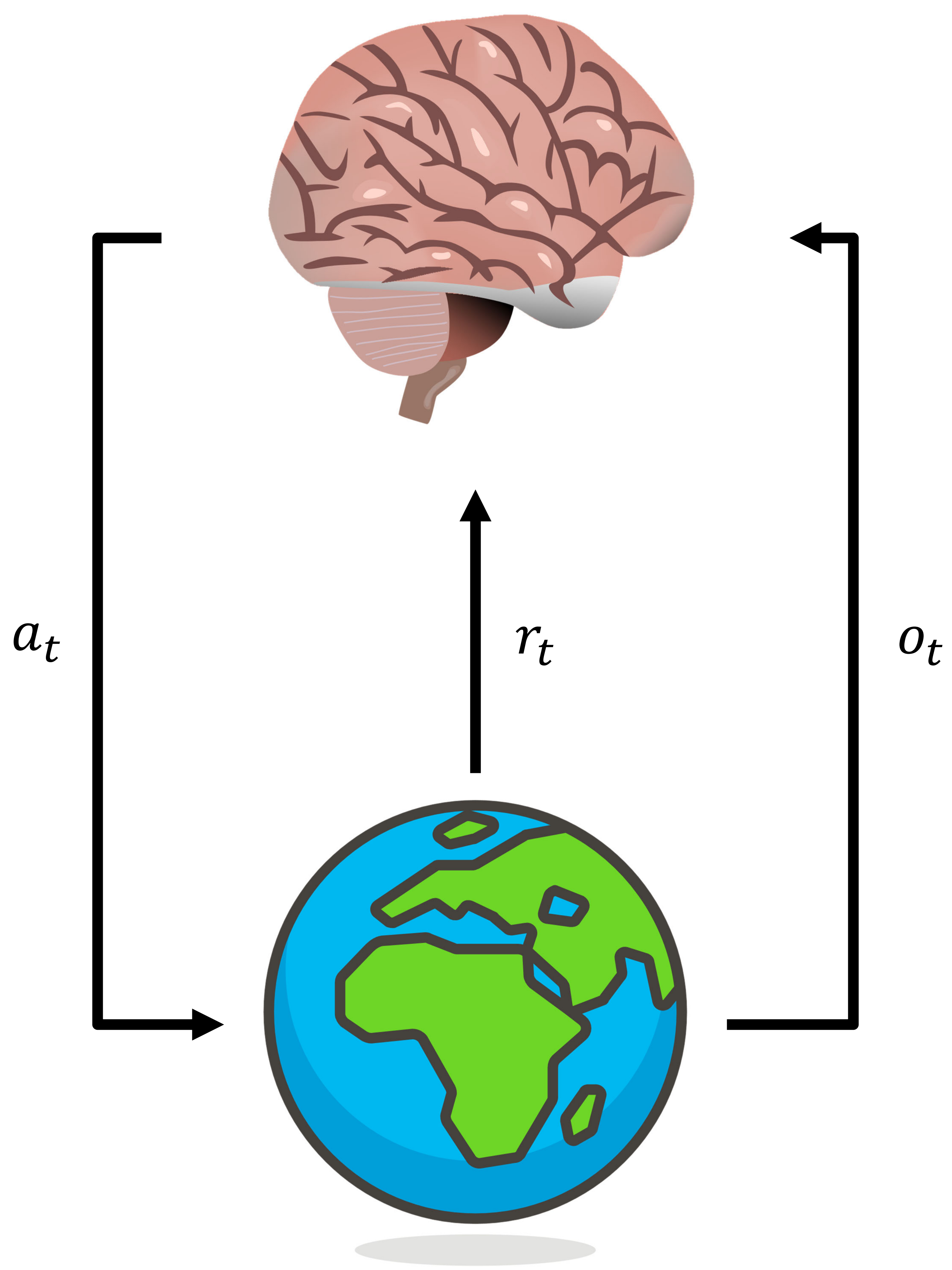
Observation:
Agent perceives the ***State*** of the environment after taking actions

Agent:
Takes actions

Environment:
The world in which the agent exists and operates

https://commons.wikimedia.org/wiki/File:201704_brain.svg
https://commons.wikimedia.org/wiki/File:Globe-Africa-and-Europe%26Asiaparts.svg

Reinforcement Learning



Reinforcement Learning is an iterative process

At each step *t* the agent:

- Receives observation o_t
- Receives scalar reward r_t
- Executes action a_t

The environment:

- Receives action a_t
- Emits observation o_{t+1}
- Emits scalar reward r_{t+1}

The next time step starts with an incremented *t*

Reinforcement Learning

UNIVERSITÄT BAYREUTH

RL is essentially trial-and-error learning

The environment is initially unknown. By interacting with the environment, the agent improves its policy

Exploration & exploitation need to be balanced

- Exploration finds more information about the environment
- Exploitation makes use of known information to maximise reward

Example:

Restaurant Selection

Stick with your favourite – *exploitation* – or try a new one – *exploration*?

Online Banner Advertisements

Show the most successful banner – *exploitation* – or show a different one – *exploration*?

Major Parts of RL Agents

Policy function (often denoted as π)

An agent‘s behaviour function. Essentially a map from state to action.

Can be represented deterministically or stochastically

Value function (often denoted as v_π)

Prediction of future reward. Evaluates goodness/badness of states

Selects between actions based on a particular policy

Model

Agent‘s representation of the environment. How the environment may work.

Predicts the next state and the next (immediate) reward

Approaches to Reinforcement Learning

Policy-based RL

- Search for the *optimal policy*
- Achieves maximum future reward from every state

Value-based RL

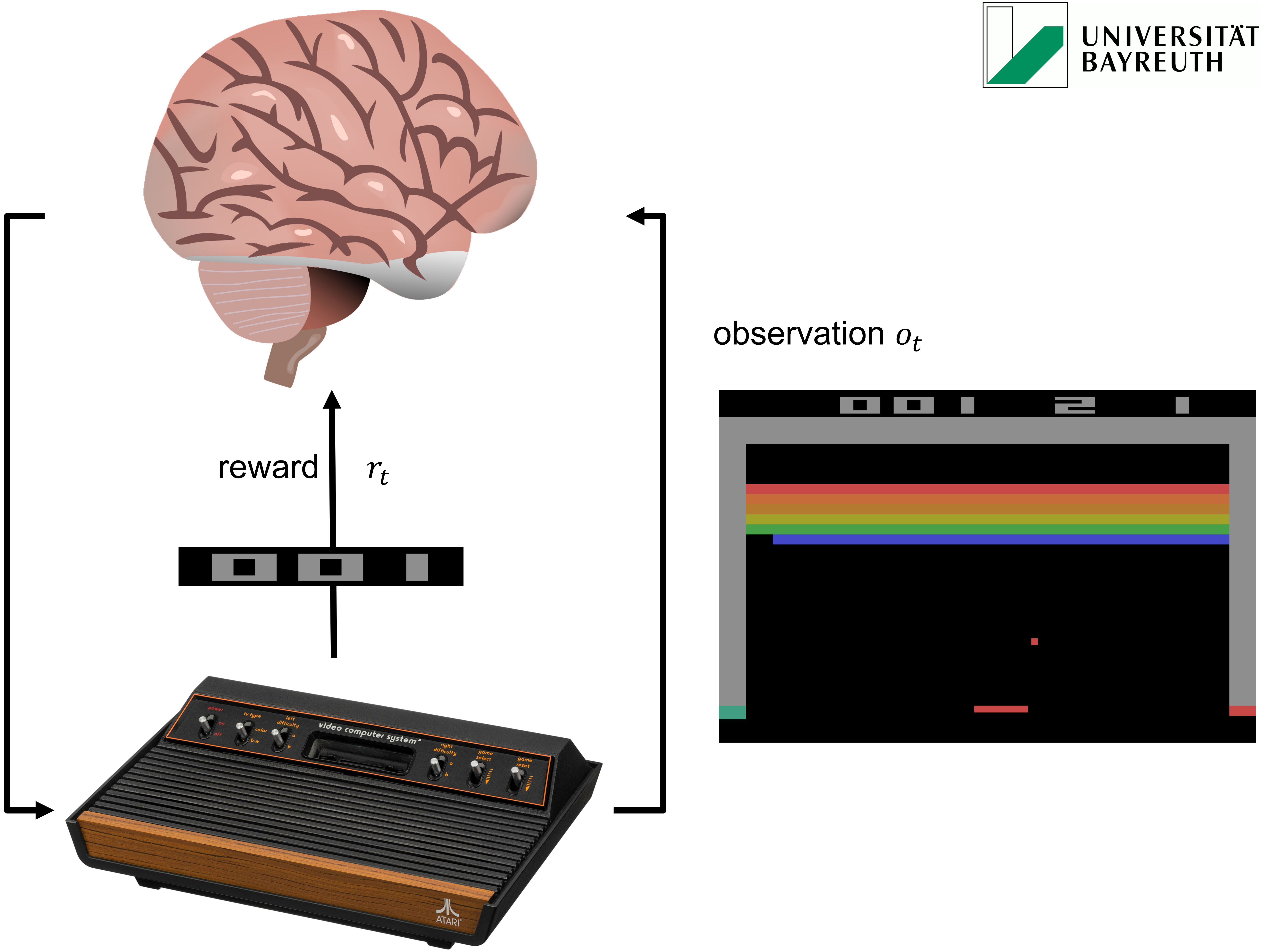
- Estimate the *optimal value function*
- Takes into account all possible ways to behave in particular situations (policies)

Model-based RL

- Build a transition model of the environment
- Plan using the model

Breakout Example

- Rules of the game are unknown
- Learn directly from interactive game-play
- Action set determined by joystick
- Pick actions on joystick
- Observe pixels and see scores



https://commons.wikimedia.org/wiki/File:Atari-2600-Joystick.jpg
https://en.wikipedia.org/wiki/File:Breakout2600.svg
Mnih, V., et al. (2013). Playing atari with deep reinforcement learning. *NIPS Deep Learning Workshop*, 2013.
Mnih, V., Kavukcuoglu, K., Silver, D. et al. Human-level control through deep reinforcement learning. Nature 518, 529–533 (2015).

Breakout Example

https://www.youtube.com/watch?v=V1eYniJ0Rnk

Mnih, V., Kavukcuoglu, K., Silver, D. et al. Human-level control through deep reinforcement learning. Nature 518, 529–533 (2015).

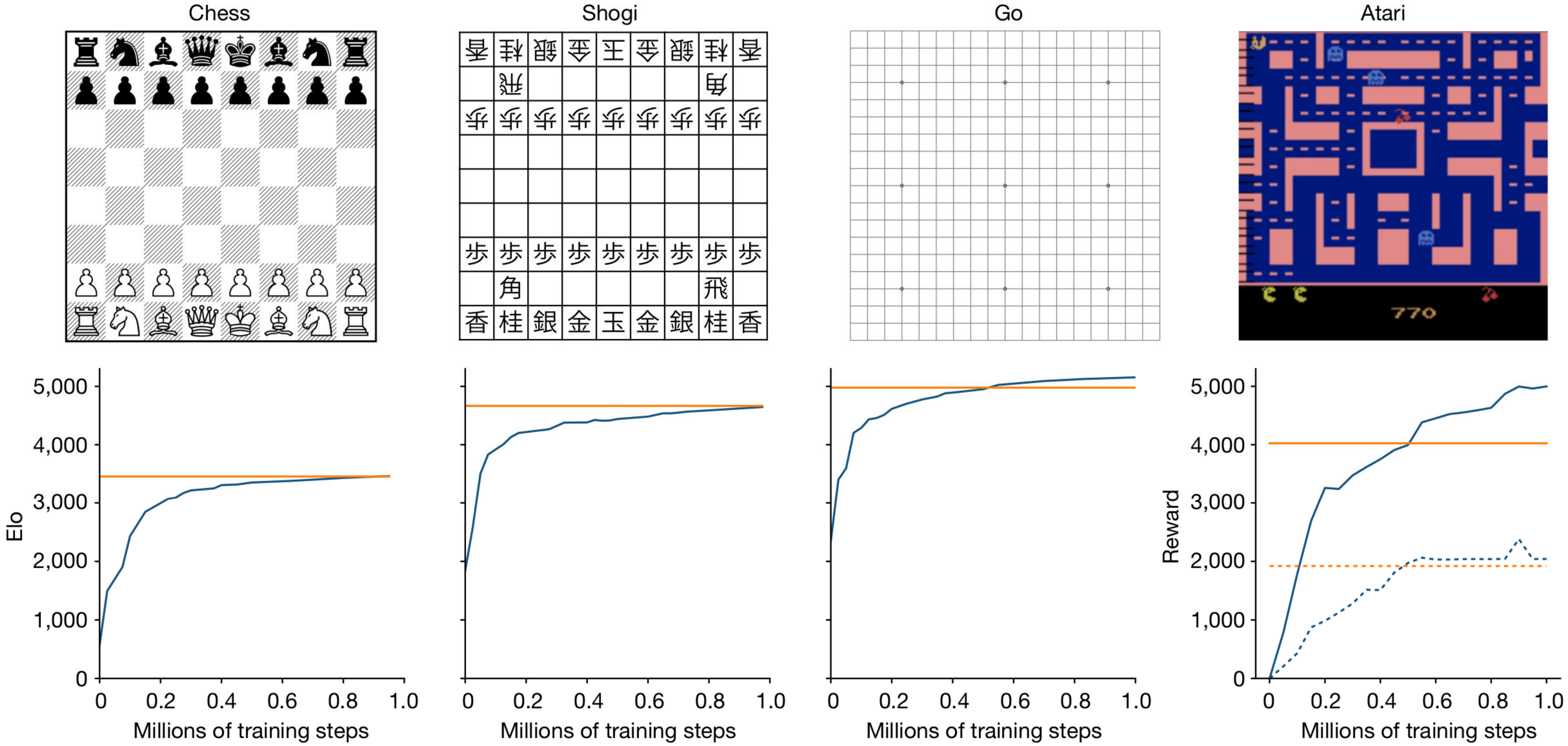
Deep Reinforcement Learning

Use a deep neural network to represent RL‘s value function / policy / model

Not necessarily limited to a fixed data set

Learns how to best accomplish its goal in diverse settings

Without any human supervision or guidance



Mnih, V., Kavukcuoglu, K., Silver, D. et al. Human-level control through deep reinforcement learning. Nature 518, 529–533 (2015).
Schrittwieser, J., Antonoglou, I., Hubert, T. et al. Mastering Atari, Go, chess and shogi by planning with a learned model. Nature 588, 604–609 (2020).

Deep Reinforcement Learning in NLP

UNIVERSITÄT BAYREUTH

Deep RL models are increasingly used in Natural Language Processing (NLP) tasks:

- Article summarization
- Question answering
- Dialogue generation
- Machine translation
- Text generation
- …

These models need to be BIG!

GPT-3:

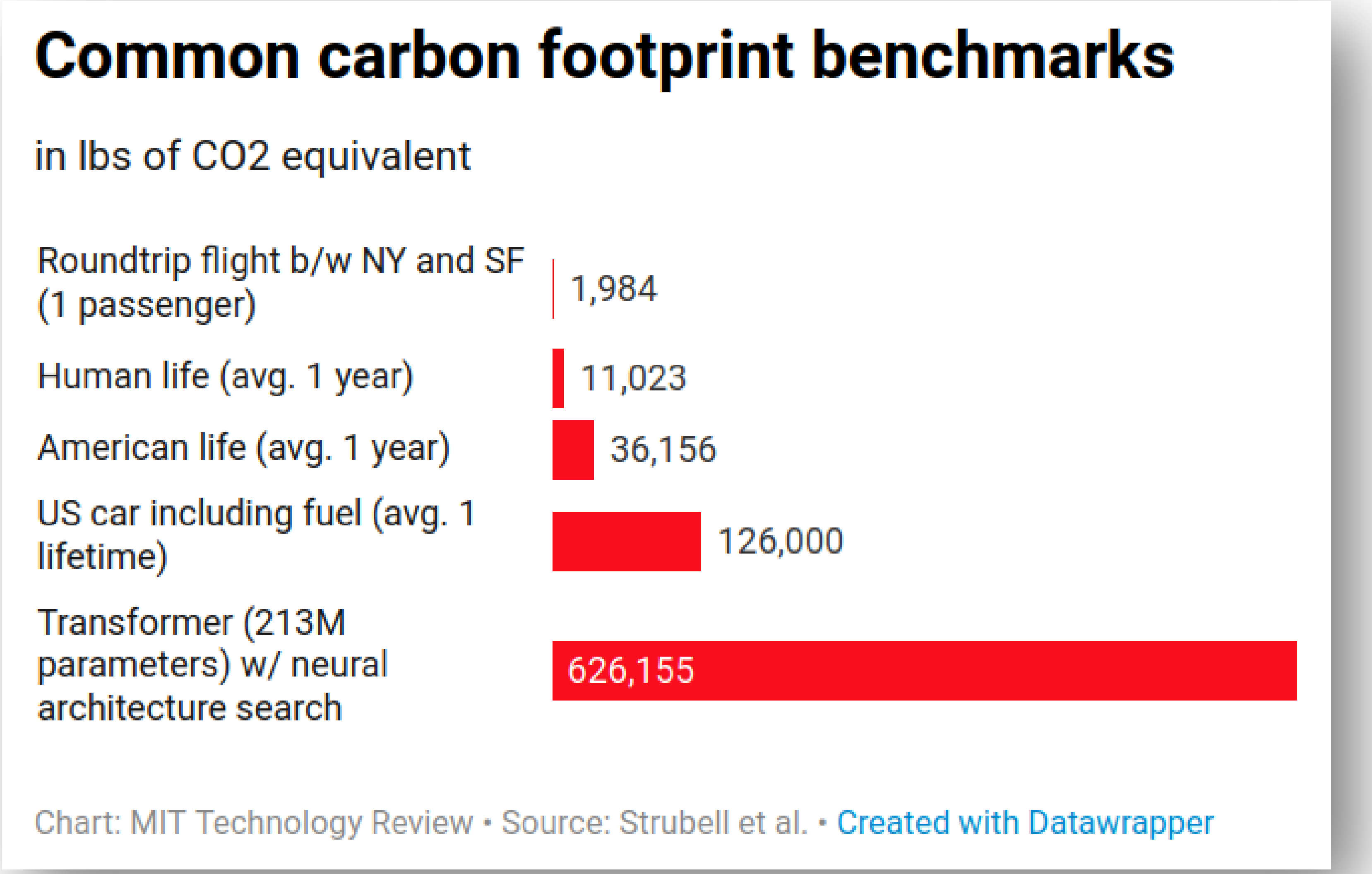
Trained on 45TB of text data. Has about 175 Billion parameters.

Only for English language

Large Models… why should we care?

UNIVERSITÄT BAYREUTH

Training large models consumes a lot of electricity.



Training one version of Google‘s language model, BERT, produced 1438 pounds of CO_2 – roughly a flight NY-SF-NY

Of course, models are trained and retrained many times over in practice.

At the same time,

- It is hard to audit training data checking for embedded biases
- It is even hard to prevent contamination of training & test data
- (Language) models don‘t actuall *understand* (language)

Is this *misdirected research effort*?

https://www.technologyreview.com/2020/12/04/1013294/
Strubell, E., Ganesh, A., & McCallum, A. (2019, July). Energy and Policy Considerations for Deep Learning in NLP. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics
Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜.
Brown, Tom B., et al. "Language models are few-shot learners." arXiv preprint arXiv:2005.14165 (2020).

Ethics in
Machine Learning & Artificial Intelligence

Why this matters

Goal of ML & AI models is to change people‘s behaviour

e.g. in recommendation settings where the goal is to make people buy more stuff

Creating these models is more than optimization & improving predictive accuracy

Technical design decisions suddenly have ethical implications for people‘s every day lifes

These ethical issues are complex and often not easy to answer

You won‘t find any answers in this section either ☺

Bias vs. Variance

We need to make assumptions to build effective machine learning algorithms
(remember the no-free-lunch theorem?)

Making assumptions leads to *bias* built into algorithms

Expected prediction error = bias^2 + variance + noise

- Bias: average prediction error over all data sets
- Variance: variation between solutions for different data sets (stability)
- Noise: deviation of measurements from the true value (unavoidable error)

Bias vs. Variance

UNIVERSITÄT BAYREUTH

Problem:

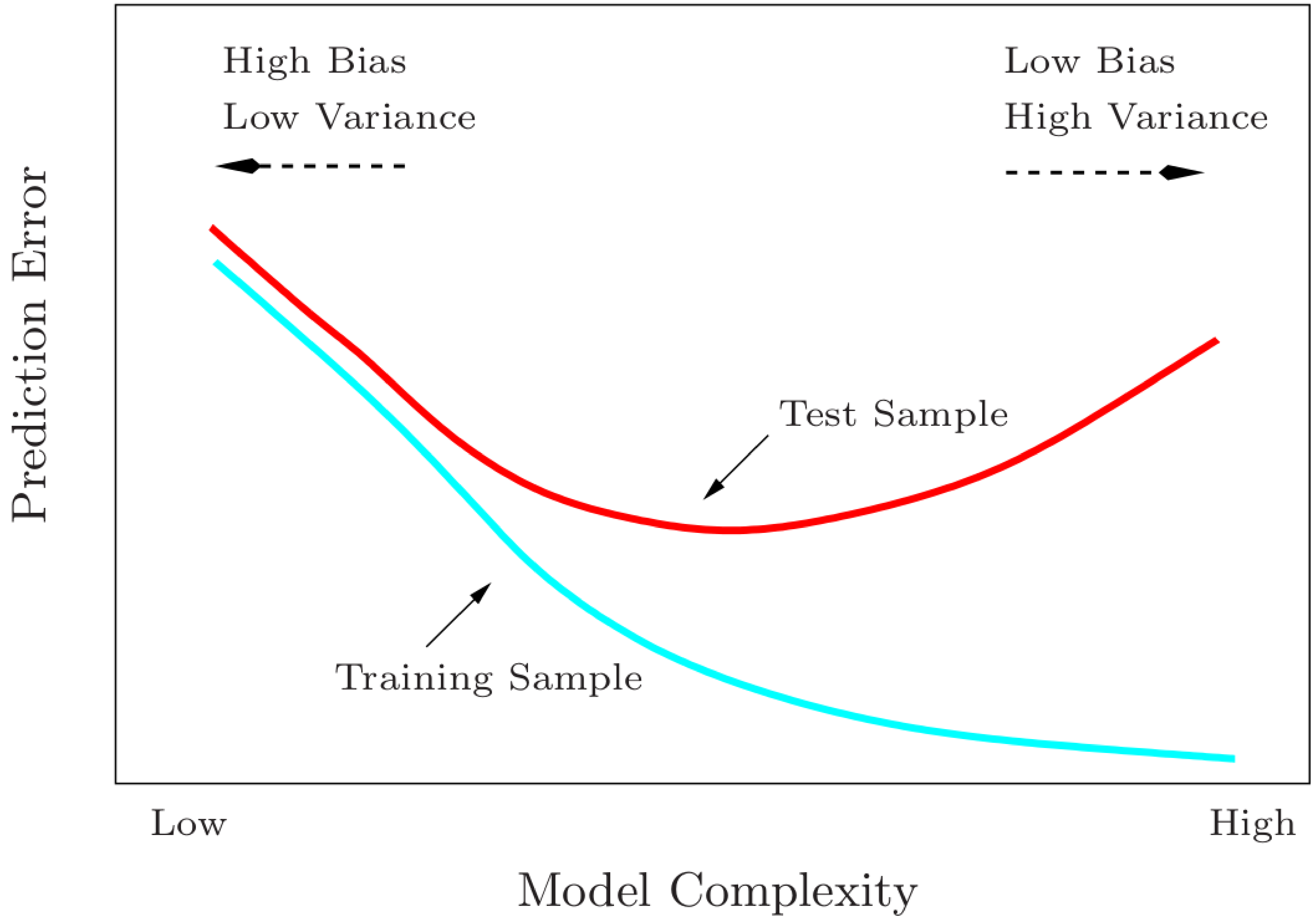
Low bias comes with high variance,
Low variance comes with high bias.

Bias too high:

Data isn‘t fit well, solutions too restricted

Bias too low:

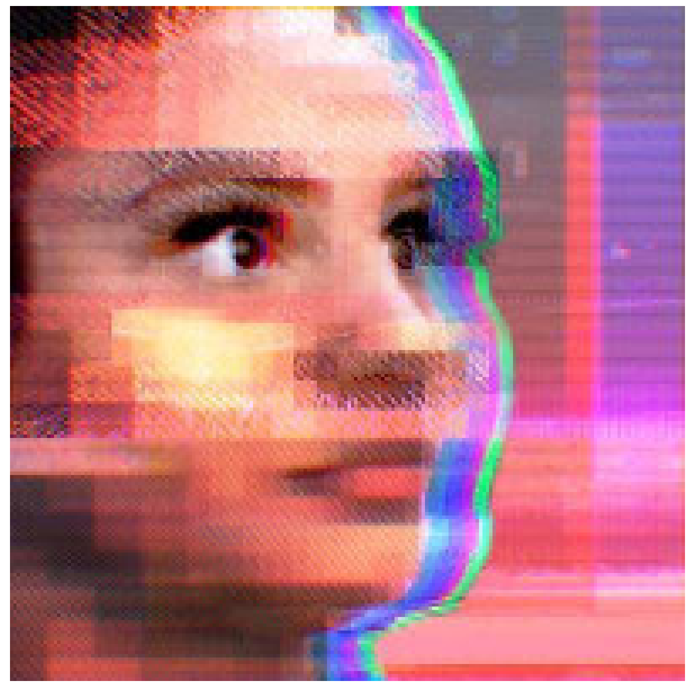
Variance too high, overfitting.



Hastie, T., Tibshirani, R. & Friedman, J. (2017). The Elements of Statistical Learning

Different types of bias

- Data not representative
- Data may have missing parts
- Training data may not reflect objectives
- Look at wrong metric
- Observing low bias by chance



> The fundamental assumption of every machine learning algorithm is that the past is correct, and anything coming in the future will be, and should be, like the past. This is a fine assumption to make when you are Netflix trying to predict what movie you'll like, but is immoral when applied to many other situations.
>
> Anthony Garvan

Garvan, A. (2016). Hey Microsoft, the Internet Made My Bot Racist, Too.
https://medium.com/@anthonygarvan/hey-microsoft-the-internet-made-my-bot-racist-too-d897fa847232

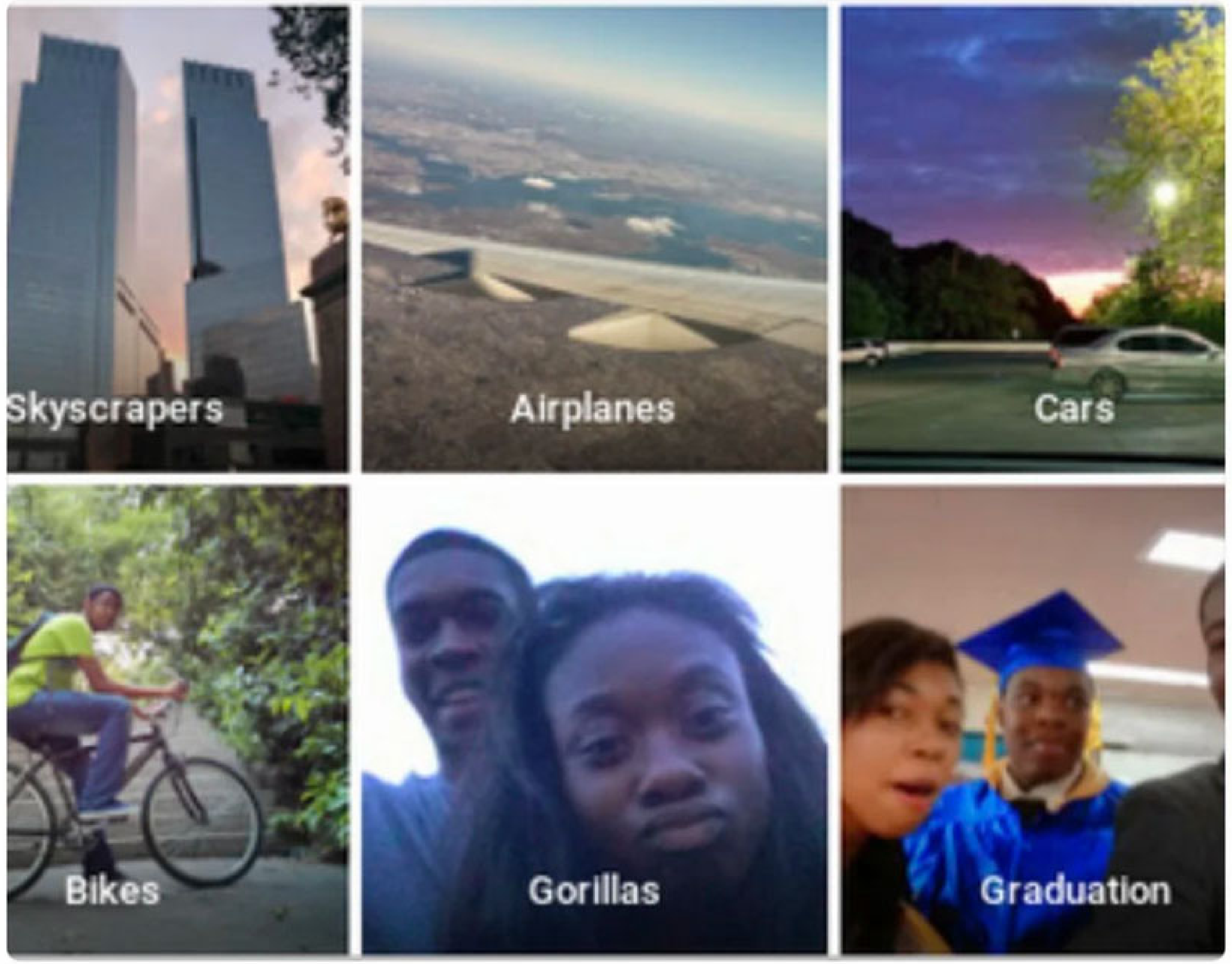
Terrance AB Johnson
@tweeterrance

#faceapp isn't't' just bad it's also racist...🔥 filter=bleach my skin and make my nose your opinion of European. No thanks #uninstalled

8:38 PM · Apr 19, 2017 · Twitter for iPhone

Jacky Alciné
@jackyalcine

Google Photos, y'all fucked up. My friend's not a gorilla.



6:22 am · 29 Jun. 15

UNIVERSITÄT BAYREUTH

Johanna Järvelä
@johannajarvela

In Finnish we have only one pronoun for third person regardless of the gender.

If you copy-paste the sentence below to google translate (or just click open original post for English translation), you see how the algorithm has learnt to be sexist.

8:12 AM · Mar 9, 2021 · Twitter Web App

Hän on journalisti. Hän on johtaja. Hän on uupunut. Hänellä on lapsenlapsi. Hän tekee töitä. Hänellä on päänsärkyä. Hänellä on hieno auto. Hän hoitaa lasta. Hän hoitaa hommat.

Kamera Keskustelu Litteroi

ENGLANTI

He is a journalist. He is a leader. She is exhausted. She has a grandchild. He works. She has a headache. He has a great car. She is taking care of the child. He takes care of things.

UNIVERSITÄT BAYREUTH

Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match

A New Jersey man was accused of shoplifting and trying to hit an officer with a car. He is the third known Black man to be wrongfully arrested based on face recognition.

PRO PUBLICA

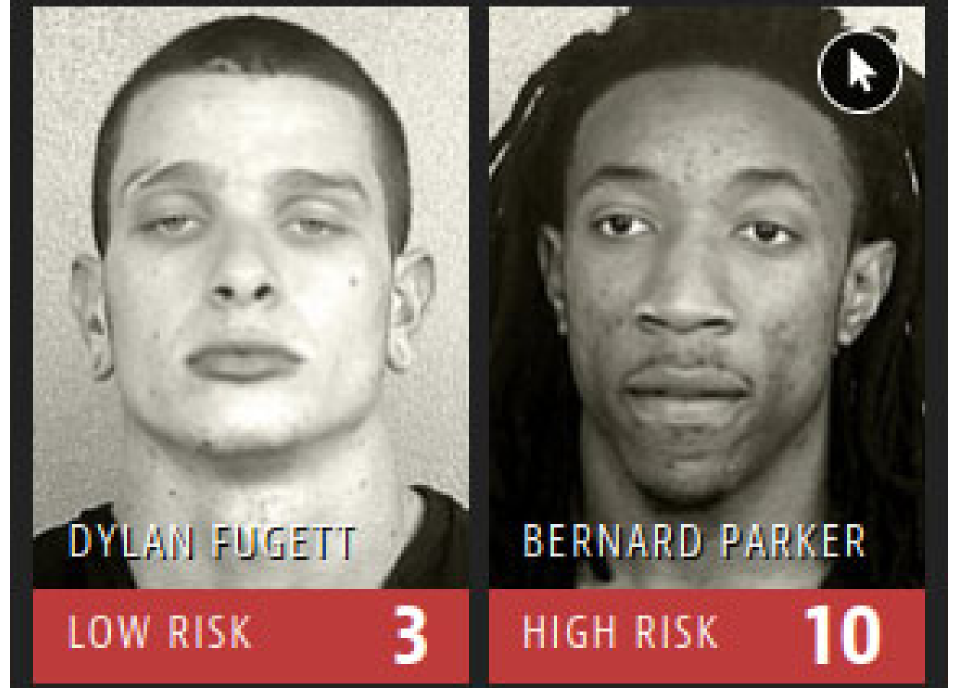
Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Two Drug Possession Arrests



Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html
https://www.nytimes.com/2019/12/19/technology/facial-recognition-bias.html
https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

A War of Words Puts Facebook at the Center of Myanmar's Rohingya Crisis

By Megan Specia and Paul Mozur
Oct. 27, 2017

Across Myanmar, Denial of Ethnic Cleansing and Loathing of Rohingya

By Hannah Beech
Oct. 24, 2017

"Kalar are not welcome here because they are violent and they multiply like crazy, with so many wives and children," he said.

Mr. Aye Swe admitted he had never met a Muslim before, adding, "I have to thank Facebook because it is giving me the true information in Myanmar."

Facebook fires human editors, algorithm immediately posts fake news

Facebook makes its Trending feature fully automated, with mixed results.

ANNALEE NEWITZ - 8/29/2016, 8:20 PM

Facebook admits it was used to 'incite offline violence' in Myanmar

6 November 2018

Rohingya sue Facebook for $150bn over Myanmar hate speech

7 December

Social Media platforms are not neutral

- Revenue model is based on clicks/impressions
- Involves experiments with content, recommendations, …
- Controls and filters available to users & advertisers

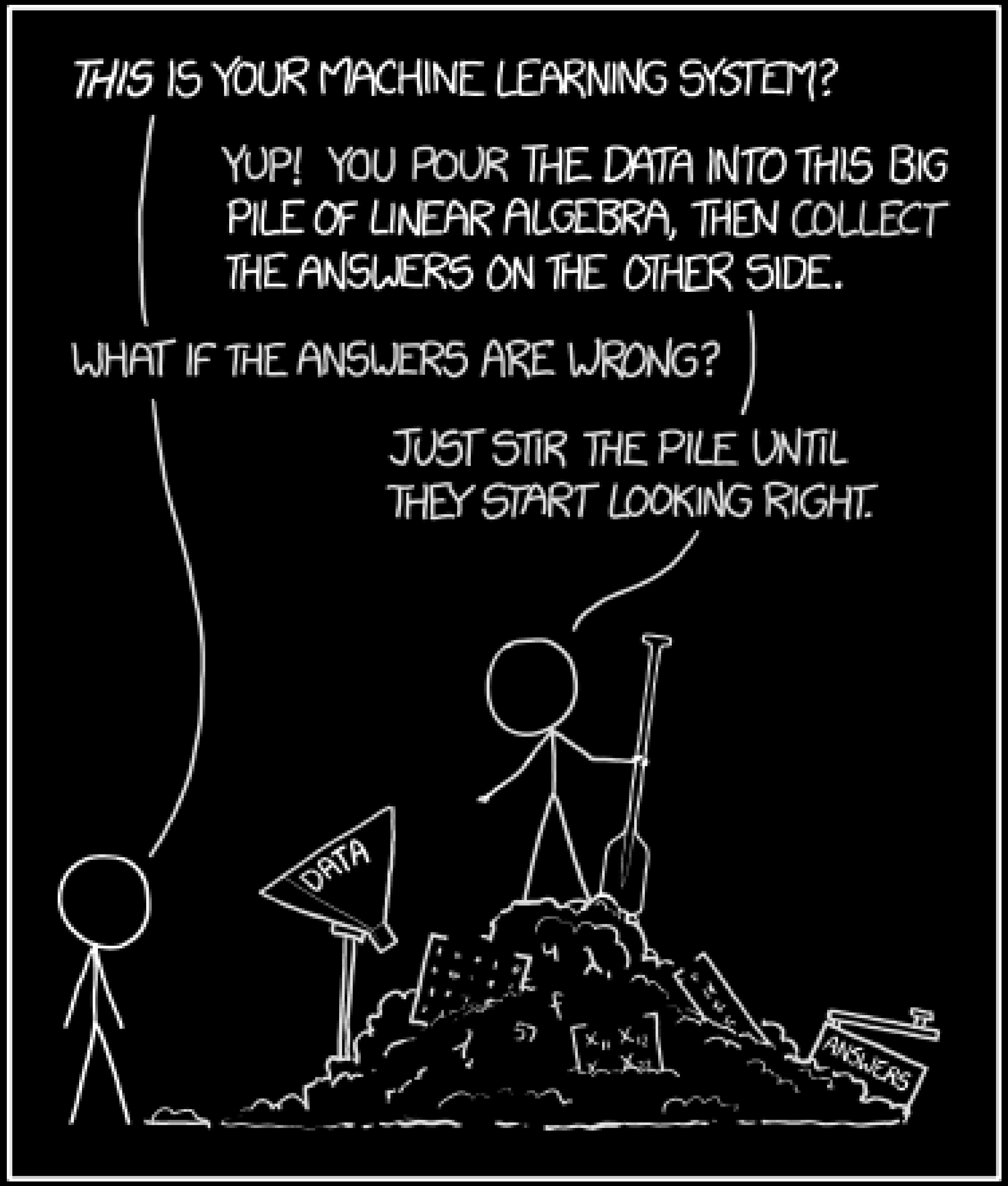
https://www.nytimes.com/2017/10/27/world/asia/myanmar-government-facebook-rohingya.html
https://www.nytimes.com/2017/10/24/world/asia/myanmar-rohingya-ethnic-cleansing.html
https://arstechnica.com/information-technology/2016/08/facebook-fires-human-editors-algorithm-immediately-posts-fake-news/
https://www.bbc.com/news/world-asia-46105934
https://www.bbc.com/news/world-asia-59558090

Questions to keep in mind

- What bias may be in the data?
- How diverse is the team that built it?
- What are error rates for different sub-groups?
- What is the accuracy of a simple rule-based alternative?
- How are appeals or mistakes being handled?

"We have an ethical obligation to not teach machines to be prejudiced.

Evan Estola, 27.05.2016
https://www.youtube.com/watch?v=MqoRzNhrTnQ



https://xkcd.com/1838/

Thanks.

mirco.schoenfeld@uni-bayreuth.de