Introduction to AI, machine learning and neural networks

Dr. Sieuwert van Otterloo

Stan Meyberg

Piet Snel

July 12-16 2021

https://utrechtsummerschool.nl/courses/engineering-and-technology/introduction_to_artificial_intelligence_machine_learning_and_neural_networks

Agenda

Monday Jul12: Data science	Tue Jul13: machine learning	Wed Jul14: Standard neural networks	Thu Jul15: complicated neural networks	Fri Jul16: other AI algorithms
Data exploration and visualisation	Decision trees and regression	Prediction with neural networks	Image recognition	Evolutionary algorithms
History of AI	AI and ethics (Ethics Inc)	AI validation / medical AI	Neural network types	Business process mining

Structure of each day

Time	Content	Remarks
8.45-9.05*	Walk-in and coffee	
9.05 – 9.30*	Recap and questions	Discuss previous day. On day 1: check if people have practical questions
9.30 – 10.30	Theory	Presentation by lecturer of key concepts
10.30 - 10.45	Coffee break	
10.30 – 11.45	Practical session	Working on assignements, individual or in groups
11.45 – 12.15	Discuss practice results, conclusion	
12.15 – 13.15	Lunch	
13.15 – 14.30	Theory	Presentation by lecturer of key concepts
14.30 – 14.45	Coffee break	
14.45 – 15.45	Practical session	Working on assignements, individual or in groups
15.45 – 16.15	Discuss practice results, conclusion	
16.15 – 16.30	Time for individual questions	Lecturer is available for individual questions

* Day one will start later at 9.30

Wednesday Jul14: Standard neural networks

Programme:

Morning theory A

- Neural network structure
- Neural network training
- Introduction to Bias

Morning practical

Morning theory B

- Understanding and correcting bias in neural networks: a credit default prediction case study – Piet Snel

practical

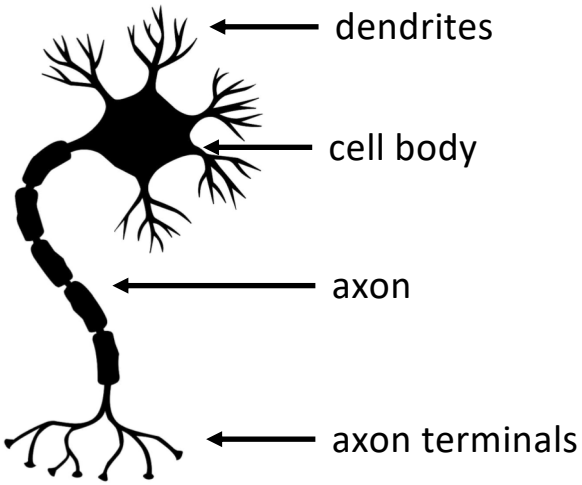
Afternoon theory

- Medical AI risks
- More on bias
- AI validation

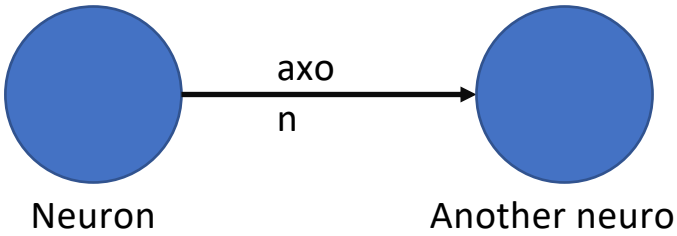
Afternoon practical

- Validation exercises

What are neural networks?



Symbolic illustration of actual neuron

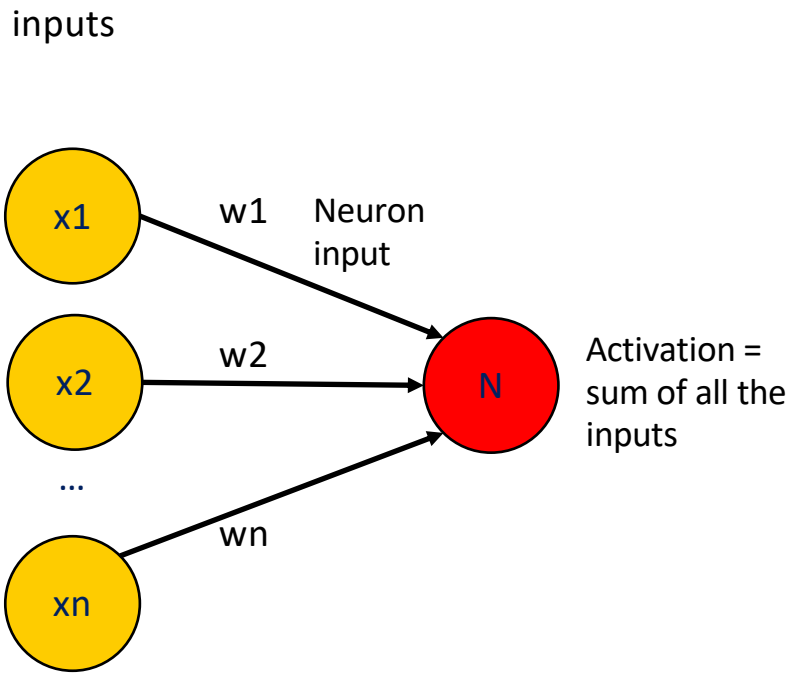


Neurons are cells that are part of the nervous system (brain, spinal cord, sensory organs, and nerves). They are often depicted as a core cell with many connections (axons/tendons) to other neurons.
The neurons thus form a neural network.

- Some neurons respond to external signals, e.g. light or sound
- Some neurons control muscle cells
- Other neurons respond to signals from other neurons. Once they are agitated by other signals, they will 'fire' and send a signal to other neurons

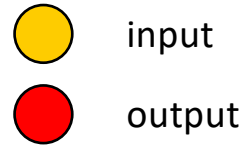
Illustration credit: Cenira Costa

How artificial, linear neural networks work



InputN = x1*w1 + … + xn*wn

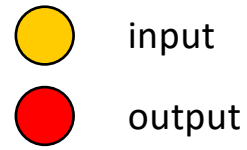
(linear combination, inproduct)

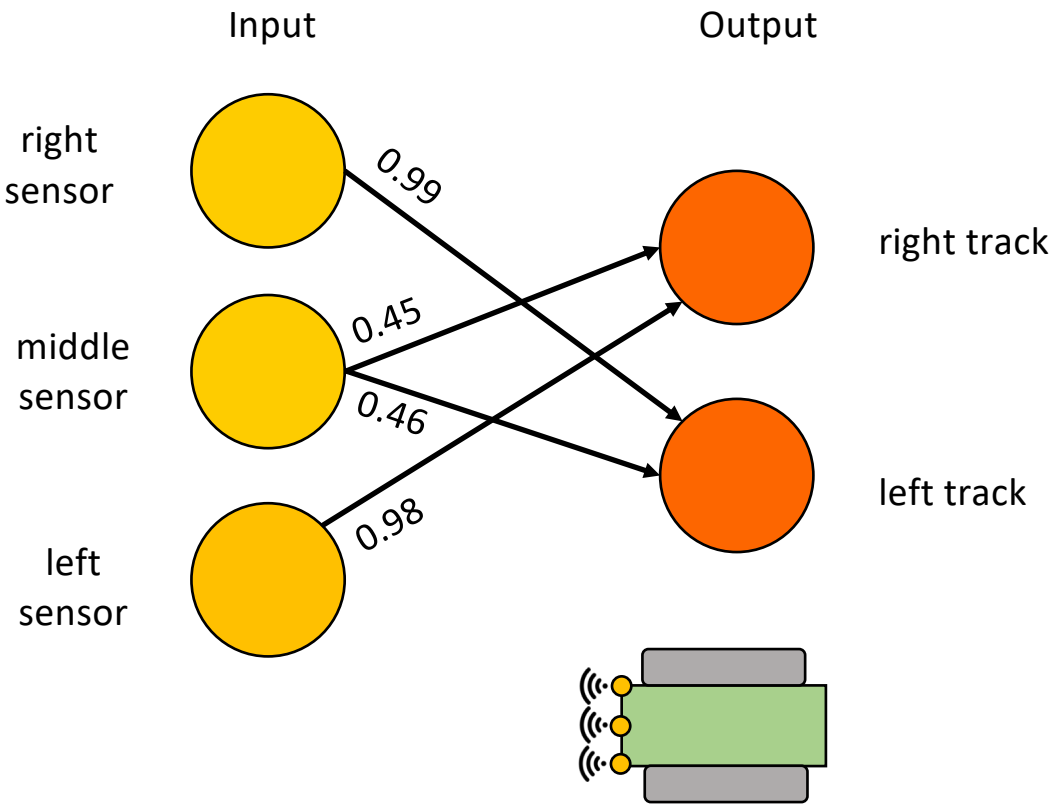


Linear neural networks

OutputN = InputN

Artificial neural networks

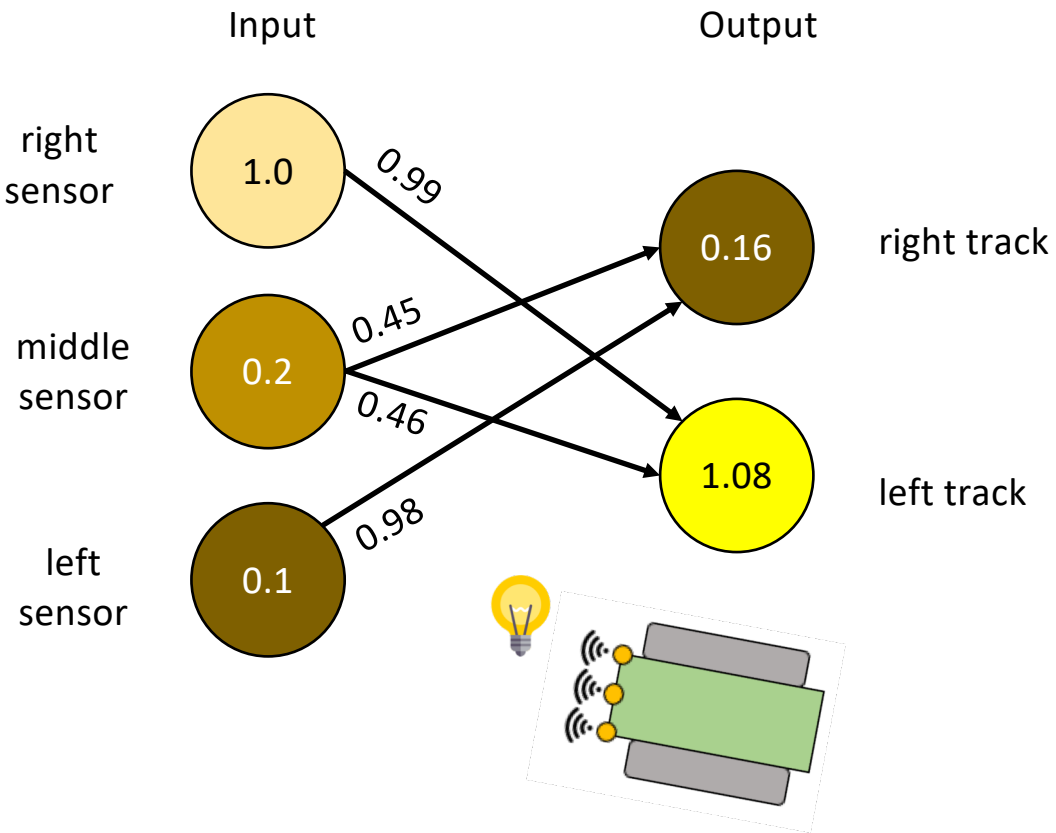




The following simple network can be used for making tiny robots that either move towards or away from the light.

- Each neuron has an ‘activation’ (0, 1 or any value in between)
- Each link has a ‘weight’, typically between -1.0 and 1.0

Artificial neural networks - example



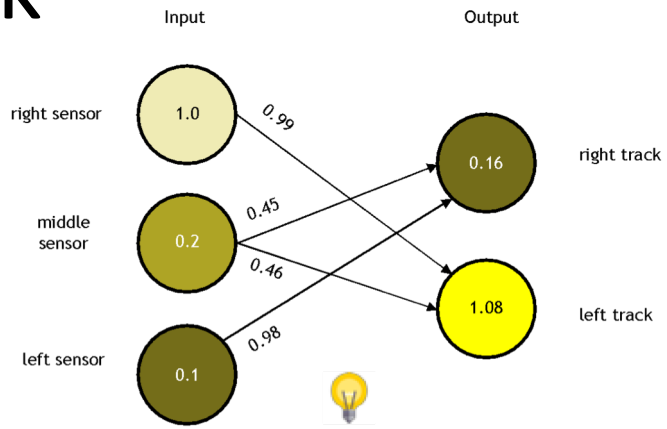
If there is a light on the right side of the robot, the left track gets the strongest signal.
The robot will turn to the right, so towards the light.

- For each neuron, you add up the sum of the input-neurons * weight
- If the sum is above the threshold (e.g. 1), the neuron is activated

Input Activation (left track) =
1.0 * 0.99. + 0.2 * 0.46 =
0.99 + 0.09 = 1.08

Exercise : Neural network

1. Write down the matrix for the neural network weights and connections shown in Python.



1. Compute the activations of the outputs for the following input vectors using a simple function:
   - [0.9,0.0,0.9]
   - [0.1,0.8,0.2]
   - [0.0,0.1,0.0]

1. Create a single new matrix such that the following input provide the specified output.
2. Is it possible to find a matrix such that the robot moves forward if either left or right is lit more than the other, but is still or goes backwards if both are lit? Please explain your answer.

Input Q3	Output Q3
(0.8,0.4,0.0)	(1.0,0)
(0.0,0.4,0.8)	(0,0.9)

Preprocessing with one hot encoding

Original data

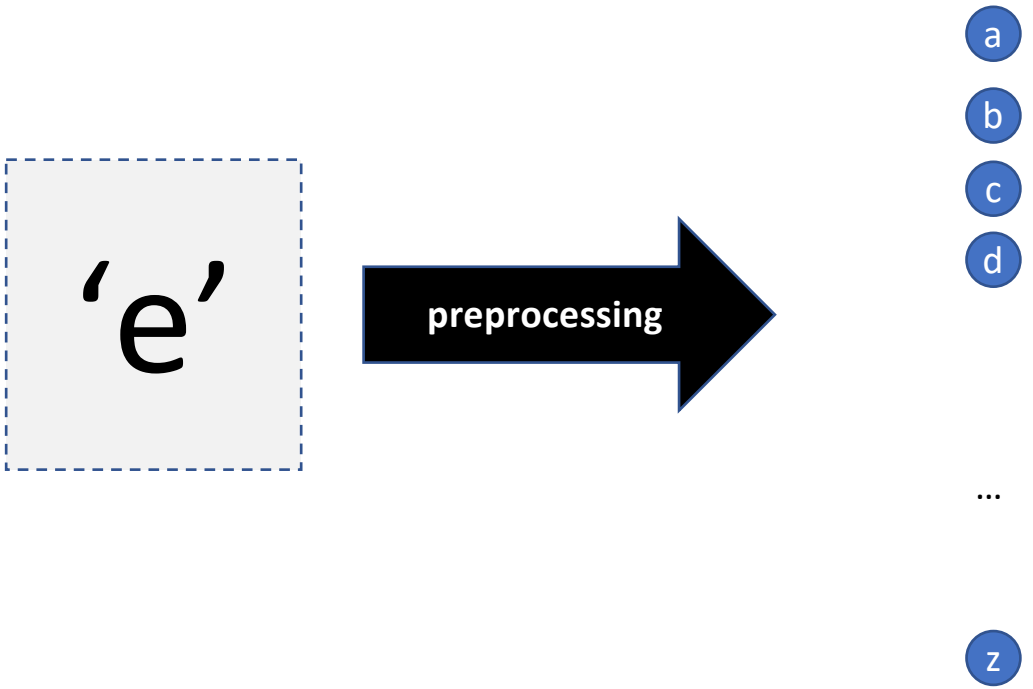
ID	Zipcode
29	3500
30	3520
31	3800

Preprocessed data

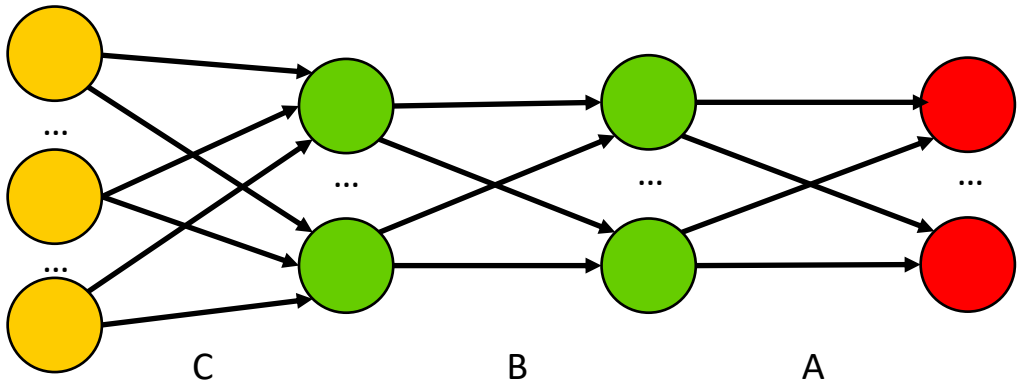
ID	Zip3500	Zip3520	Zip3800
29	1.0	0.0	0.0
30	0.0	1.0	0.0
31	0.0	0.0	1.0

Neural network input neurons only work with single-value inputs. You need to preprocess or ‘encode’ input to make it accessible for neural networks

How would you encode a character?



Multilayer linear neural networks are not useful



When computing power was scarce, AI researchers used linear neural networks: the activation of a neuron is the weighted sum of the input like we have seen before.

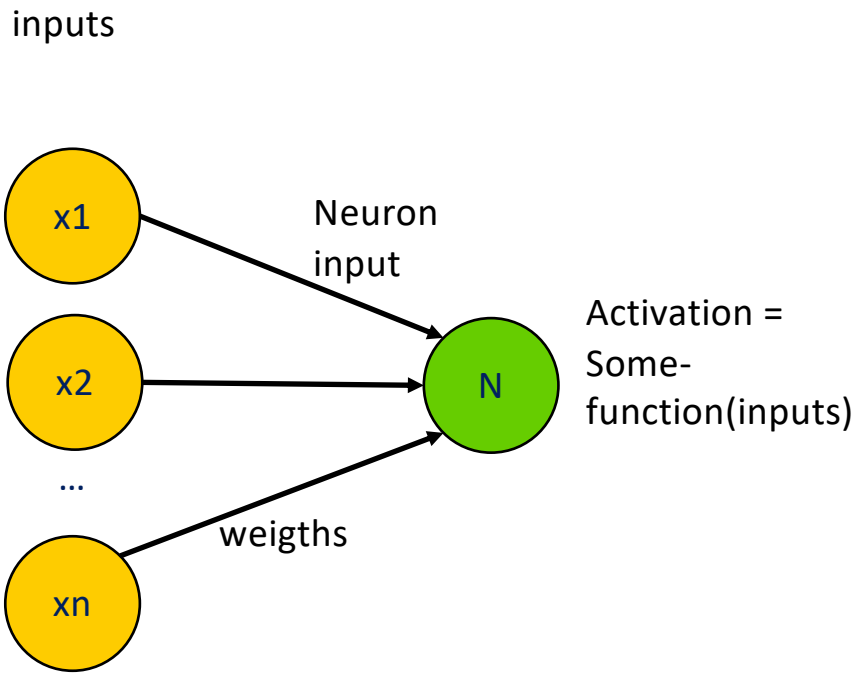
To try to make the networks able to decide complex issues, they added hidden layers of neurons. Each layer transition is modelled via a matrix.

Output = A(B(C(input)))

With linear functions you prove that this does not work, since you can prove that a smaller matrix D=A*B*C that will give exactly the same responses. A small network will do the same

As a result, all neural networks use non-linear activation functions: a threshold or S-curve.

How nonlinear neural networks work



InputN = x1 * w1 + … + xn * wn

(linear combination, inproduct)

Nonlinear neural networks:

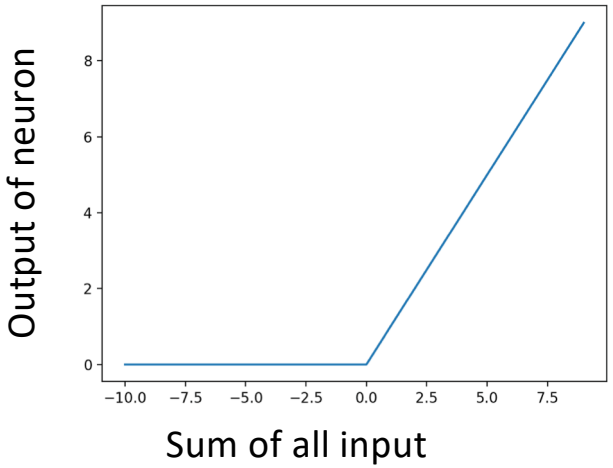
OutputN = Somefunction (sum-of-inputs)

The ‘somefunction’ can be:

- Treshold
- ReLU
- Sigmoid

Commonly used output functions: ReLU

Rectified linear unit (ReLU)



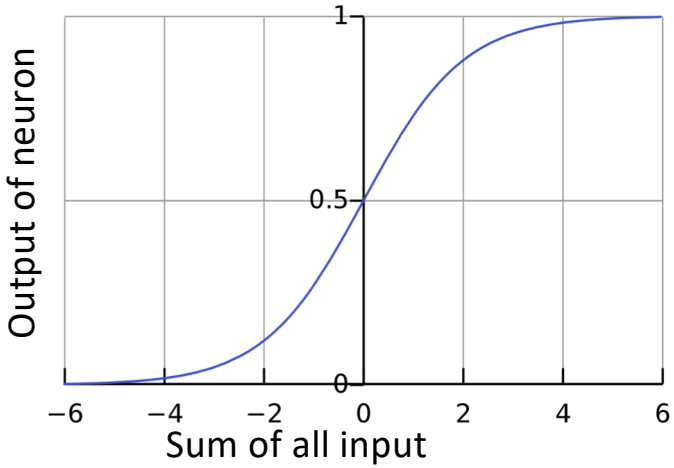
ReLU neurons can be used for the following effects:

- The first bathroom does not add value, but the second one adds € 2000
- If xcoor > 5700, subtract € 100 for each increase (away from center)
- If gardensize <100, subtract € 500 for each. Square meter less.

Commonly used output functions: Sigmoid

Logistic function

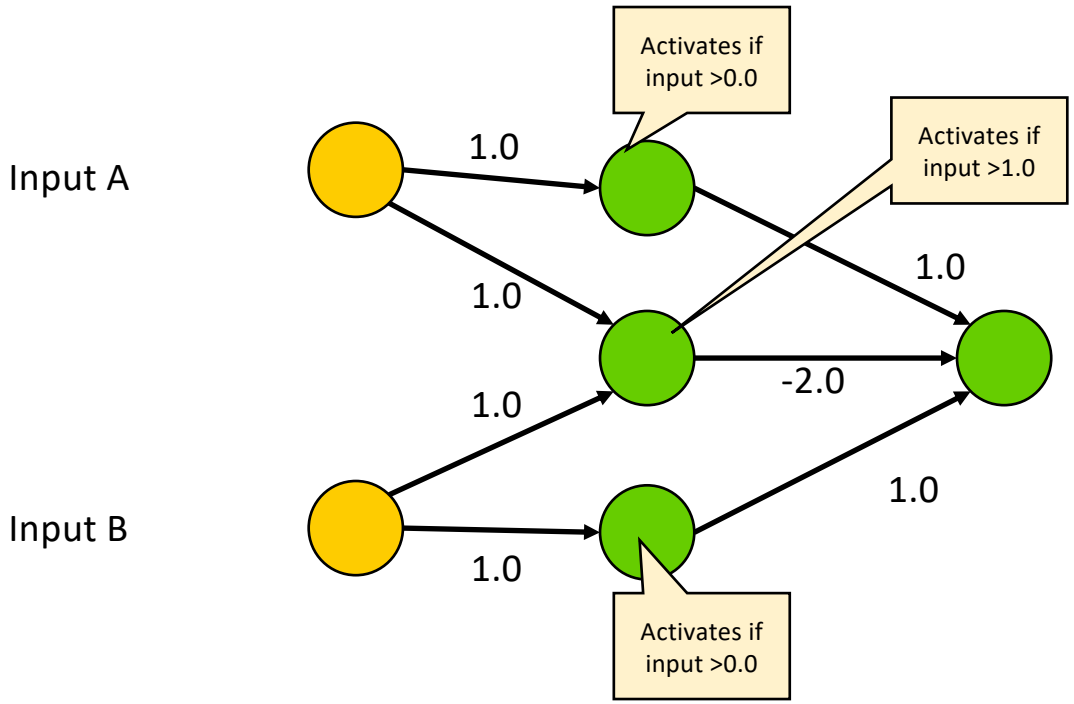
$$S(x) = \frac{1}{1 + e^{-x}}$$



Sigmoid functions are functions that approximate a step function but in a smooth way.

This function is slower to compute that the ReLU function.
It does have a non-zero derivative /gradient everywhere. This makes some optimisation algorithms work better.

Solving XOR with Relu



Neural network pros and cons

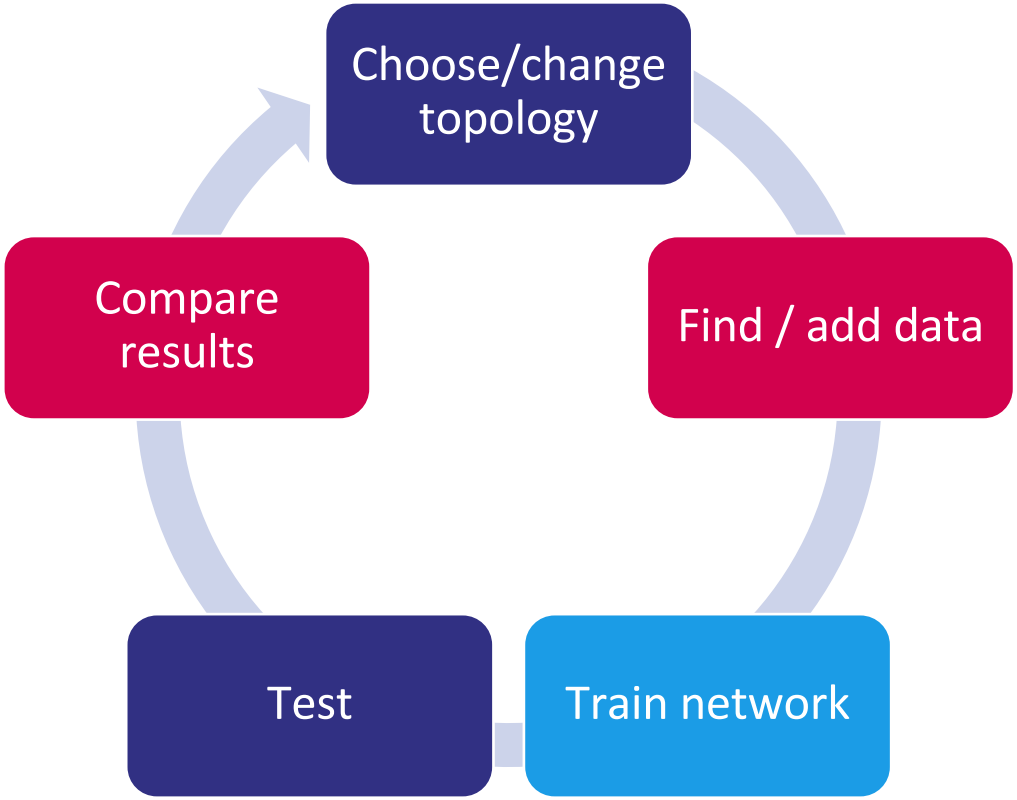
Advantages

- Can solve many problems that simpler algorithms cannot
- Results improve with data collection, even without software development
- Neural networks are fast in making predictions

Disadvantage

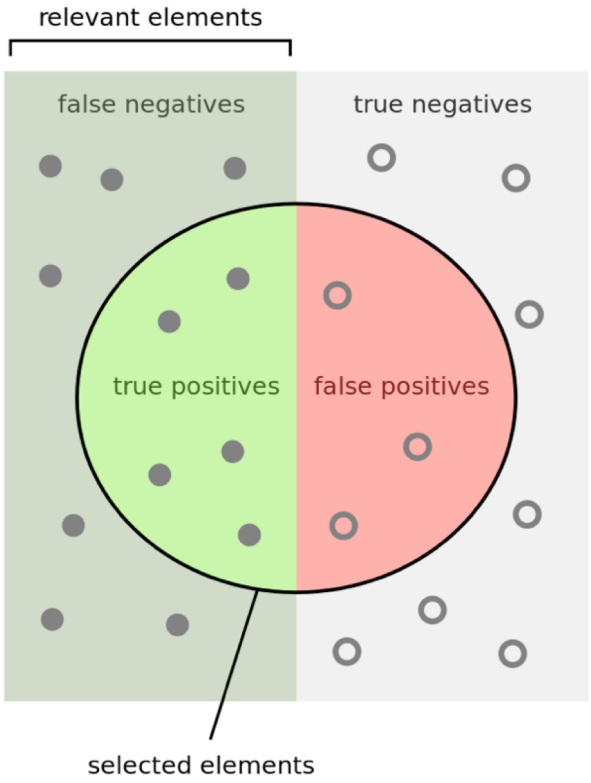
- Neural networks are black boxes: the results are not easily explainable and often not understood by the researchers
- Neural network implementation can have hidden mistakes, biases or discrimination
- Neural network training can take a long time and require lost of data

Neural network training cycle



- A data scientist will repeatedly train and test neural networks to find the best network topology and configuration
- Larger topologies are only better if you have enough data. You need to experiment to find the best combination
- ‘Finding data’ is an ongoing problem. “Data is the new oil” – a very scarce resource

Definition accuracy



The accuracy of a prediction algorithm is a percentage, defined as

#<correct answers> / #< answers>

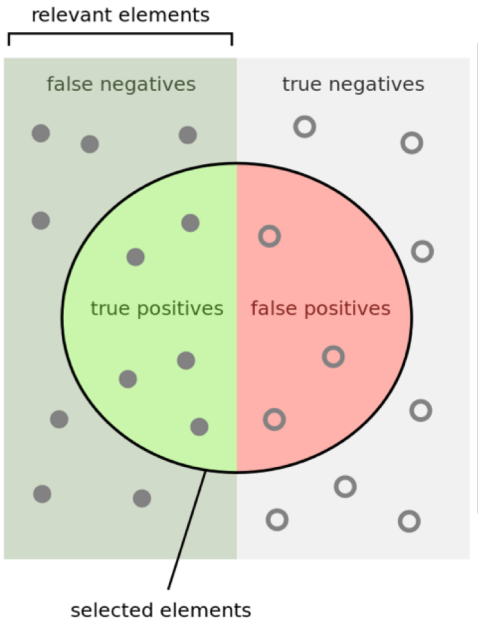
Answers can be incorrect in two ways:

- False negative: algorithm says ’no’, actual answer is yes
- False positive: algorithm says ’yes’, actual answer is no

Similarly, there are two types of correct answers:

- True negative: algorithm says ’no’, actual answer is no
- True positive: algorithm says ’yes’, actual answer is yes

Other metrics: Precision and recall



How many selected items are relevant?

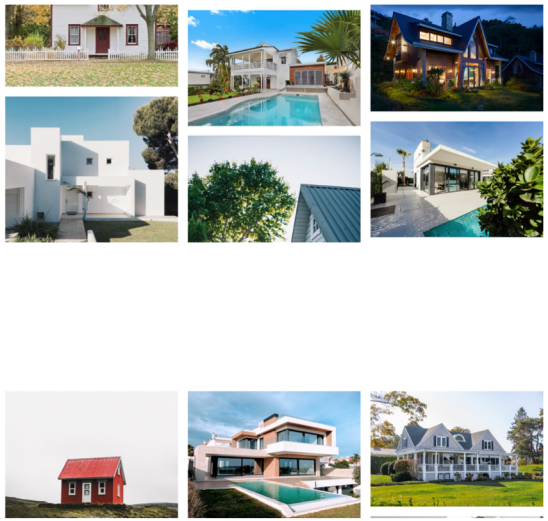
$$\text{Precision} = \frac{\text{true positives}}{\text{selected elements}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{relevant elements}}$$

- Precision is important if the cost of a false positive is high. This is for instance the case in hiring for popular positions, or in case of high risk treatment in non-urgent cases
- Recall is important if the cost of false negatives are high

Train vs. test data

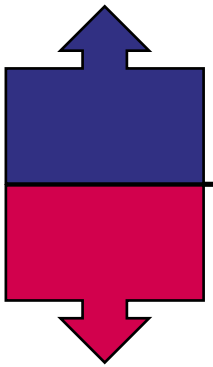


- Neural networks are known to ‘overfit’: they have perfect accuracy on data that has been used to train the network, since they ‘recognize’ the data based on trivial features
- To properly evaluate a trained NN, you should test it using data that was not used in training. You therefor split your dataset into a training part and a testing part
- You can decide how to split:
  - Classically schooled researchers split 50%-50% because of symmetry;
  - Young mavericks often split 80% training -20% testing to make best use of data

Dataset 1: Training and testing

All houses

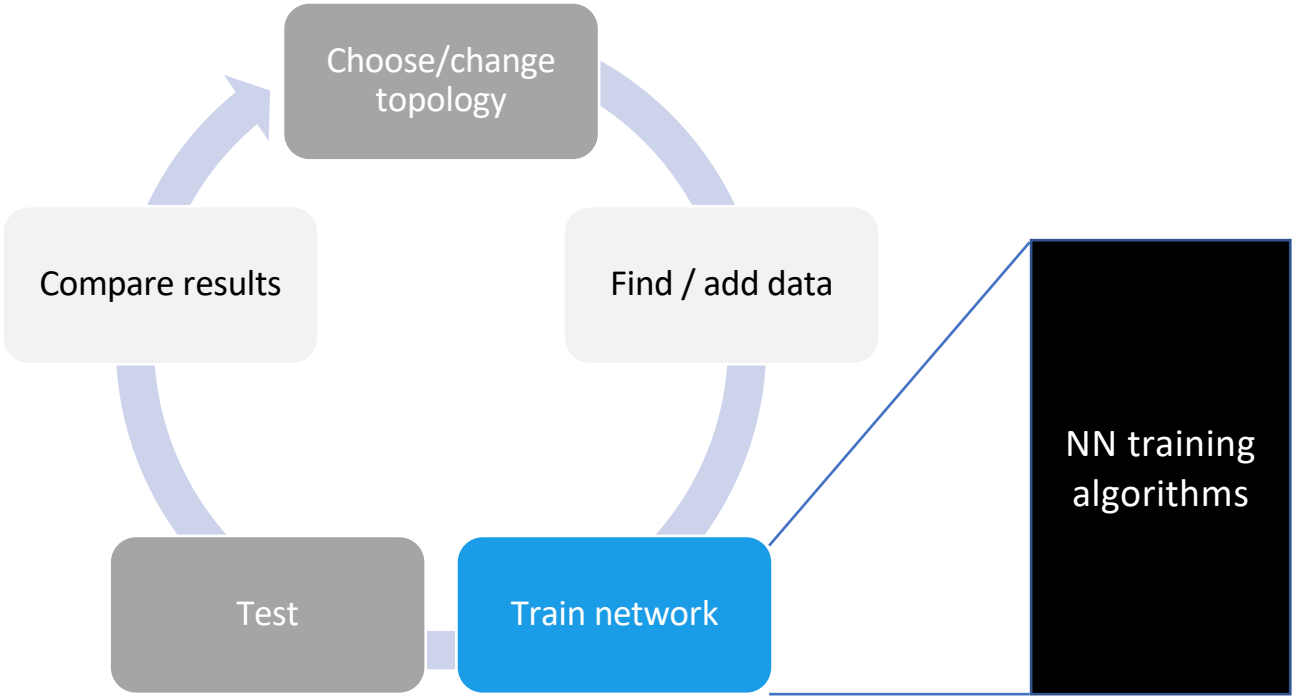
80 data points



Training data

Test data

Training

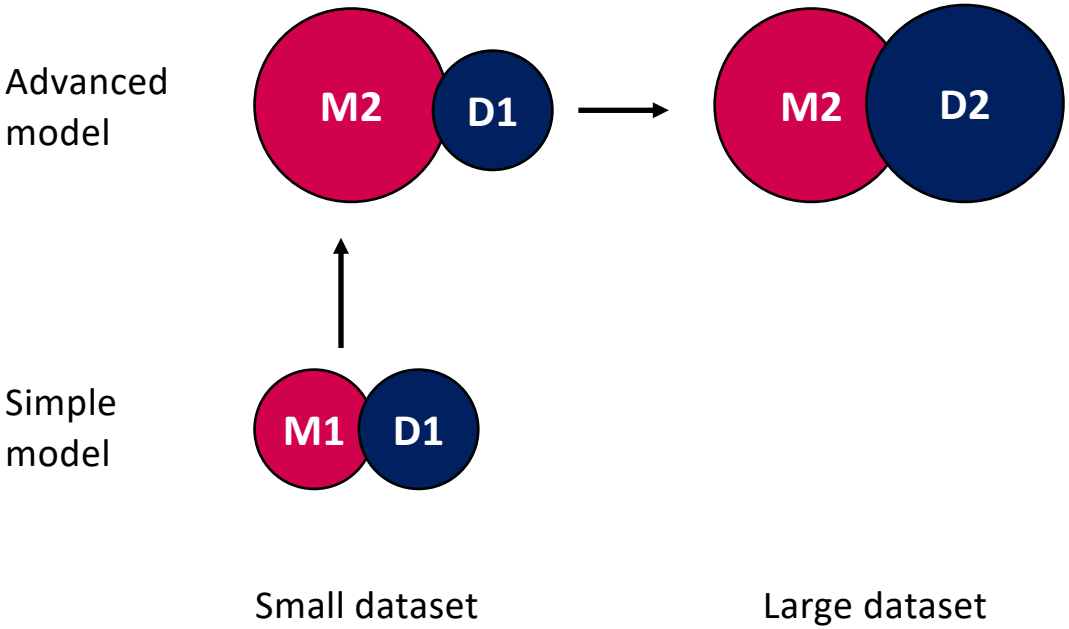


- Neural network training algorithms are a black box even for most data scientists.
- They are optimisation algorithms so presumably they are similar to gradient search.
- ‘Adam’ is the algorithm we use. It is very fast if you have an expensive computer

A Exercise

- Train the neural network to estimate house value based on location (xoor, ycoor) with a simple topology and 50-50 train-test-split. What is the accuracy?
- Train the neural network to estimate house value based on location (xoor, ycoor) with a simple topology and 80-20 train-test-split. What is the accuracy?
- Train the neural network to estimate house value based on location (xoor, ycoor) and tax value with a simple topology. What is the accuracy?
- Train the neural network to estimate house value based on location (xoor, ycoor) and tax value with a topology with more layers. What topology gives the best results?

Playing with models and data



- Initial results are often insufficient. You need to repeat training and testing to get high accuracy
- When repeating, you can use a different topology, or a better dataset.

B Exercise

- Train your last neural network again, but with the larger dataset. Do the results improve?
- Which topology works best for your dataset?

Common problems during training

Symptom	Possible root cause	Next step
Low accuracy on training set, low accuracy on test set	Not enough training	Increase the number of training rounds
Persistent low accuracy on training set, low accuracy on test set	The network topology is too simple to learn this concept, or we have bad data	Change the topology by adding more layers, or clean up the data
High accuracy on training set, low accuracy on test set	'Overfitting': the network has learning to recognize the training set	Simplify the topology, or add more data.
Training time too long, training accuracy does not converge	The network topology is too complex to learn this concept	Simplify the topology
High accuracy on training set, high accuracy on test set	It works!	Stop and save the trained network
Low accuracy on training set, high accuracy on test set	This is not possible.	Manually inspect your code and data

Explainability in AI is an active research area

Technical research

- Generate explanations (LIME, SHAP)
- Combine NN with more explainable technologies

NN as a black box

Social research

- Investigate whether people need explanation in specific cases
- Investigate what type of explanation people need
- Test what bias/mistakes people find acceptible

Definition of Bias

- In algorithm is ***biased*** if it produces certain outcomes more ofte, and other outcomes less, than is warranted by the input data or solution requirements

Bias example

- Nolan Bushell called his company Atari since he noticed that companies whose name starts with ‘A’ get more calls than other companies. Apparently people call companies in alphabetical order.
- Steve Jobs took notice and also believed in this effect, and called his company Apple so that it would be listed before Atari
- Activision was started by disgruntled Atari employees, who choose their name so that it would appear before Atari

Forms of bias you should consider

As an AI researcher, you should know the following types of bias. Discussing the full history and overall impact on society on these biases is not in scope of today's lecture.

Traditional bias

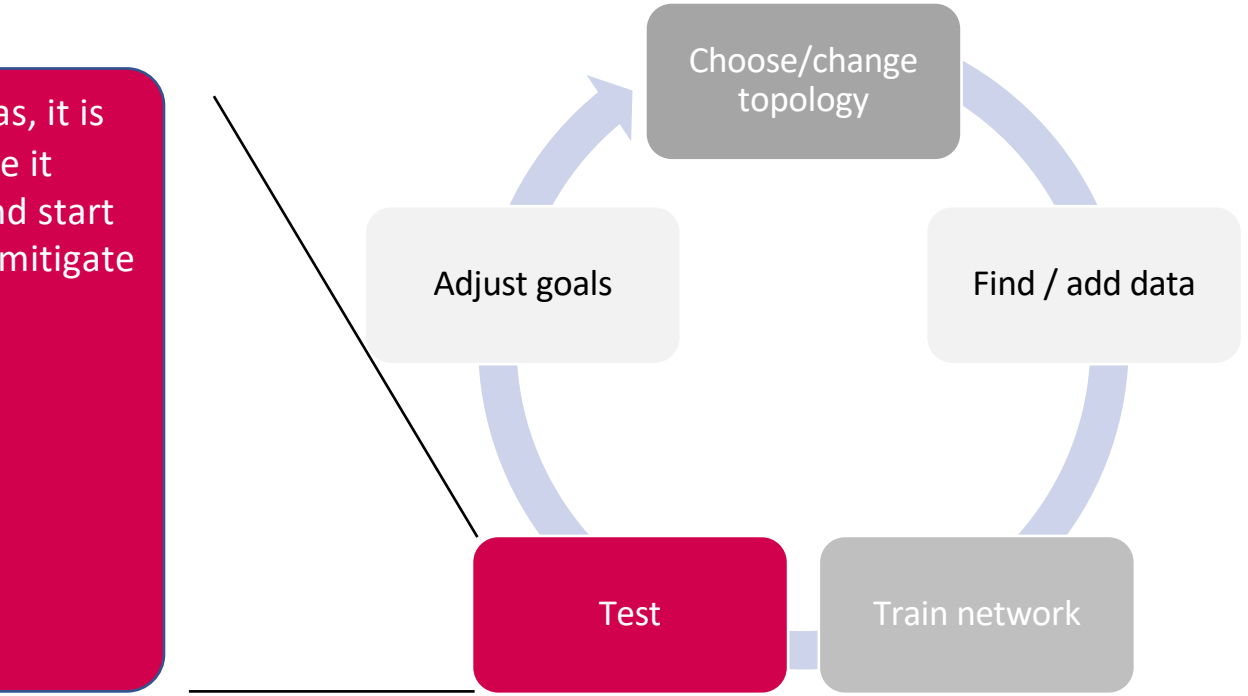
- **Gender bias**: people have overlooked qualified women for positions
- **Age bias**: by using age as a substitute for experience, candidates with non-traditional career paths are excluded from suitable positions
- **Cultural background**: people have overlooked qualified people with non-western-European backgrounds.

Bias in AI

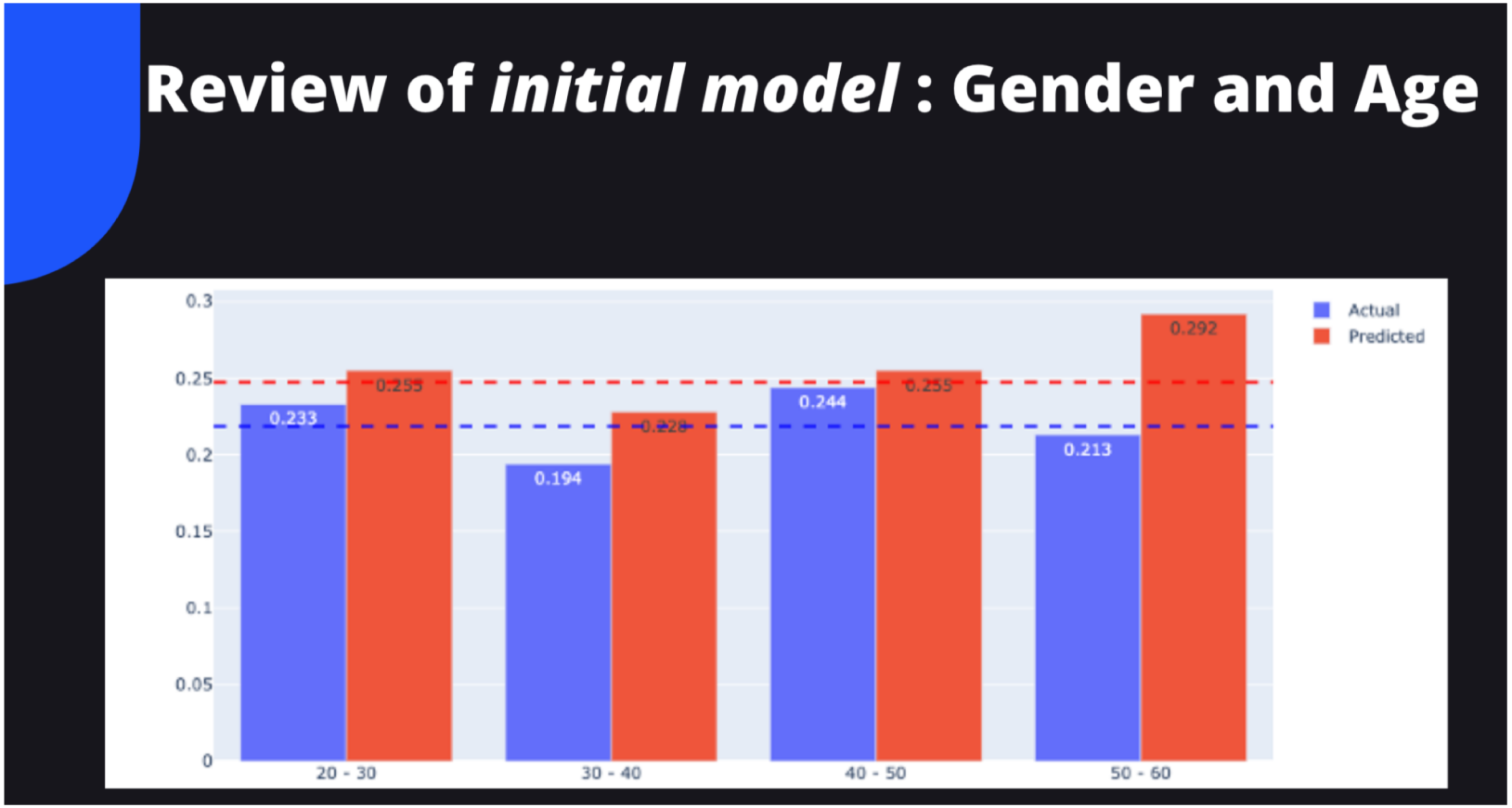
- **Gender, age bias**: Algorithms are often only tested on AI grad students. This subgroup is not representative for society as a whole
- **Skin colour bias**: Visual recognition algorithms are often not tested on all skin tone variations.
- **More gender, age and health bias**: For reasons, medical research is often only conducted with healthy (young) male students

How to address bias?

- Once you start testing for bias, it is easy to discover and measure it
- If you can measure it, you and start searching for adjustment to mitigate bias



Piet Snel: measuring and correcting bias



Credit default dataset

ID: ID of each client
LIMIT_BAL: Amount of given credit in NT dollars (includes individual and family/supplementary credit
SEX: Gender (1=male, 2=female)
EDUCATION: (1=graduate school, 2=university, 3=high school, 4=others, 5=unknown, 6=unknown)
MARRIAGE: Marital status (1=married, 2=single, 3=others)
AGE: Age in years
PAY_0: Repayment status in September, 2005 (-1=pay duly, 1=payment delay for one month, 2=payment delay for two months, … 8=paym
…
PAY_6: Repayment status in April, 2005 (scale same as above)
BILL_AMT1: Amount of bill statement in September, 2005 (NT dollar)
BILL_AMT2: Amount of bill statement in August, 2005 (NT dollar)
…
BILL_AMT6: Amount of bill statement in April, 2005 (NT dollar)
PAY_AMT1: Amount of previous payment in September, 2005 (NT dollar)
PAY_AMT2: Amount of previous payment in August, 2005 (NT dollar)
..
PAY_AMT6: Amount of previous payment in April, 2005 (NT dollar)
default.payment.next.month: Default payment (1=yes, 0=no)

https://www.kaggle.com/uciml/default-of-credit-card-clients-dataset

Bias questions

1. Load the credit-default-dataset and make a correlation matrix
2. Train a model1 to predict default-probability based on all available inputs
3. How much gender bias is present in model1? Is it more favorable to men or women?
4. Can you use the code provided to make an adjustment? What adjustment is needed to correct the gender bias?
5. Can you make a script to show 60- , 60+ age bias in a chart?
6. Can you make a script with five age categories and estimate bias?

Bias questions

1. Plot a correlation matrix and increase its size to make it easier to interpret
2. Make the calculation of 'default_count' and 'no_default_count' for Actual instead of Predicted
3. Plot actual and predicted default rate for the ‘AgeGroupSex’ column
4. Apply a treshold correction for females of age 50 - 60 and compare results prior to and after running this function