

# Machine Learning: An Introduction

Joseph Bullock

*Institute of Particle Physics Phenomenology*



Science & Technology  
Facilities Council

# Bad Press

BBC Your account News Sport Weather iPlayer TV Radi

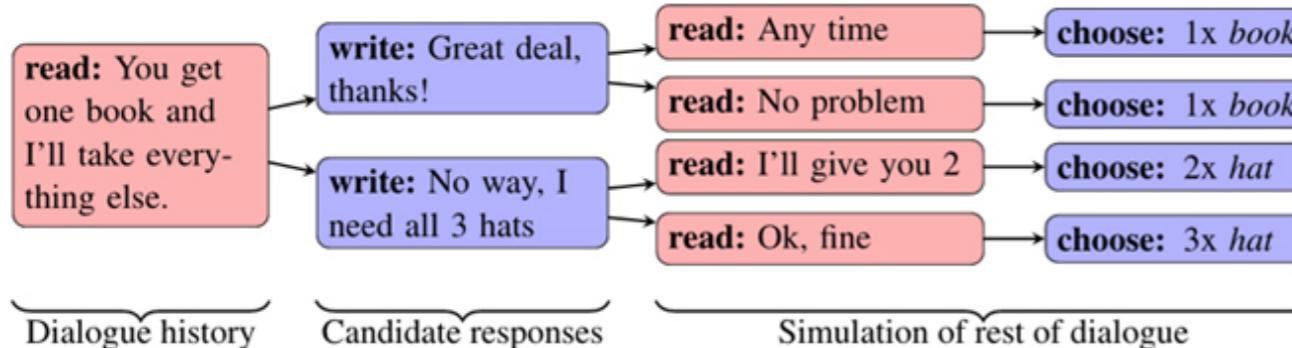
## NEWS

Home | UK | World | Business | Politics | Tech | Science | Health | Family & Education

### Technology

## The 'creepy Facebook AI' story that captivated the media

By Chris Baraniuk  
Technology reporter



## Deal or No Deal? End-to-End Learning for Negotiation Dialogues

Mike Lewis<sup>1</sup>, Denis Yarats<sup>1</sup>, Yann N. Dauphin<sup>1</sup>, Devi Parikh<sup>2,1</sup> and Dhruv Batra<sup>2,1</sup>  
<sup>1</sup>Facebook AI Research      <sup>2</sup>Georgia Institute of Technology  
{mikelewis, denisy, ynd}@fb.com    {parikh, dbatra}@gatech.edu

BBC Your account News Sport Weather iPlayer TV Radi

## NEWS

Home | UK | World | Business | Politics | Tech | **Science** | Health | Family & Education

### Technology

## Stephen Hawking warns artificial intelligence could end mankind

By Rory Cellan-Jones  
Technology correspondent

“...such articles are actually rather impressive, because they succinctly summarise the scenario that AI researchers don’t worry about. That scenario combines as many as three misconceptions: concern about *consciousness, evil and robots*”

Max Tegmark, Future of Life Institute



# Google Photos

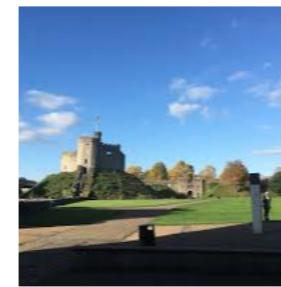
◀ **llama** ...

Thu, 7 Sep 2017



◀ **castle** ...

Fri, 27 Oct 2017



Thu, 7 Sep 2017

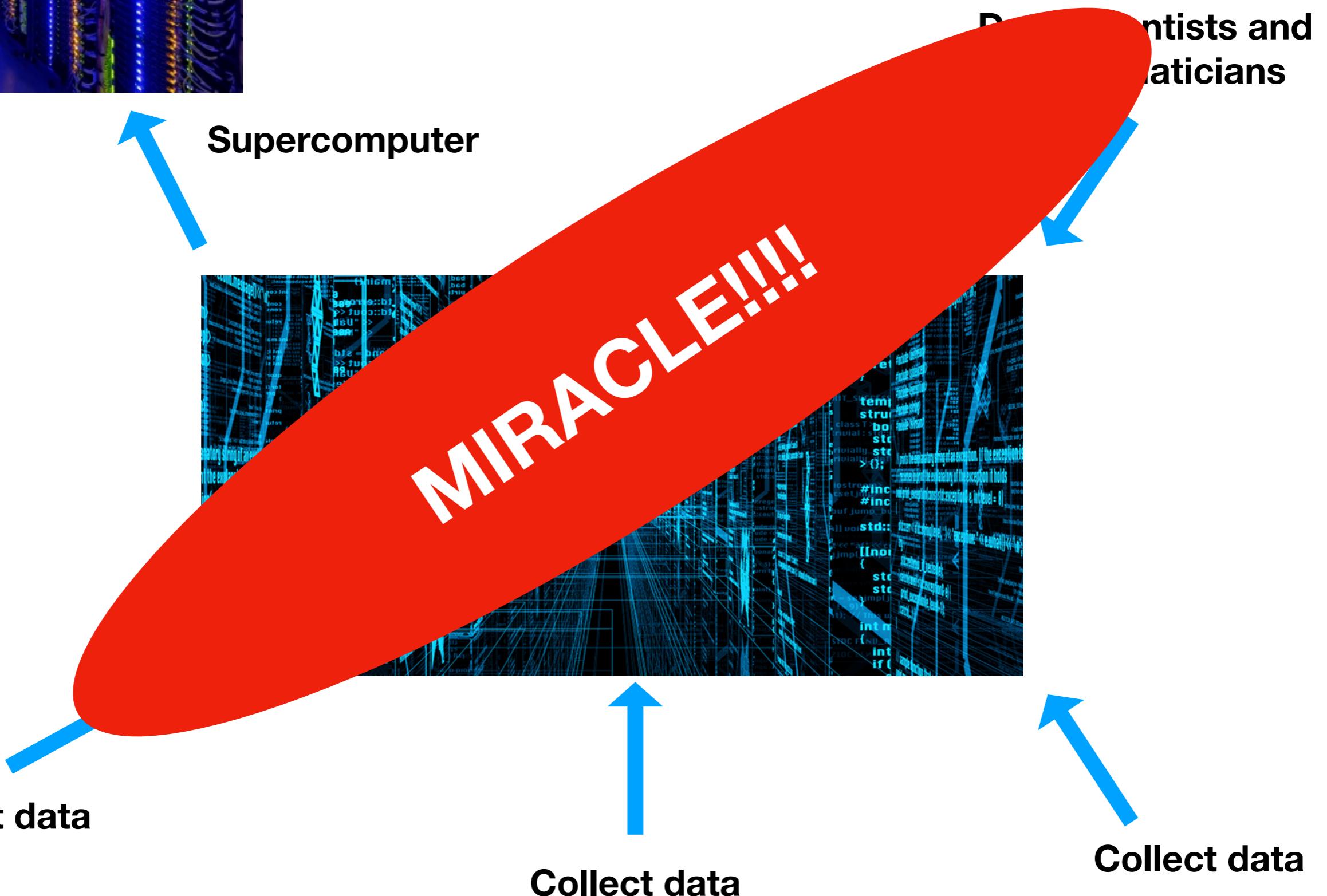


Wed, 16 Aug 2017





# Industry and Political view



[Top image: [Inside HPC](#); Middle image: [Numeric Analytics](#)]

# Types of Machine Learning

## **Supervised Learning**

- Training sets with labelled samples

## **Unsupervised Learning**

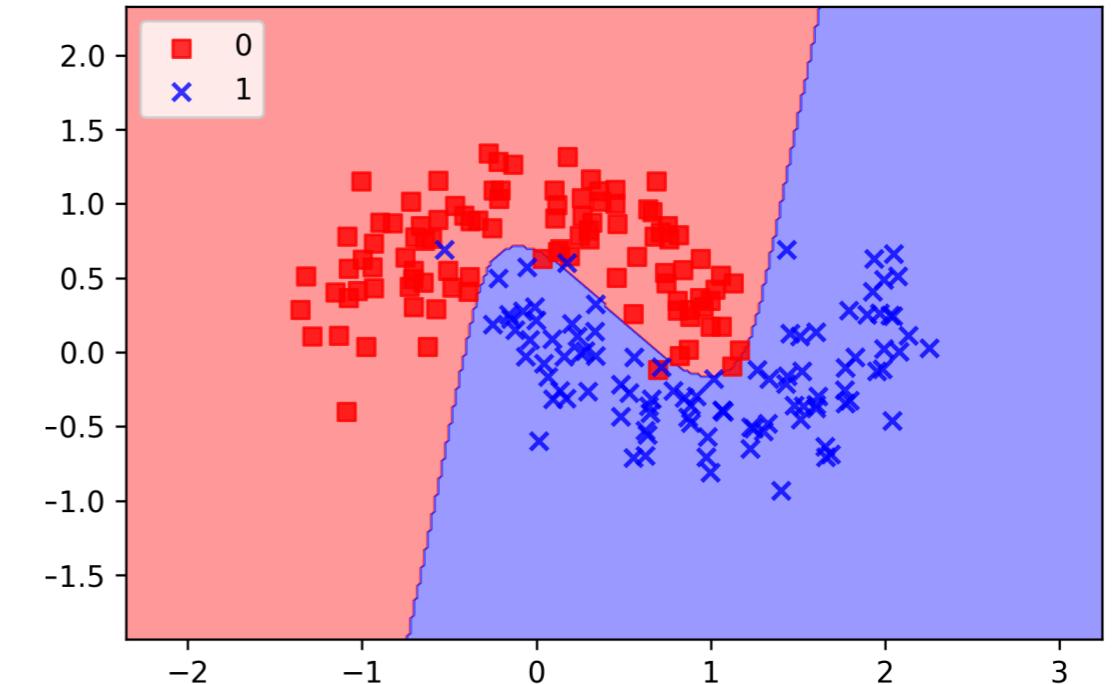
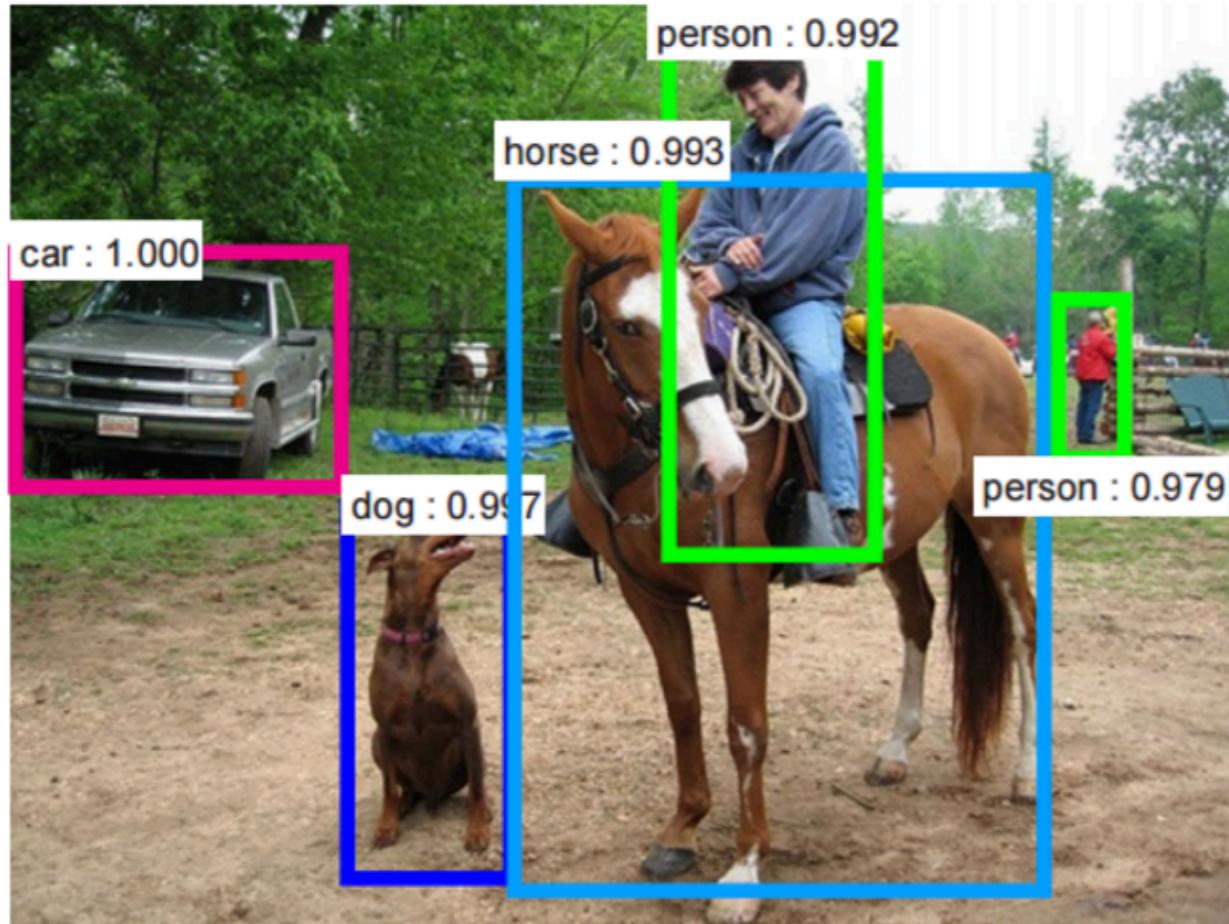
- No labels
- Model has to find features
- Can pre-classify before supervised learning

## **Deep Learning**

- Any system which layers the above styles of learning
- Neural Networks

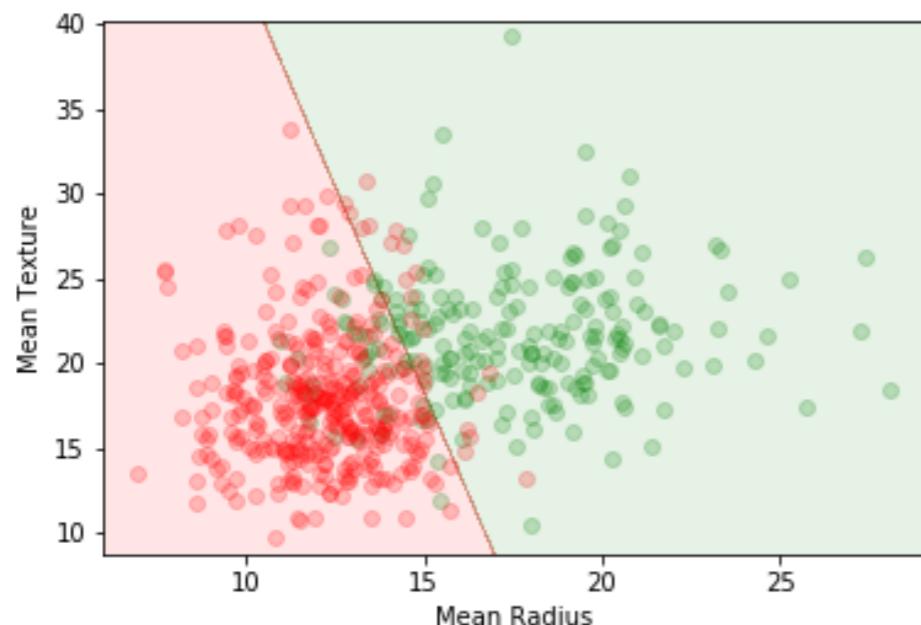
# What Really is Machine Learning?

## Classification problem



llama

...



Thu, 7 Sep 2017



Left top: [Data Science Blog](#); Left bottom: adapted from [Maire '18](#)

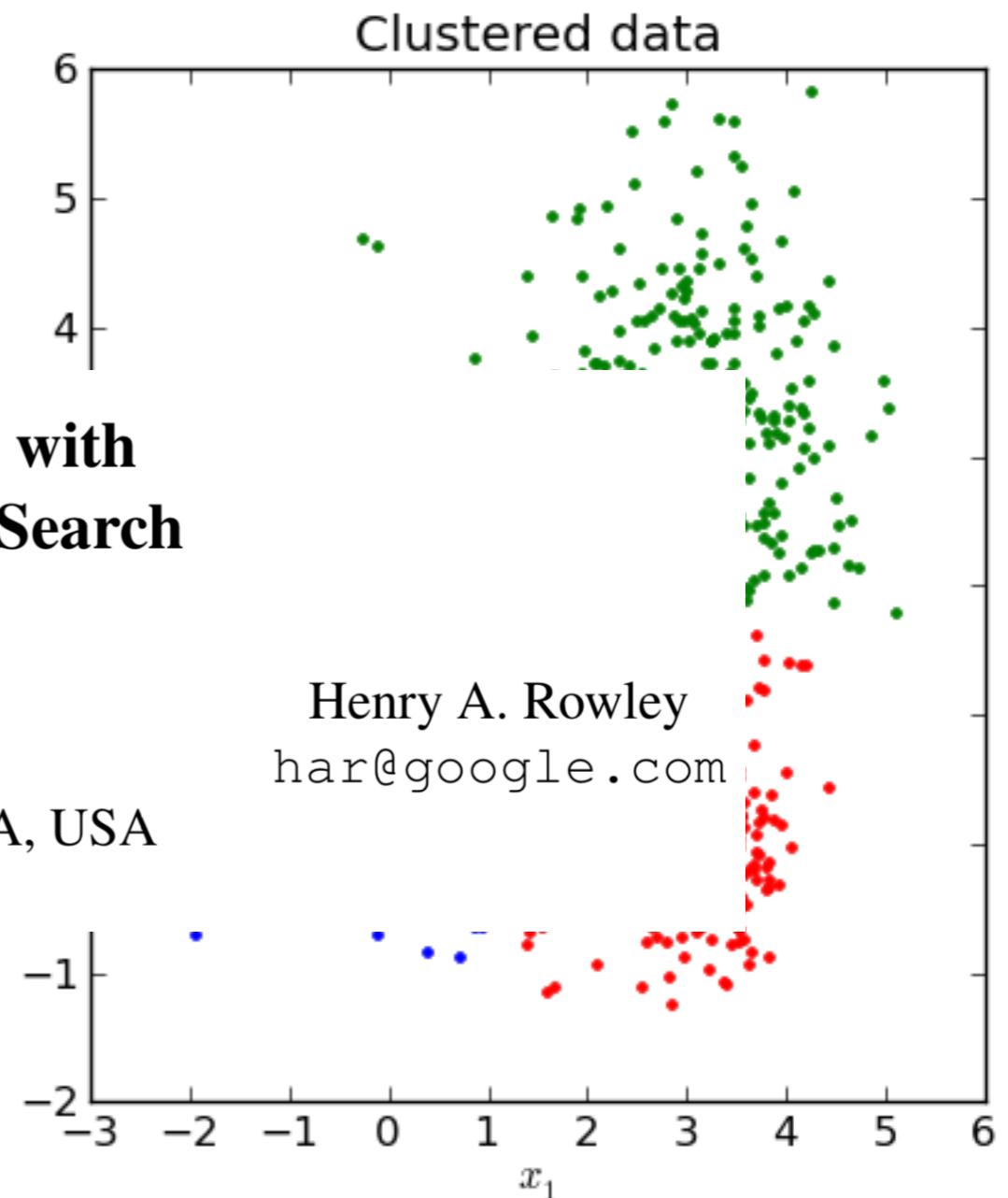
# Supervised Learning - K Nearest Neighbours

## Clustering Billions of Images with Large Scale Nearest Neighbor Search

Ting Liu  
[tingliu@google.com](mailto:tingliu@google.com)

Charles Rosenberg  
[chuck@google.com](mailto:chuck@google.com)  
Google Inc., Mountain View, CA, USA

Henry A. Rowley  
[har@google.com](mailto:har@google.com)

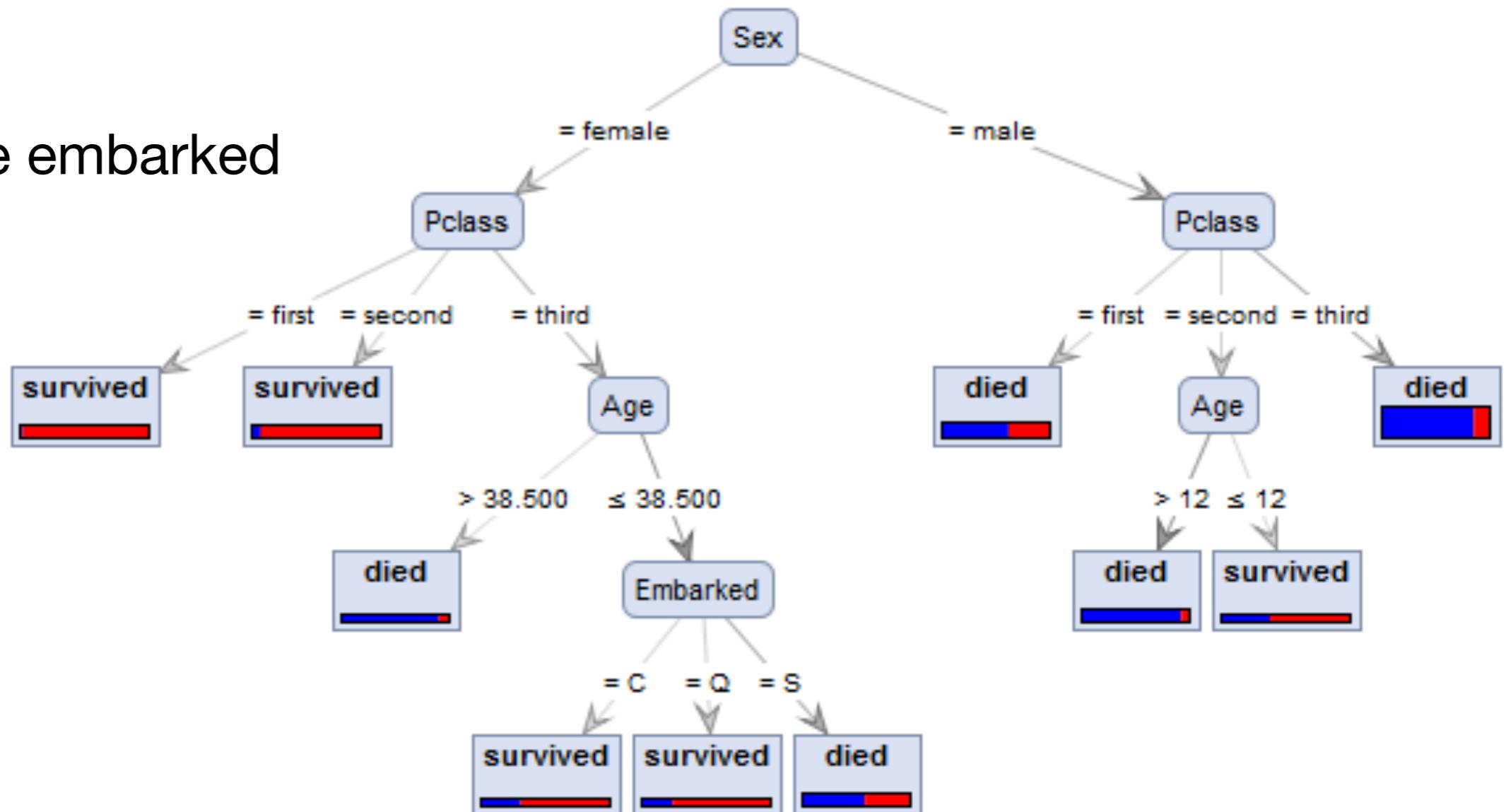


[Graph: [Mubaris '17](#)]

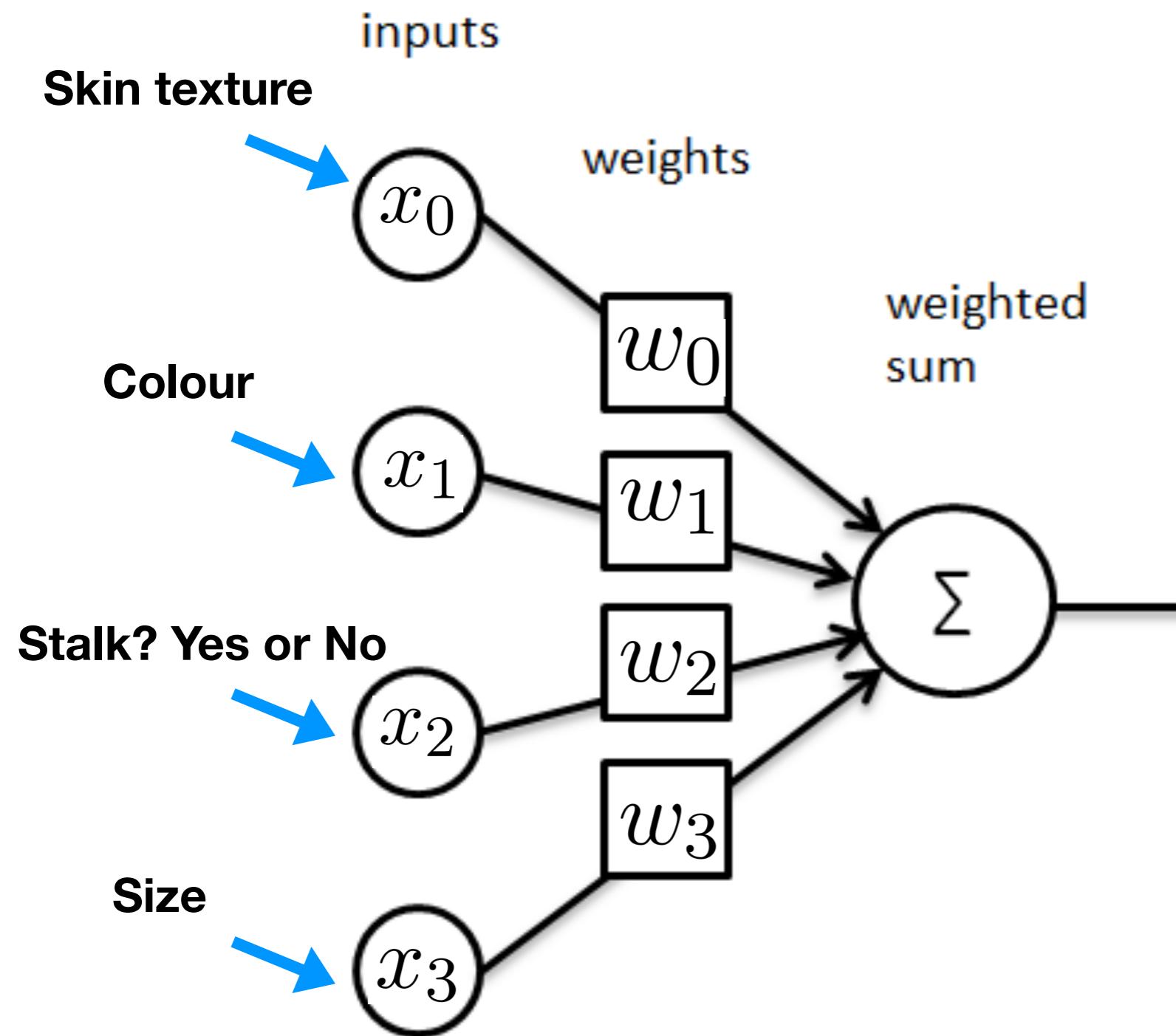
# Supervised Learning - Decision Tree

## Input parameters:

- Sex
- Age
- Class
- Where embarked



# Supervised Learning - Perceptron



# The Iris Dataset

sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
5.9	3.0	4.2	1.5
5.8	2.6	4.0	1.2
6.8	3.0	5.5	2.1
4.7	3.2	1.3	0.2
6.9	3.1	5.1	2.3
5.0	3.5	1.6	0.6
5.4	3.7	1.5	0.2
5.0	2.0	3.5	1.0
6.5	3.0	5.5	1.8
6.7	3.3	5.7	2.5



Setosa

0

Versicolor

Virginica



2



1

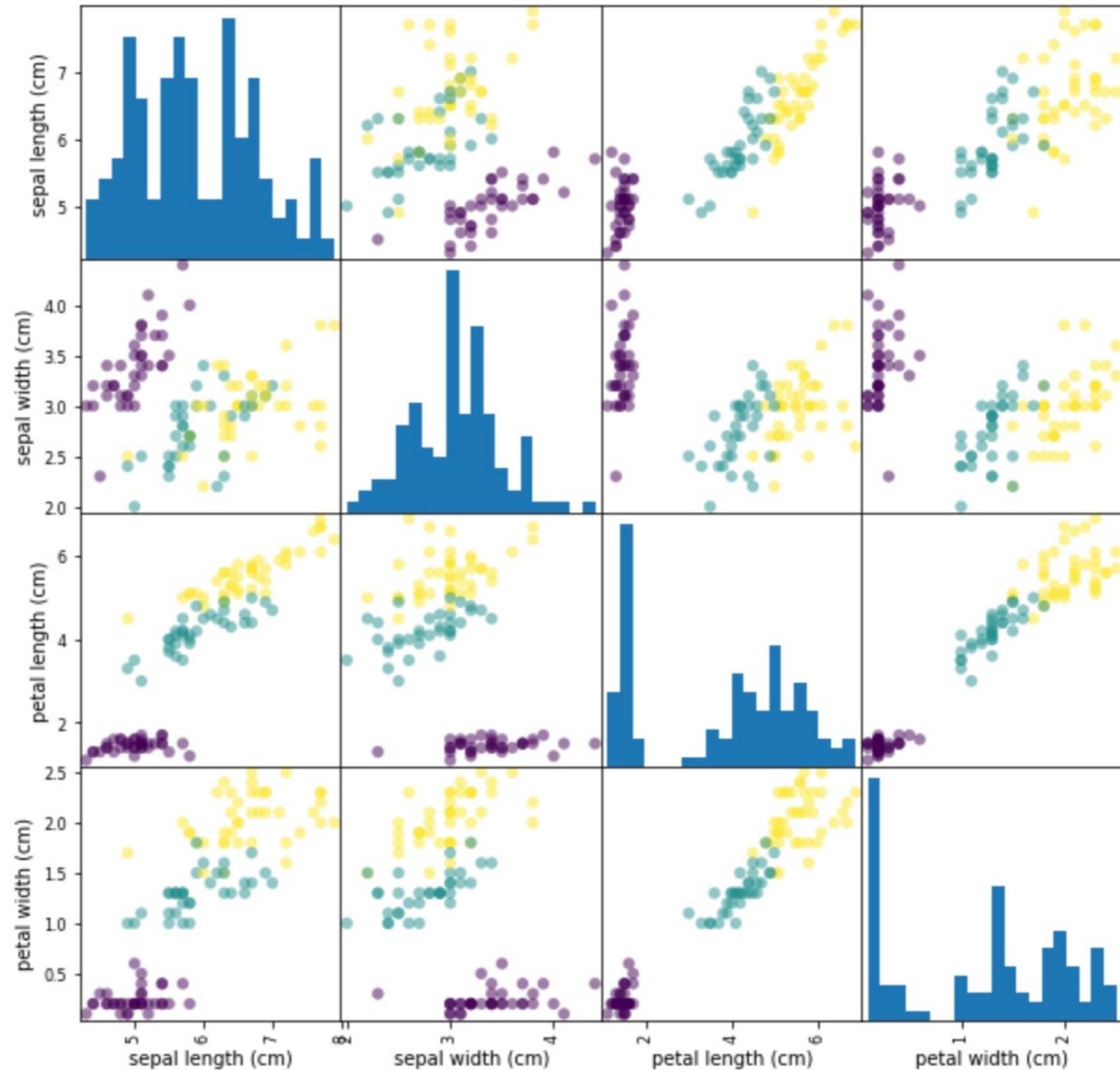
[Fisher '36]

# The Machine Learning Process?

## Data

sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
5.9	3.0	4.2	1.5
5.8	2.6	4.0	1.2
6.8	3.0	5.5	2.1
4.7	3.2	1.3	0.2
6.9	3.1	5.1	2.3
5.0	3.5	1.6	0.6
5.4	3.7	1.5	0.2
5.0	2.0	3.5	1.0
6.5	3.0	5.5	1.8
6.7	3.3	5.7	2.5

Understand correlation



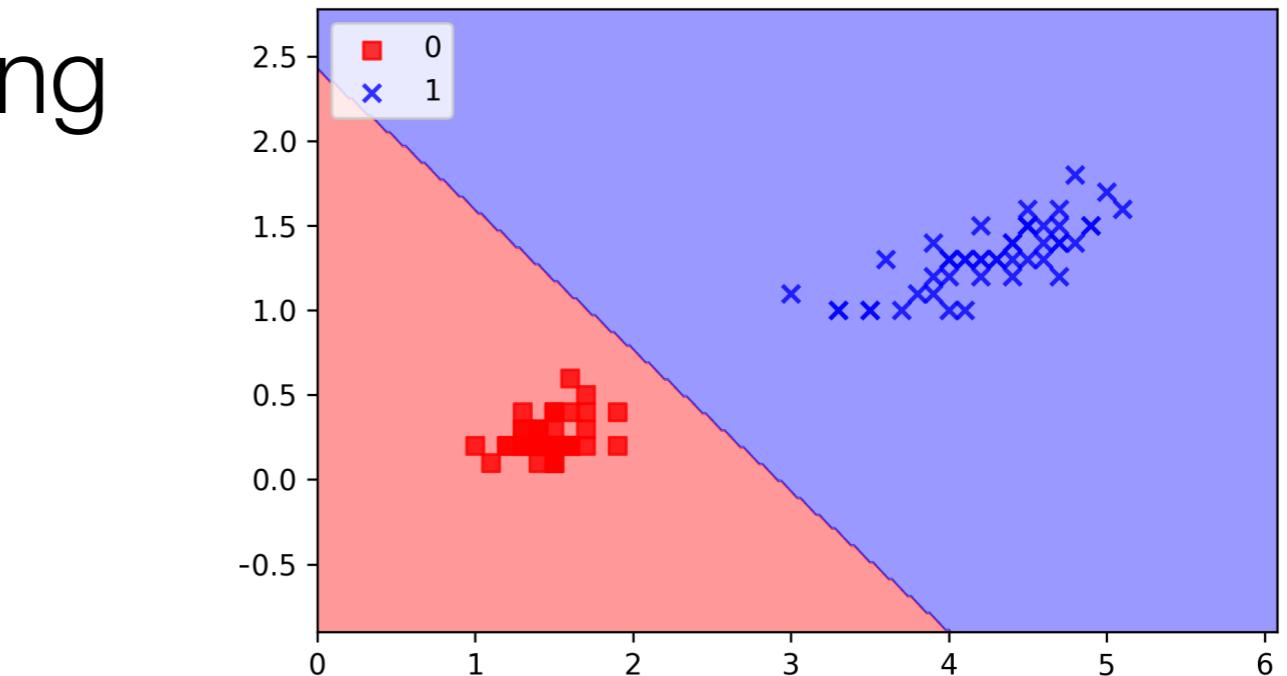
# Supervised Learning

## Linear Models

### Linear / logistic regression

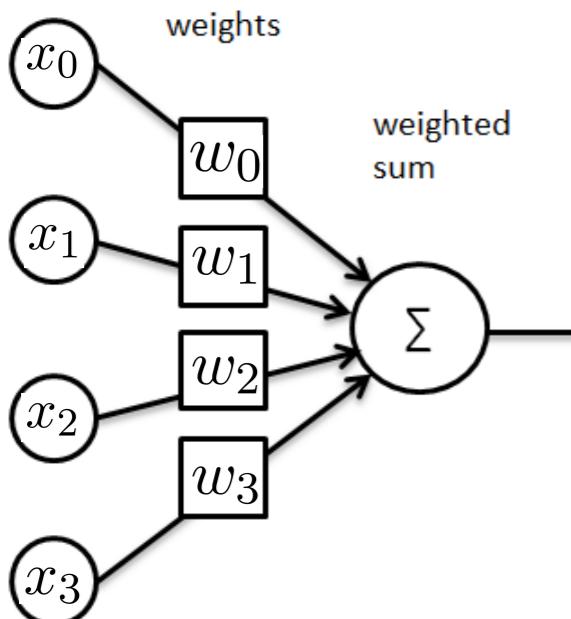
#### 1. Data

sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
5.9	3.0	4.2	1.5
5.8	2.6	4.0	1.2
6.8	3.0	5.5	2.1
4.7	3.2	1.3	0.2



→  $x_{ij} = \begin{pmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \\ x_{20} & x_{21} \\ x_{30} & x_{31} \end{pmatrix}$  →  $y_i = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix}$

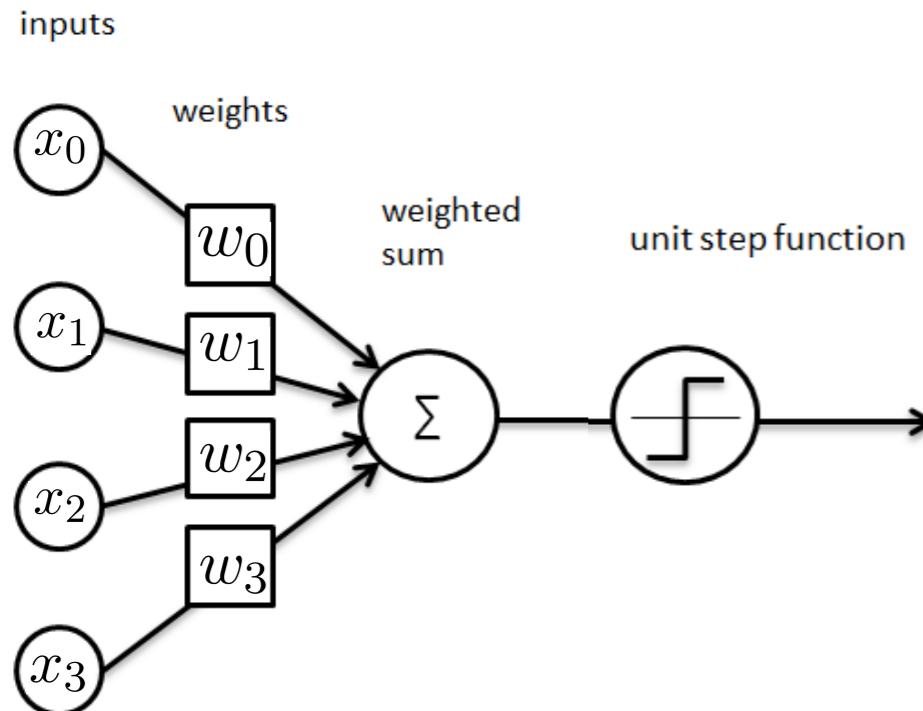
inputs



#### 2. Initialise weights and bias

$$w_j = (w_0, w_1) \quad b$$

# Supervised Learning

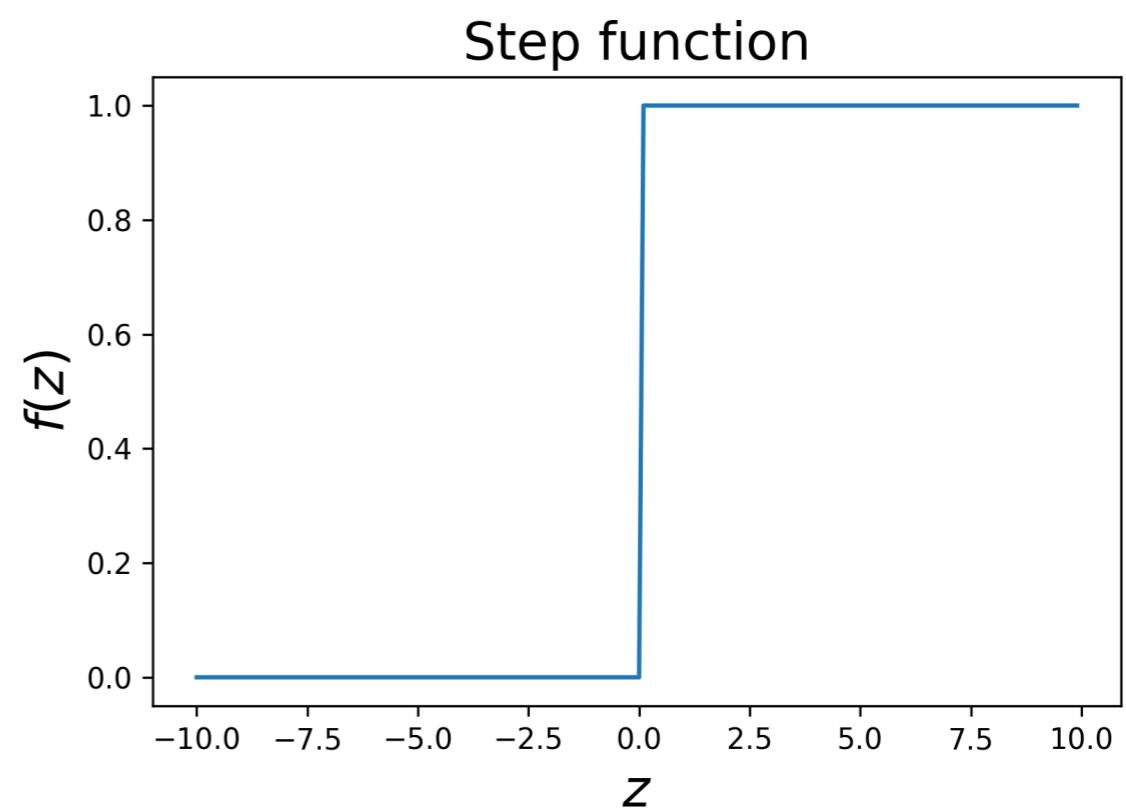


## 3. Calculate the **input**

$$\begin{aligned}z^{(i)} &= w \cdot x^{(i)} + b \\&= w_0 x_{i0} + w_1 x_{i1} + b\end{aligned}$$

## 4. Pass the input through an **activation function**

$$f(z^{(i)}) = \begin{cases} 0, & \text{if } z^{(i)} < 0 \\ 1, & \text{otherwise} \end{cases}$$



# Supervised Learning

## Training

1. Best fit the line / Minimise the cost function

$$J(z^{(i)}) = \frac{1}{2} |f(z^{(i)}) - y^{(i)}|^2$$

Prediction

Target

2. Update the weights

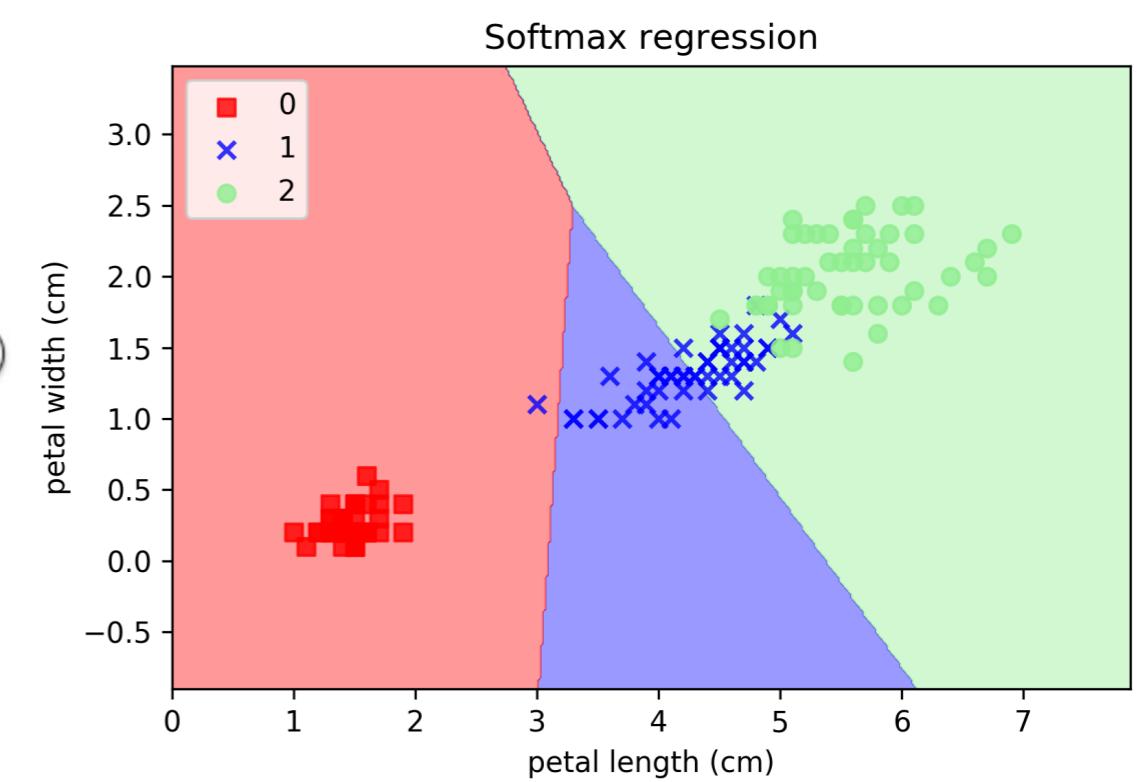
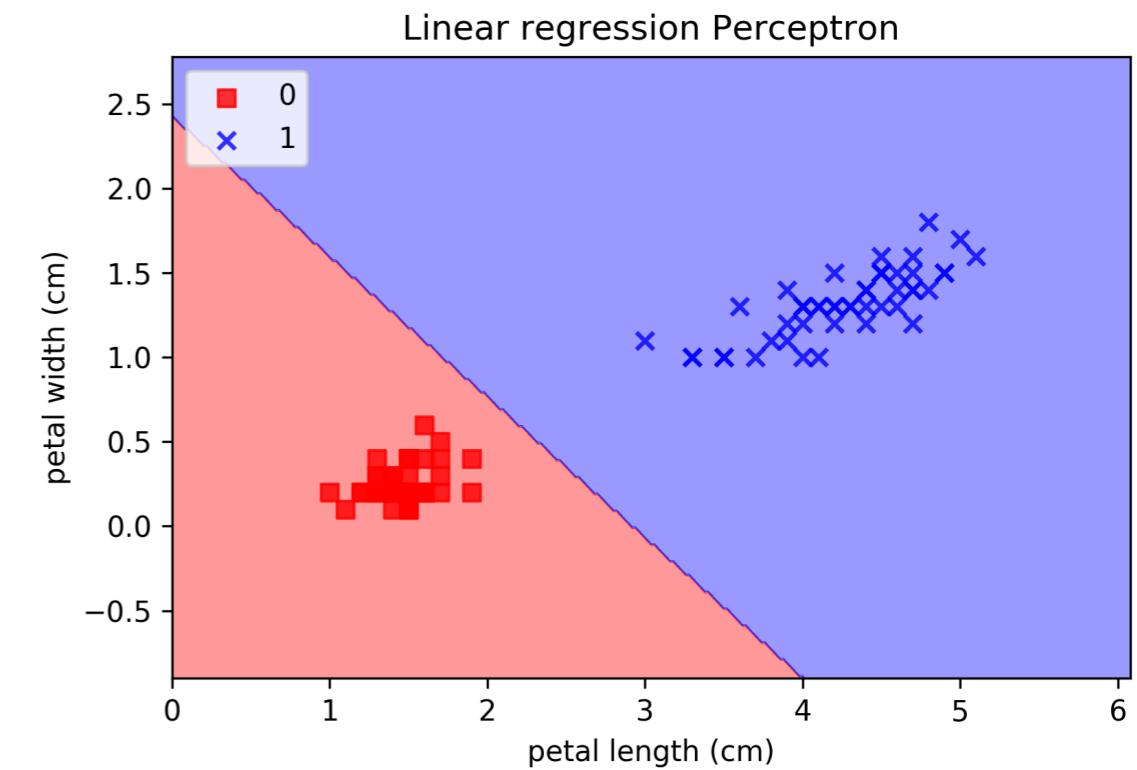
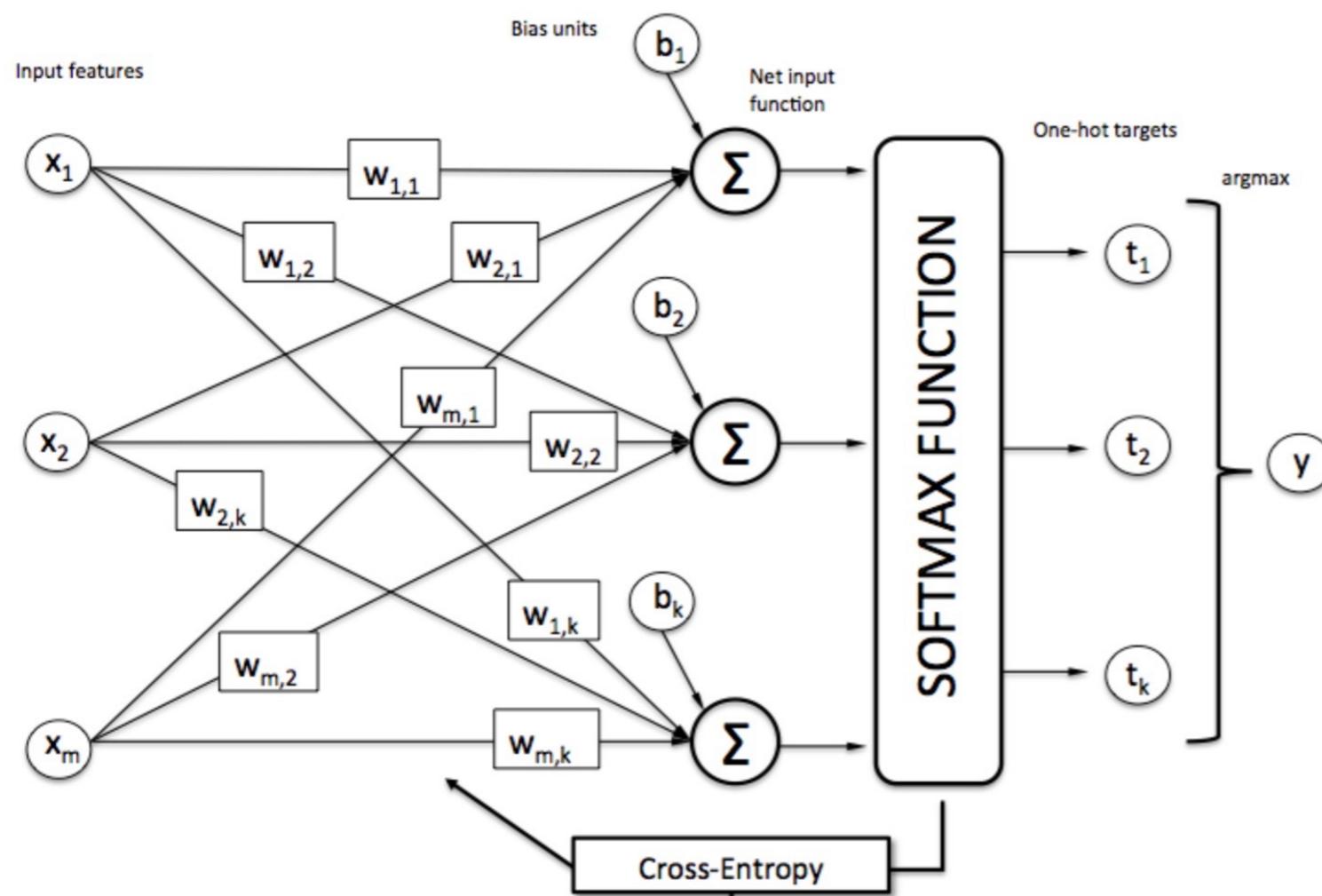
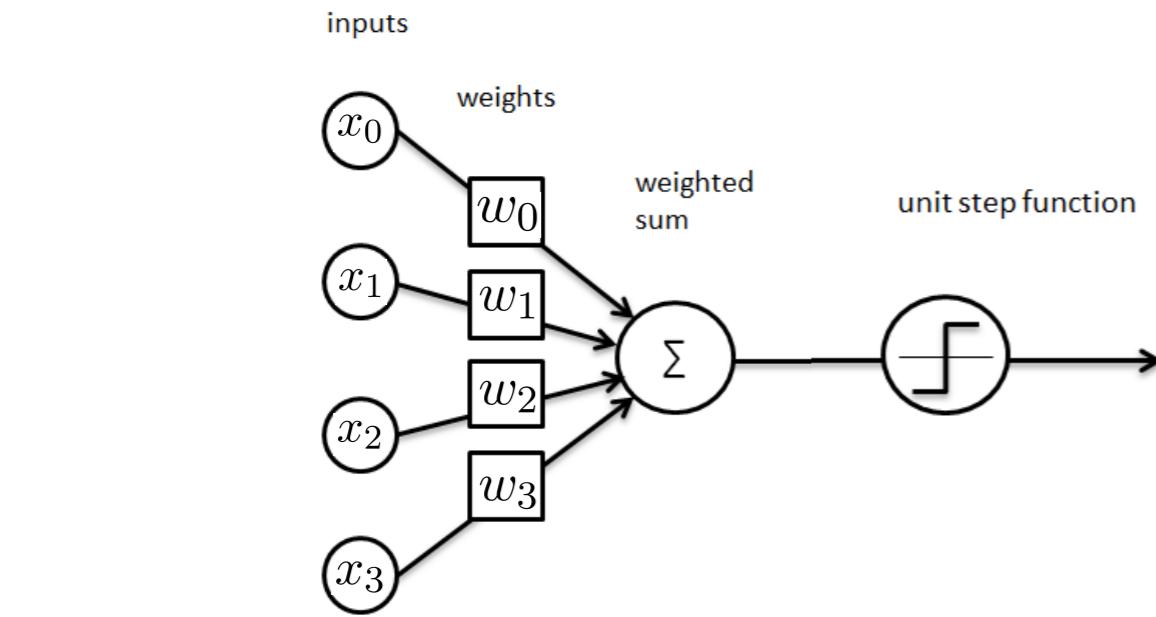
$$\vec{w} = \vec{w} - \eta \nabla_{\vec{w}} J(\vec{w}, \vec{b})$$

3. Repeat over all data points

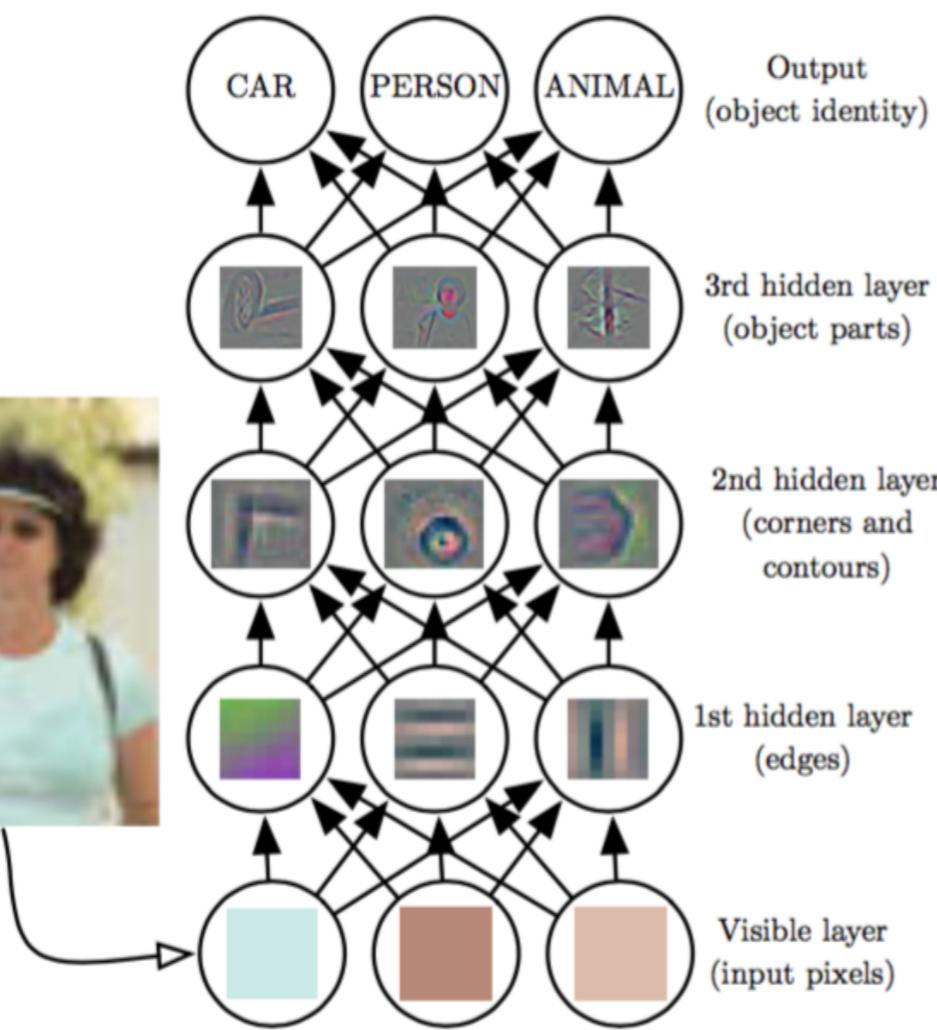
4. Repeat over a given number of epochs

[Bottom left image: Raschka '16]

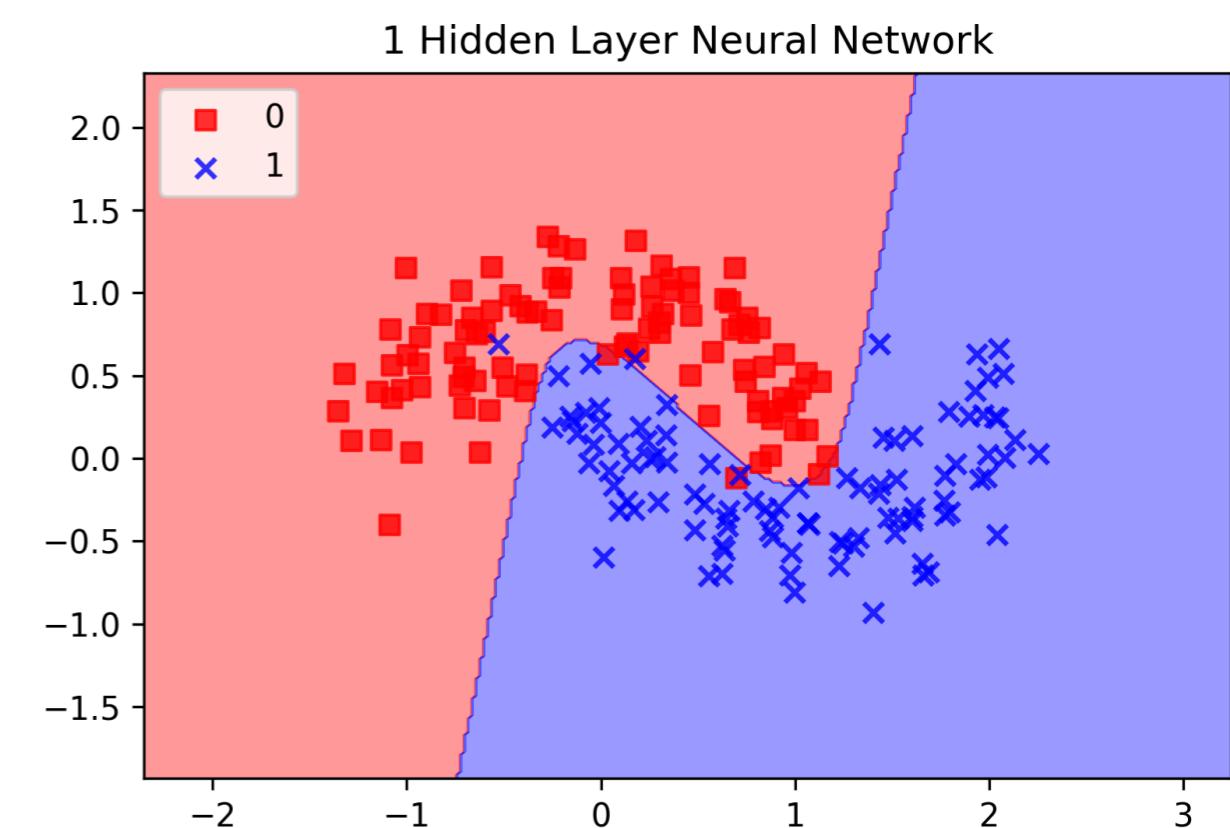
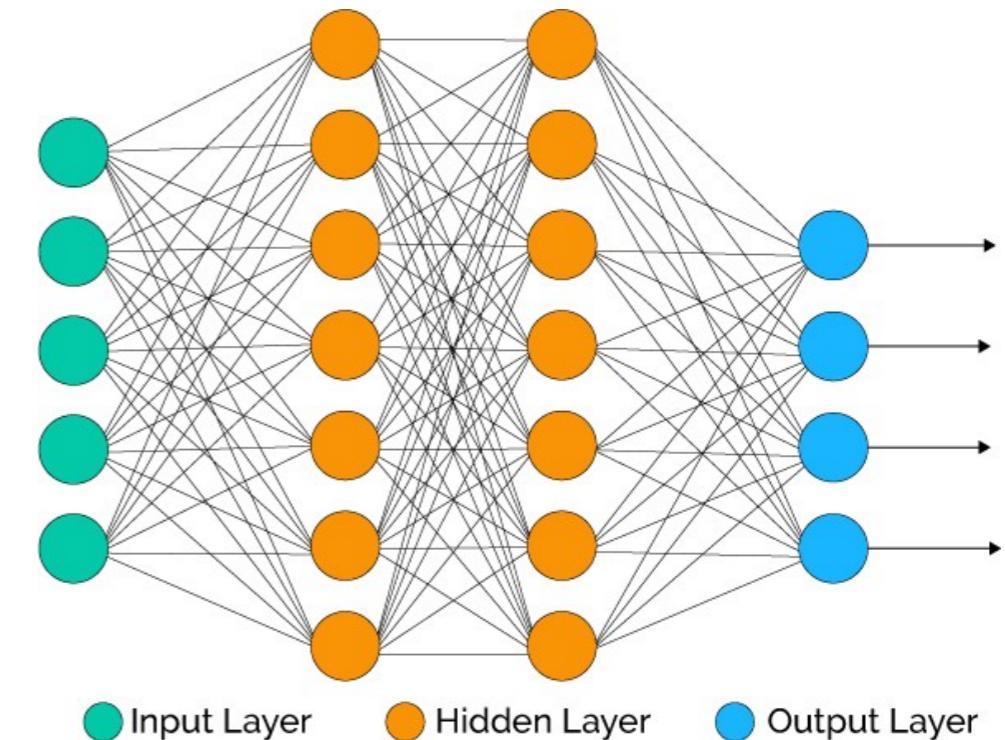
# Supervised Learning



## Neural Networks



# Deep Learning





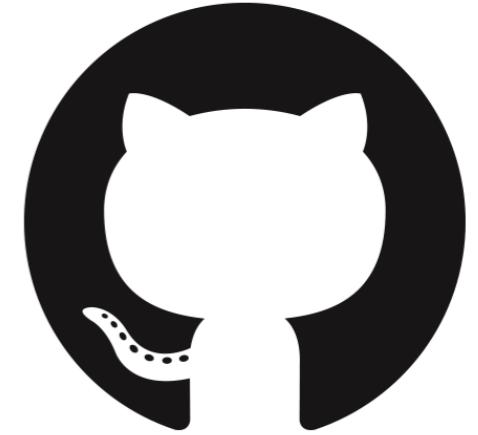
DeepMind



# Examples and code

## Supervised Learning:

- [My GitHub repository](#)



## Neural Networks:

- [TensorFlow Playground](#)
- [ConvNetJS](#)

## Other code sources:

Overview: “[Getting started with machine learning](#)”

[Python Machine Learning book repository](#)



# Summary

- Machine Learning is a classification problem
- The problem revolves around determining how important different input parameters are in order to get an accurate classification
- It is not a ‘black box’ where we cannot get information about what’s going on inside the algorithms
- Machine Learning is not the answer to all problems, the techniques should be applied intelligently and appropriately, and there may well be other more classical statistical techniques that are more appropriate

