Introduction to Machine Learning

Zifan Jiang 07/12/2022

University of Neuchâtel

About Me

Universität Zürich[UZH]

Institut für Computerlinguistik

Institut • **Staff** • Studium • Forschung • Stichwortverzeichnis • Unsere Zentren

- **Research Team and Doctoral Candidates**
  - **Text Technologies**
    - **Zifan Jiang**
    - Phonetics & Speech Sciences
    - Digital Linguistics
  - Former Staff Members and Alumni
  - Student Assistants
  - Institutsleitung

Zifan Jiang
PhD Student
Tel.: +41763376465
Raumbezeichnung: AND-2-18
→ jiang@cl.uzh.ch
→ Webseite

Hi, my name is Zifan [tsž: fǽn] Jiang [tɕiaŋ]

I am a PhD student interested in machine translation and natural language processing, supervised by → Prof. Dr. Rico Sennrich and → Prof. Dr. Martin Volk. I am currently working on automatic sign language translation as part of the → Flagship IICT Project led by → Dr. Sarah Ebling.

I did my master thesis project with → Dr. Mathias Müller and ↗ Amit Moryossef on ↓ Machine Translation between Spoken Languages and Signed Languages in Written Form (PDF, 2 MB).

Before intrigued by languages and linguistics ...

I have a solid background in computer science and software engineering. For some years I was mainly engaged in Web development, which makes me experienced in topics including computer network, human-computer interaction, programming languages, Web accessibility, and rich-text editors.

Teaching

Spring 2021	Teaching Assistant for → Informatics II (data structures and algorithms), with → Qing Chen

https://www.cl.uzh.ch/de/people/team/compling/jiang.html

About Me

Universität Zürich[UZH]

Institut für Computerlinguistik

Institut • Staff • Studium • Forschung • Stichwortverzeichnis • Unsere Zentren

- Research Team and Doctoral Candidates
  - Text Technologies
    - Zifan Jiang
  - Phonetics & Speech Sciences
  - Digital Linguistics
- Former Staff Members and Alumni
- Student Assistants
- Institutsleitung

Zifan Jiang
PhD Student
Tel.: +41763376465
Raumbezeichnung: AND-2-18
→ jiang@cl.uzh.ch
→ Webseite

Hi, my name is Zifan [tsž: fǽn] Jiang [tɕäŋ]

I am a PhD student interested in machine translation and natural language processing, supervised by → Prof. Dr. Rico Sennrich and → Prof. Dr. Martin Volk. I am currently working on automatic sign language translation as part of the → Flagship IICT Project led by → Dr. Sarah Ebling.

I did my master thesis project with → Dr. Mathias Müller and ↗ Amit Moryossef on ↓ Machine Translation between Spoken Languages and Signed Languages in Written Form (PDF, 2 MB).

Before intrigued by languages and linguistics ...

I have a solid background in computer science and software engineering. For some years I was mainly engaged in Web development, which makes me experienced in topics including computer network, human-computer interaction, programming languages, Web accessibility, and rich-text editors.

Teaching

Spring 2021	Teaching Assistant for → Informatics II (data structures and algorithms), with → Qing Chen

https://www.cl.uzh.ch/de/people/team/compling/jiang.html

Université de Neuchâtel › Faculté des sciences › Institut de biologie › Laboratory of language evolution › Team › Zifan Jiang

unine Université de Neuchâtel

Faculty of Science
Institute of Biology
Laboratory of Language Evolution

Current site search
Rechercher

Directory

EN FR

- Home
- Team
  - Steven Moran
  - Tiena Danner
  - Marco Maiolini
  - Zifan Jiang
  - Carlos Silva
- Publications
- Open positions
- News

Zifan Jiang

Zifan Jiang
Scientific programmer

› Print

Top of the page

https://www.unine.ch/evolang/home/team/zifan-jiang.html

What We Do …

By leveraging data science and machine learning techniques

Machine translation

Image classification

Natural language processing

Machine learning

Human pose estimation

Computational biology

Evolution of phonetics and phonology (EVOPHON)

Audio synthesis / reconstruction

What is Machine Learning: One Example

Classify images of horse vs. zebra …

What is Machine Learning: One Example

Classify images of horse vs. zebra …

- Classical Approach: manual rules
  - IF there are black and white stripes THEN “zebra” ELSE “horse”

What is Machine Learning: One Example

Classify images of horse vs. zebra …

- Classical Approach: manual rules
  - IF there are black and white stripes THEN “zebra” ELSE “horse”
- Machine learning: **automatic** discovery of rules from training data (examples)

What is Machine Learning: One Example

Classify images of horse vs. zebra … how about this one?

https://www.smithsonianmag.com/smart-news/scientists-dressed-horses-zebras-determine-purpose-stripes-180971540/

What is Machine Learning: One Definition

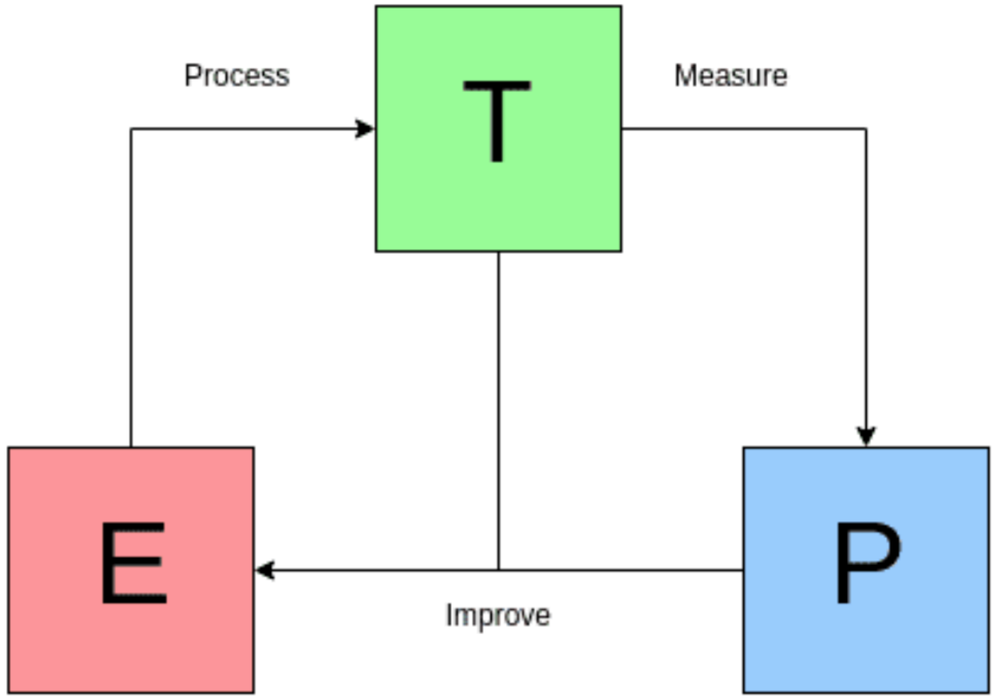
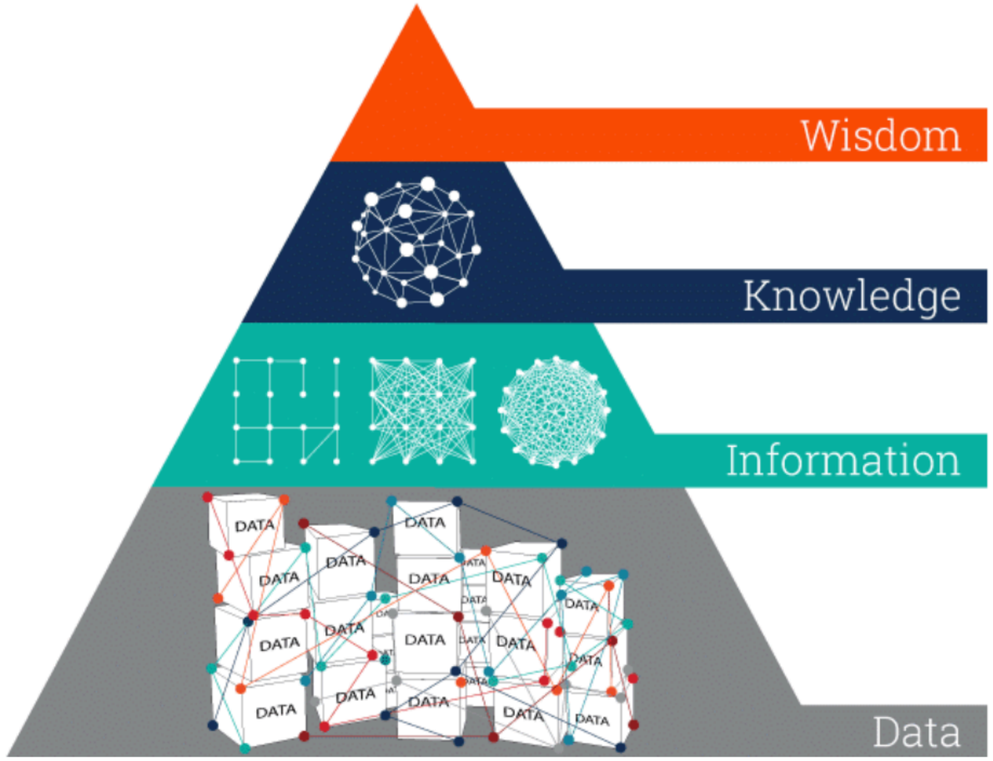


Figure 1. The Mitchell Paradigm, visualized.

What is Machine Learning: the Role

Machine Learning plays an important role in this pyramid



Data, Information, Knowledge, Wisdom (DIKW) Pyramid

What is Machine Learning: the Role

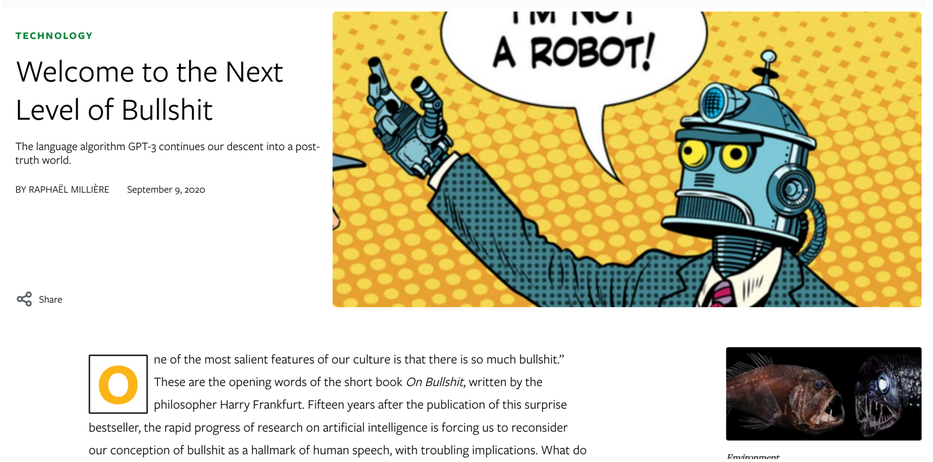
Machine Learning plays an important role in this pyramid

Data, Information, Knowledge, Wisdom (DIKW) Pyramid

What is Machine Learning: the Role

Machine Learning plays an important role in this pyramid

Data, Information, Knowledge, Wisdom (DIKW) Pyramid



Outline

- Supervised Learning
  - Regression
  - Classification
- Data Representation
  - Text
  - Audio
  - Image
- Unsupervised Learning
  - Clustering
  - Dimension Reduction
  - Autoencoders
- Deep Learning
  - Power of Nonlinearity
  - Neural Network
  - Universal Approximation Theorem

Supervised Learning

$$f : X \rightarrow Y$$

Supervised Learning - Regression

Goal: Predict real valued labels

Example: X = thigh circumference, Y = body fat percentage

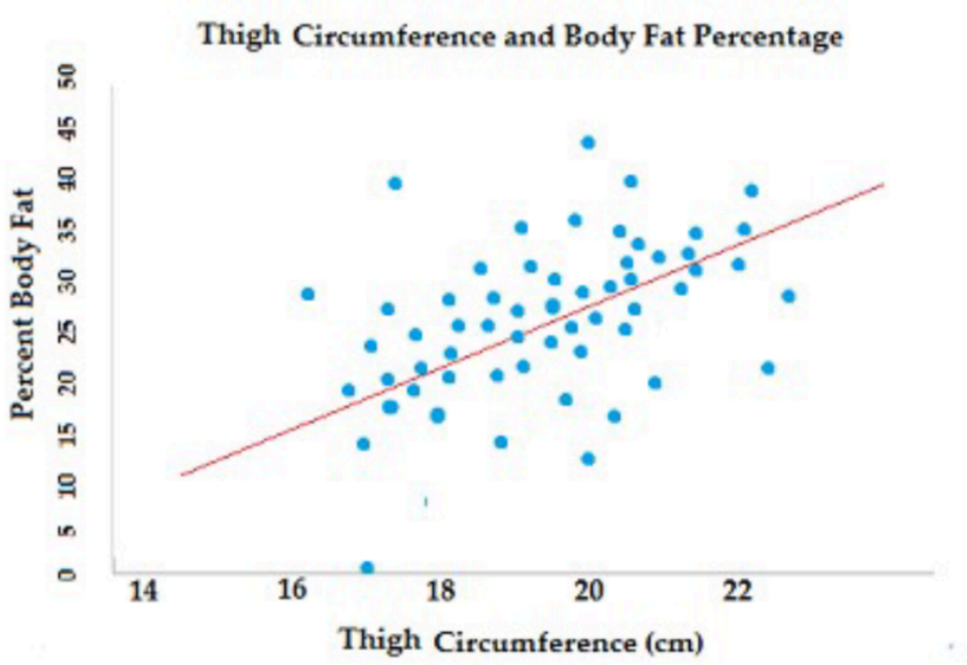


Figure from lecture Linear models I

Supervised Learning - Regression

Goal: Predict real valued labels

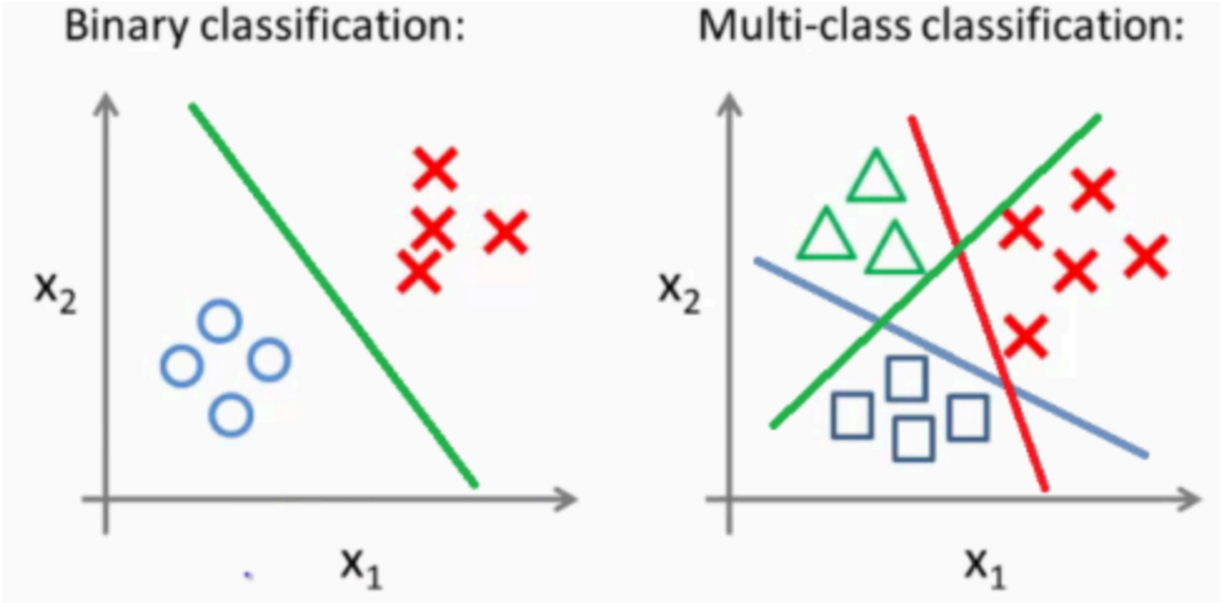
Summary

Model	Variables	Distribution	R code
Linear Regression	$Y = b_0 + b_1x$	Normal	*lm(formula, data)*
General Linear Models	$Y = b_0 + b_1x_1 + b_2x_2 + ...$	Normal	*lm(formula, data)*
Generalized Linear Models (GLM)	$Y = b_0 + b_1x_1 + b_2x_2 + ...$	Any	*glm(formula, family, data)*

Figure from lecture Linear models I

Supervised Learning - Classification

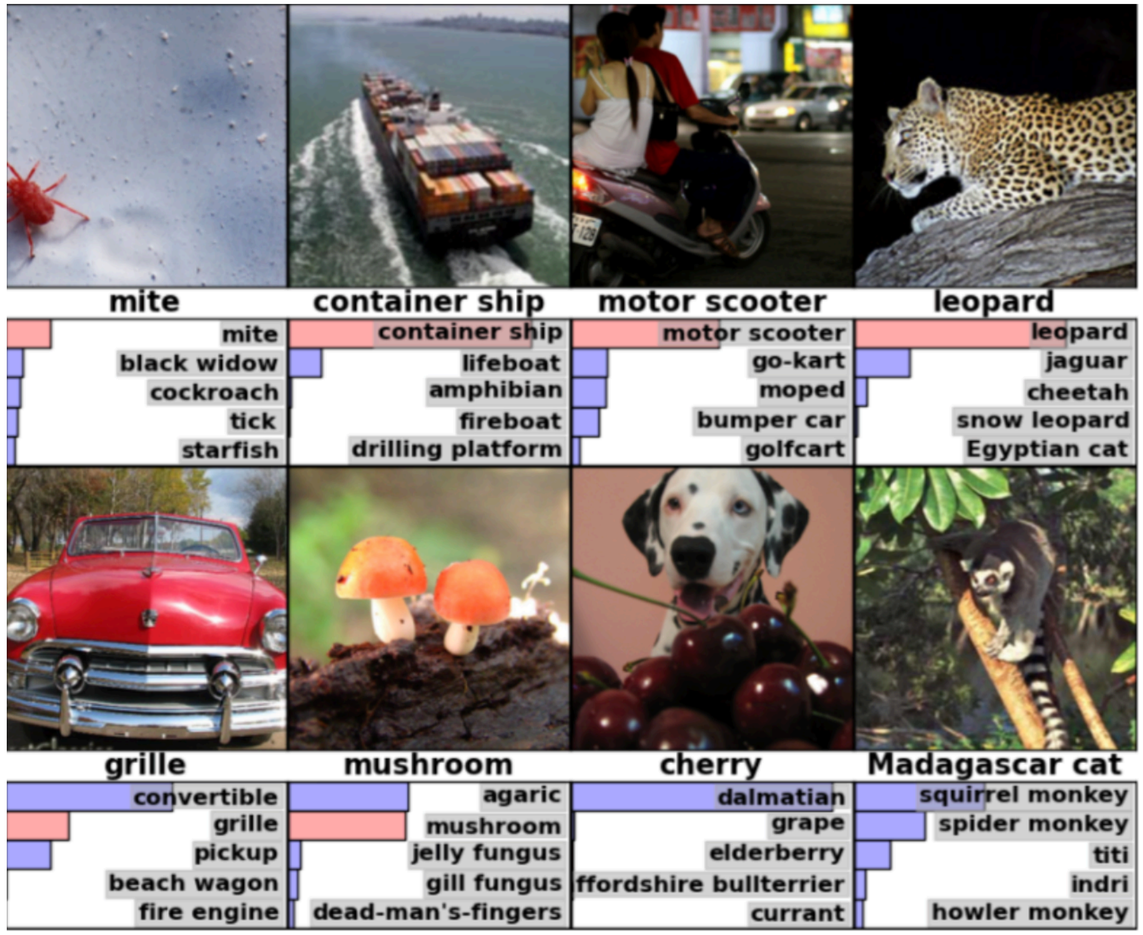
Goal: Predict categorical labels



https://medium.com/swlh/classification-79288caf6338

Supervised Learning - Classification

Example: image classification - ImageNet



https://blog.acolyer.org/2016/04/20/imagenet-classification-with-deep-convolutional-neural-networks/

Supervised Learning - Classification

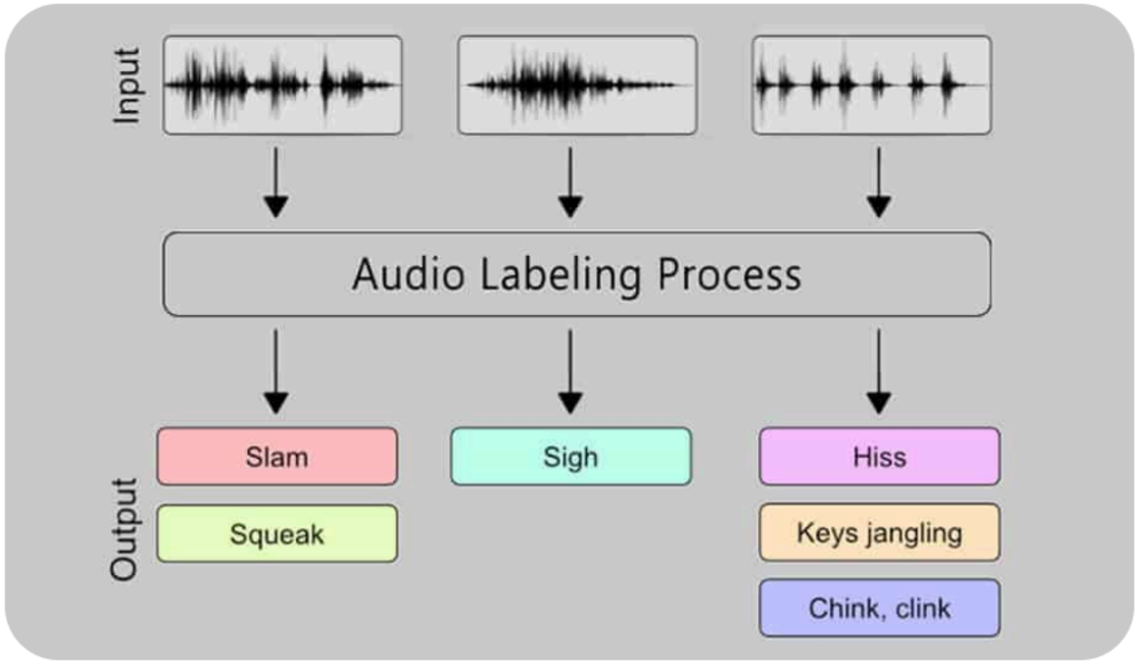
Example: image classification - Chihuahua (dog) or muffin?

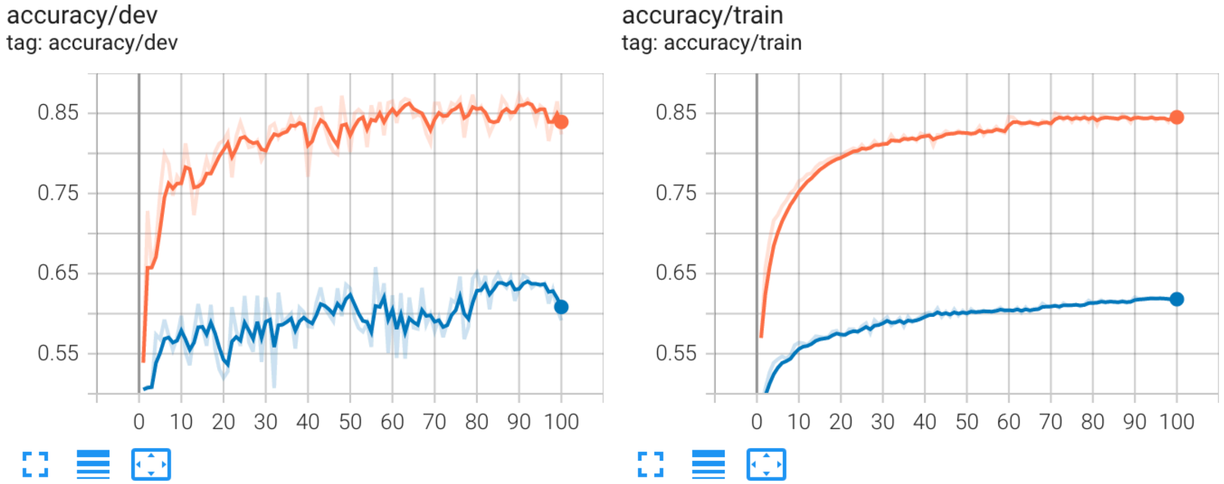
https://www.freecodecamp.org/news/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d/

Supervised Learning - Classification

Example: extract and classify monkey calls from raw audio

X = audio waveform, Y = call types (or non-call)



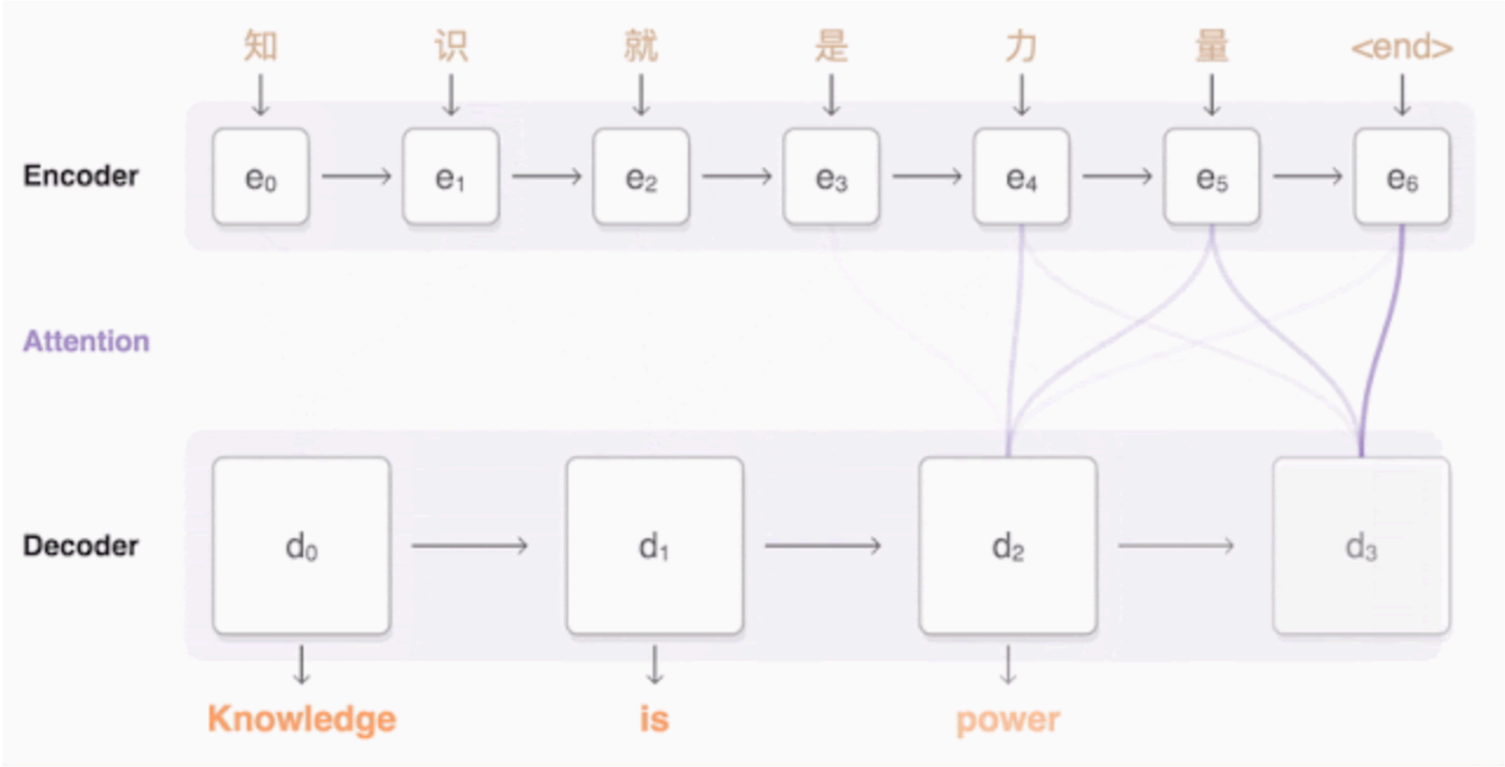


Statistics from our research

Supervised Learning - Classification

Example: neural machine translation

X = tokens in source language, Y = tokens in target language



https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html

Supervised Learning - Pipeline / Paradigm

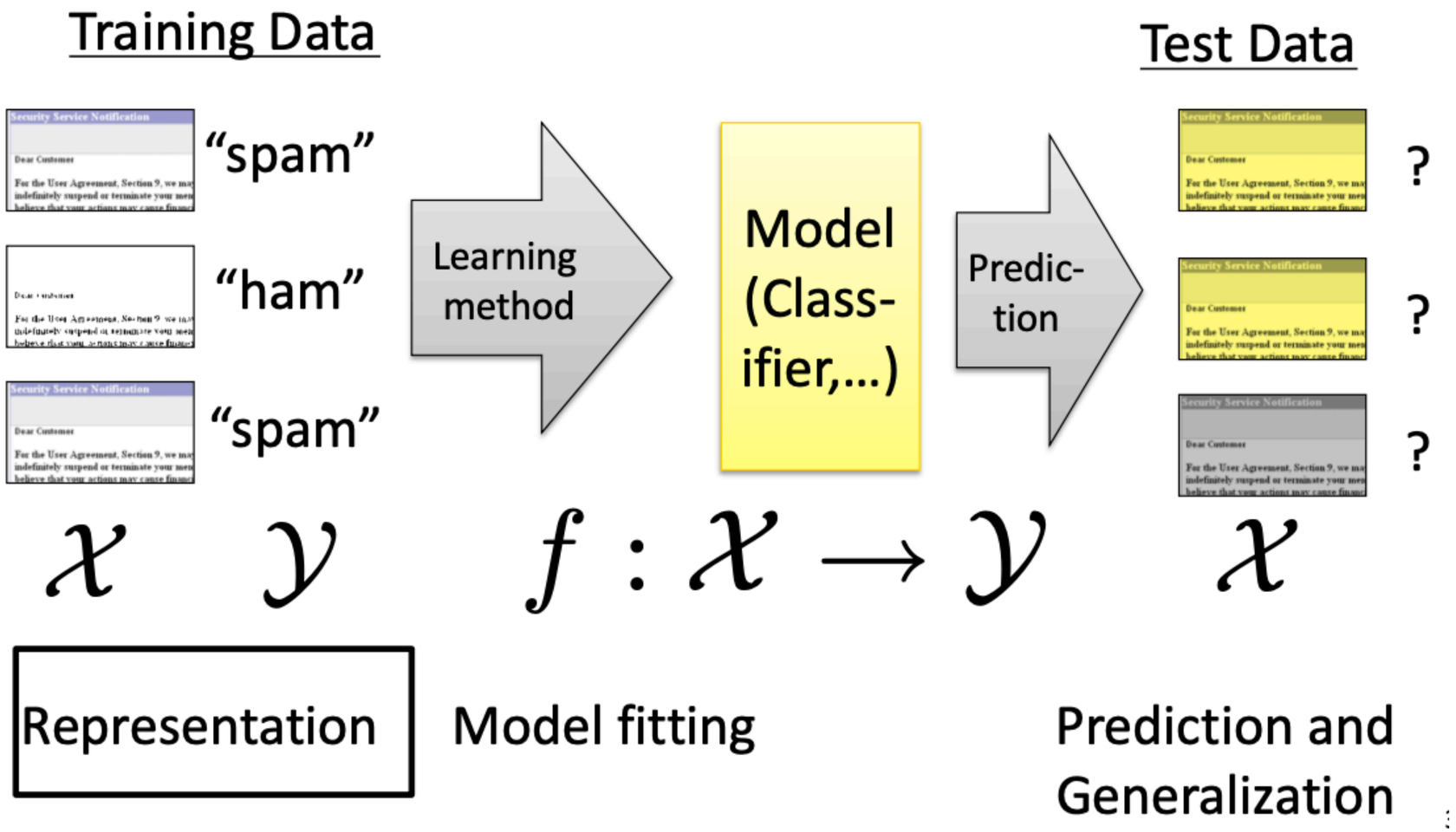


Figure from course https://las.inf.ethz.ch/teaching/introml-s20

Data Representation

How to represent data in a machine-understandable fashion?

The quick brown fox jumps over the lazy dog ...

Figure from course https://las

Data Representation

How to represent data in a machine-understandable fashion?

The quick brown fox jumps over the lazy dog ...

[0 1 0 0 0 3 2 0 1 0 0 0]

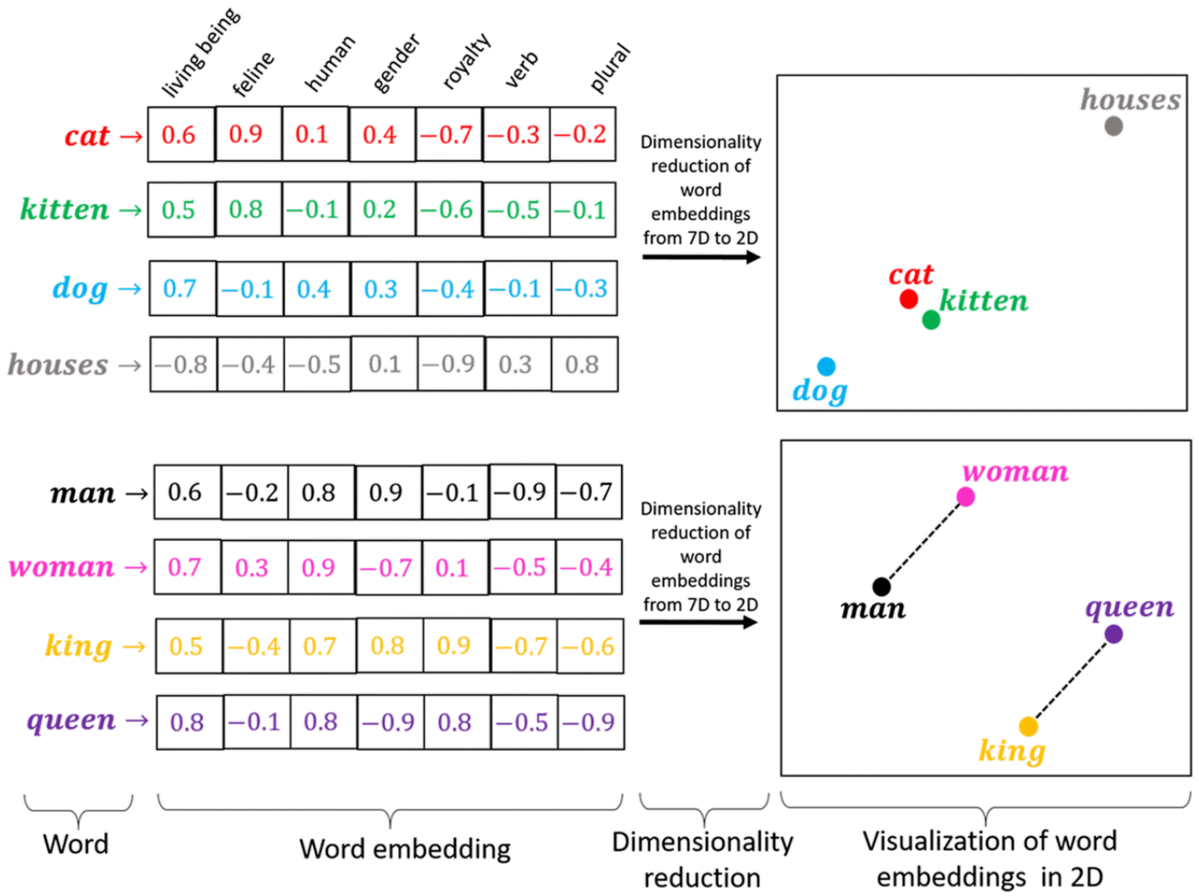
[.3 .01 .1 2.3 0 0 1.1 ...]

Vectors!

Figure from course https://las.inf.ethz.ch/teaching/introml-s20

Data Representation - Text

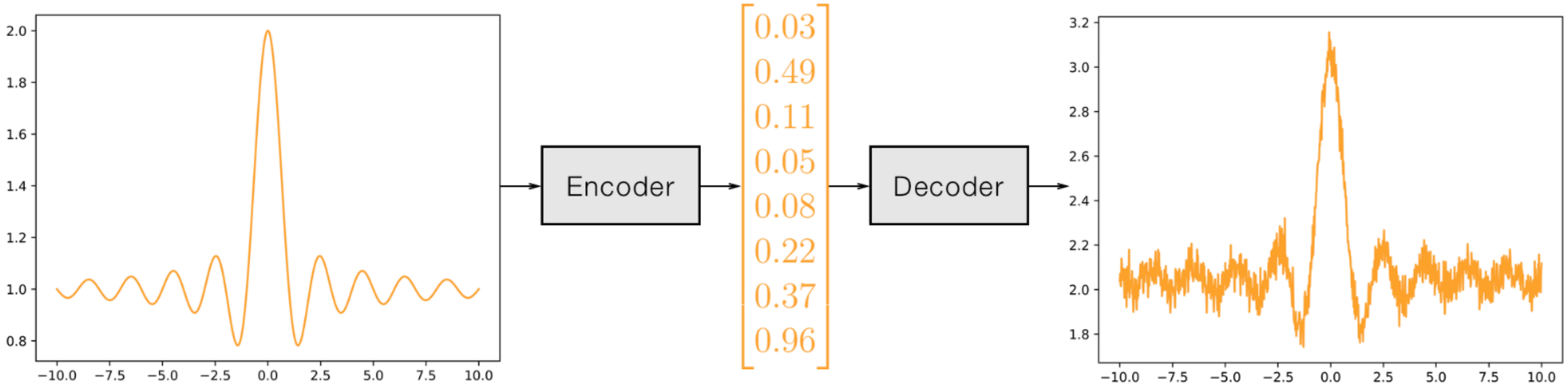
How to represent data in a machine-understandable fashion?



https://medium.com/@hari4om/word-embedding-d816f643140

Data Representation - Audio

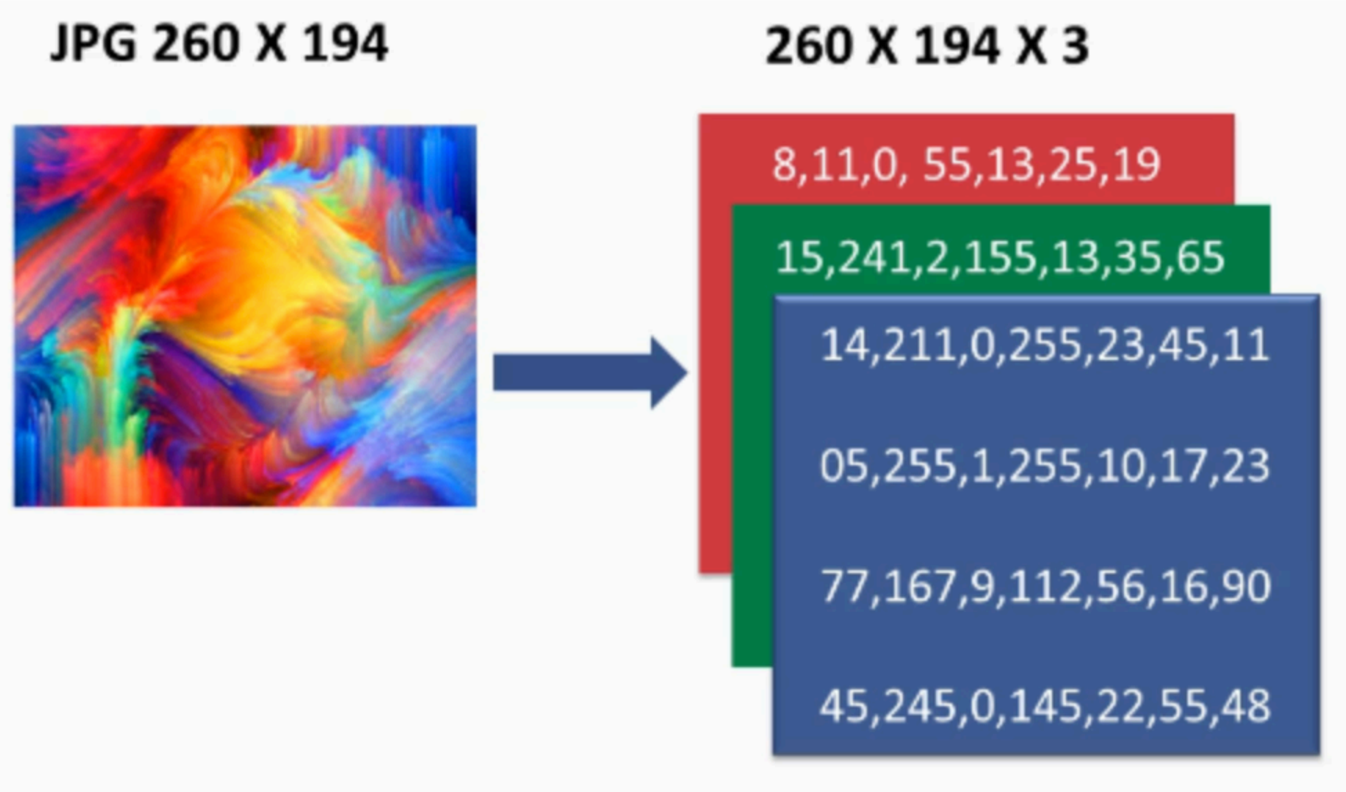
How to represent data in a machine-understandable fashion?



https://github.com/shobrook/sequitur

Data Representation - Image

How to represent data in a machine-understandable fashion?



Unsupervised Learning - Pipeline / Paradigm

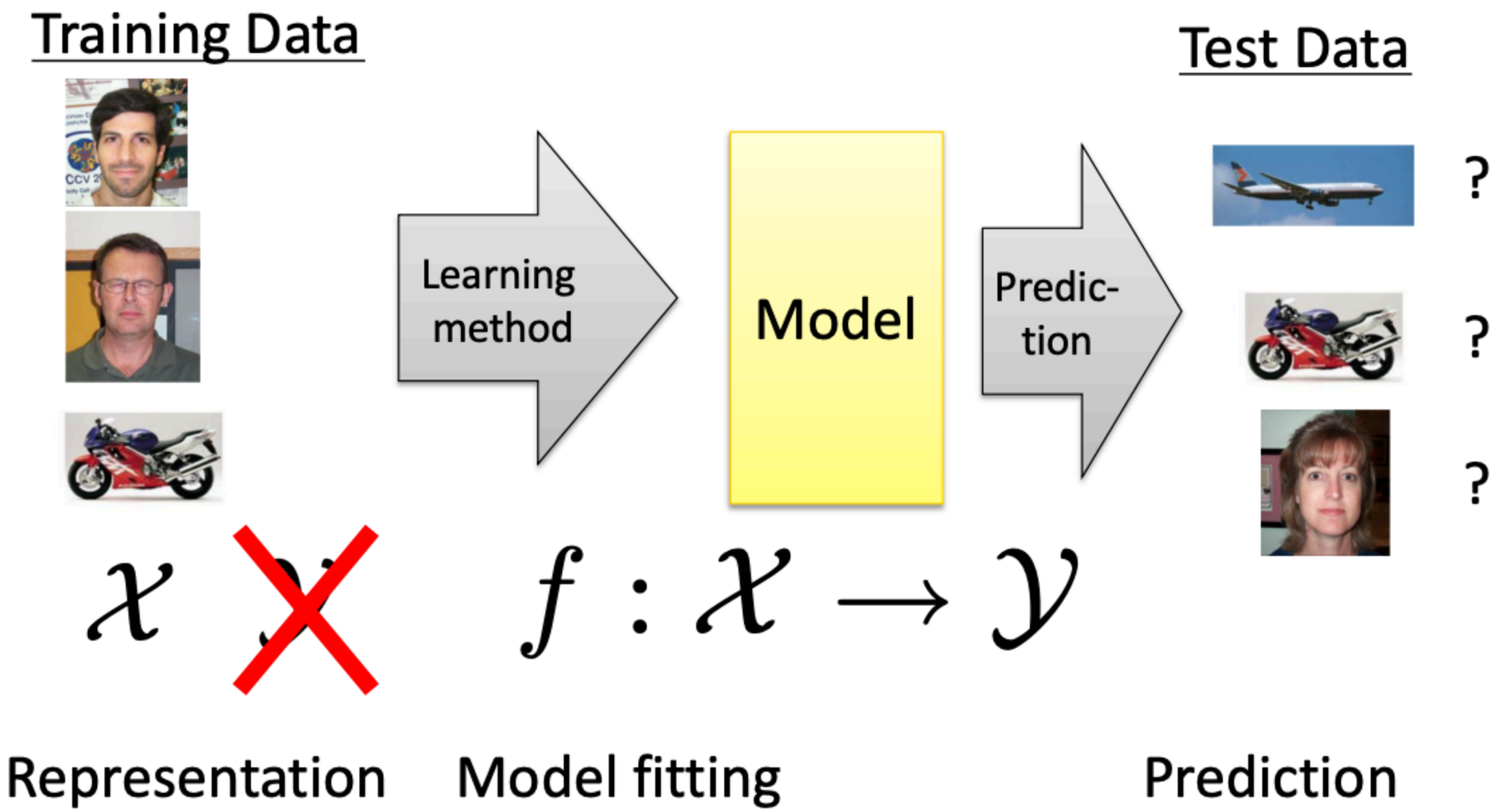


Figure from course https://las.inf.ethz.ch/teaching/introml-s20

Unsupervised Learning - Clustering

Unsupervised classification

Example: K-means clustering

Figure from lecture Dimensionality reduction and clustering II

Unsupervised Learning - Dimension Reduction

Unsupervised regression

Example: principal component analysis (PCA)

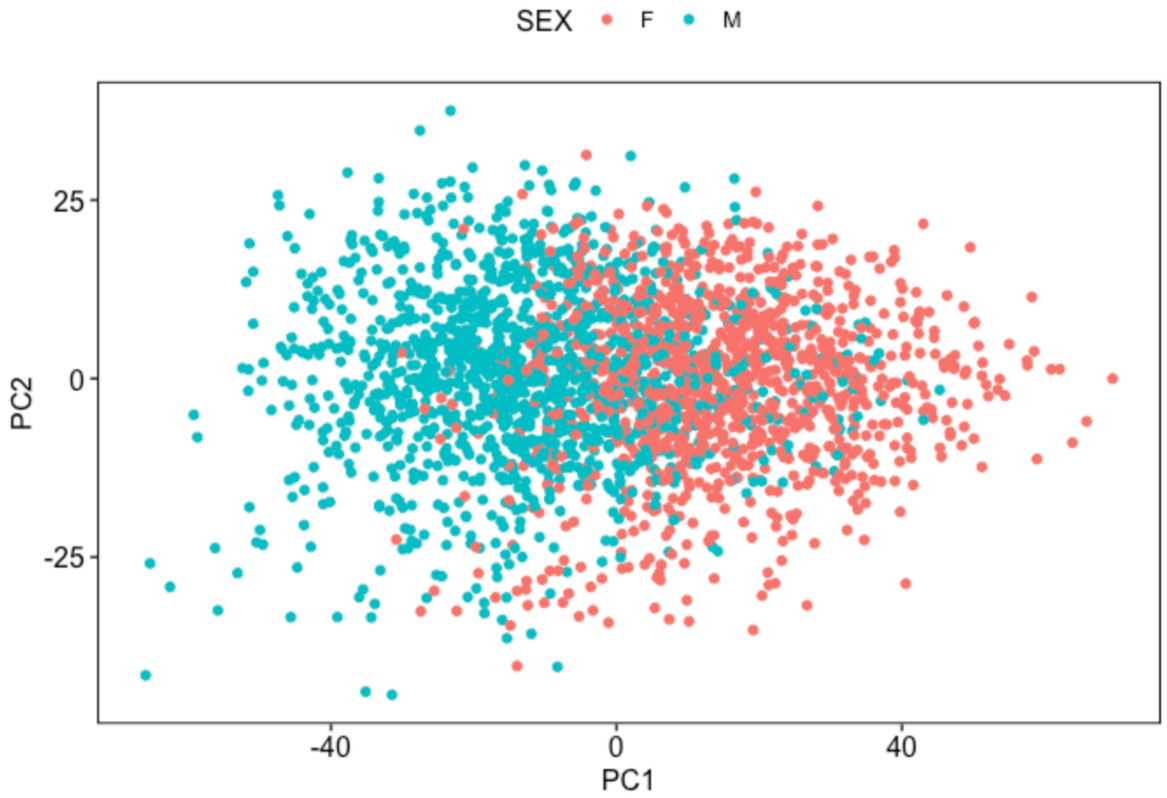
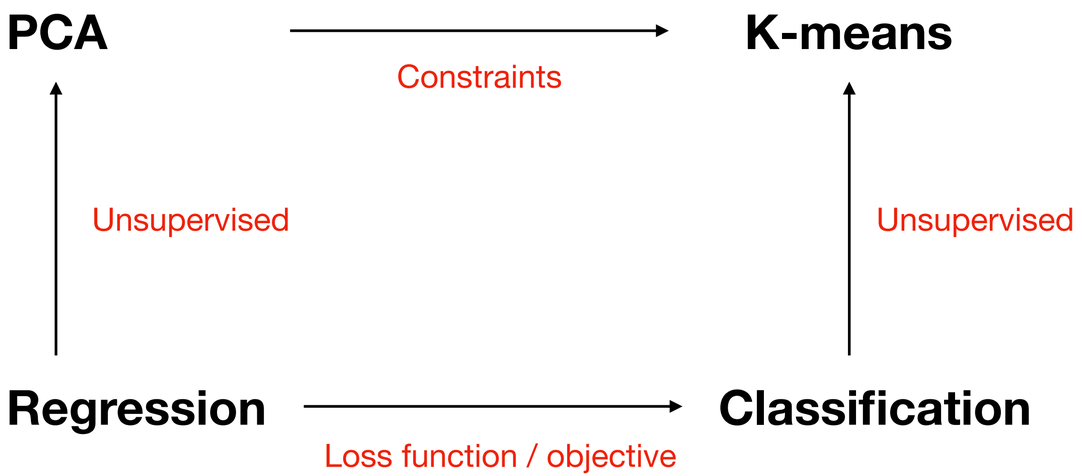


Figure from lecture Dimensionality reduction and clustering I

Unsupervised Learning - Comparison

Look at everything so far in a big piecture …



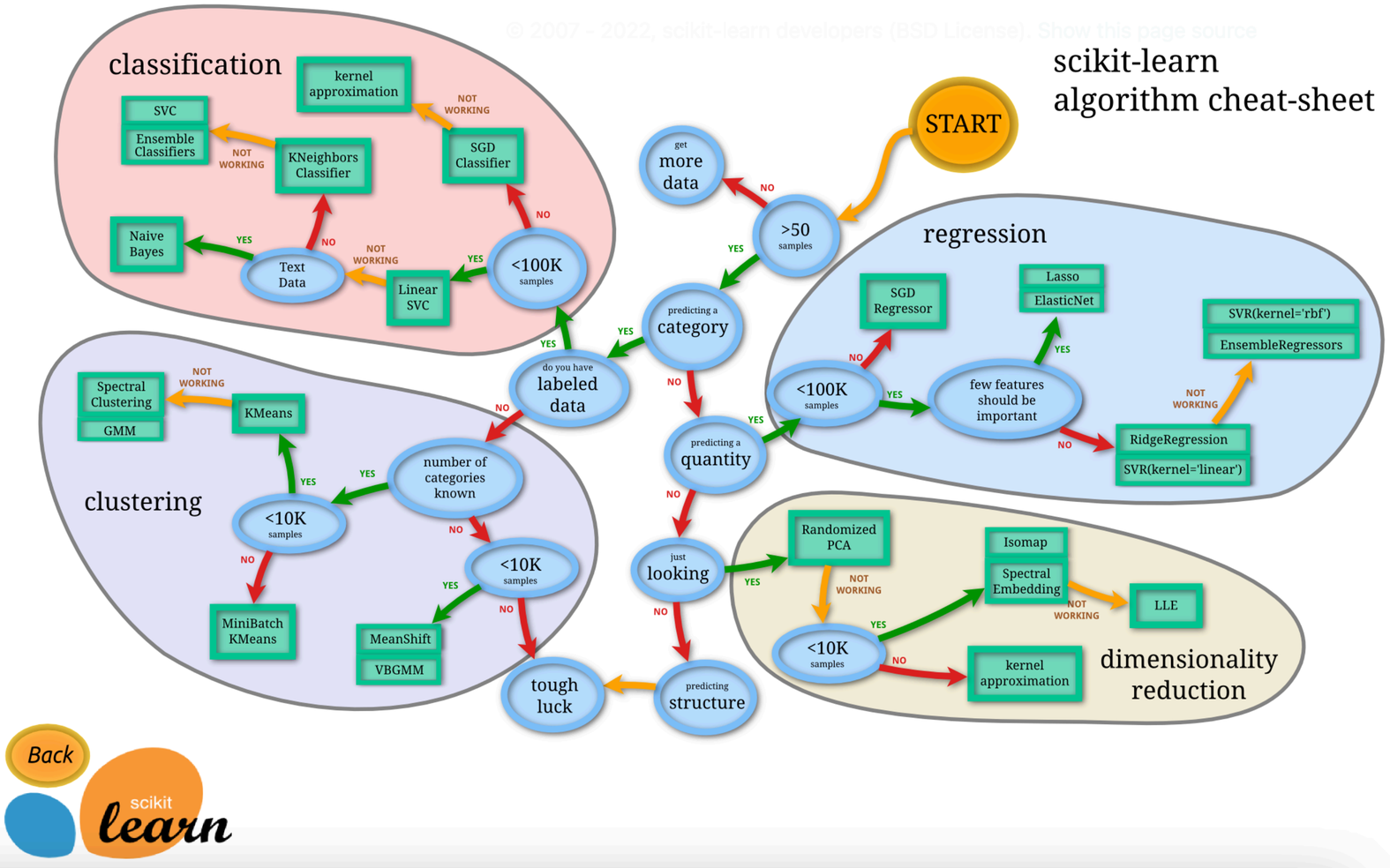
Unsupervised Learning - Comparison

Look at everything so far in a big piecture …

PCA → Constraints

↑ Unsupervised

Regression → Loss function / objective



https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Unsupervised Learning - Autoencoders

First reduce the dimension, then reconstruct to the original …

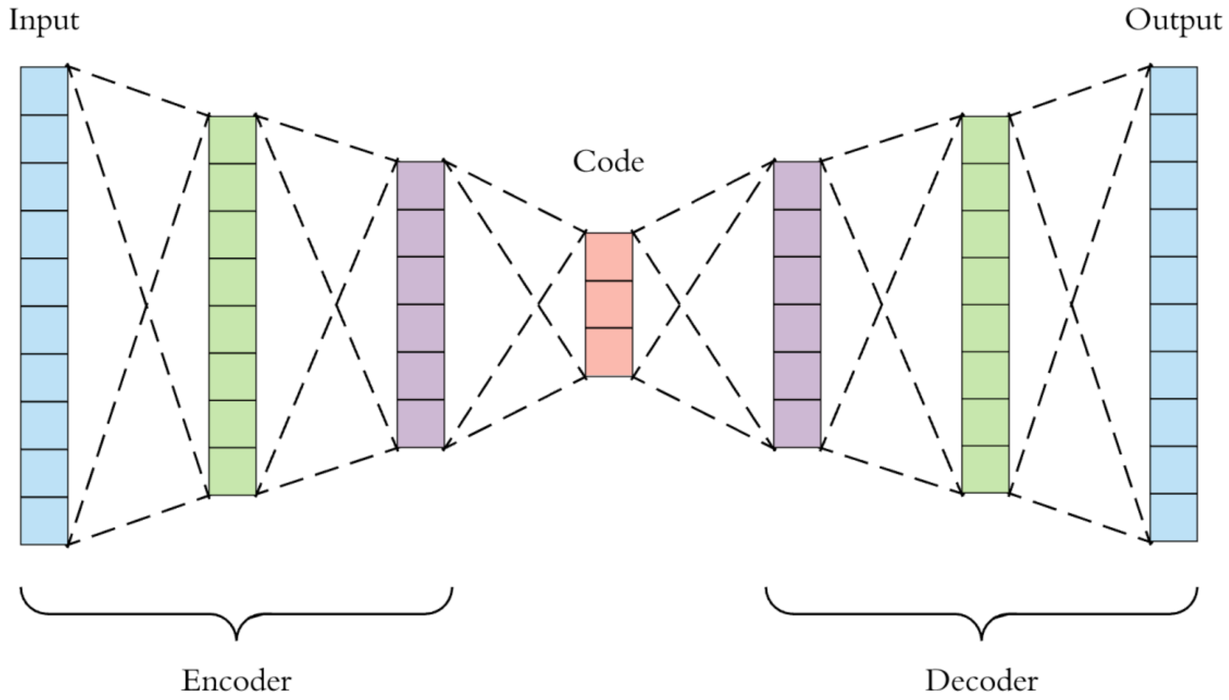
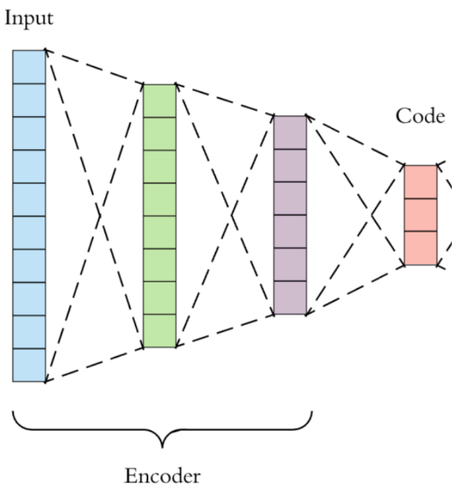


Figure from course http://www.da.inf.ethz.ch/teaching/2022/CIL/

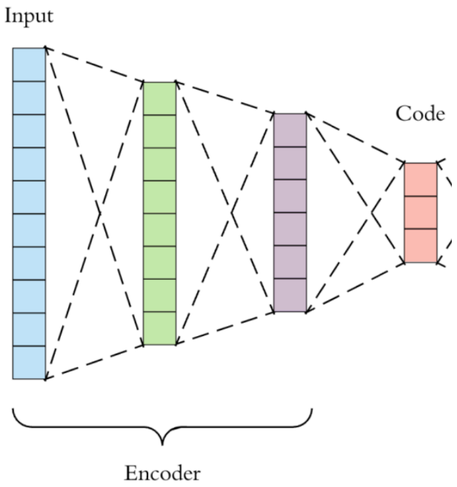
Unsupervised Learning - Autoencoders

Example: reconstruct word sound from thousands of years ago …

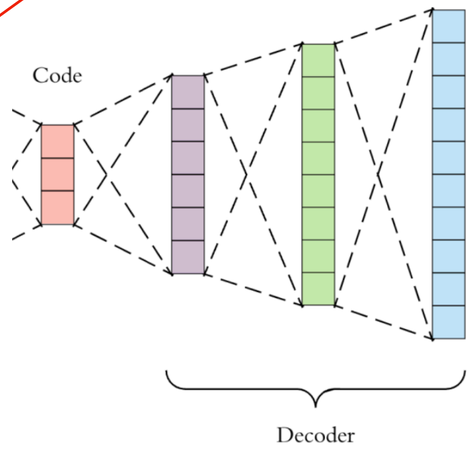
French three “trois”



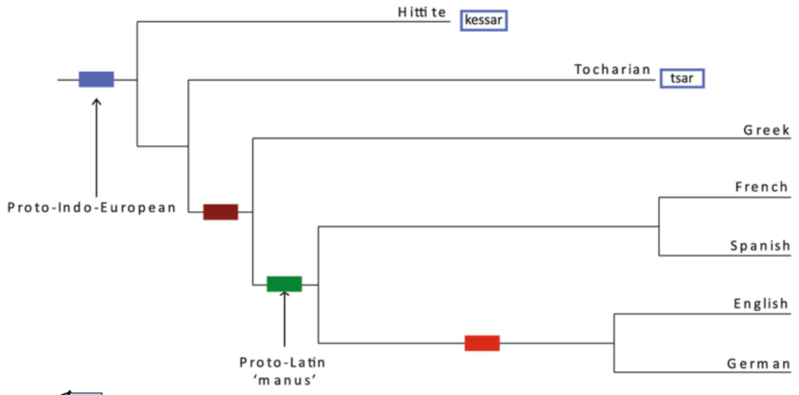
Italian three “tre”



Combine!

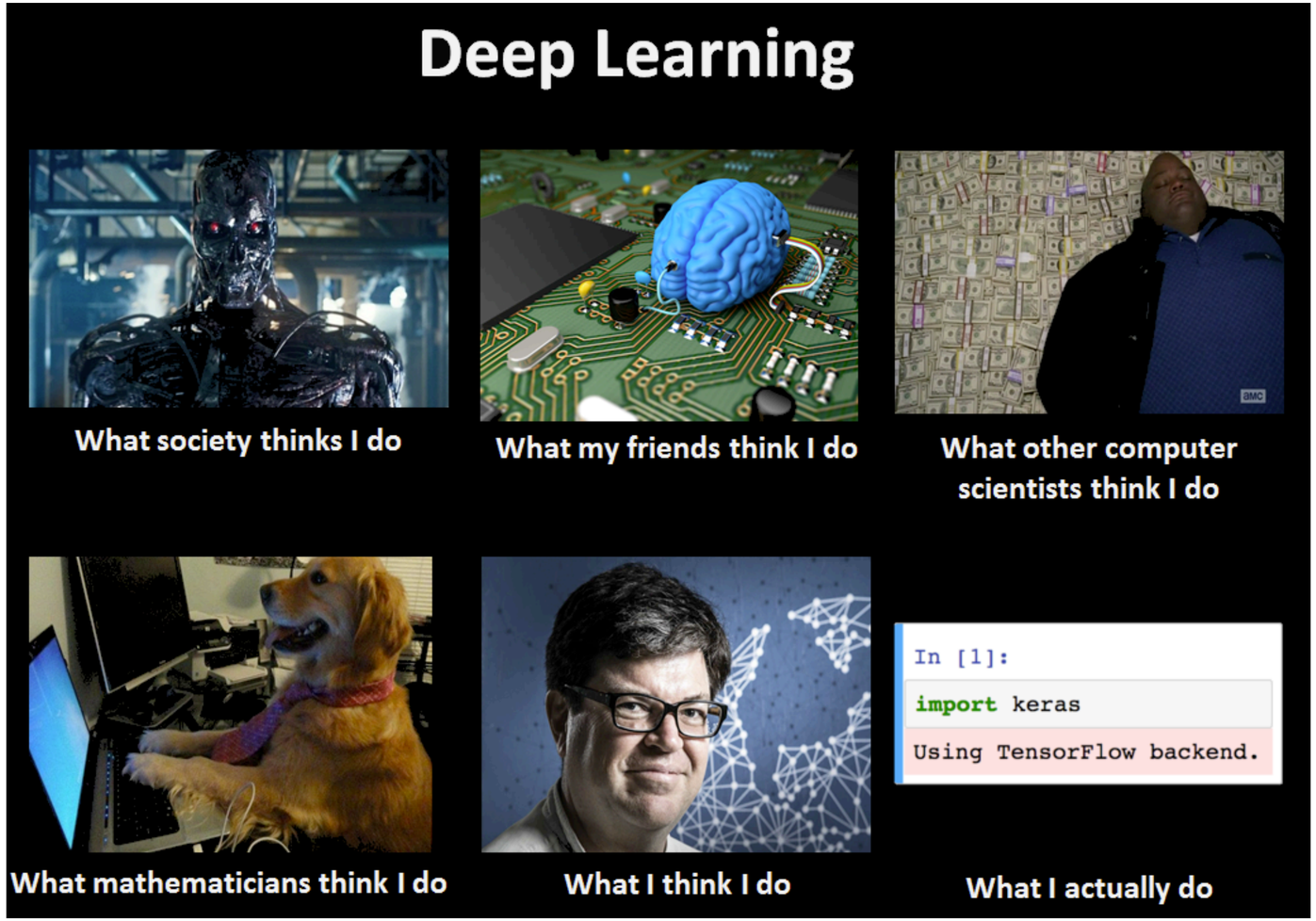


Latin three “???”

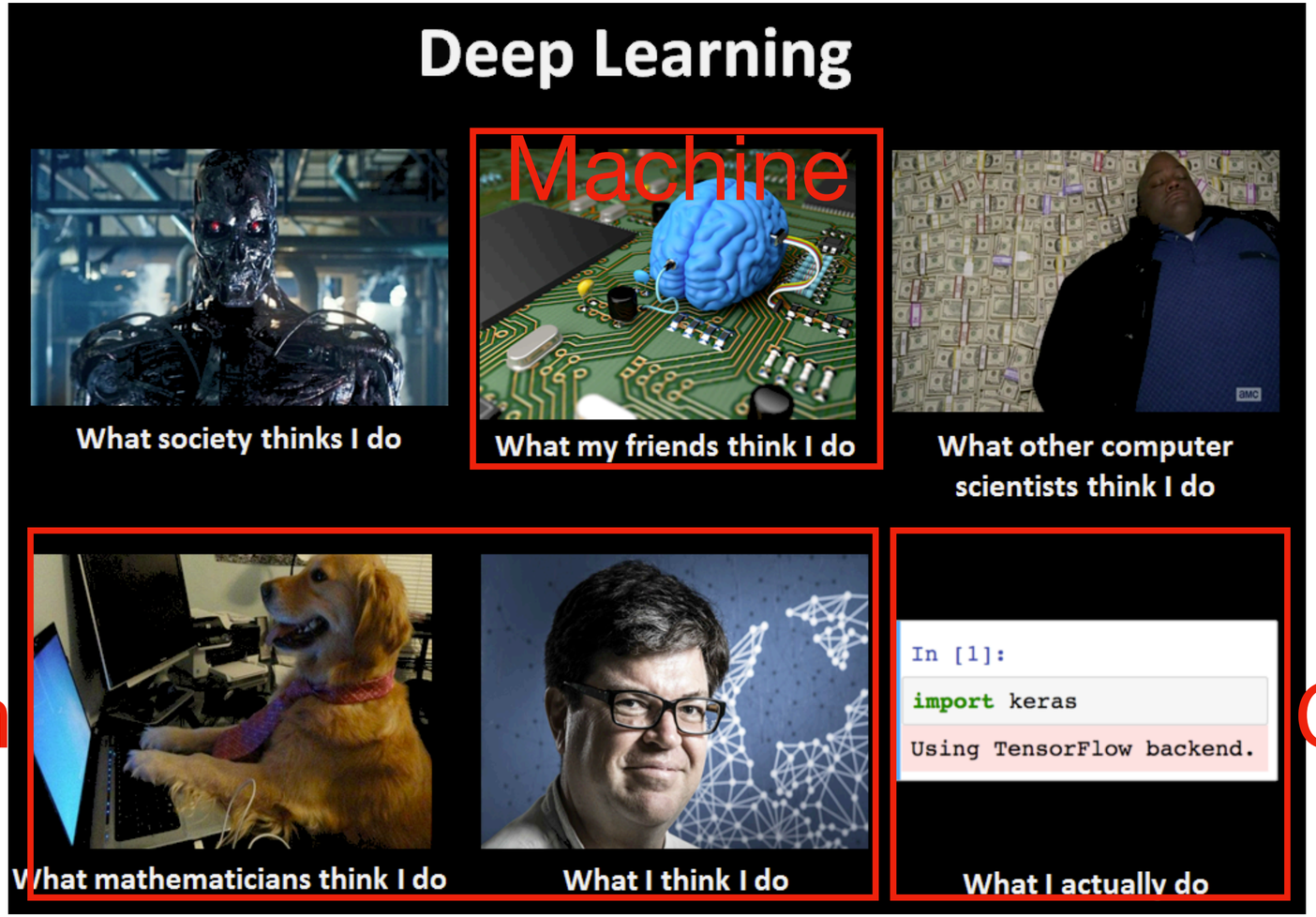


Example: a phylogenetic tree

Deep Learning



Deep Learning



Math

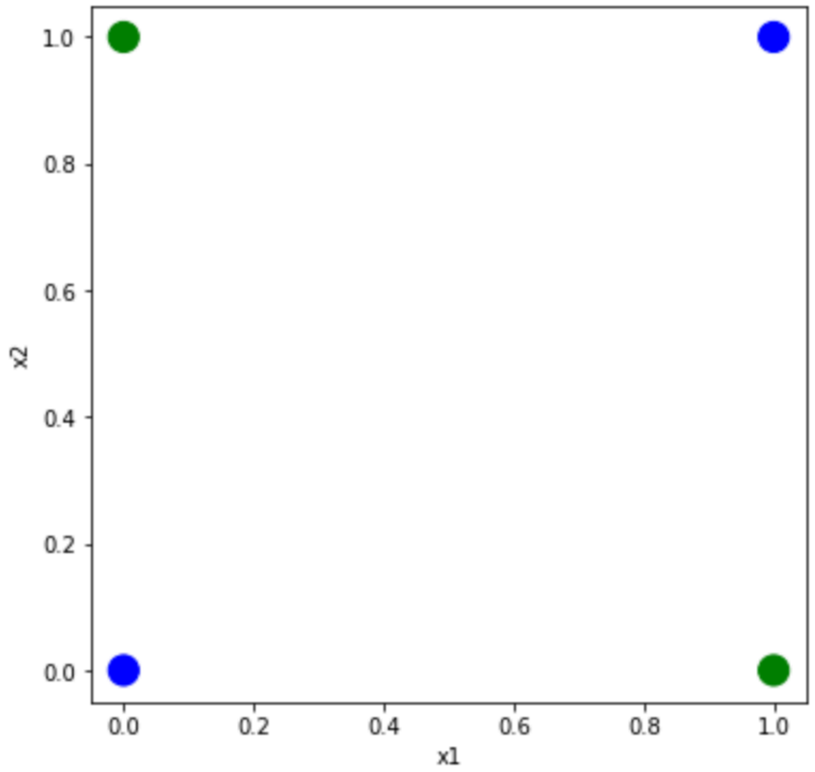
Code

Deep Learning - Power of Nonlinearity

Example: XOR problem

x1	x2	y
0	0	0
0	1	1
1	0	1
1	1	0

Truth table



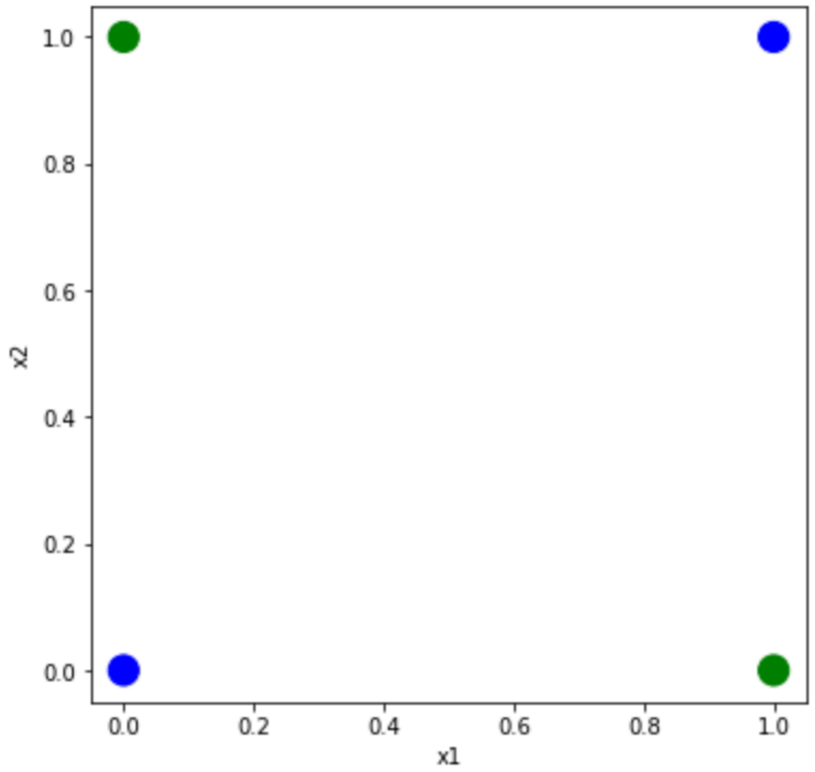
Visualization - how to draw a desision line?

Deep Learning - Power of Nonlinearity

Example: XOR problem

x1	x2	y
0	0	0
0	1	1
1	0	1
1	1	0

Truth table



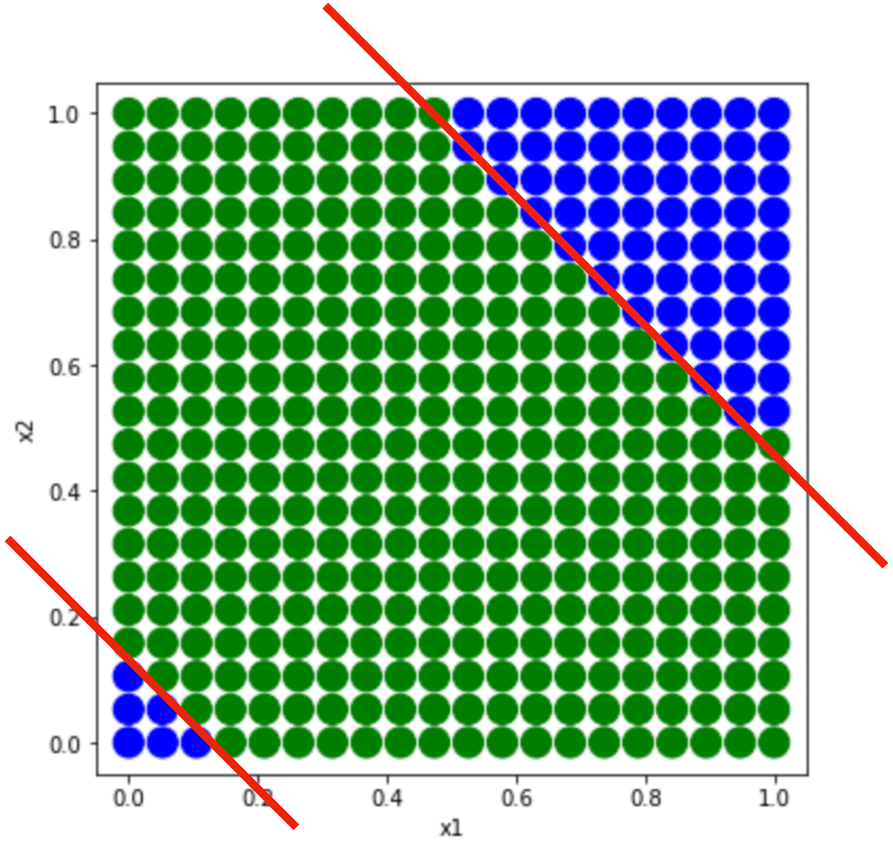
Visualization - how to draw a desision line?

Deep Learning - Power of Nonlinearity

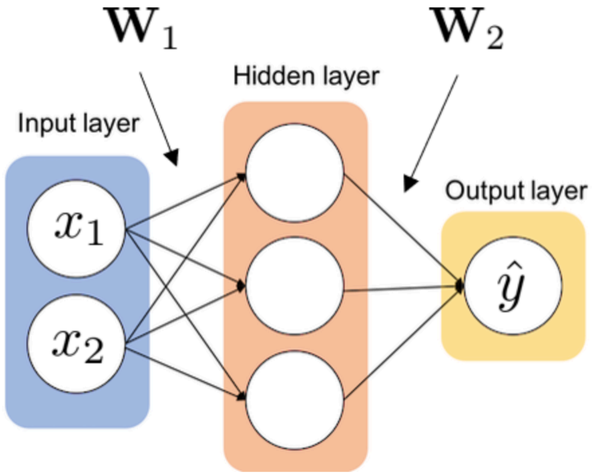
Example: XOR problem

x1	x2	y
0	0	0
0	1	1
1	0	1
1	1	0

Truth table

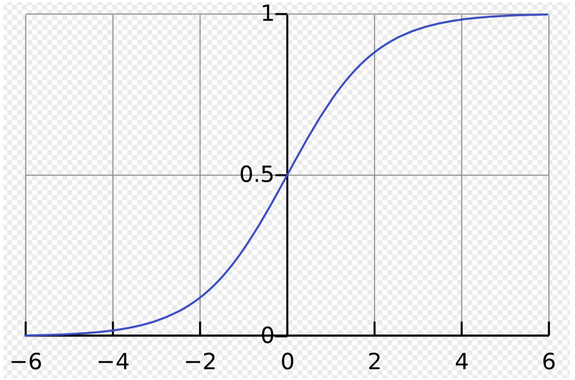


Visualization - nonlinear desision boundary



$$\mathbf{a} = g(\mathbf{h}) = g\left(\mathbf{x}^{\mathrm{T}}\mathbf{W}_1\right)$$

$$\hat{y} = \mathbf{a}\mathbf{W}_2$$



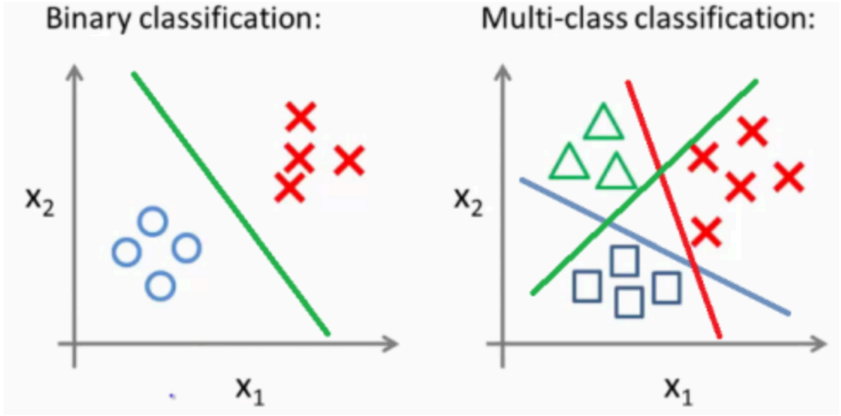
$$S(x) = \frac{1}{1+e^{-x}}$$

Solution: g is a sigmoid activation function

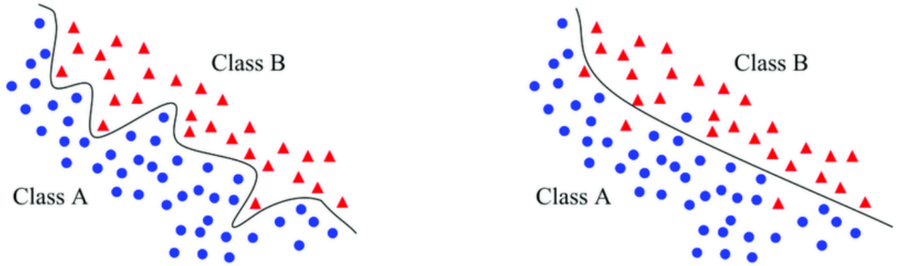
https://colab.research.google.com/drive/15p575yLuGuT1DO2cWiW1QR6Ujw4TX4zN?usp=sharing#scrollTo=L3oW6EGn1qIg

Deep Learning - Power of Nonlinearity

From linear to nonlinear decision boundaries …

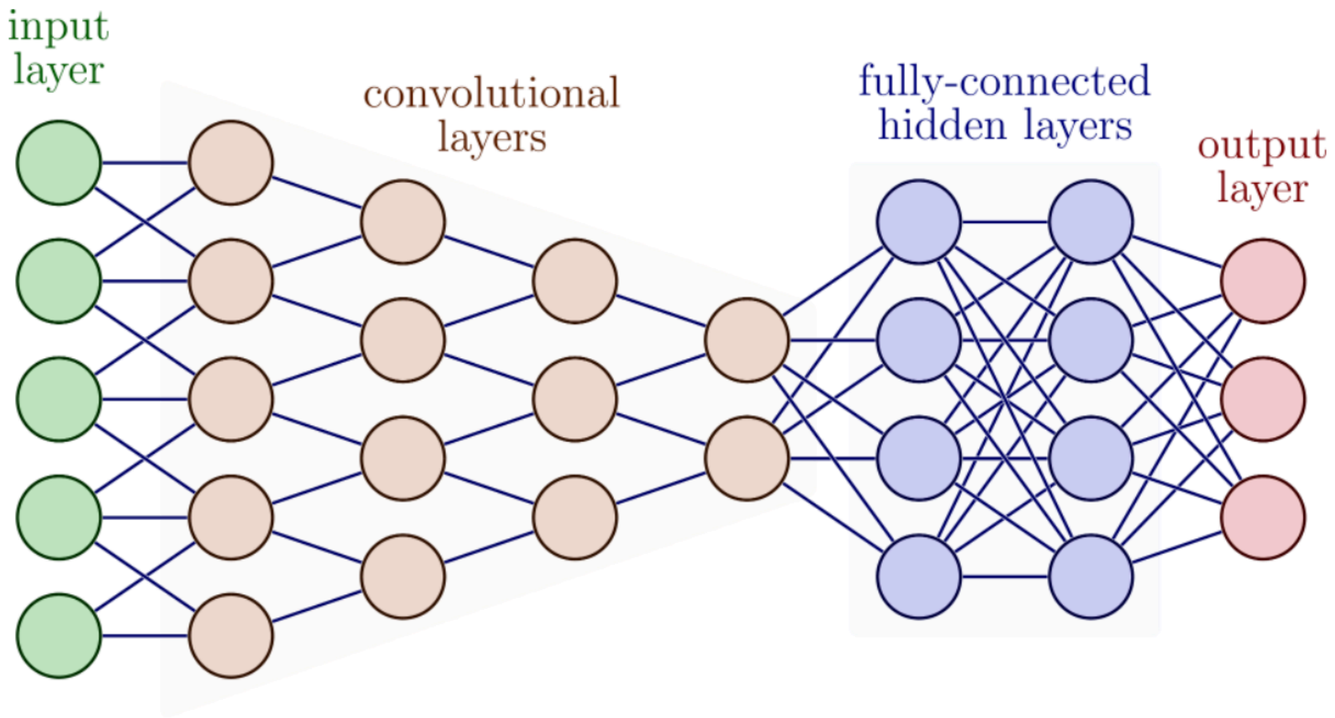


Linear decision boundaries



Nonlinear decision boundaries

Deep Learning - Neural Network



https://tikz.net/neural_networks/

Deep Learning - Universal Approximation Theorem

Nonlinearity + enough neurons = any continuous function

Universal Approximation Theorem

A neural network with one hidden layer containing a sufficient but finite number of neurons can approximate any continuous function to any reasonable accuracy.

Requirements

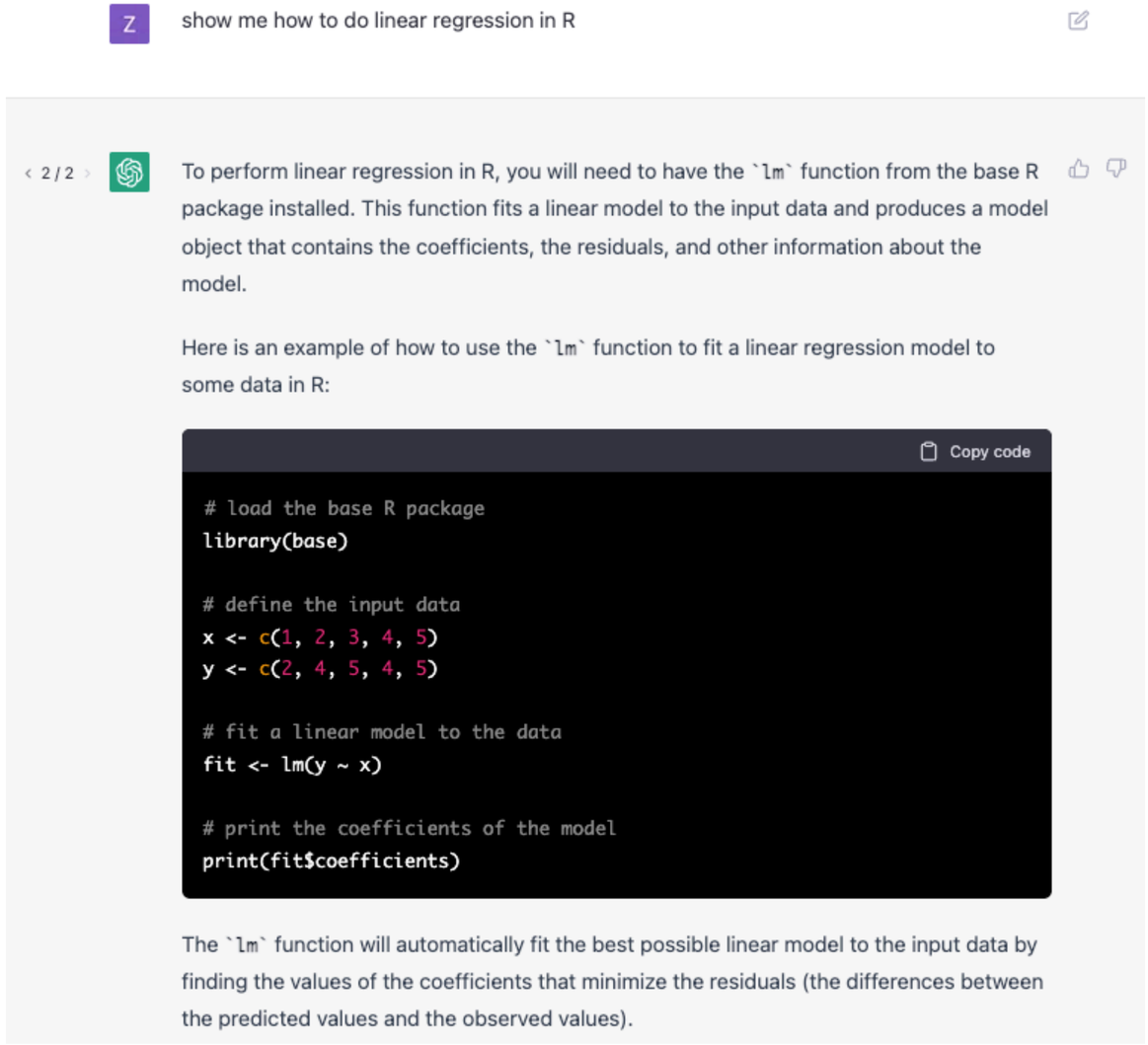
- Non-linear activation function
- Sufficient number of neurons
  - $\rightarrow$ Might be exponential in N

Caveats

- "Approximate", not "Learn"
- Overfitting to training samples
- Deeper architectures better

Figure from course https://www.ifi.uzh.ch/en/aiml/teaching/Lecture-Deep-Learning.html

Deep Learning - Demo



https://openai.com/blog/chatgpt/

Fin

Slides Credit to

- Course Introduction to Machine Learning @ETH Zürich (https://las.inf.ethz.ch/teaching/introml-s20)
- Course Computational Intelligence Lab @ETH Zürich (http://da.inf.ethz.ch/teaching/2022/CIL/)
- Course Deep Learning @UZH (https://www.ifi.uzh.ch/en/aiml/teaching/Lecture-Deep-Learning.html)
- Interesting work @Department of Computational Linguistics UZH and @UniNE with Prof. Steven Moran