Artificial Neural Network: from zero-to-hero

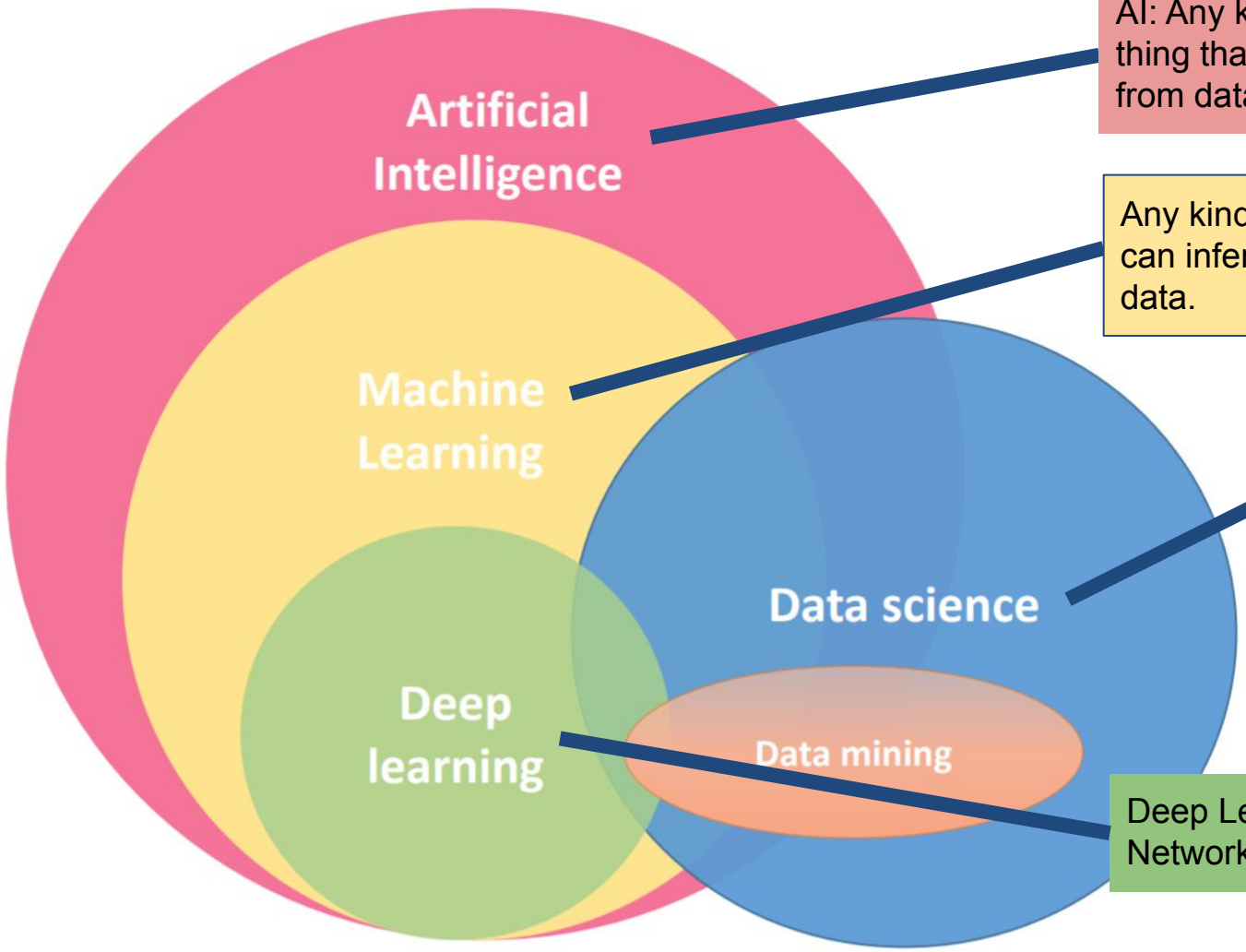
Mohamadou Diallo

OUTLINE

1. Demystifying Machine Learning (ML) Methods
2. The Machine Learning Landscape: *What, Why and How?*
3. Example: Artificial Neural Network (ANN)
   - *Data and Architecture*
   - *Forward Propagation*
   - *Cost Function*
   - *Gradient Descent*
   - *Stochastic (Numerical, Batch) Gradient Descent*
   - *Backpropagation*
   - *Training, Testing and Cross-Validating*
   - *Overfitting and Regularisation*
   - *Visualisation Tools*
4. Learning Resources
5. Example ML Papers
6. Ethical Questions

Demystifying Machine Learning Approaches

Machine Learning in the Field of Data Science



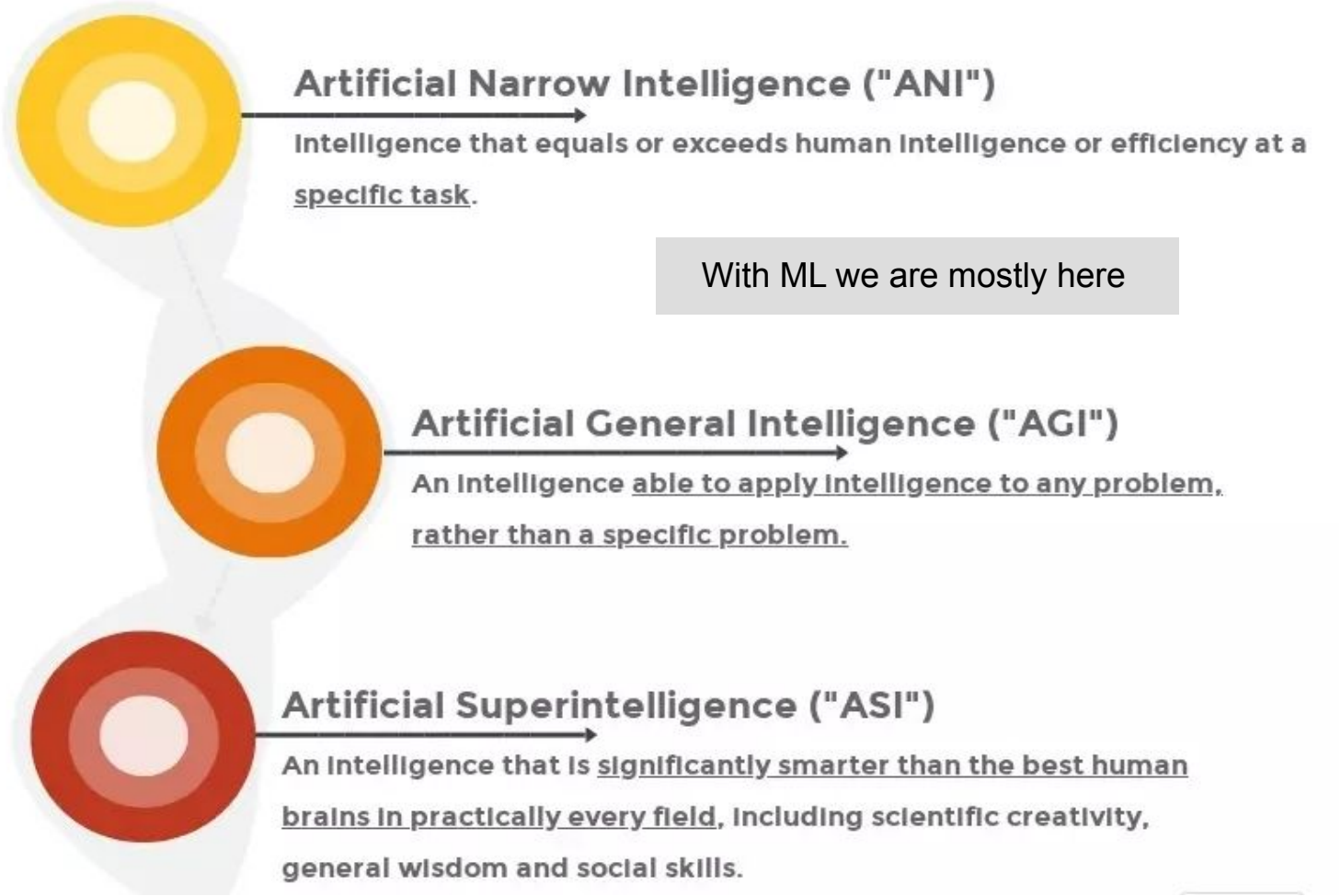
AI: Any kind of man made thing that can infer conclusions from data

Any kind of algorithm that can infer conclusions from data.

Any kind of thing that a human can do with data, e.g. also visualization, storage management.

Deep Learning: Neural Networks with a lot of Layers

Demystifying the Artificial Intelligence (AI) Methods



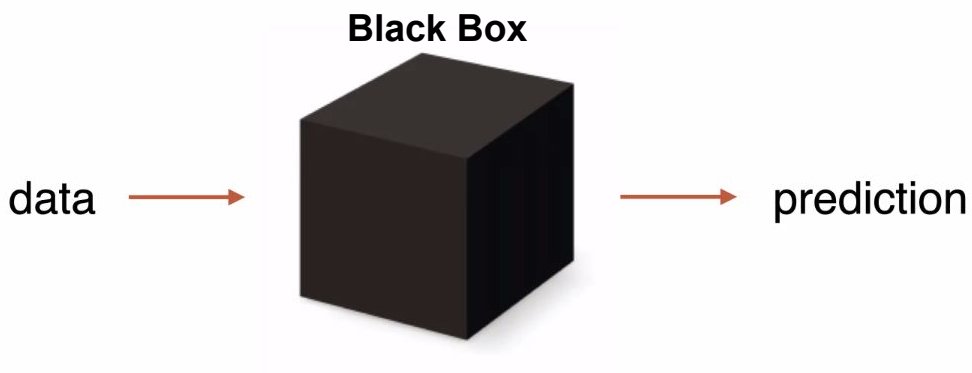
Myth 1: ML *can already solve general problems more efficient than humans: "Just plug it in and it solves everything mentality"*

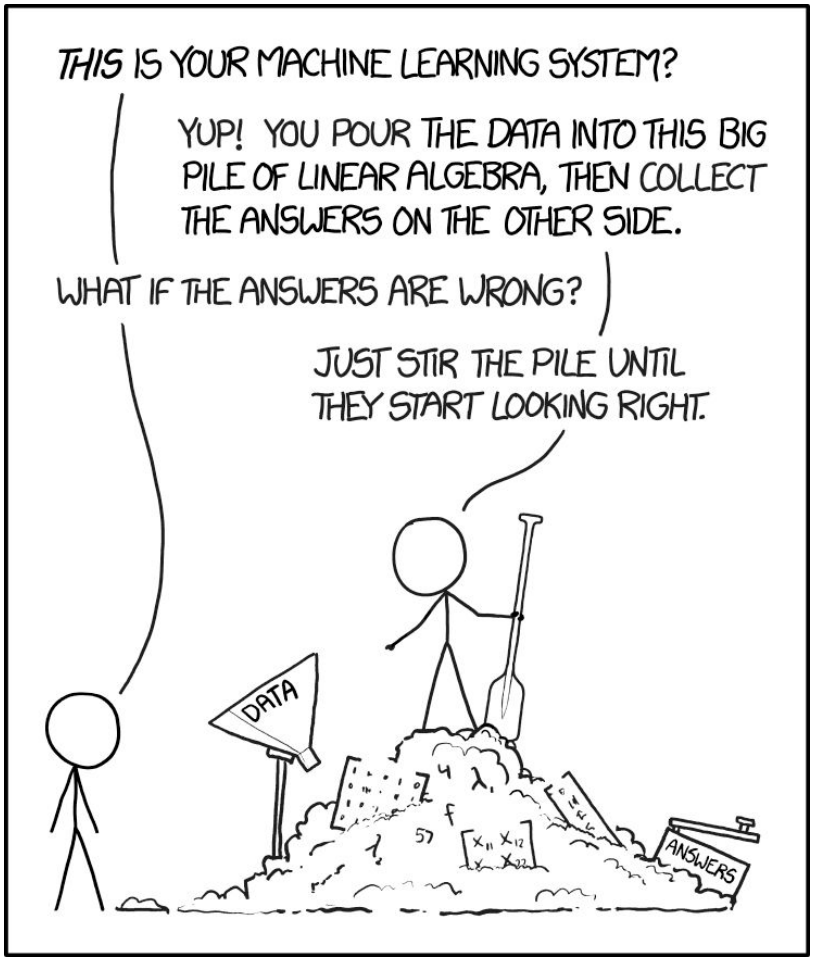
The ML Landscape: What, why and How?

What Is Machine Learning?

"ML is the science and art of programming computers to learn from data and to automatically improve through experience."

- A. Samuel, 1959 & T. Mitchell, 1997





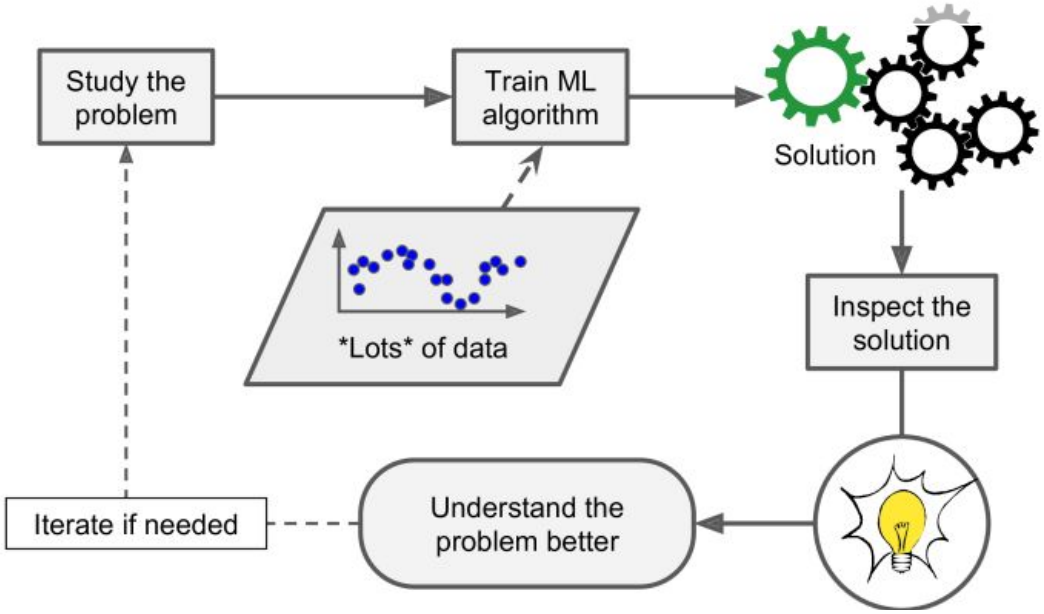
Myth 2: *The general idea of ML is an incredible difficult thing to understand.*

Why Use Machine Learning?

- ML algorithm can often *simplify code*, *perform better*, *find a solution/insights* and *adapt to new data*:

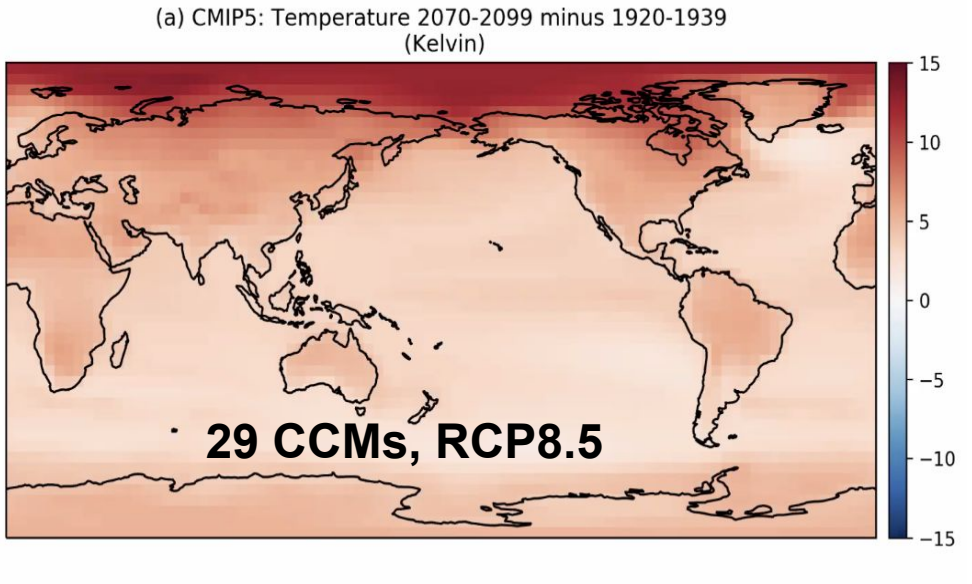
  a) Complex problems with a lot of hand-tuning or long lists of rules or *no good solution at all using a traditional approach,*
  b) *Fluctuating environments,*
  c) *Large amounts of data.*

- ML can help human to *learn and discover patterns* that were not immediately apparent.



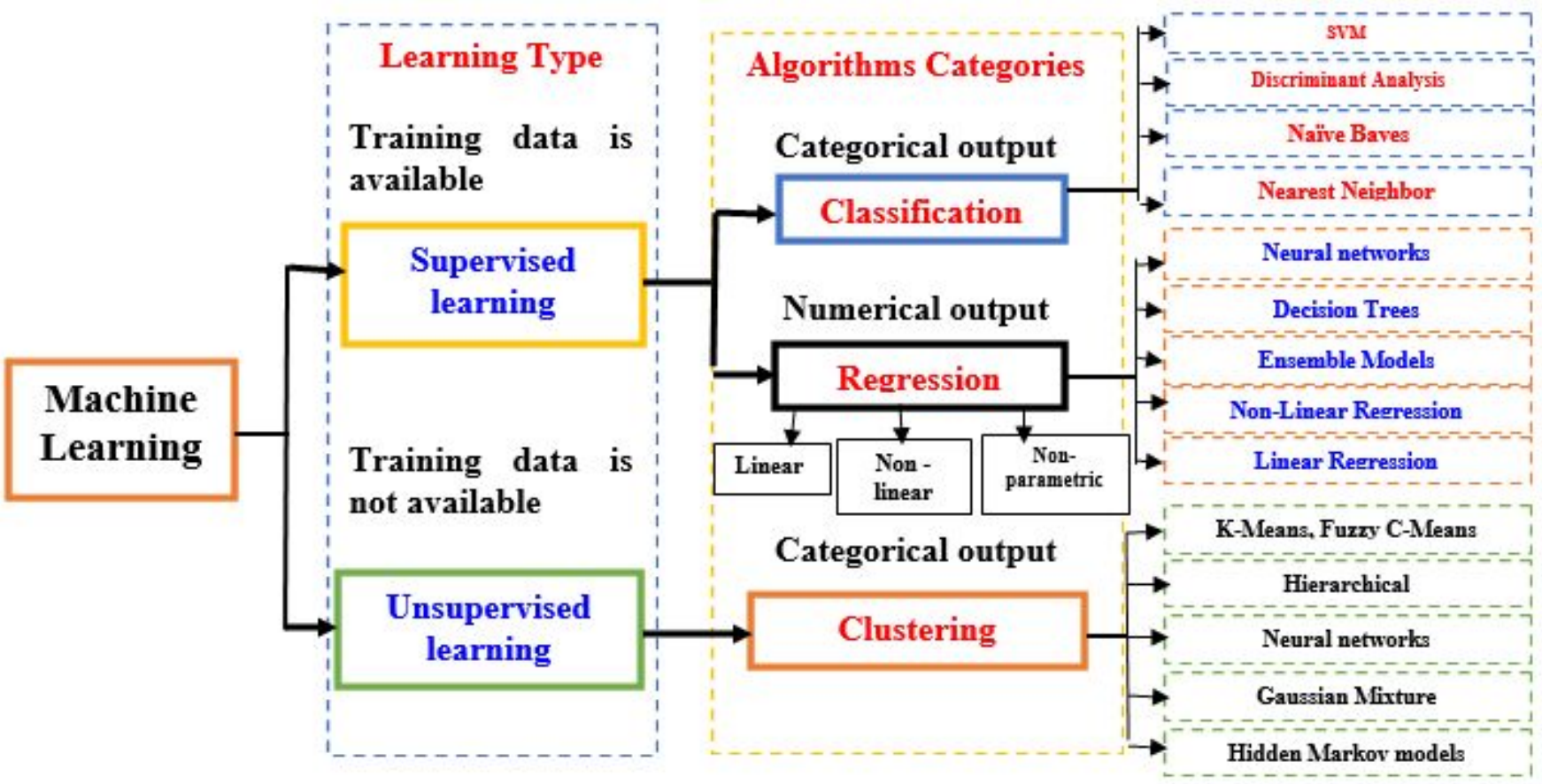
Why Use Machine Learning in Climate Science?

- The Earth atmosphere is *noisy* across timescales.
- Separating the climate change ***signal*** from the ***noise*** in our observed Earth system is not easy
  - weather noise can interfere with understanding teleconnections on weekly timescales,
  - yearly to multi-decadal variability can interfere with identifying the human impact on climate change
- Two sources of models' uncertainty:
  - Structural model biases *(i.e. simulating the Physics)*
  - Internal variability *(i.e. climate noise)*
- How can we tell which ***changes*** are ***Climate signal*** and which are ***Chaotic noise*** in our observed Earth System?



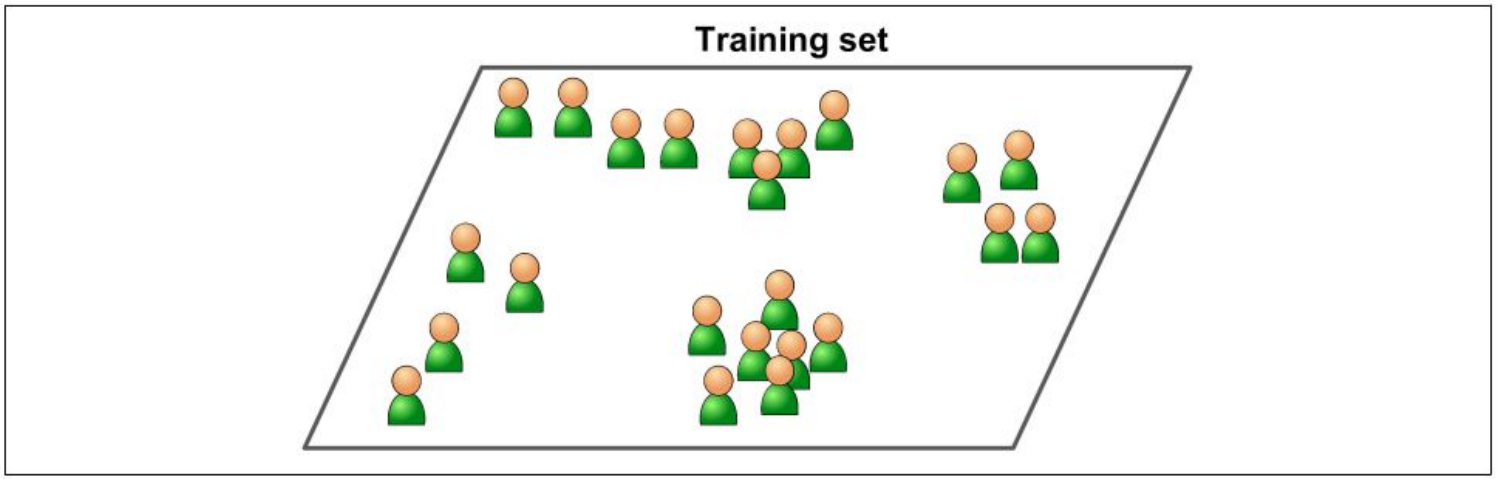
Types of Machine Learning Systems

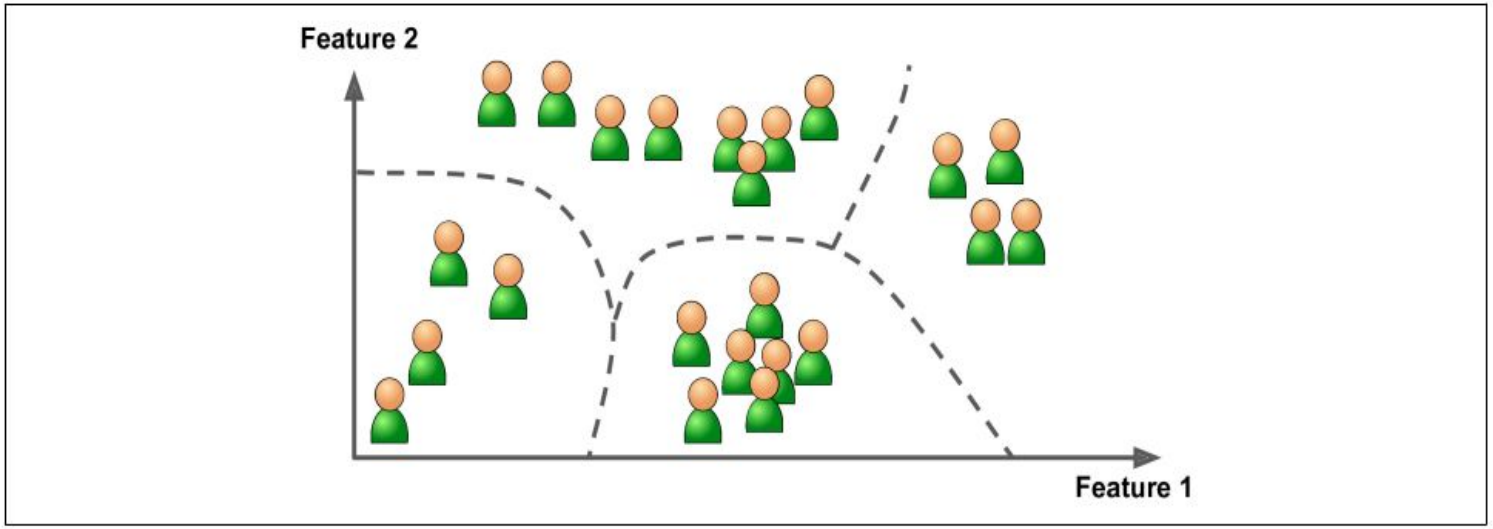
- ***No-free-lunch-theorem***: There is no one best ML method for all problems.



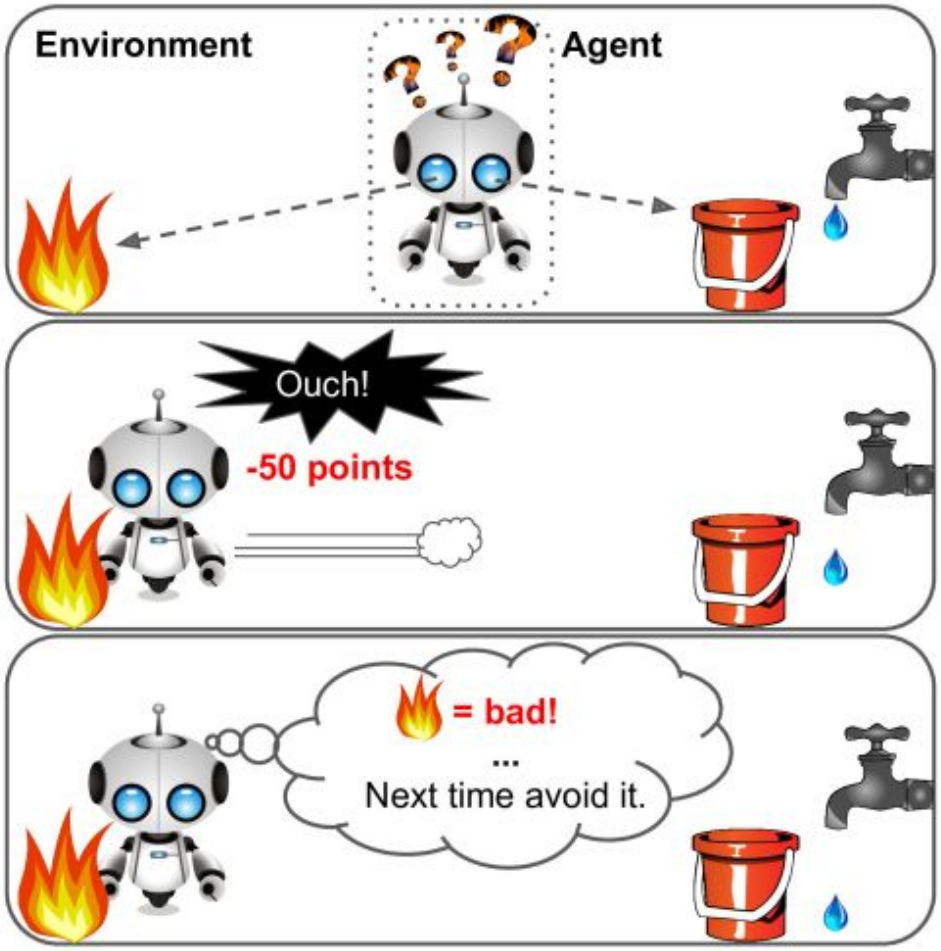
Supervised Learning Example: *Regression*

Unsupervised Learning Example: *Clustering*





Reinforcement Learning Example: *Walking Robot and AlphaGo*



1. Observe
2. Select action using policy
3. Action!
4. Get reward or penalty
5. Update policy (learning step)
6. Iterate until an optimal policy is found

Early usage of Machine Learning in Climate Science

- The application of ML for atmospheric science dates back to 1960's.

An Application of Adaptive Logic to Meteorological Prediction

1964

H. R. GLAHN

U. S. Weather Bureau, Washington, D. C.

(Manuscript received 12 May 1964, in revised form 25 July 1964)

New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Longwave Radiation in a Climate Model

2004

VLADIMIR M. KRASNOPOLSKY

NOAA/NCEP/SAIC, Camp Springs, and Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

MICHAEL S. FOX-RABINOVITZ AND DMITRY V. CHALIKOV

Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

(Manuscript received 26 April 2004, in final form 25 October 2004)

Nonlinear Principal Component Analysis by Neural Networks: Theory and Application to the Lorenz System

1999

ADAM H. MONAHAN

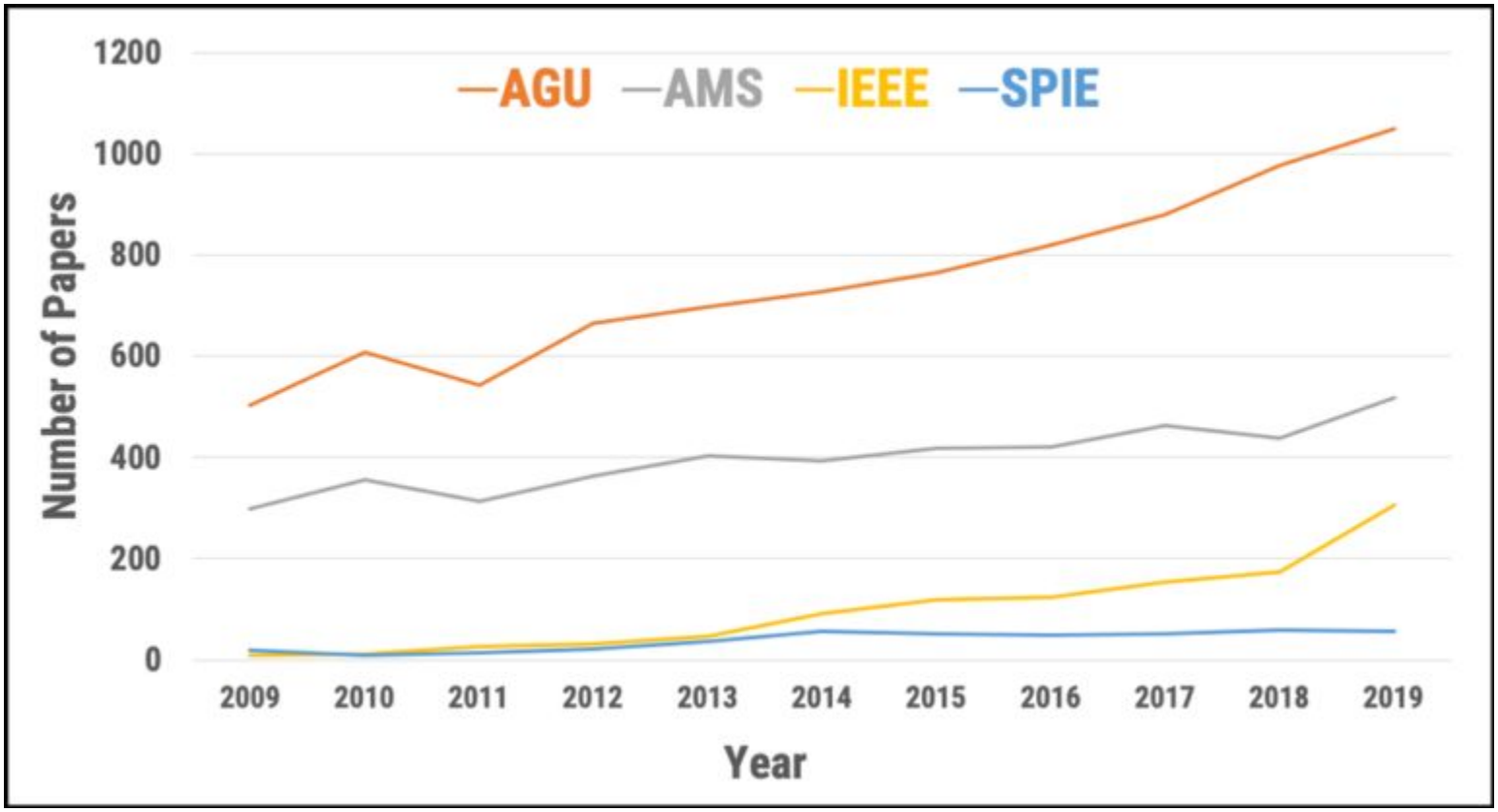
Oceanography Unit, Department of Earth and Ocean Sciences, and Crisis Points Group, Peter Wall Institute for Advanced Studies, University of British Columbia, Vancouver, British Columbia, Canada

(Manuscript received 28 October 1998, in final form 7 May 1999)

Myth 3: *ML is a brand new thing of the last decade.*

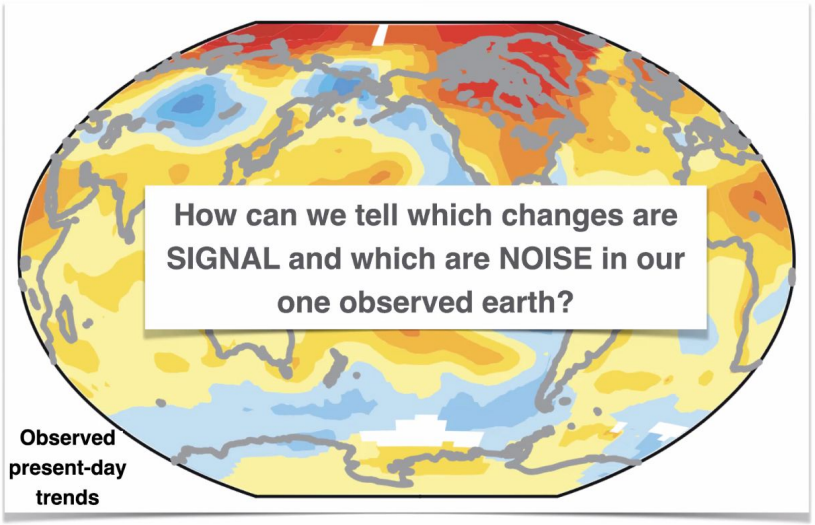
Evolution in using Machine Learning in Atmos. Science

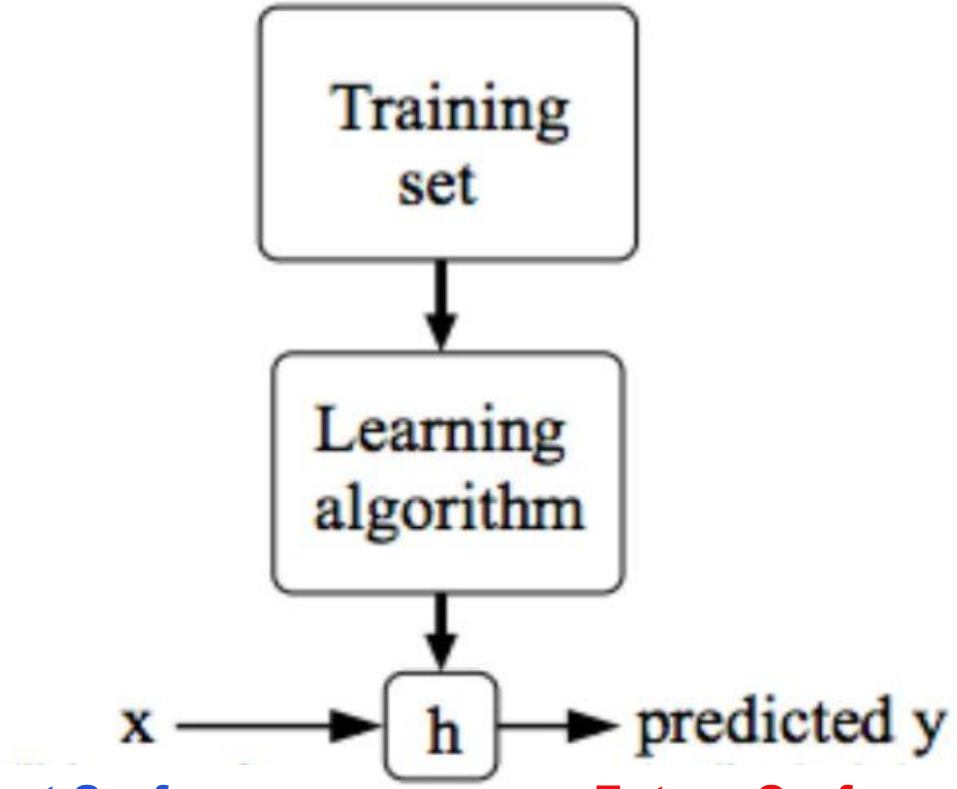
- The application of ML for atmospheric science has increased faster in the 3 last years.



Artificial Neural Network (ANN)

Mathematical Formulation of a Causality Problem





Present Surface Temperatures

Future Surface Temperature Trends

Algorithm: *Example of a Linear Formulation*

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

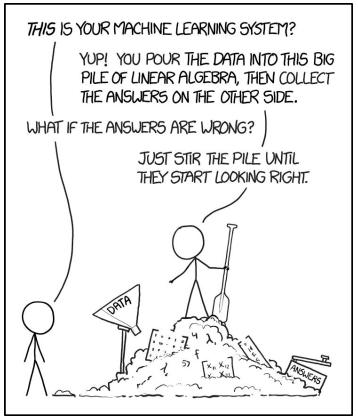
Parameters: θ_0, θ_1

Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$

Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

- In a classical regression problem, the ***minimization*** is done by calculating the derivative of the ***cost function (J)*** toward $\boldsymbol{\Theta_0}$ and $\boldsymbol{\Theta_1}$.
- This is ending up on solving Linear Algebra problem (Sometimes Very Expensive).

How to Solve the same Problem with ML?



Have some function $J(\theta_0, \theta_1)$

Want $\min_{\theta_0, \theta_1} J(\theta_0, \theta_1)$

Outline:

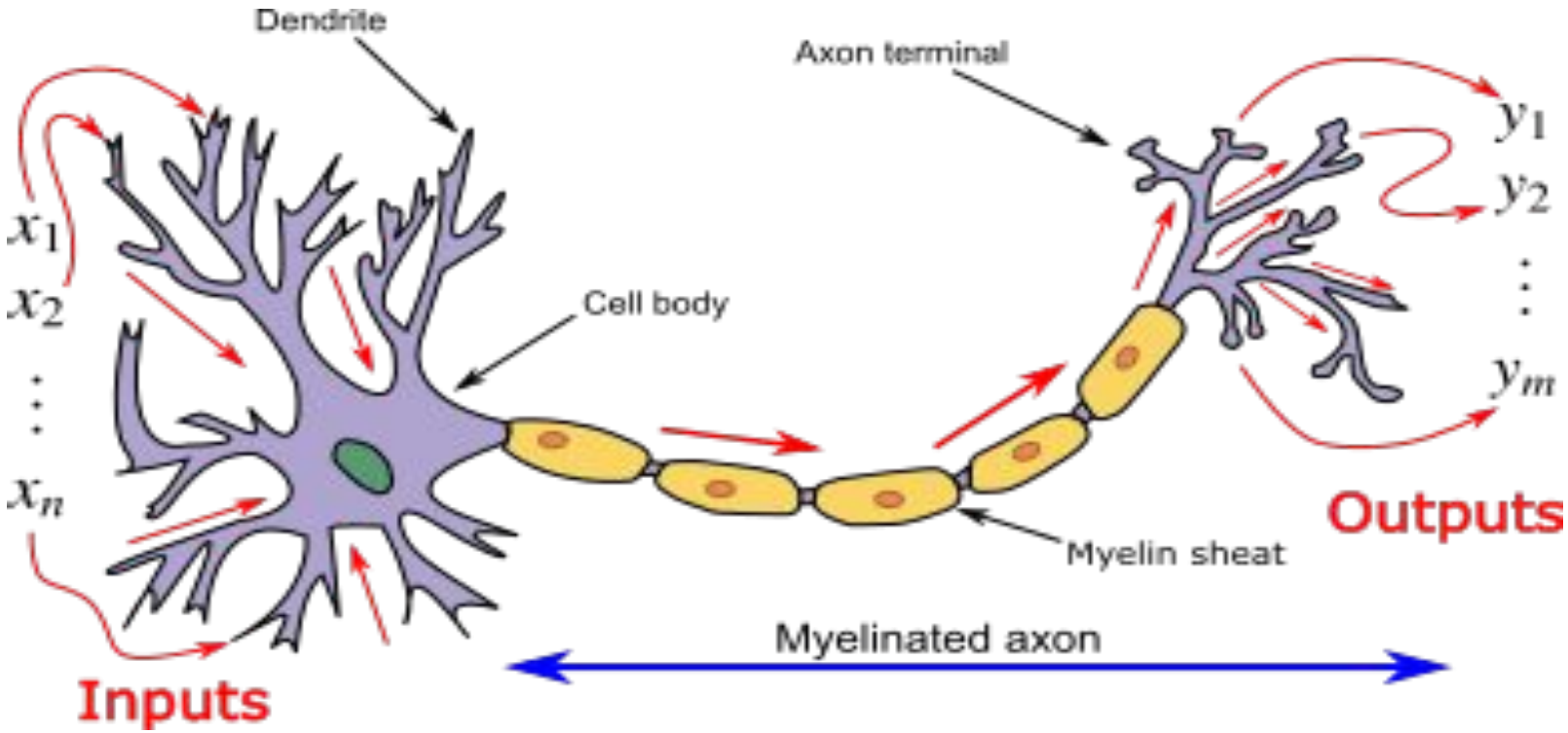
- Start with some θ_0, θ_1
- Keep changing θ_0, θ_1 to reduce $J(\theta_0, \theta_1)$

  until we hopefully end up at a minimum

Andr

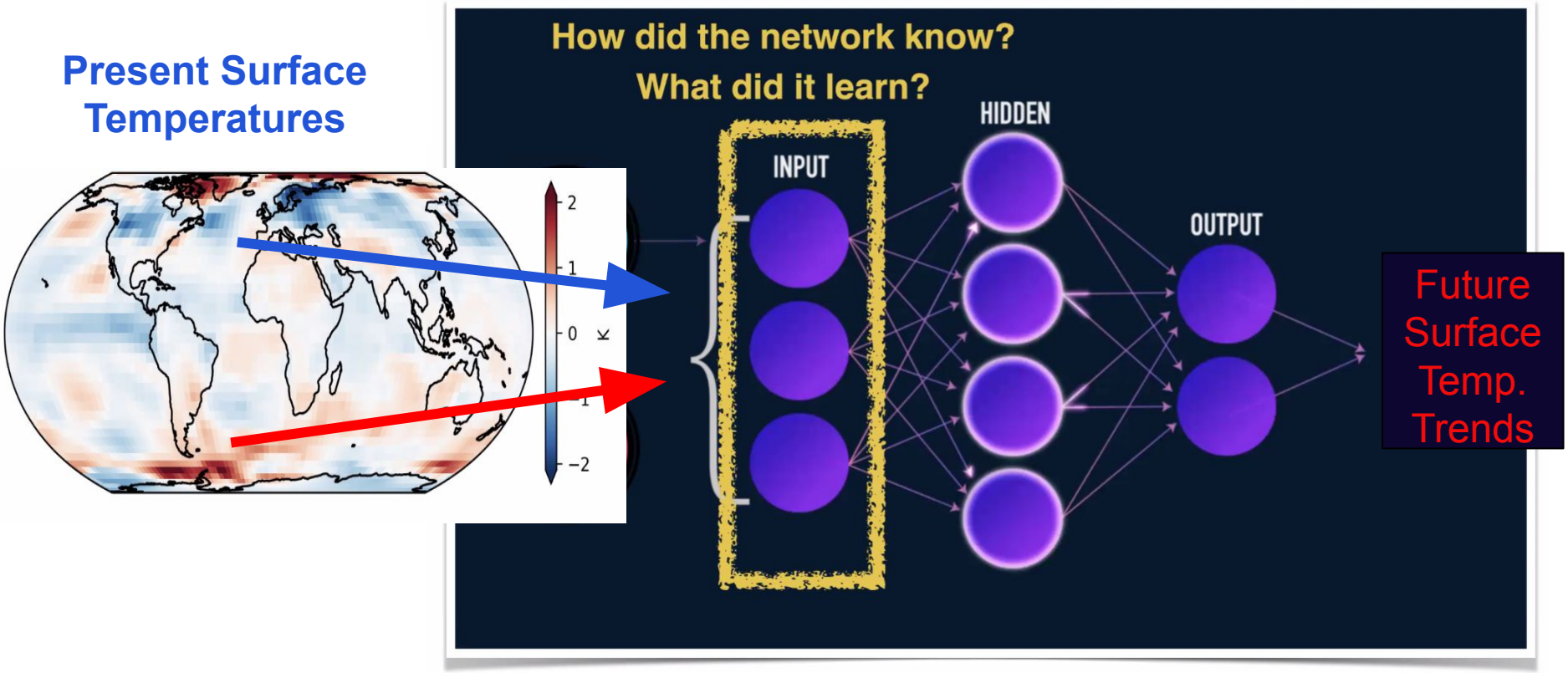
Analogous formulation of cartoon joke :))

ANN Analogous: *What is a Neuron? How Does it Work?*



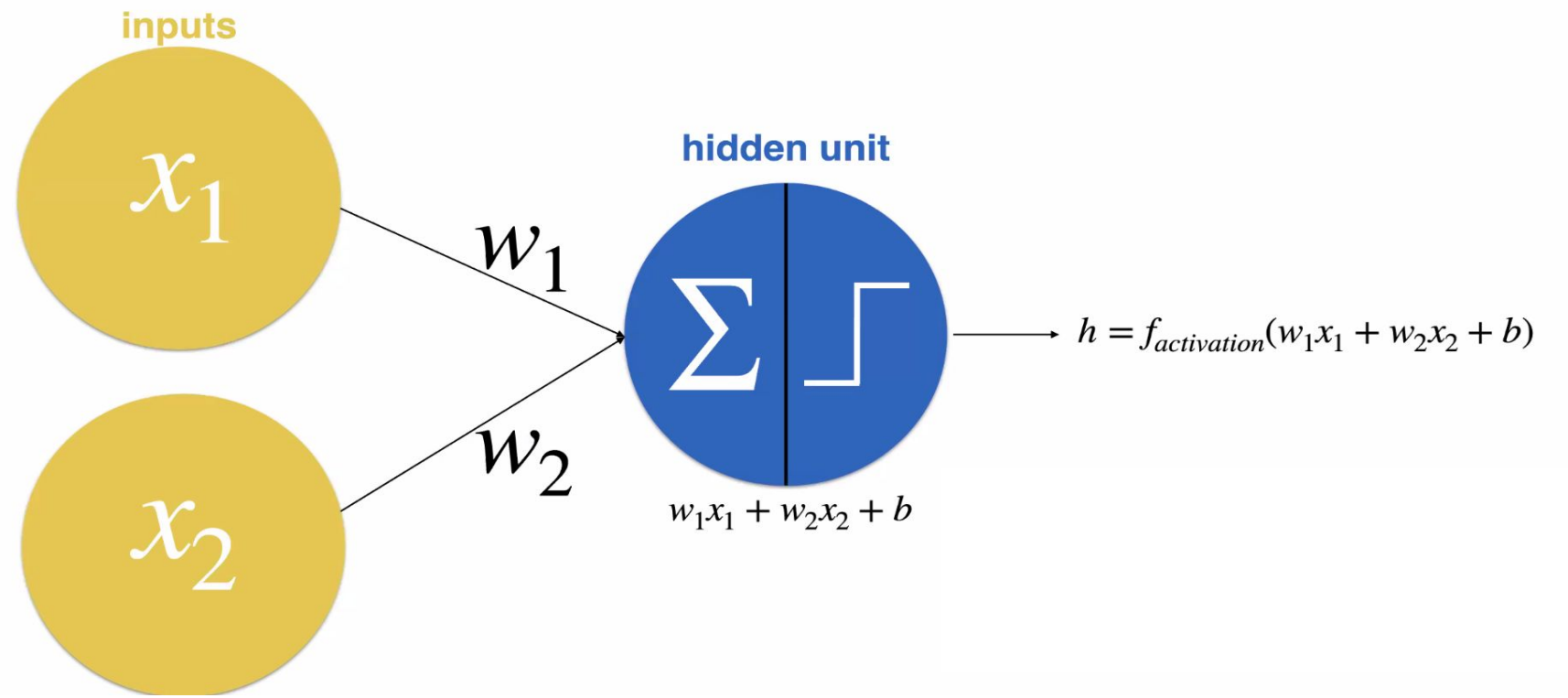
ANN for Climate Change: *How does it Work? What does it learn?*

Present Surface Temperatures

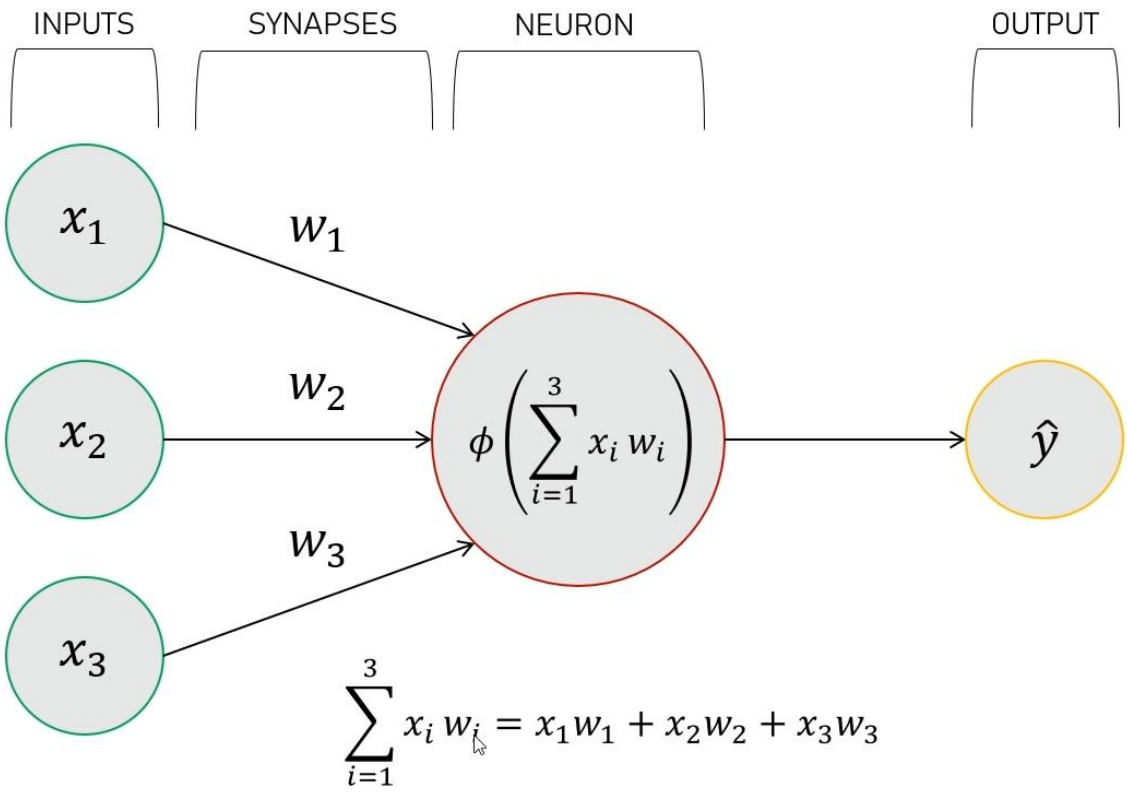


ANN for Climate Change: *How does it Work? What does it learn?*

- Linear regression with non-linear mapping by an "***activation function***".
- Training of the network is merely determining the ***weights*** "w" and ***bias/offset*** "b".



ANN with single Hidden Layer' Perceptron: *Forward Propag.*

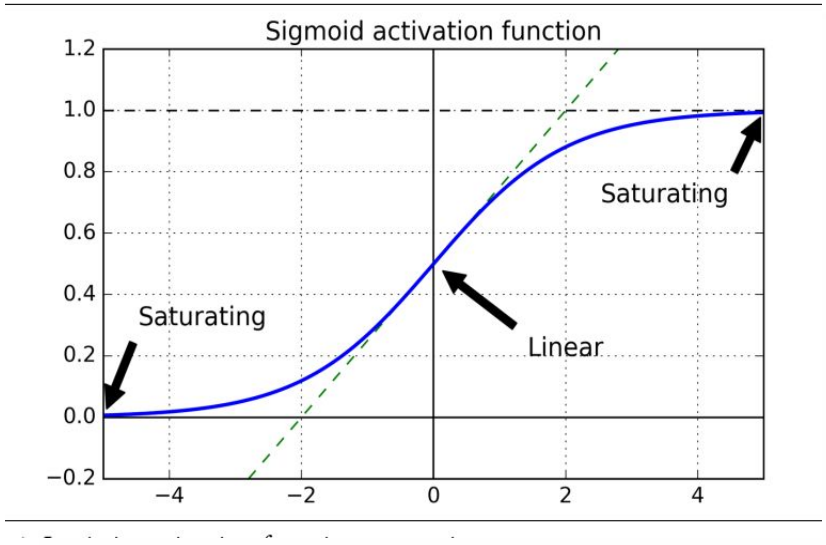


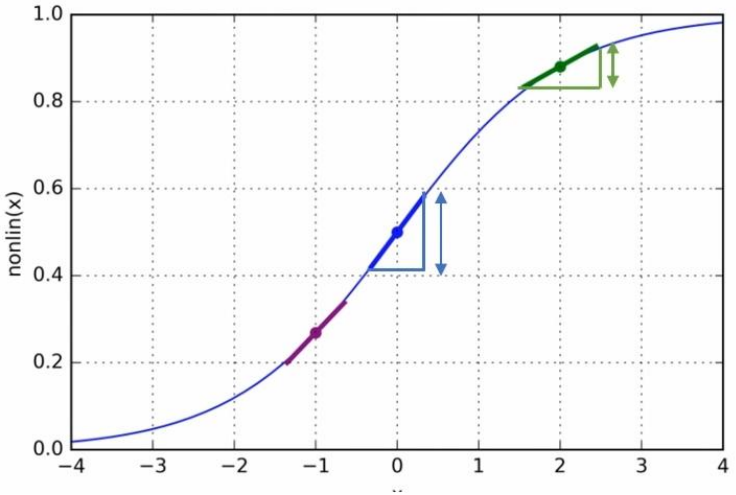
yhat = h_θ (x)

ANN with single Hidden Layer' Perceptron: *Forward Propag.*

- Using Sigmoid as an activation function, which normalizes the input.

$$\phi(x) = \frac{1}{1 + e^{-x}} \longrightarrow \phi(x) = \frac{1}{1 + e^{-\sum_{i=1}^{3} x_i w_i}}$$

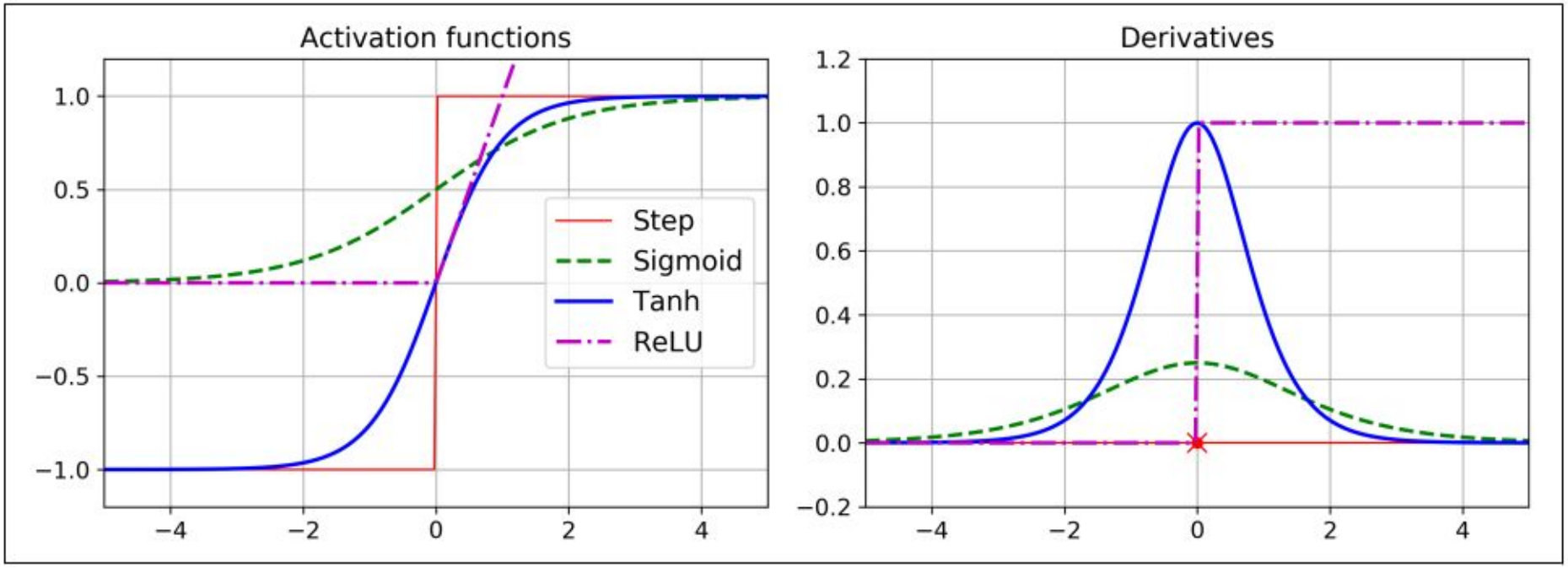




$$\phi(x) = \frac{1}{1 + e^{-x}} \longrightarrow \phi'(x) = x \,.\, (1 - x)$$

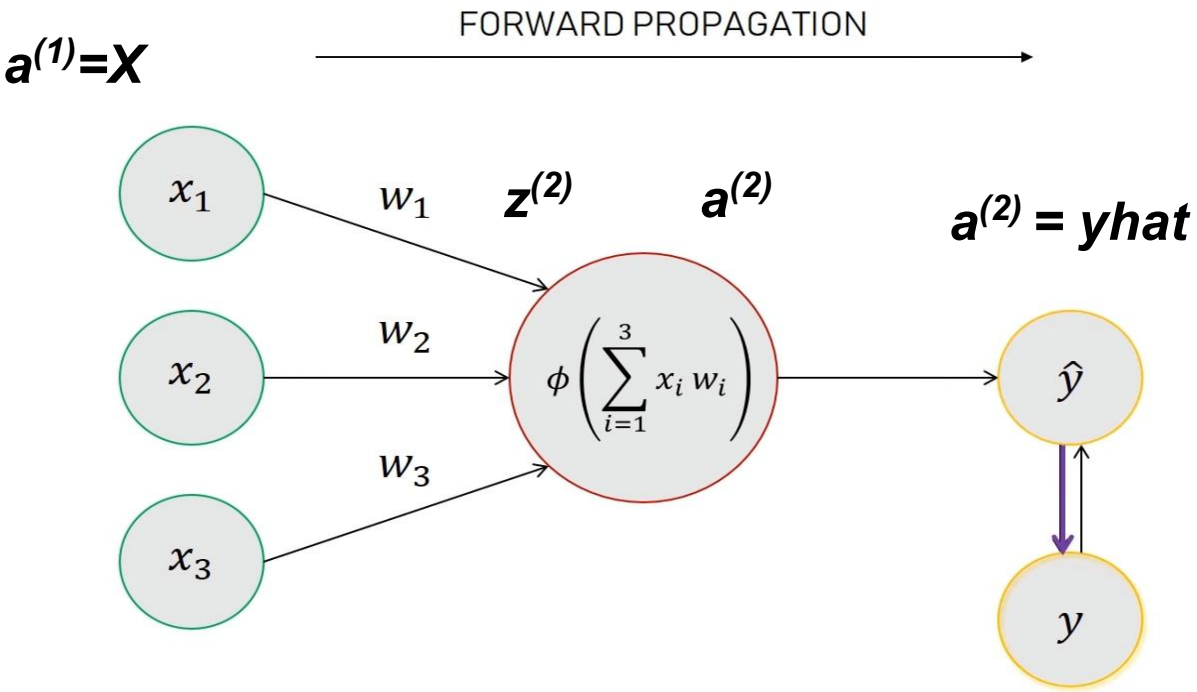
ANN with single Hidden Layer' Perceptron: *Forward Propag.*

- Others Activation function: tanh, ReLU, Softmax,...



ANN with single Hidden Layer' Perceptron: *Forward Propag.*

- Difference steps to predict the solution ***yhat*** given the input data sets.



Forward Propagation

$$W^{(1)^T} X = z^{(2)}$$

$$a^{(2)} = g(z^{(2)})$$

$$a^{(1)} = X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

ANN with single Hidden Layer' Perceptron: *Cost Function (J)*

- The square difference between the predicted ***yhat = $h_\theta(x)$*** and the observerd y is equal to the sum of the square error (SSE).

- "Best Fit" means difference between the predicted ***yhat = $h_\theta(x)$*** values and the observed y values are a minimum.

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

Parameters: θ_0, θ_1

Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

ANN with single Hidden Layer' Perceptron: *GD*

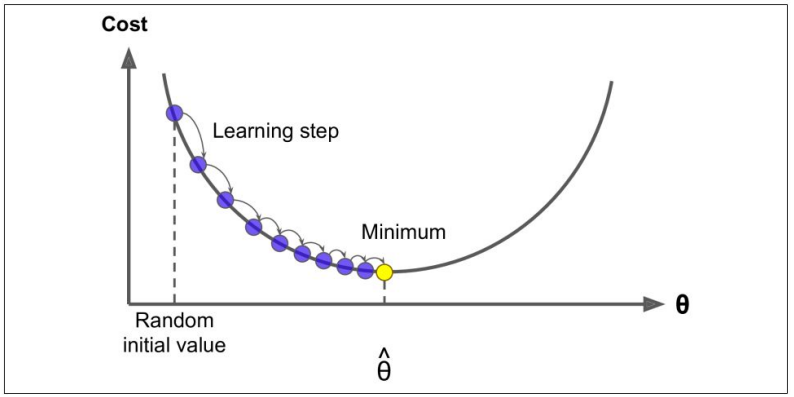
repeat until convergence {

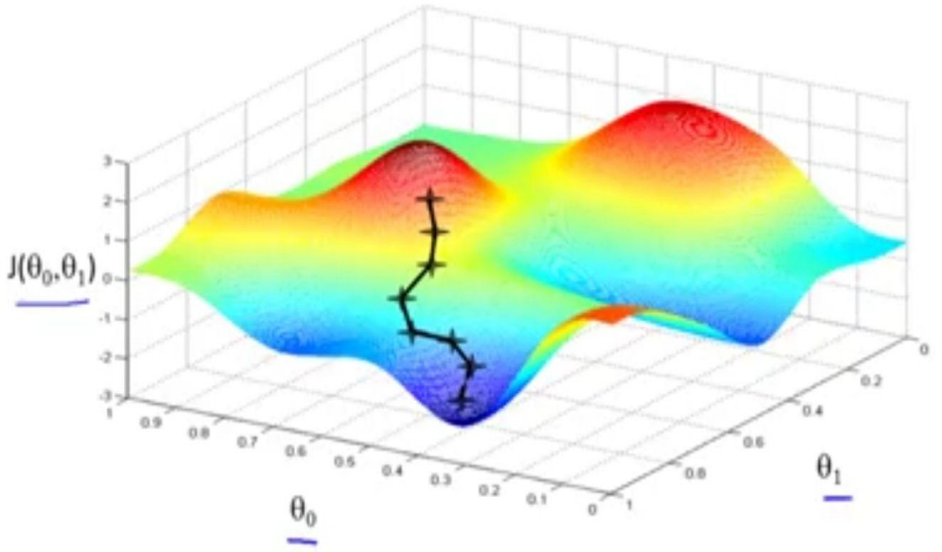
$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \qquad (\text{for } j = 0 \text{ and } j = 1)$$

}

Correct: Simultaneous update

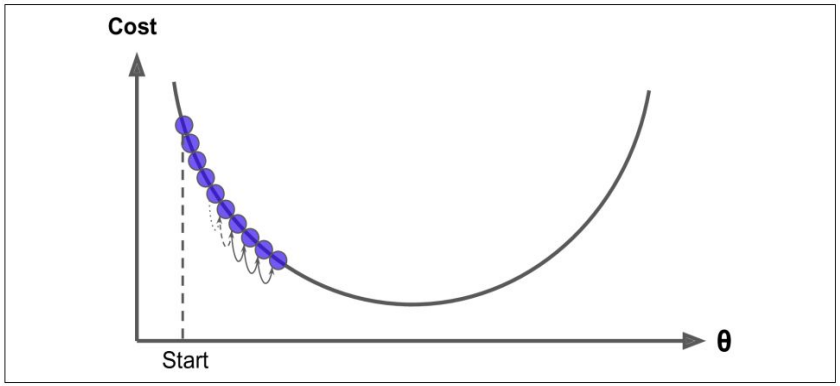
$$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$
$$\text{temp1} := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$
$$\theta_0 := \text{temp0}$$
$$\theta_1 := \text{temp1}$$

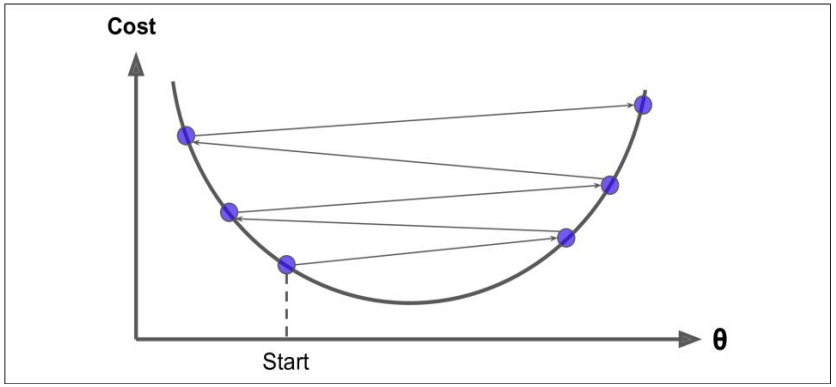


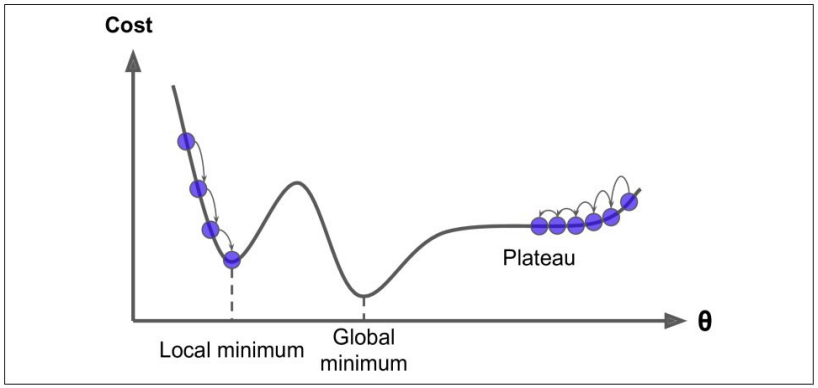


ANN with single Hidden Layer' Perceptron: *GD*

- Learning rate too small
- Learning rate too large
- Issues of local minimum







ANN with single Hidden Layer' Perceptron: *GD*

- Issues of learning with large data (m) using Gradient Descent.

$$h_\theta(x) = \sum_{j=0}^{n} \theta_j x_j$$

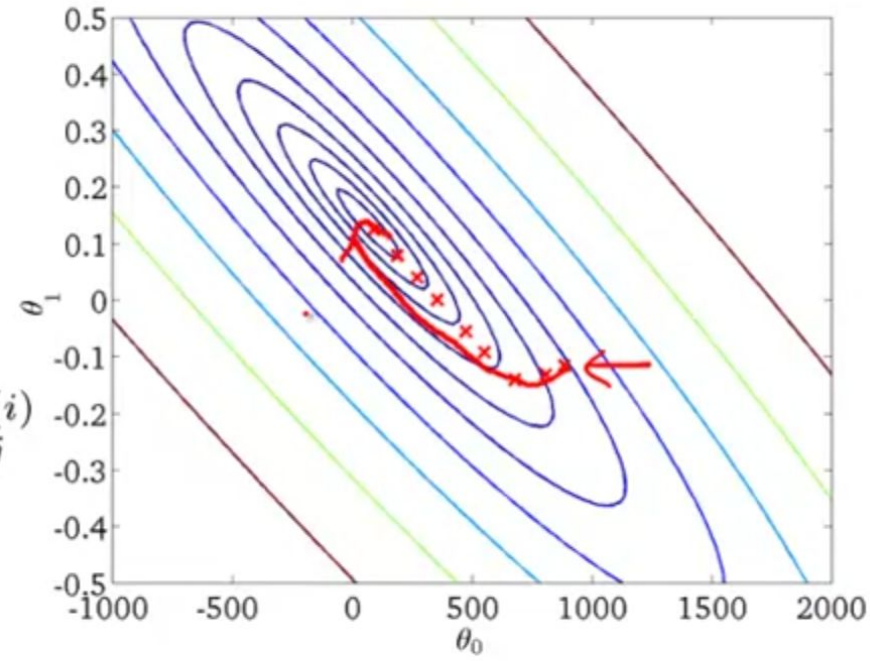
$$J_{train}(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

Repeat {

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

(for every $j = 0, \ldots, n$)

}

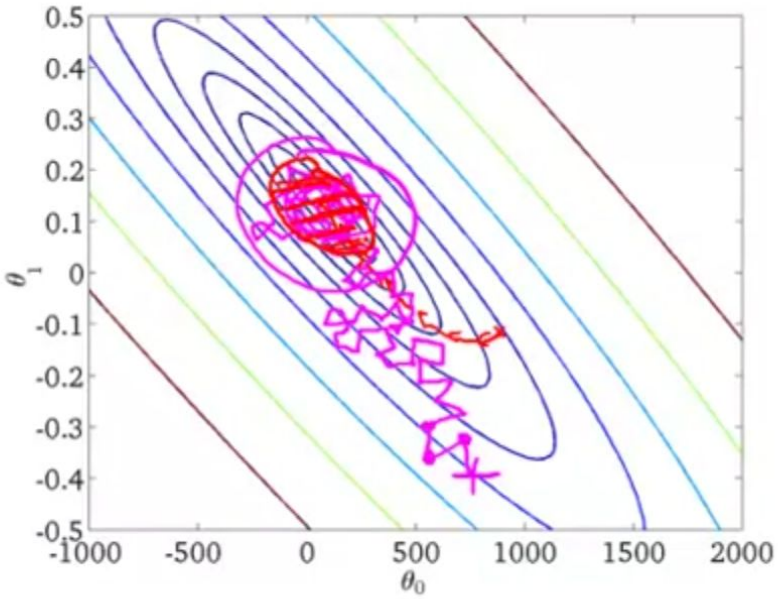


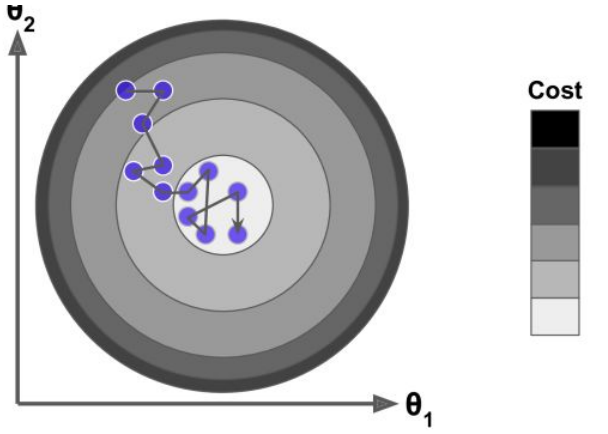
ANN for Climate Change: *Train,Testing & Cross-validation*

- Speeding up the learning with large data using Stochastic Gradient Descent.

1. Randomly shuffle (reorder) training examples

2. Repeat {
   for $i := 1, \ldots, m$ {
   $\theta_j := \theta_j - \alpha(h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)}$
   (for every $j = 0, \ldots, n$)
   }
}





ANN with single Hidden Layer' Perceptron: *Batch/SGD*

$$\phi' = \phi(1-\phi)$$

$$\frac{\partial J}{\partial w_1^{(3)}} = \frac{\partial}{\partial w_1^{(3)}}\left[\frac{1}{2}(y-\hat{y})^2\right]$$

$$= -(y-\hat{y}) \cdot \frac{\partial \hat{y}}{\partial w_1^{(3)}}$$

$$= -(y-\hat{y}) \cdot \frac{\partial}{\partial w_1^{(3)}}\left[\phi(z^{(4)})\right]$$

activation function

$$\hat{y} = \phi(z^{(4)})$$

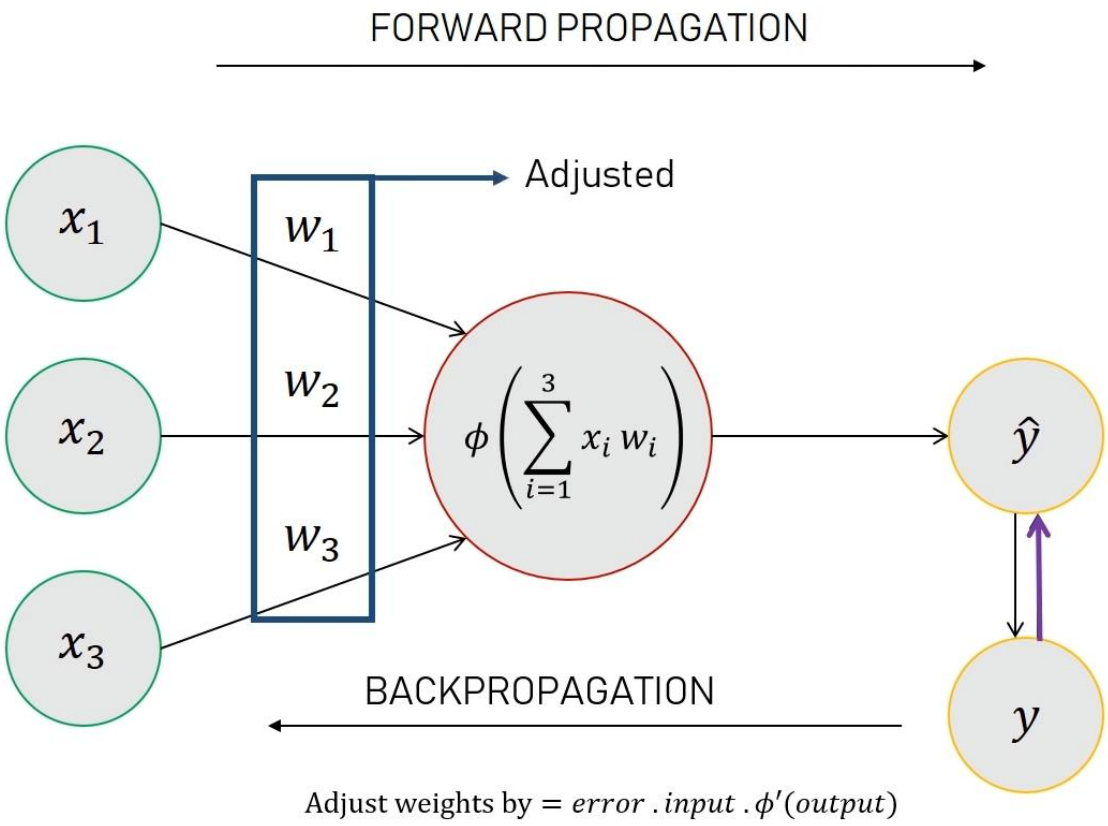
$$= -(y-\hat{y}) \cdot \frac{\partial \hat{y}}{\partial z^{(4)}} \cdot \frac{\partial z^{(4)}}{\partial w_1^{(3)}}$$

$$\frac{\partial J}{\partial w_1^{(3)}} = \underbrace{-(y-\hat{y}) \cdot \phi'(z^{(4)})}_{\delta^{(4)}} \cdot a^{(3)}$$

$$\frac{\partial J}{\partial w_1^{(3)}} = \text{error} \cdot \phi'(\text{input}) \cdot \text{input}$$

ANN with single Hidden Layer' Perceptron: *BackProp*

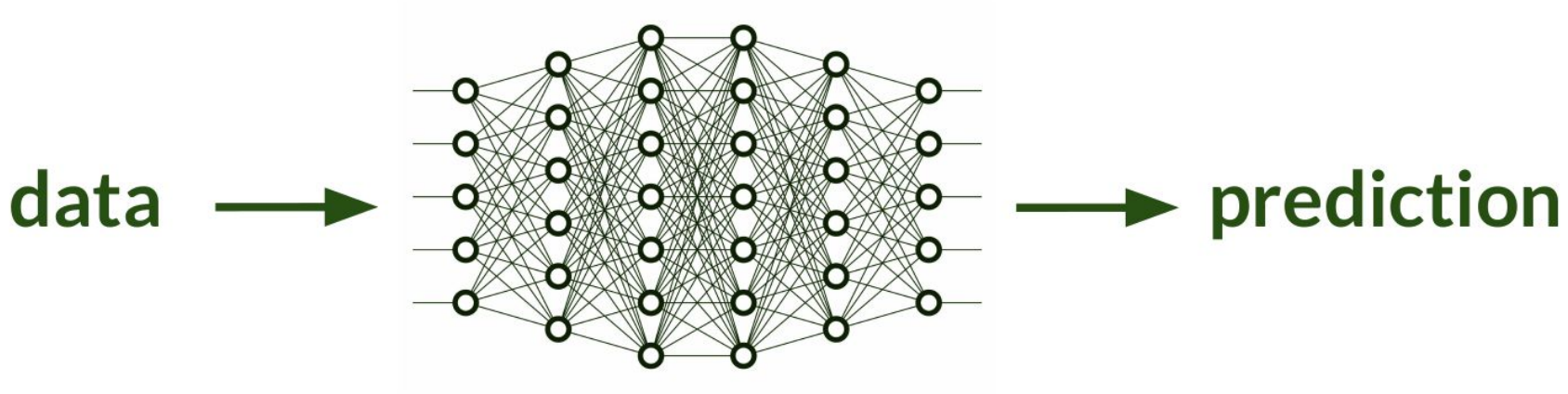
- Convergence to best fit due to the ***adjusted weight*** by BackPropagation



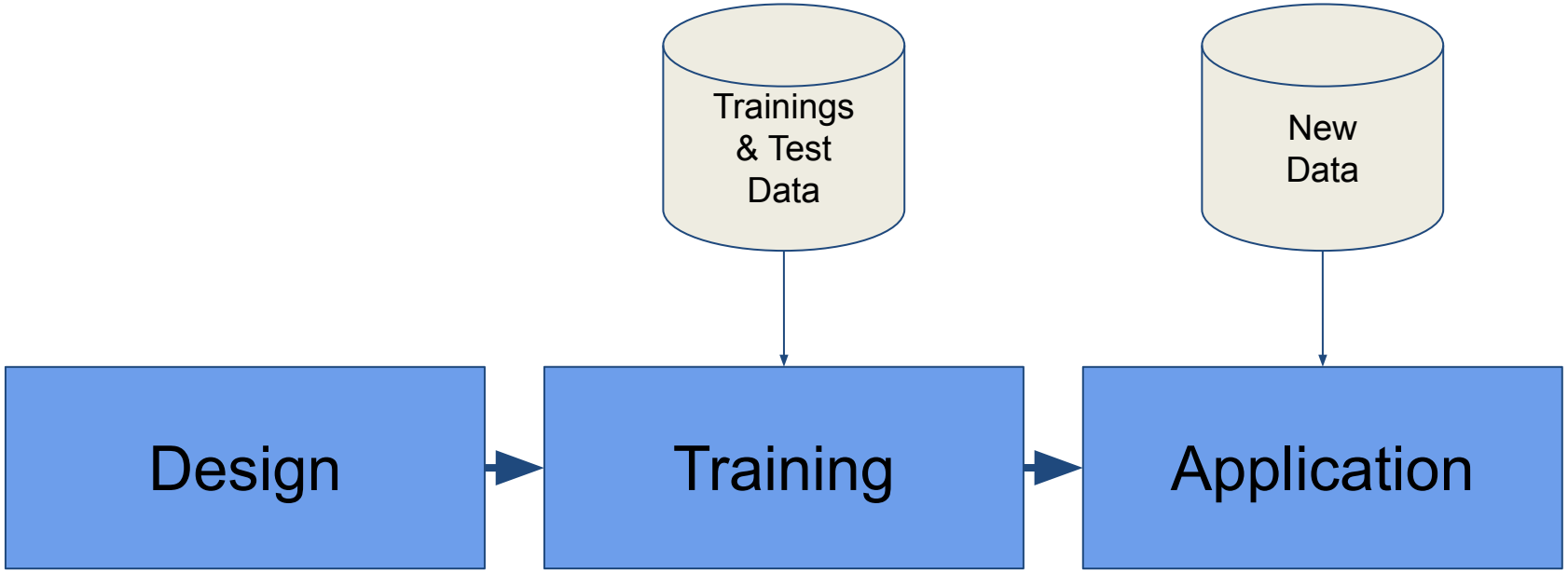
ANN Complexity and Nonlinearity

- ***Complexity*** and ***nonlinearities*** of the ANN allow it to ***learn many different pathways*** of predictable behaviour.

- Once trained, ***ANN*** has an array of ***weights*** and ***biases***, which can be used for ***prediction on new data sets.***



What Is the ML Workflow?



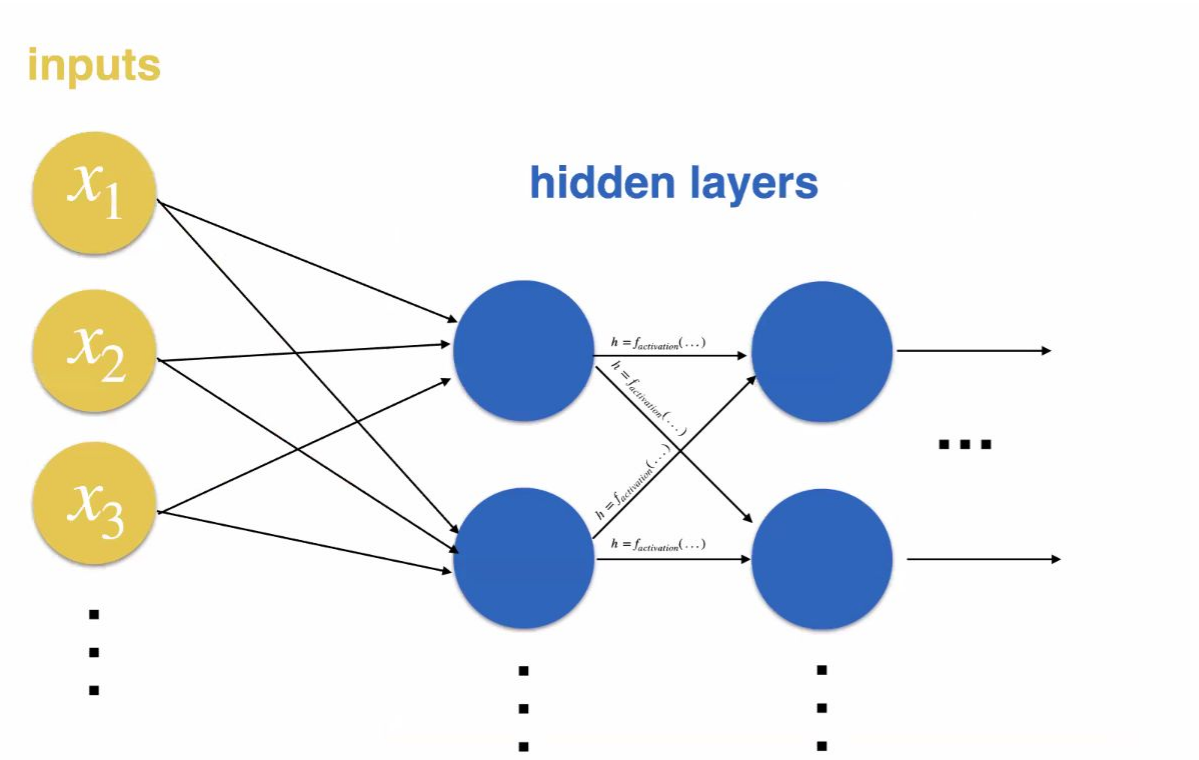
What method? How many layers? What type of layers? What is the input? What is the output? What activation Function? What Lossfunction?

What training data can be used? Does the training succeed ? Do we have overfitting? Does the ML learn generalized or just memorizes?

Is the new data similar enough to the trainings data or are additional processes involved? Can possible errors be tolerated?

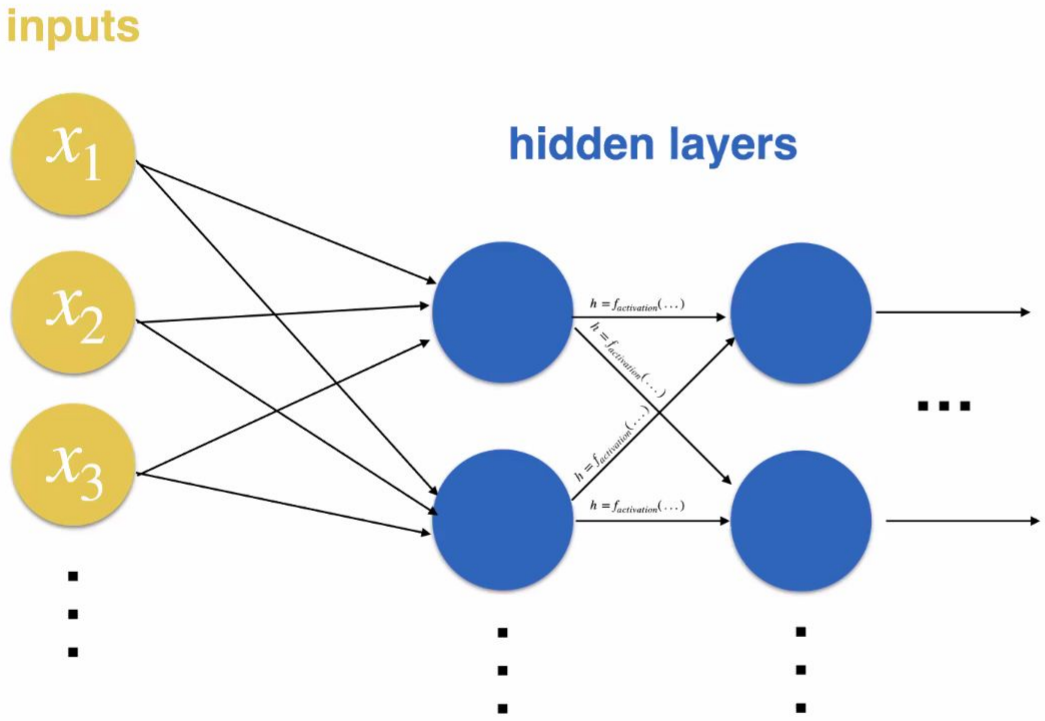
ANN for Climate Change: *Train,Testing & Cross-validation*

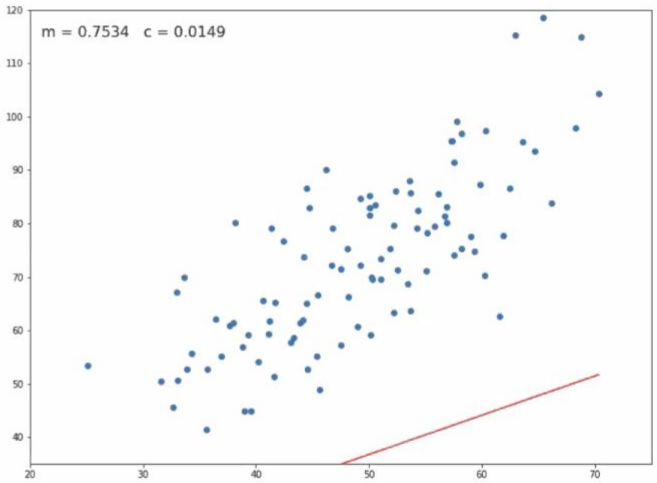
- Networks can become quite nonlinear and complex.



ANN for Climate Change: *Train,Testing & Cross-validation*

- Learning from the data
- Adjusting the weights and biases

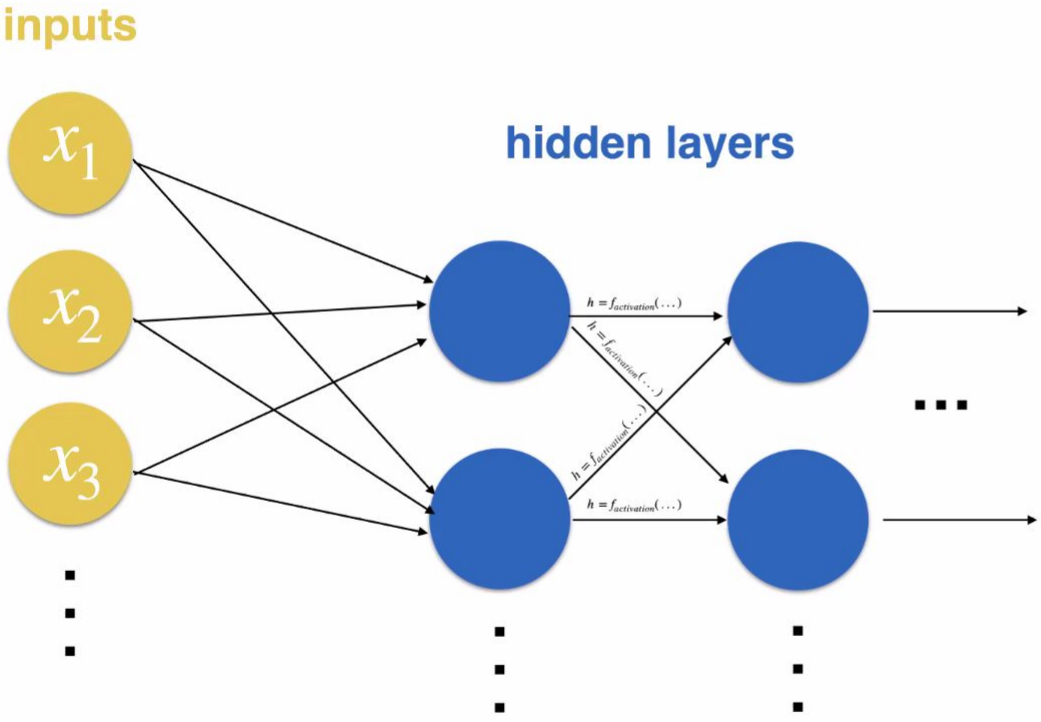


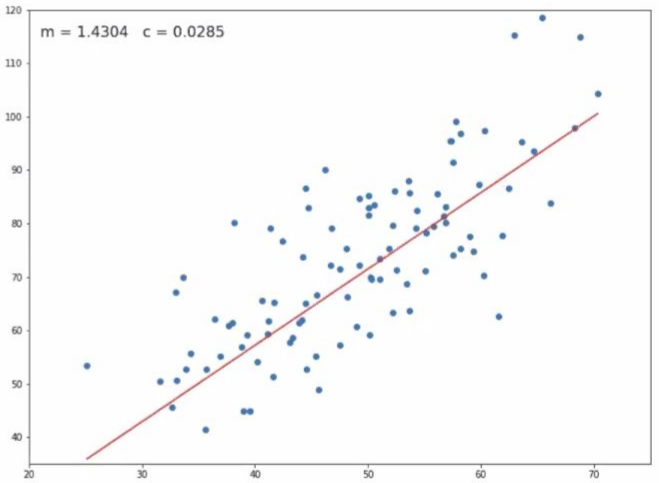


search iteratively for
optimal weights and biases

ANN for Climate Change: *Train,Testing & Cross-validation*

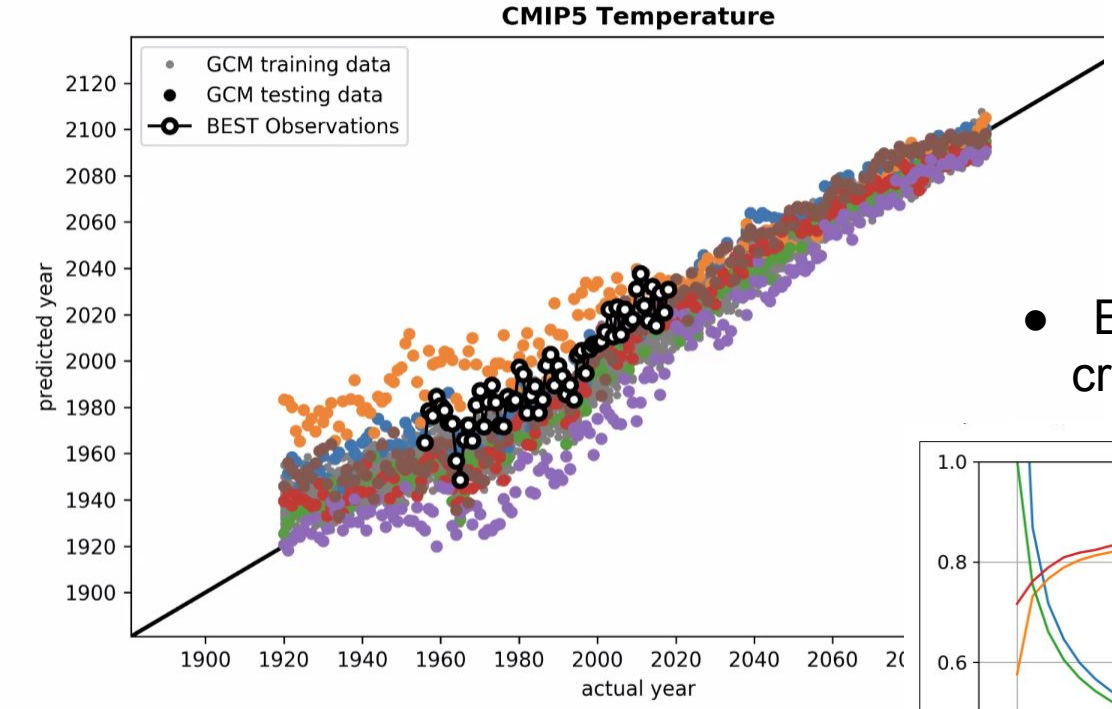
- Best Fit after training



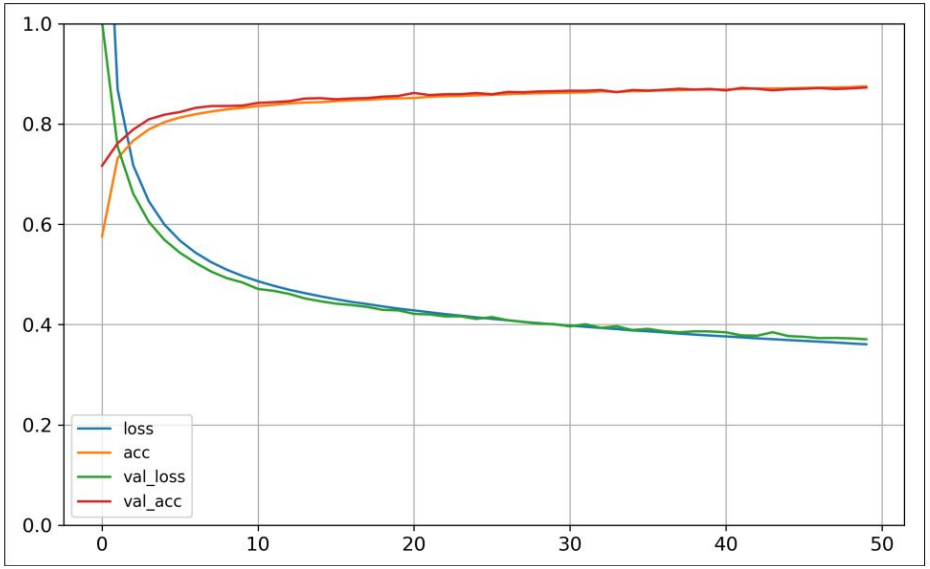


search iteratively for
optimal weights and biases

ANN for Climate Change: *Train,Testing & Cross-validation*



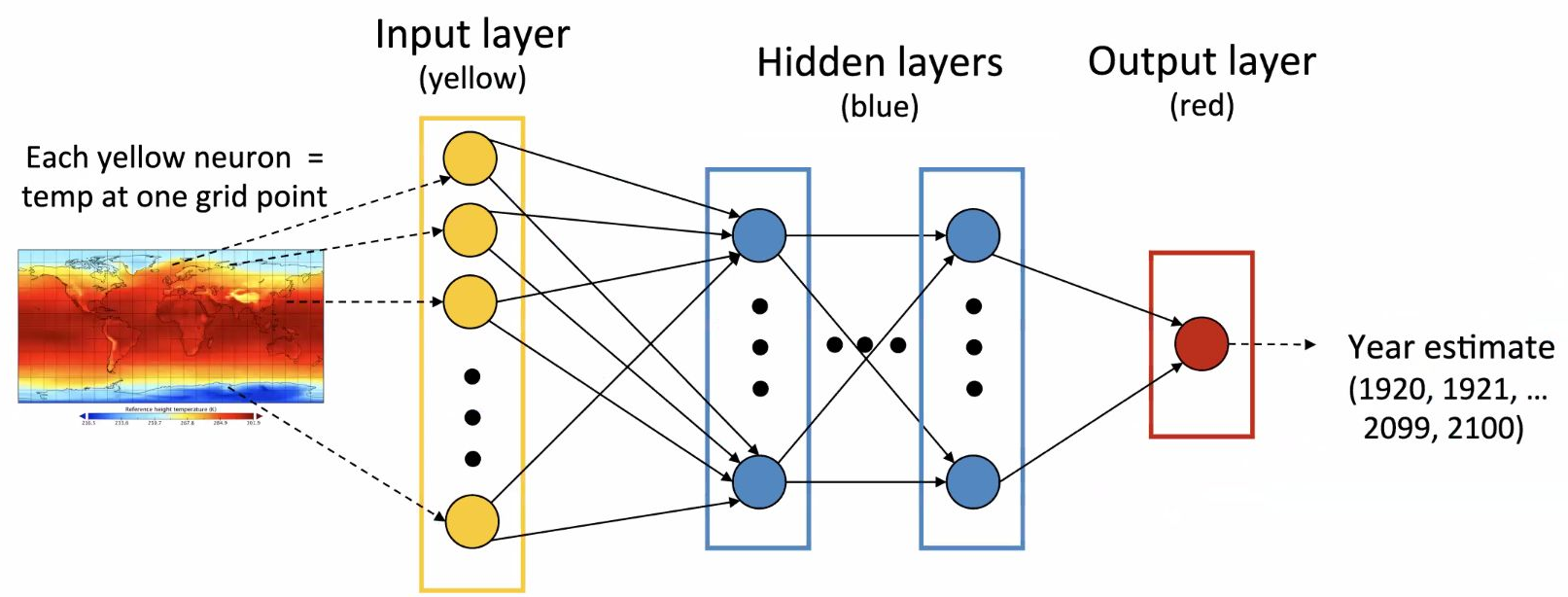
- Estimating the accuracy using cross-validation



- Cross-Validation of the Best Fit after training

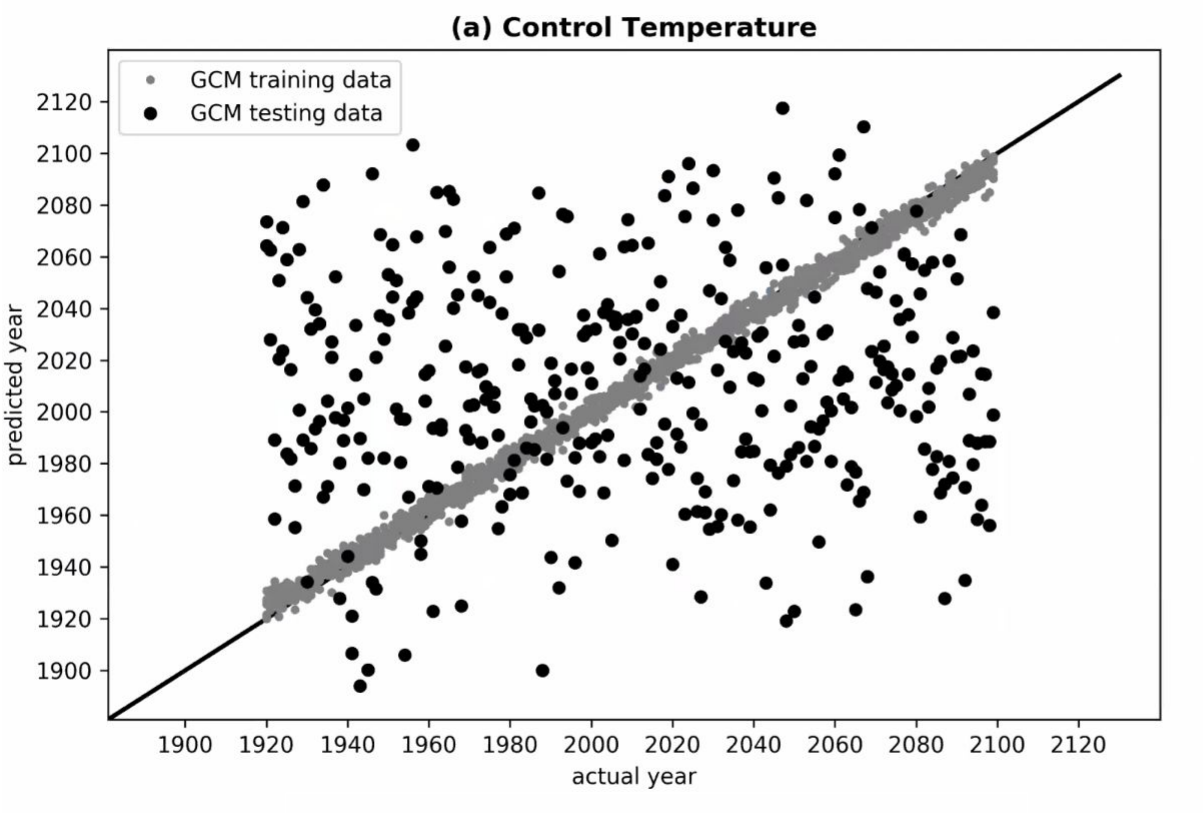
ANN for Climate Change: *Prediction of Temperature Trend*

- The trained ***ANN*** uses the ***weights*** and ***biases*** for ***predicting temperatures on new data sets:***



ANN for Climate Change: *High Bias versus High Variance*

- Most of the time, it won't converge to "Best Fit" at first place.
- Diagnostics are need to figure out what's going one: High Bias or High Variance.



- Failure of the Fit after training:
  High Bias or High Variance

ANN for Climate Change: *High Bias versus High Variance*

Evaluating your hypothesis

Dataset:

	Size	Price	
60%	2104	400	Training set
	1600	330	
	2400	369	
	1416	232	
	3000	540	
	1985	300	
20%	1534	315	Cross validation (CV) set
	1427	199	
20%	1380	212	test set
	1494	243	

$$(x^{(1)}, y^{(1)})$$
$$(x^{(2)}, y^{(2)})$$
$$\vdots$$
$$(x^{(m)}, y^{(m)})$$

$$(x_{cv}^{(1)}, y_{cv}^{(1)})$$
$$(x_{cv}^{(2)}, y_{cv}^{(2)})$$
$$\vdots$$
$$(x_{cv}^{(m_{cv})}, y_{cv}^{(m_{cv})})$$

m_{cv} = no. of cv example $(x_{cv}^{(i)}, y_{cv}^{(i)})$

$$(x_{test}^{(1)}, y_{test}^{(1)})$$
$$(x_{test}^{(2)}, y_{test}^{(2)})$$
$$\vdots$$
$$(x_{test}^{(m_{test})}, y_{test}^{(m_{test})})$$

m_{test}

ANN for Climate Change: *High Bias versus High Variance*

Train/validation/test error

Training error:

$$J_{train}(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

Cross Validation error:

$$J_{cv}(\theta) = \frac{1}{2m_{cv}} \sum_{i=1}^{m_{cv}} (h_\theta(x_{cv}^{(i)}) - y_{cv}^{(i)})^2$$

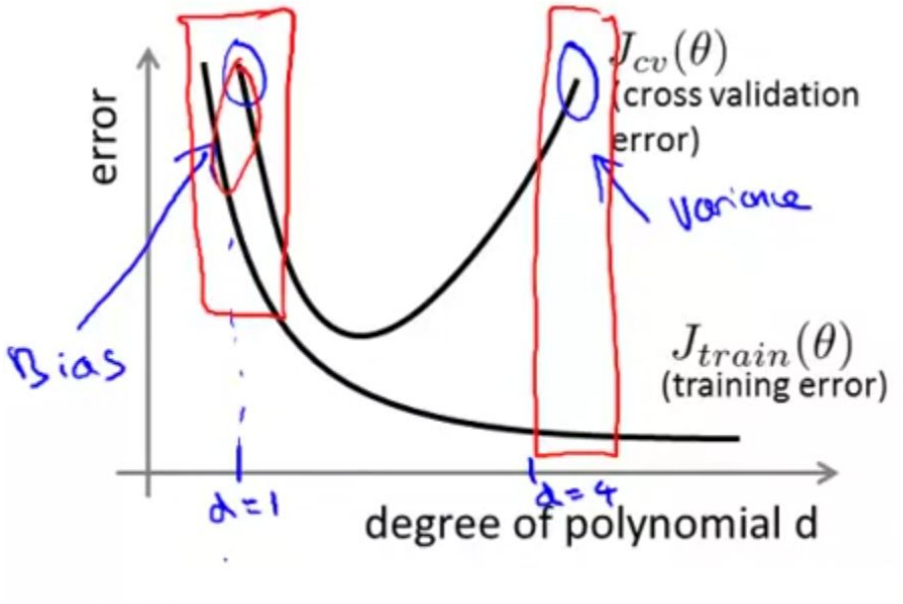
Test error:

$$J_{test}(\theta) = \frac{1}{2m_{test}} \sum_{i=1}^{m_{test}} (h_\theta(x_{test}^{(i)}) - y_{test}^{(i)})^2$$

ANN for Climate Change: *High Bias versus High Variance*

Diagnosing bias vs. variance

Suppose your learning algorithm is performing less well than you were hoping. ($J_{cv}(\theta)$ or $J_{test}(\theta)$ is high.) Is it a bias problem or a variance problem?



Bias (underfit):

→ $J_{train}(\theta)$ will be high

$J_{cv}(\theta) \approx J_{train}(\theta)$

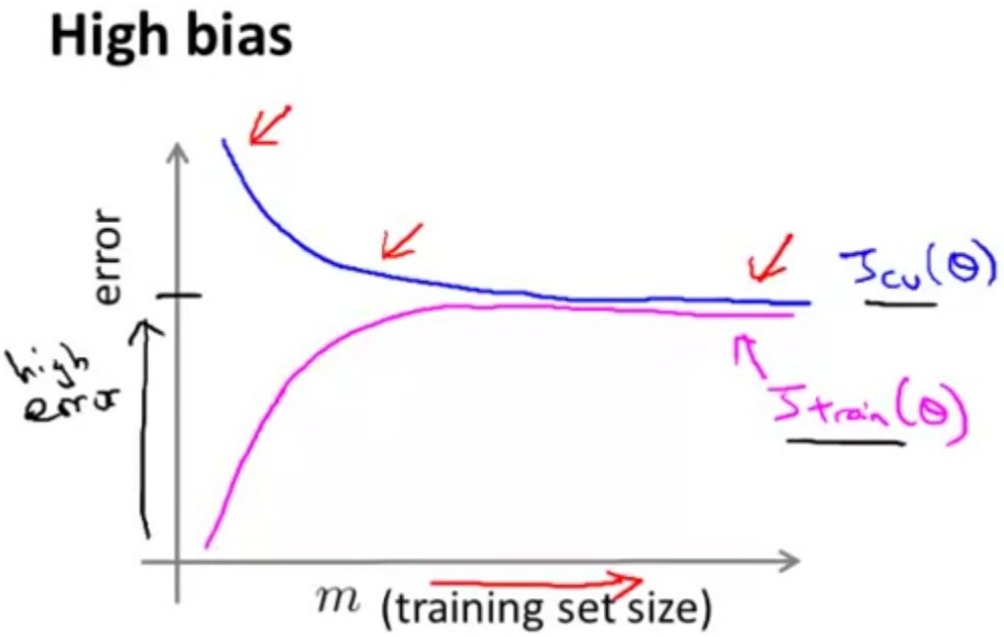
Variance (overfit):

→ $J_{train}(\theta)$ will be low

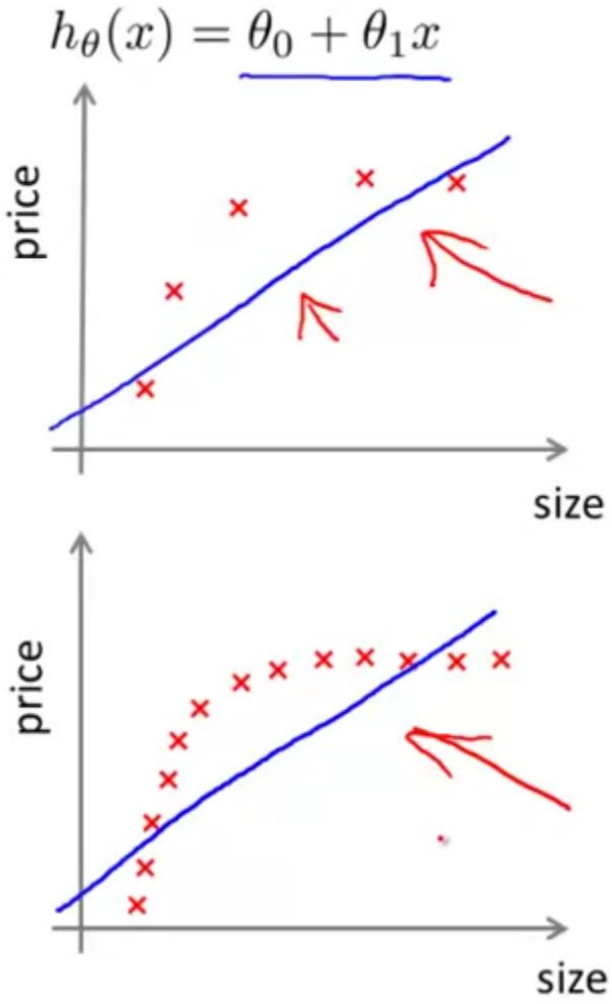
$J_{cv}(\theta) \gg J_{train}(\theta)$

- There is also the diagnostic of the learning and the regularisation factor, but tackled here.

ANN for Climate Change: *Underfitting issues*

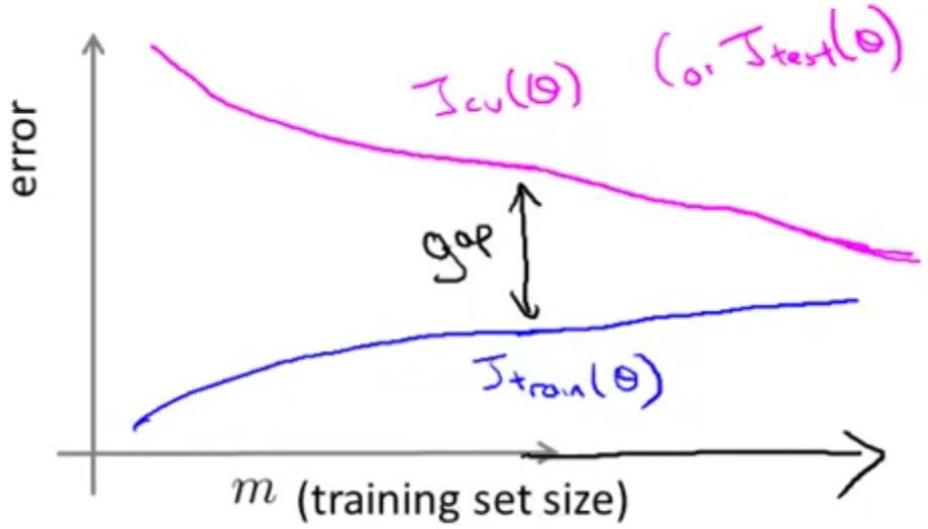


If a learning algorithm is suffering from high bias, getting more training data will not (by itself) help much.



ANN for Climate Change: *Overfitting and Regularisation*

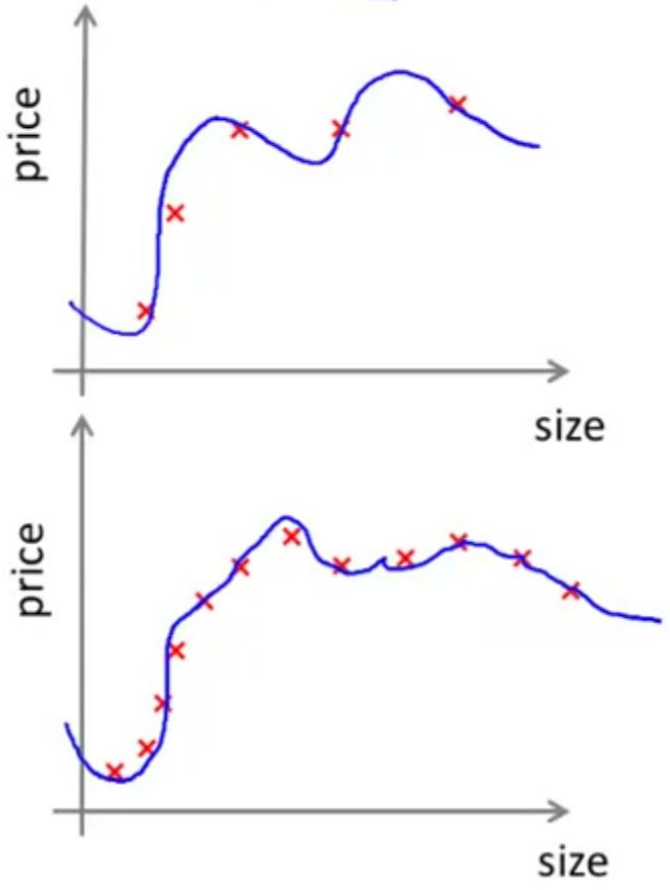
High variance



If a learning algorithm is suffering from high variance, getting more training data is likely to help.

$$h_\theta(x) = \theta_0 + \theta_1 x + \cdots + \theta_{100} x^{100}$$

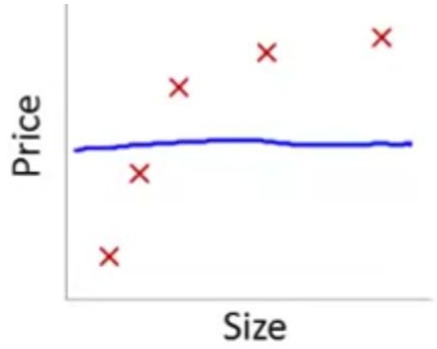
(and small λ)



ANN for Climate Change: *Overfitting and Regularisation*

- **Ridge Regression** or L2-regularisation.

$$J(\theta) = \frac{1}{2m}\left[\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda\sum_{j=1}^{m}\theta_j^2\right]$$

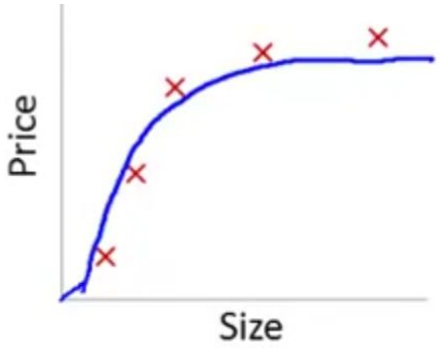


Large λ ←

⇒ High bias (underfit)

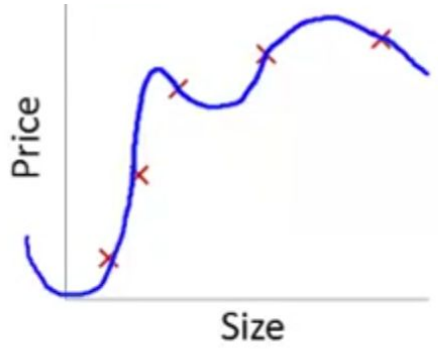
$\lambda = 10000.\ \theta_1 \approx 0, \theta_2 \approx 0, \ldots$

$h_\theta(x) \approx \theta_0$



Intermediate λ ←

"Just right"



⇒ Small λ

High variance (overfit)

→ $\lambda = 0$

Ar

ANN for Climate Change: *Overfitting and Regularisation*

Choosing the regularization parameter λ

Model: $h_\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2m}\sum_{j=1}^{m}\theta_j^2$$

1. Try $\lambda = 0$ → $\min_\theta J(\theta)$ → $\theta^{(1)}$ → $J_{cv}(\theta^{(1)})$
2. Try $\lambda = 0.01$ → $\min_\theta J(\theta)$ → $\theta^{(2)}$ → $J_{cv}(\theta^{(2)})$
3. Try $\lambda = 0.02$ → $\theta^{(3)}$ → $J_{cv}(\theta^{(3)})$
4. Try $\lambda = 0.04$
5. Try $\lambda = 0.08$ → $\theta^{(5)}$
   ⋮
12. Try $\lambda = 10$ → $\theta^{(12)}$ → $J_{cv}(\theta^{(12)})$

10.24

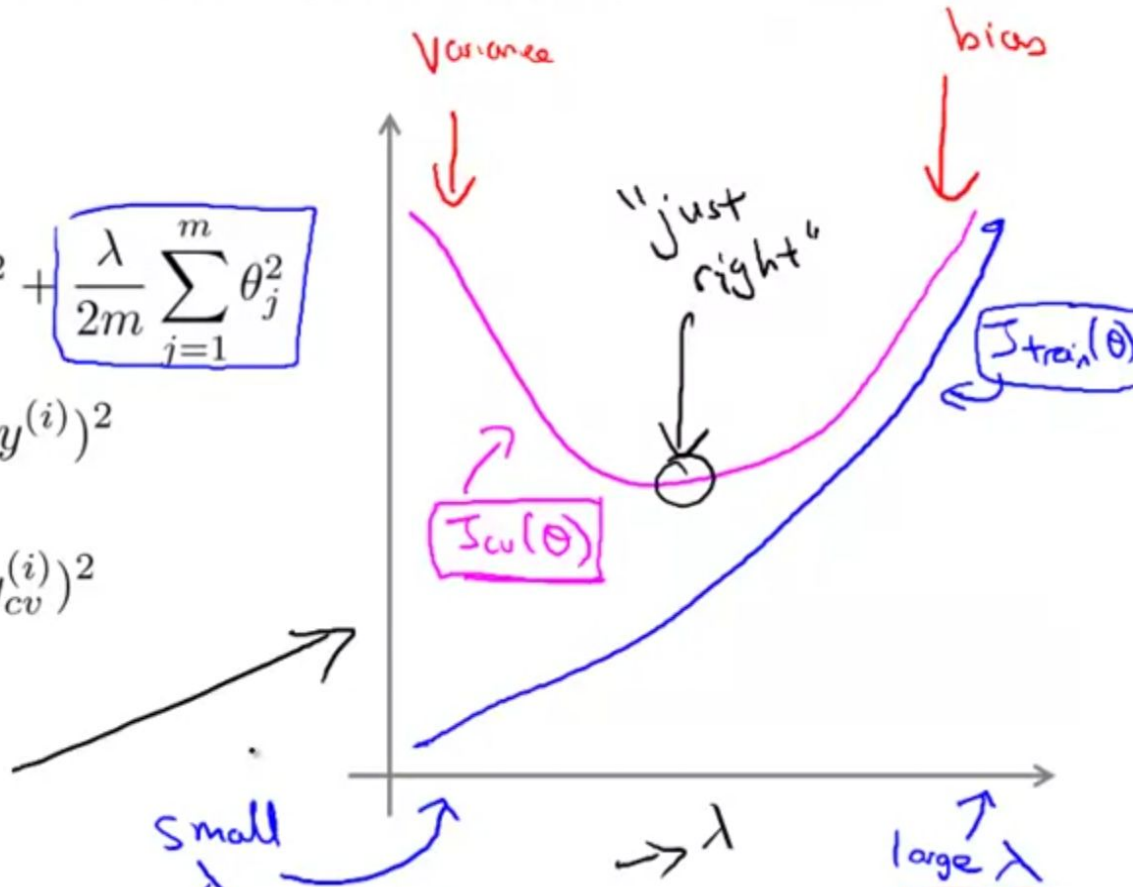
ANN for Climate Change: *Overfitting and Regularisation*

Bias/variance as a function of the regularization parameter λ

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2m}\sum_{j=1}^{m}\theta_j^2$$

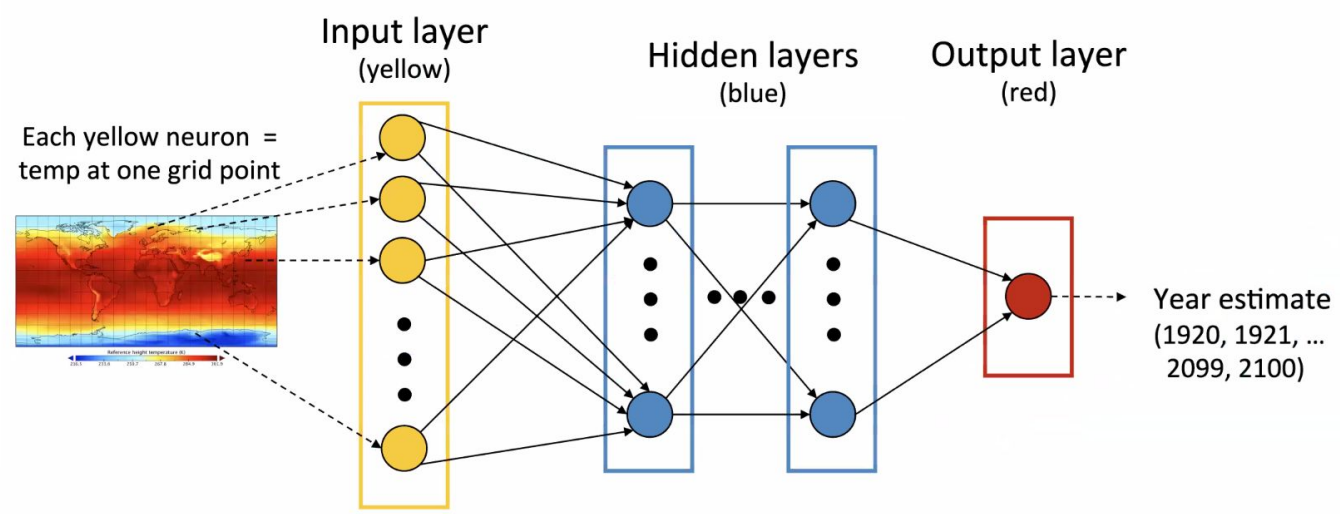
$$J_{train}(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2$$

$$J_{cv}(\theta) = \frac{1}{2m_{cv}}\sum_{i=1}^{m_{cv}}(h_\theta(x_{cv}^{(i)}) - y_{cv}^{(i)})^2$$



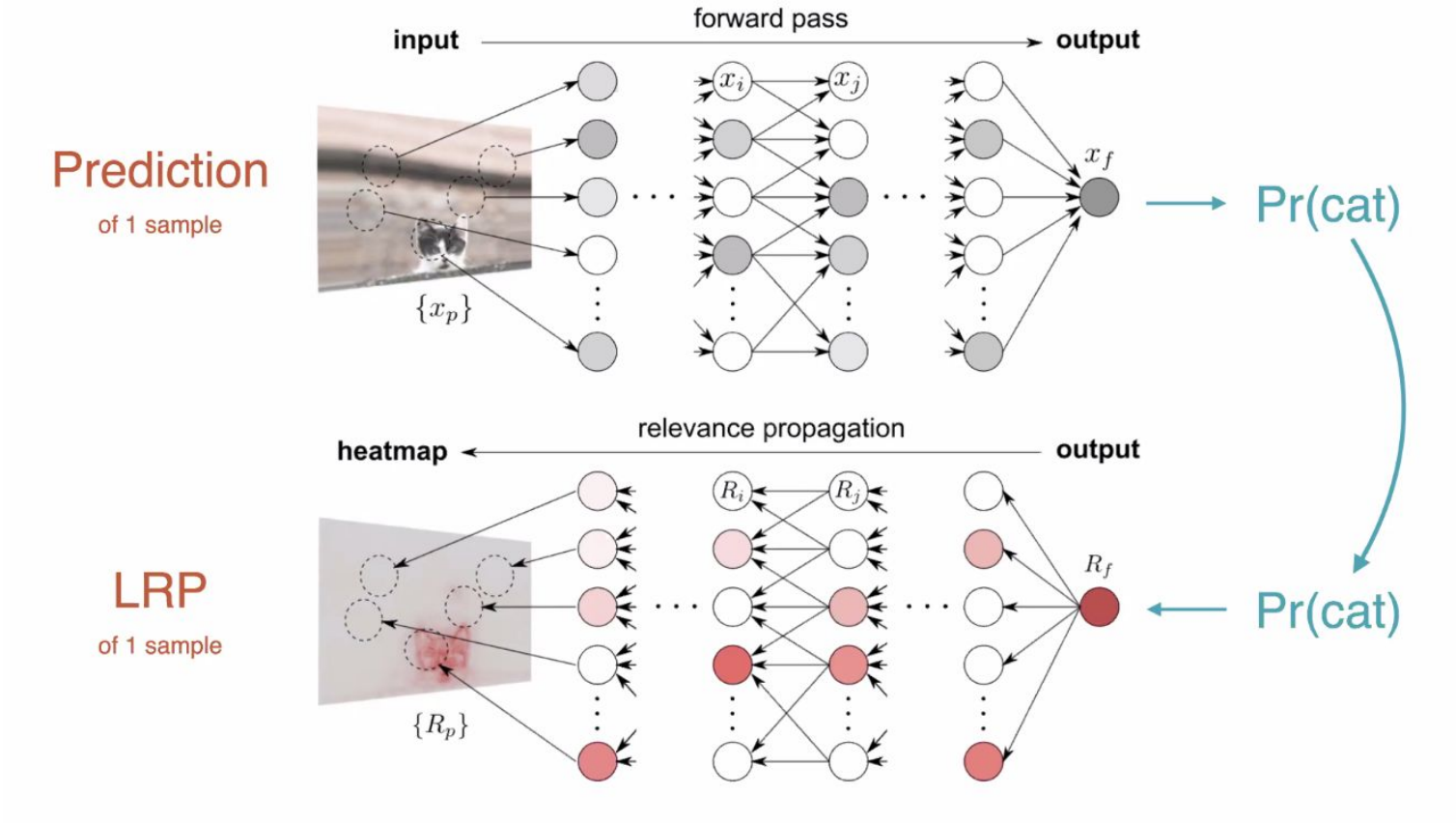
ANN Visualisation: *Layer-wise Relevance Propagation (LRP)*

- What patterns allow the ANN to make a Prediction?
- Which regions are relevant for for correctly prediction a specific year?



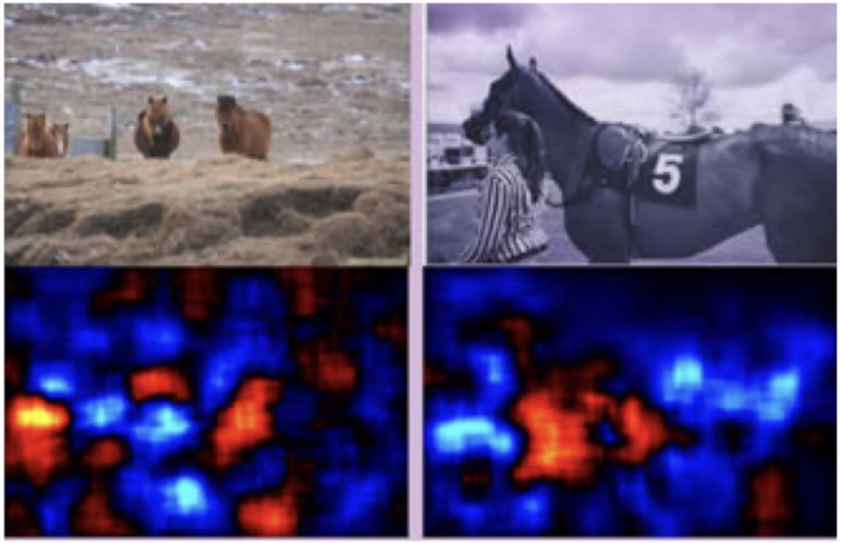
ANN Visualisation: *Layer-wise Relevance Propagation (LRP)*

- Tools to open the black box of NN using LRP: **Heat maps.**



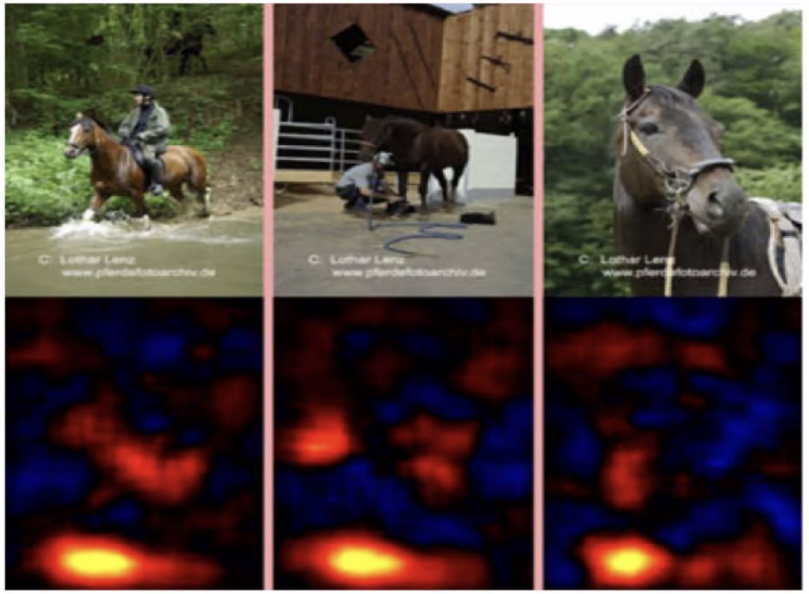
ANN Visualisation for debugging: *Heatmaps*

- Where did the ANN look to relevant patterns for Prediction determine there was a horse?



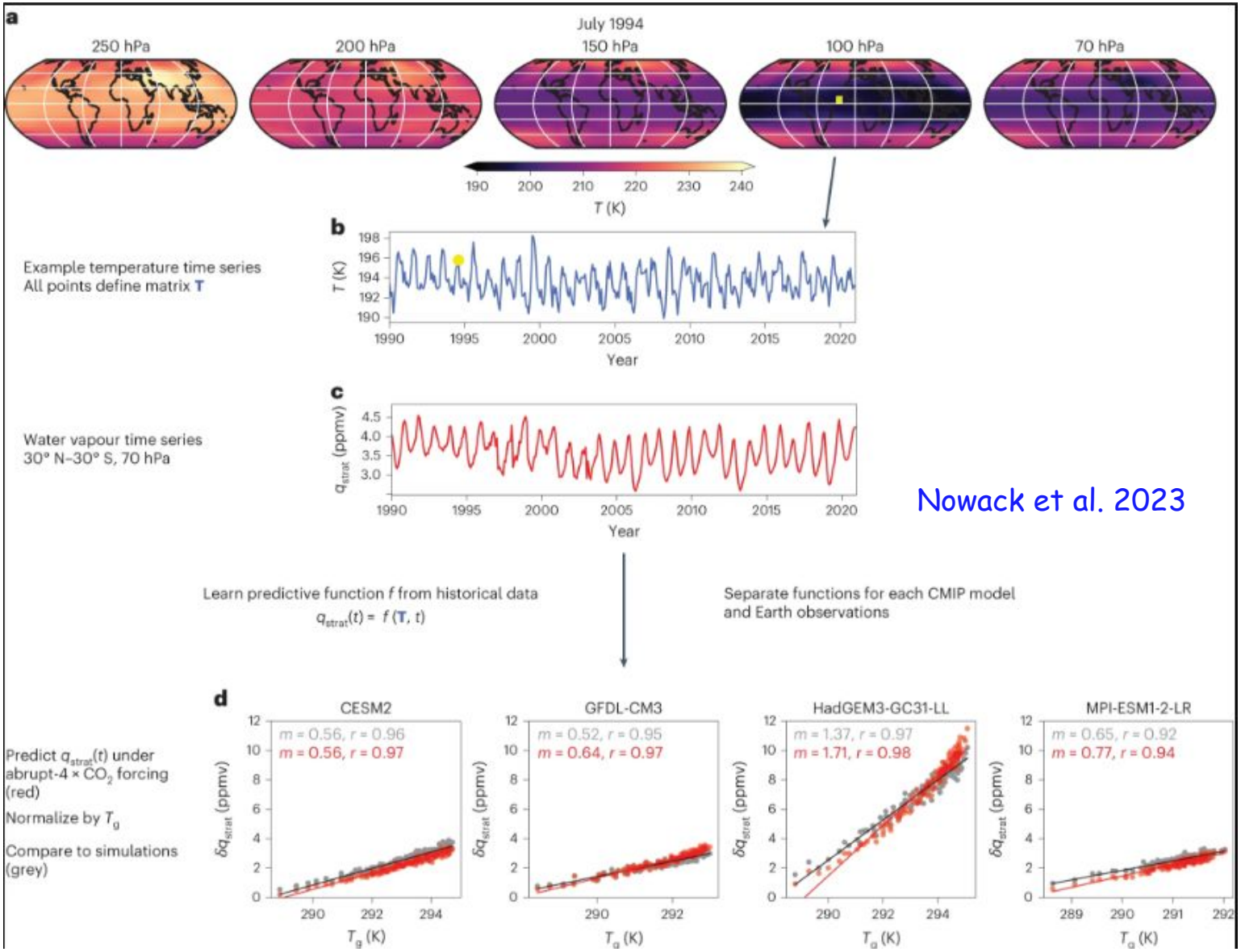
- Relevant patterns for Prediction

- Relevant patterns for Prediction



Example ML/AI applications

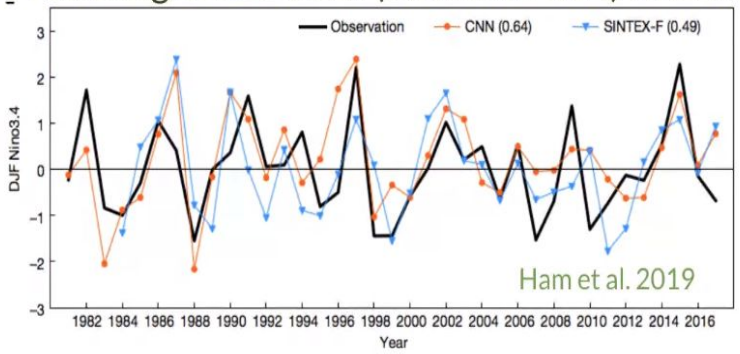
Constraining climate model projections using ML



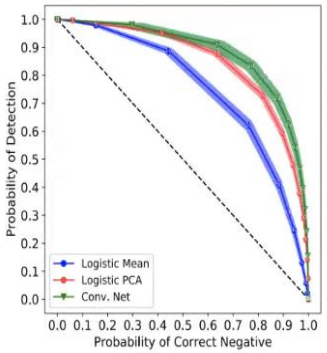
Nowack et al. 2023

Machine Learning in weather & climate

Predicting **ENSO** Index (18 month lead) with CNN



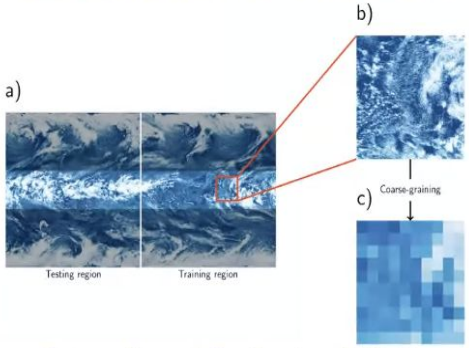
Predicting **severe hail** with a CNN



Gagne et al. 2019

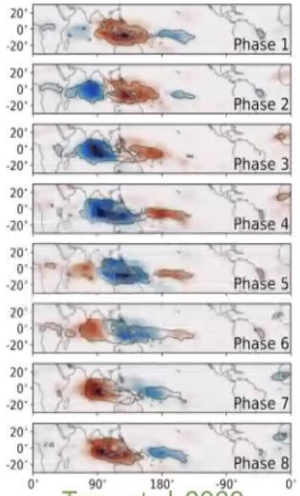
Convective **Parameterizations**

e.g. Rasp et al. (2018; PNAS); Schneider et al. (2017; GRL); O'Gorman and Dwyer (2018); Beucler et al. (2020; PRL)



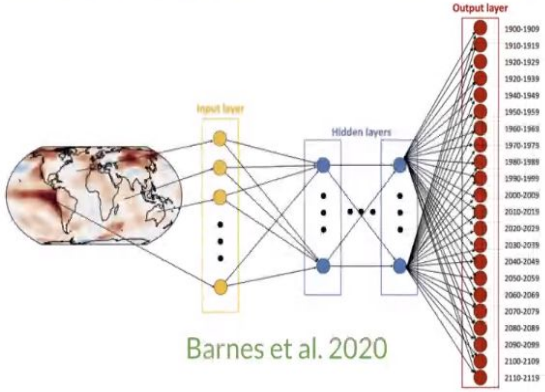
Brenowitz and Bretherton (2018)

NN-identified **MJO**



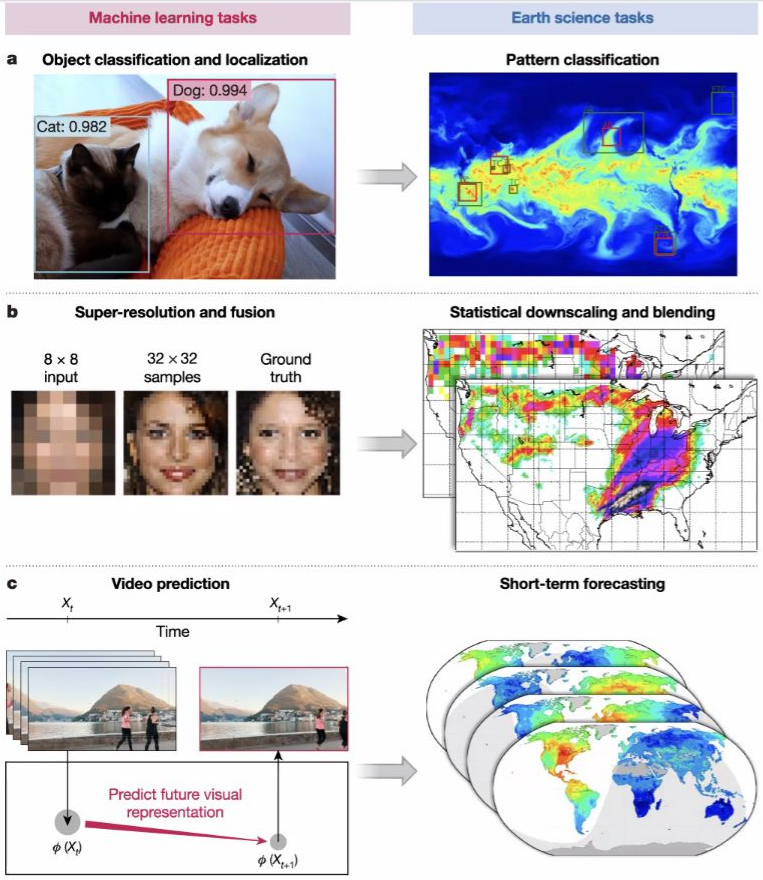
Toms et al. 2020

Predicting **the year** of 2-m temp map



Barnes et al. 2020

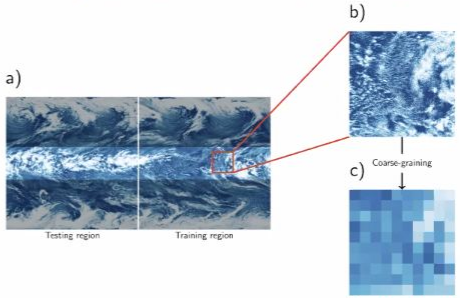
Where ML is already Being Used



Reichstein et al. (2019; SCI)

Convective parameterizations

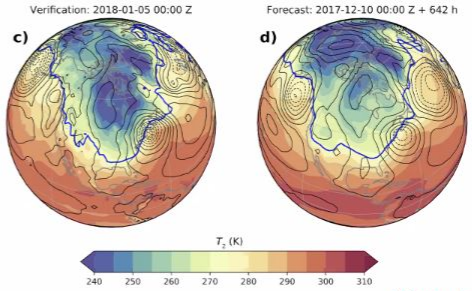
e.g. Rasp et al. (2018; PNAS); Schneider et al. (2017; GRL); O'Gorman and Dwyer (2018); Beucler et al. (2020; PRL)



Brenowitz and Bretherton (2018; GRL)

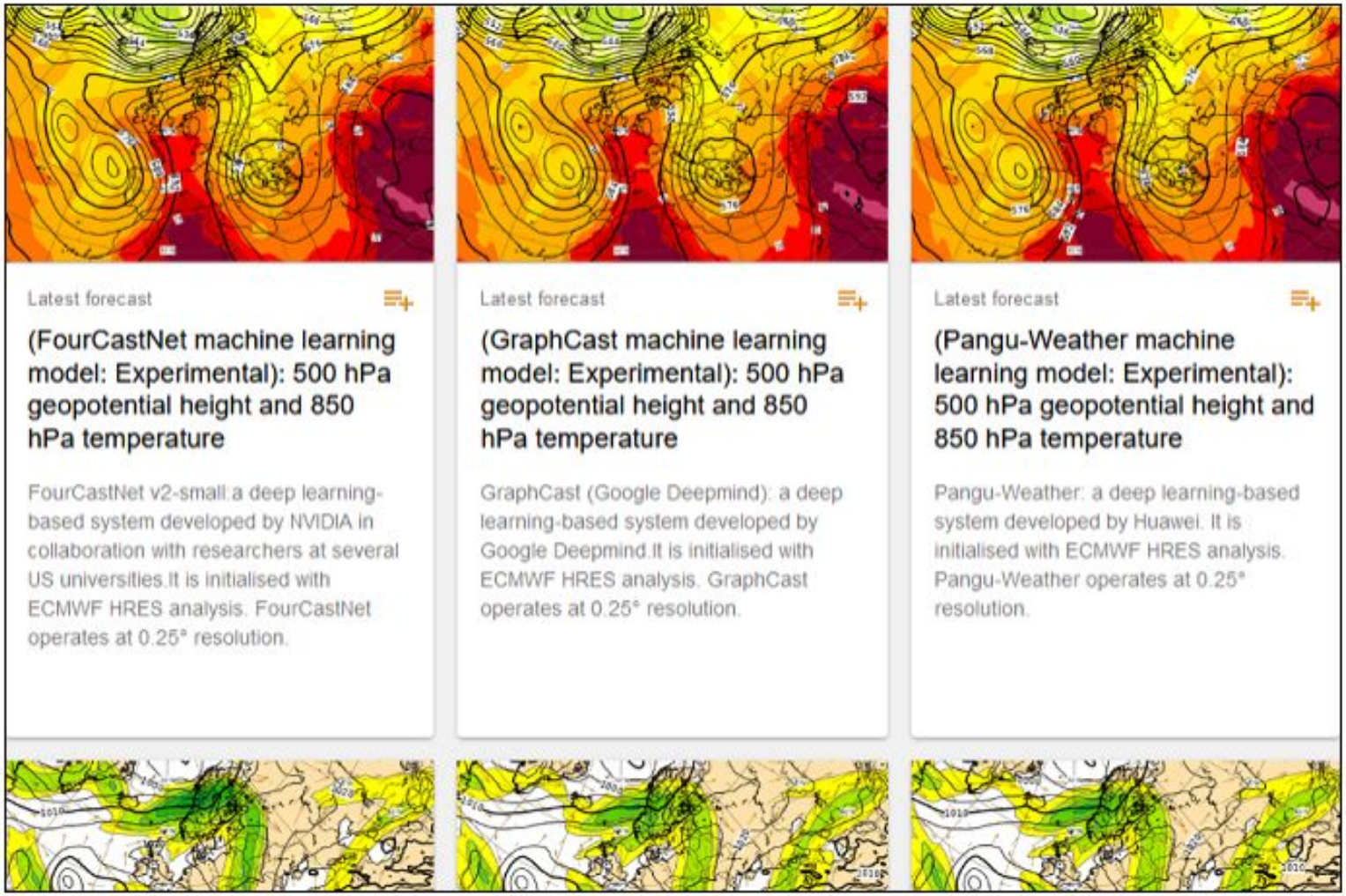
Weather Prediction

e.g. Gagne et al. (2019); Gagne et al. (2017); Chattopadhyay et al. (2019); Lagerquist et al. (2020)



Weyn et al. (2020)

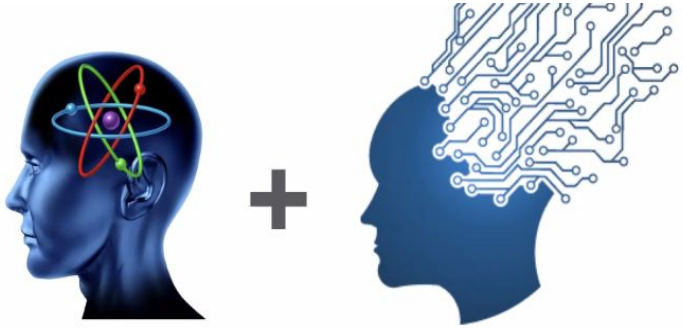
weather forecasting: a showcase of data-driven systems



Perspectives: *ML for Climate Science*

CHALLENGES

- spatio-temporal structure
- high dimensionality
- lack of concise object definitions
- paucity of ground truth (labeled data)
- rare classes/small sample size
- unlike for mainstream applications in the commercial sector, **geoscience phenomena are governed by physical laws and principles**



A way forward:

physics-guided machine learning

✓use visualization methods to see what the ANN is learnir modify the network to conform to specific physics

✓incorporate preferred physics into the model by

- penalizing for breaking physical laws
- build into the prediction itself

Perspectives: *ML for Climate Science*

- Artificial neural networks (ANNs) can **identify forced climate patterns** amidst the noise as **early as the 1960's**.

Barnes, Elizabeth A., J. W. Hurrell, I. Ebert-Uphoff, C. Anderson and D. Anderson, 2019: Viewing forced climate patterns through an AI Lens, *Geophysical Research Letters,* doi.org/10.1029/2019GL084944.

- ANNs are **no longer black boxes** - tools exist to help **visualize their decisions**. This is a "game changer" for their use in geoscience research.

Toms, Benjamin A., Elizabeth A. Barnes, and Imme Ebert-Uphoff, 2020: Physically interpretable neural networks for the geosciences: Applications to earth system variability, *JAMES, in press.*

Barnes, Elizabeth A., Benjamin Toms, J. W. Hurrell, I. Ebert-Uphoff, C. Anderson and D. Anderson, 2020: Indicator patterns of forced climate change learned by an artificial neural network, *submitted to JAMES; preprint available on arXiv.*

- ANNs can be used for more than just prediction. The ***science*** **can be what the network learns, rather than the prediction**. **Get creative!**

Learning Resources

What Softwares and Packages Help?

How to go from ANN to a much complex NN (CNN, DL)

```
from functools import partial

DefaultConv2D = partial(keras.layers.Conv2D,
                        kernel_size=3, activation='relu', padding="SAME")

model = keras.models.Sequential([
    DefaultConv2D(filters=64, kernel_size=7, input_shape=[28, 28, 1]),
    keras.layers.MaxPooling2D(pool_size=2),
    DefaultConv2D(filters=128),
    DefaultConv2D(filters=128),
    keras.layers.MaxPooling2D(pool_size=2),
    DefaultConv2D(filters=256),
    DefaultConv2D(filters=256),
    keras.layers.MaxPooling2D(pool_size=2),
    keras.layers.Flatten(),
    keras.layers.Dense(units=128, activation='relu'),
    keras.layers.Dropout(0.5),
    keras.layers.Dense(units=64, activation='relu'),
    keras.layers.Dropout(0.5),
    keras.layers.Dense(units=10, activation='softmax'),
])
```

Learning Resources and Expertise

1. **Courses:**
- Coursera: Stanford - Maschine Learning, or others …
- JSC: Introduction to Scalable Deep Learning, or others ...

1. **Youtube Lectures:**
- Parallel Computing and Scientific Maschine Learning: https://youtu.be/C3vf9ZFYbjI
- MIT - Introduction into Deep Learning: https://youtu.be/5tvmMX8r_OM

1. **Expertise:**
- Institutes Seminar: Talk at May 12th
- JSC Cross-Sectional Team Deep Learning

1. **Books**

Ethical Questions: *Trustworthy, Transparency & Reproducibility*

Ethical Questions

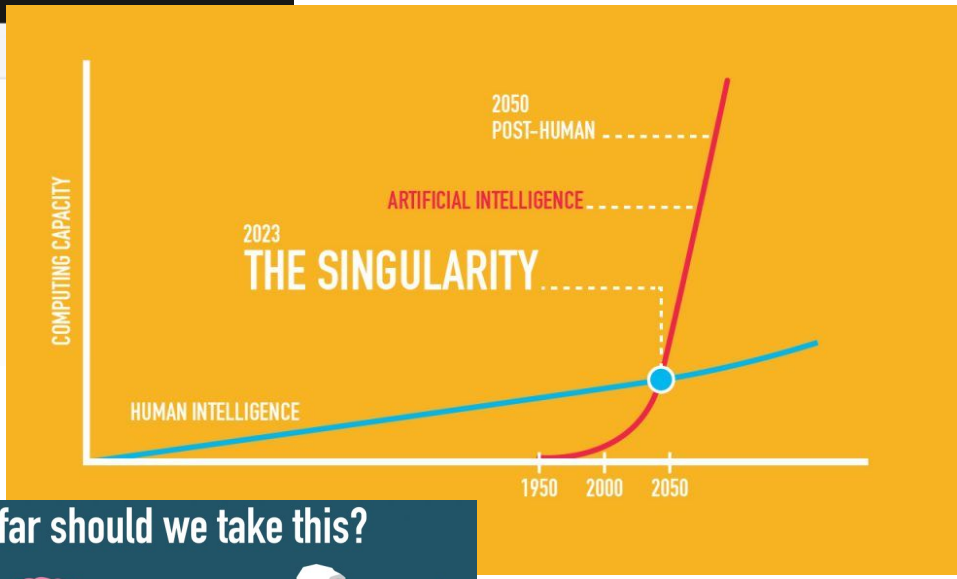
Stephen Hawking: AI will be 'either best or worst thing' for humanity

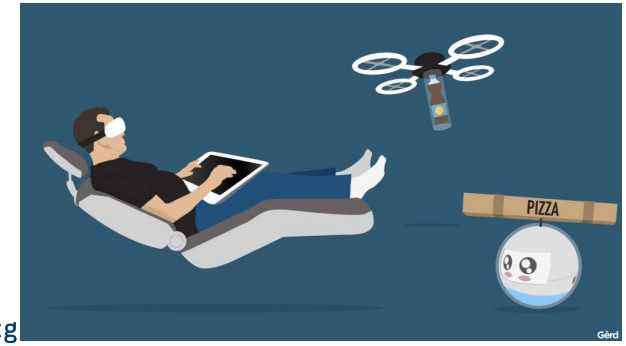
AI Could Save the World, If It Doesn't Ruin the Environment First

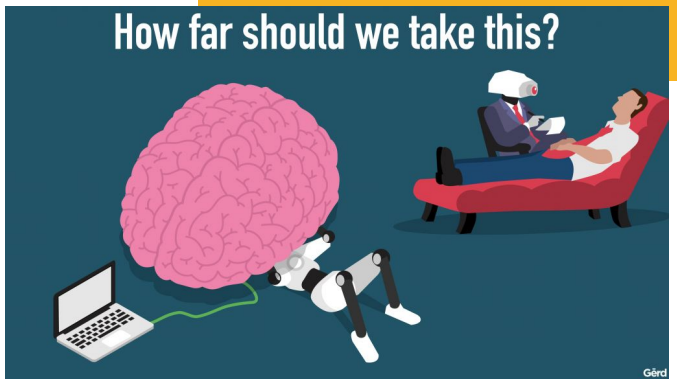
Forbes

PCMag Follow
Apr 17, 2020 · 7 min read ★

Deep Learning's Carbon Emissions Problem







Mitg

JÜLICH
Forschungszentrum