

Introduction to Machine Learning

Zifan Jiang 15/11/2023

University of Neuchâtel

About Me



Department of Computational Linguistics
Language, Technology and Accessibility

[Research](#) [Team](#) [Publications](#) [Datasets and Resources](#) [Media and Outreach](#) [News](#)

[Home](#) / [Team](#) / Zifan Jiang



Zifan Jiang, M. Sc.

PhD candidate

Phone: +41763376465

Room number: AND-2-18

> [E-Mail](#)

> [Website](#)

Hi, my name is Zifan (子凡) [tsɿ³ fan²] Jiang (蒋) [tɕjan³]

I am a Ph.D. student interested in machine translation and natural language processing, advised by > Prof. Dr. Sarah Ebling and > Prof. Dr. Rico Sennrich. I am currently working on automatic sign language processing and translation as part of the > Flagship IICT Project. I also study computer vision and sign language linguistics so as to cope with the visual-gestural nature of sign language.

<https://www.cl.uzh.ch/en/research-groups/accessibility/team/jiang.html>

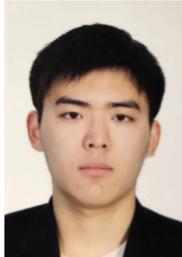
About Me



Department of Computational Linguistics
Language, Technology and Accessibility

Research Team Publications Datasets and Resources Media and Outreach News

Home / Team / Zifan Jiang



Zifan Jiang, M. Sc.

PhD candidate

Phone: +41763376465

Room number: AND-2-18

> E-Mail

> Website ↗

Hi, my name is Zifan (子凡) [tsɿ³ fan²] Jiang (蒋) [tɕjan³]

I am a Ph.D. student interested in machine translation and natural language processing, advised by > Prof. Dr. Sarah Ebli and > Prof. Dr. Rico Sennrich. I am currently working on automatic sign language processing and translation as part of the > Flagship IICT Project. I also study computer vision and sign language linguistics so as to cope with the visual-gestural nature of sign language.

<https://www.cl.uzh.ch/en/research-groups/accessibility/team/jiang.html>

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



Faculty of Science
Institute of Biology
Laboratory of Language Evolution

Current site search

Rechercher



Directory

EN FR

Home

Team

Steven Moran

Tiana Danner

Marco Maiolini

Zifan Jiang

Carlos Silva

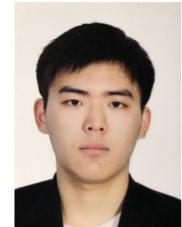
Publications

Open positions

News

Zifan Jiang

> Print



Zifan Jiang

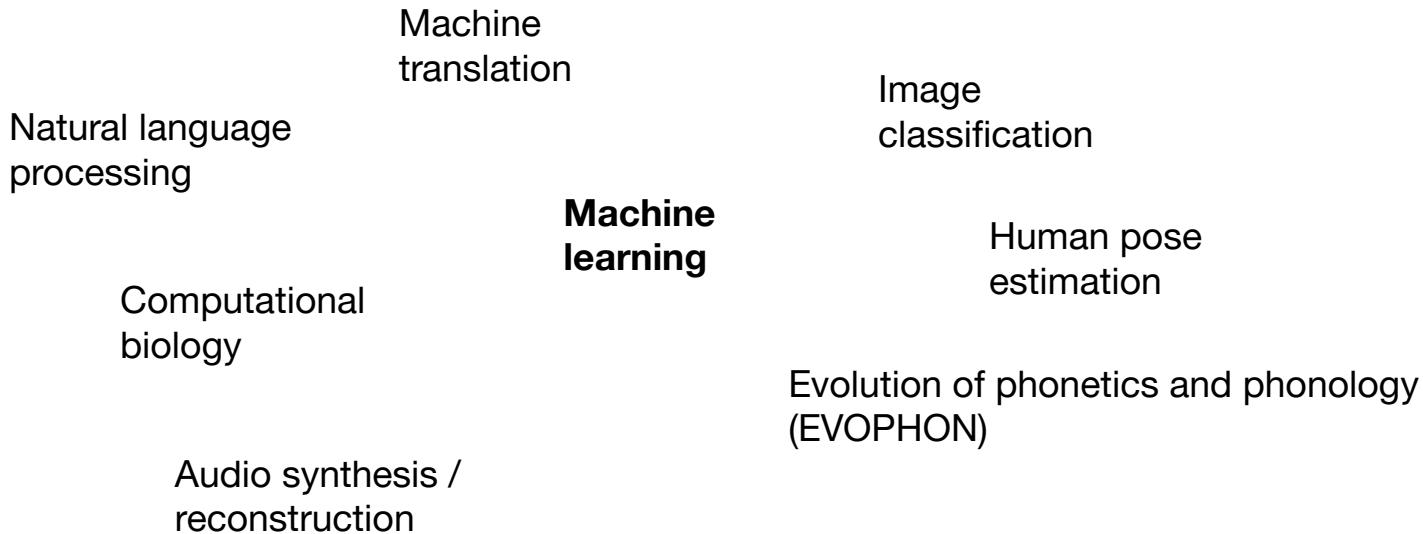
Scientific programmer

↗ 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



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

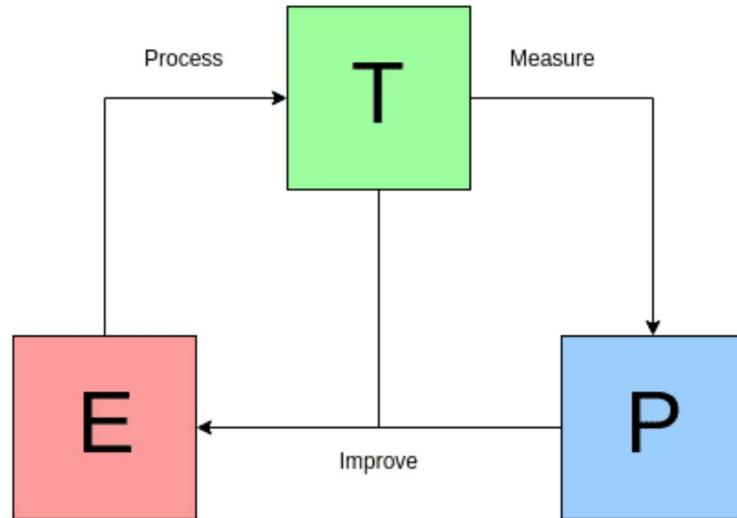
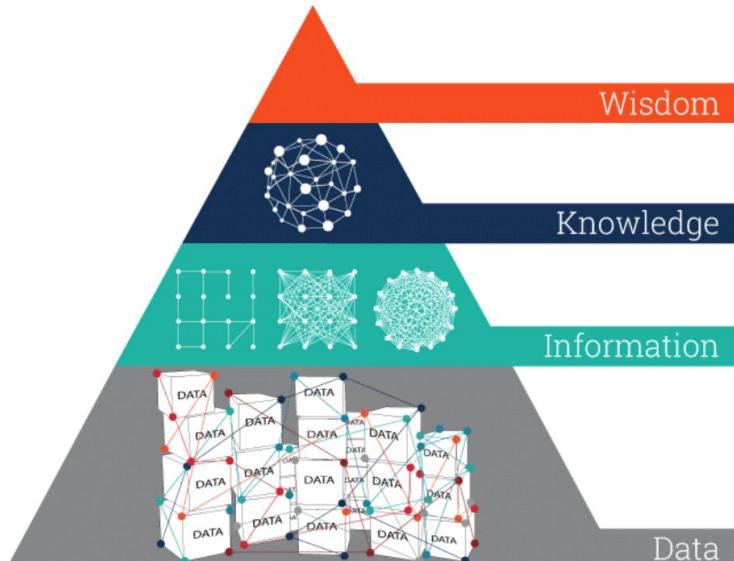


Figure 1. The Mitchell Paradigm, visualized.

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



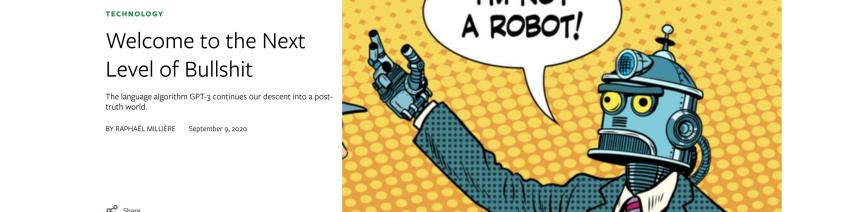
Data, Information, Knowledge, Wisdom (DIKW) Pyramid

What is Machine Learning: the Role

Machine Learning plays an important role in this pyramid



Data, Information, Knowledge, Wisdom (DIKW) Pyramid



 ne of the most salient features of our culture is that there is so much bullshit." These are the opening words of the short book *On Bullshit*, written by the philosopher Harry Frankfurt. Fifteen years after the publication of this surprise bestseller, the rapid progress of research on artificial intelligence is forcing us to reconsider our conception of bullshit as a hallmark of human speech, with troubling implications. What do



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: $X = \text{thigh circumference}$, $Y = \text{body fat percentage}$

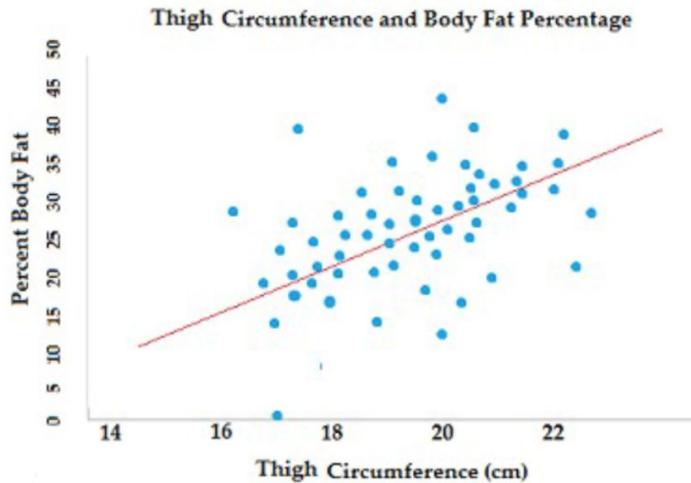


Figure from lecture [Linear models I](#)

Supervised Learning - Regression

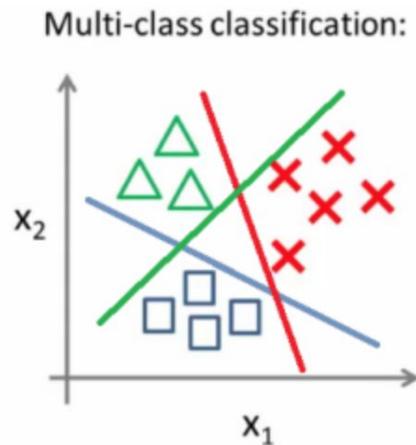
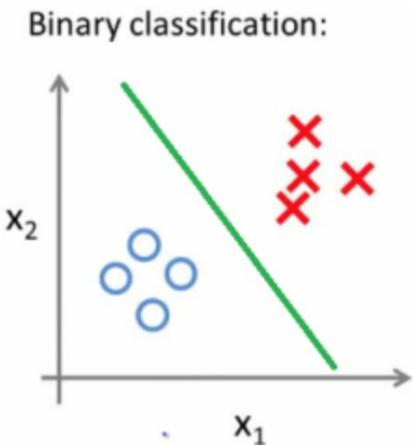
Goal: Predict real valued labels

Summary

Model	Variables	Distribution	R code
Linear Regression	$Y = b_0 + b_1 x$	Normal	<code>lm(formula, data)</code>
General Linear Models	$Y = b_0 + b_1 x_1 + b_2 x_2 + \dots$	Normal	<code>lm(formula, data)</code>
Generalized Linear Models (GLM)	$Y = b_0 + b_1 x_1 + b_2 x_2 + \dots$	Any	<code>glm(formula, family, data)</code>

Supervised Learning - Classification

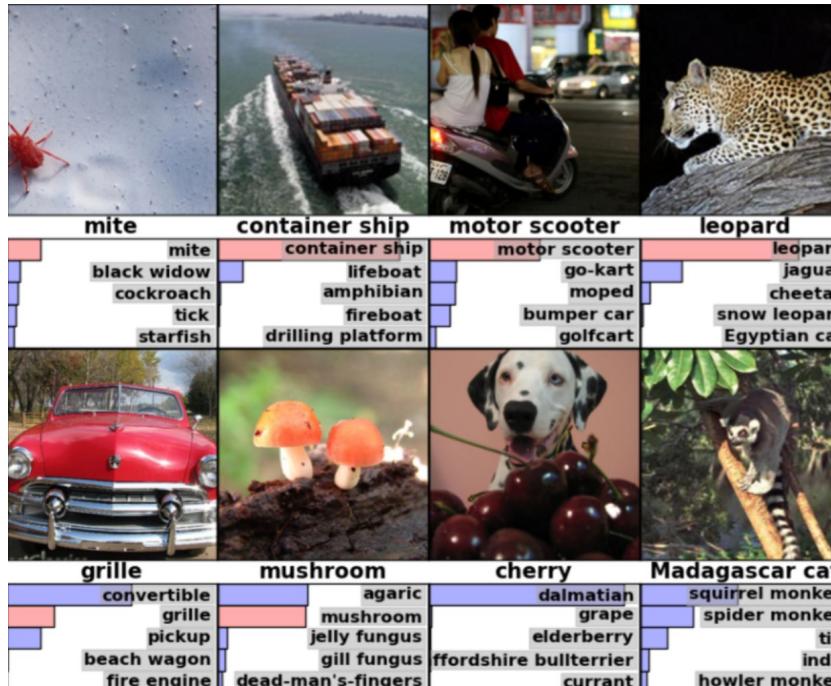
Goal: Predict categorical labels



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

Supervised Learning - Classification

Example: image classification - [ImageNet](#)



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: image classification - Chihuahua (dog) or muffin?

G GPT-4 ⓘ



Please tell me what is in each cell in a tabular format aligned with this picture grid.

Certainly! Here's a breakdown of the content of each cell in the format of a table:

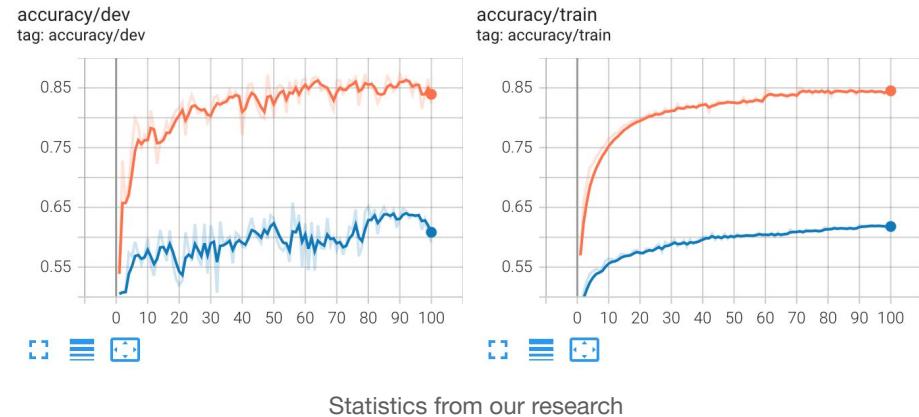
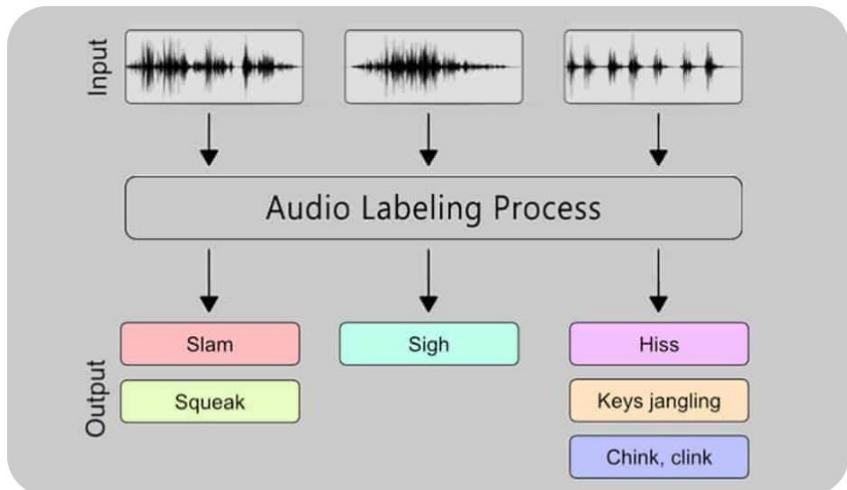
Row	Column 1	Column 2	Column 3	Column 4
1	Muffin with blueberries	Chihuahua (side profile)	Blueberry Muffin	Chihuahua (profile looking up)
2	Chihuahua (frontal face)	Cookie with chocolate chips	Chihuahua (face close-up)	Blueberry Muffin with sugar on top
3	Blueberry Muffin with a smiley face	Chihuahua (frontal face close-up)	Blueberry Muffin with blueberries on top	Chihuahua (slightly showing teeth)
4	Chihuahua (puppy face close-up)	Group of Cookies with chocolate chips	Group of Chihuahua puppies	Muffin with blueberries

知乎 @小小将

Supervised Learning - Classification

Example: extract and classify monkey calls from raw audio

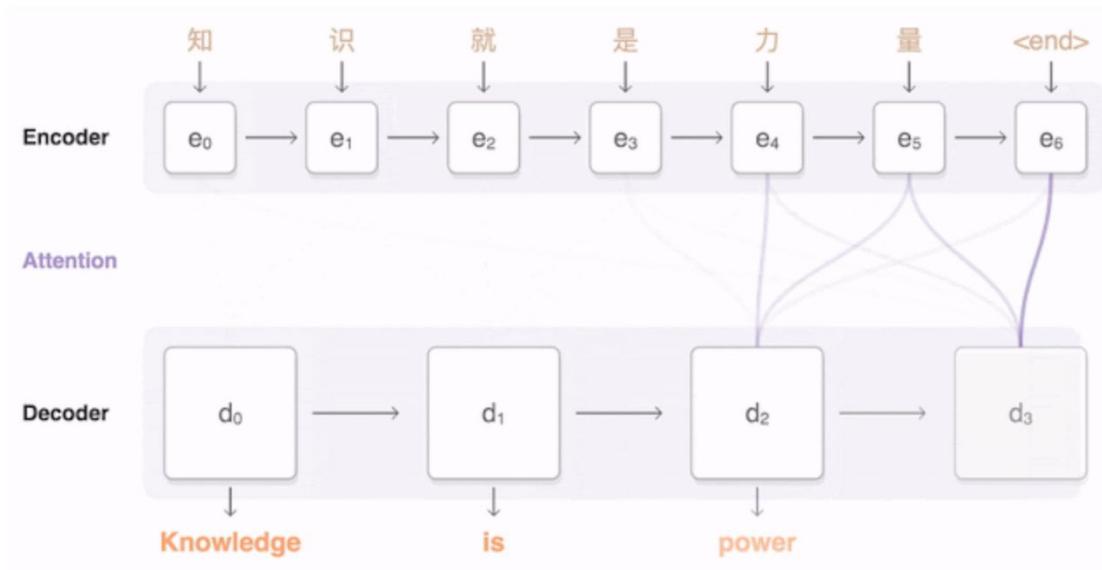
X = audio waveform, Y = call types (or non-call)



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

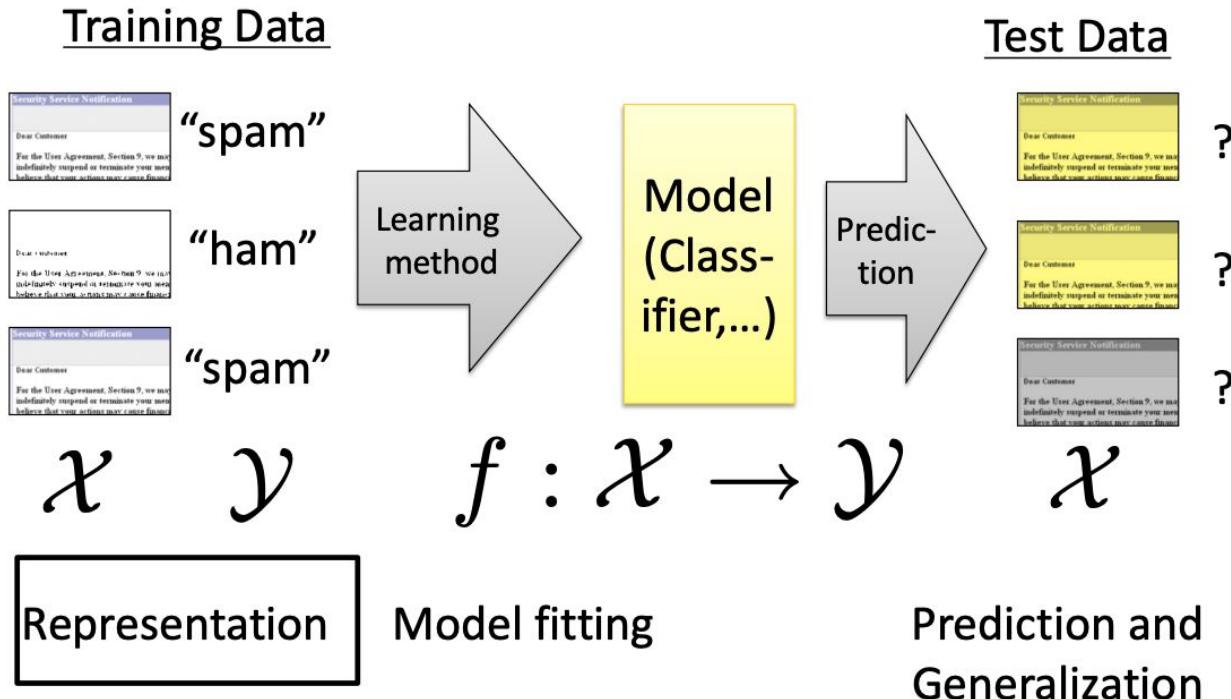
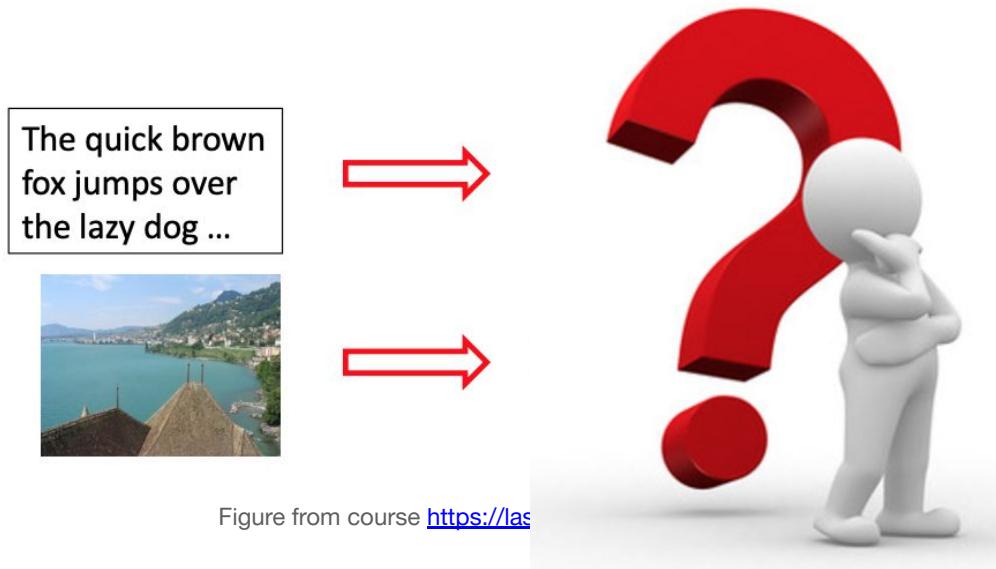


Figure from course <https://las.inf.ethz.ch/teaching/introml-s20>

Data Representation

How to represent data in a machine-understandable fashion?



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]



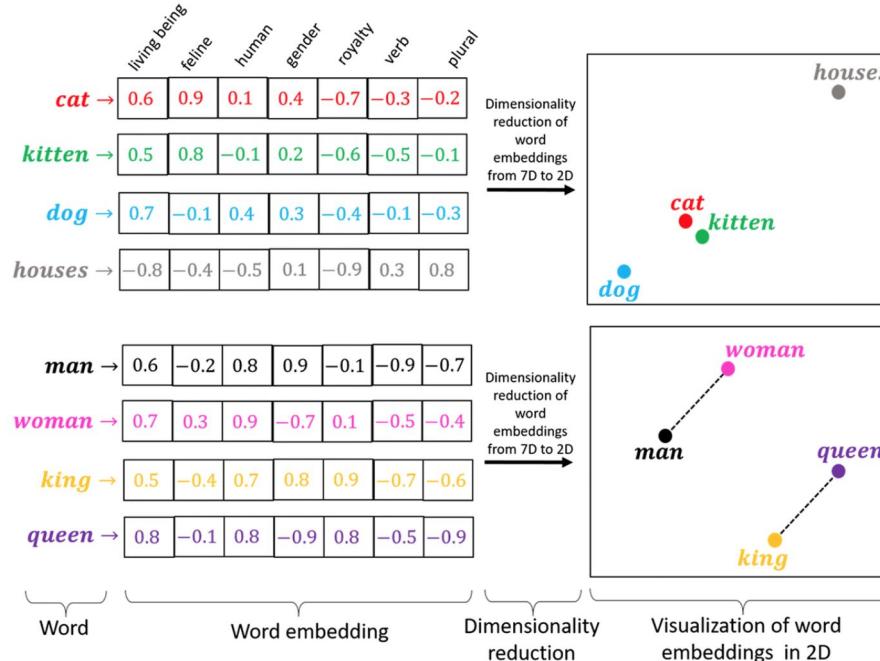
[.3 .01 .1 2.3 0 0 1.1 ...]

Vectors!

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

Data Representation - Text

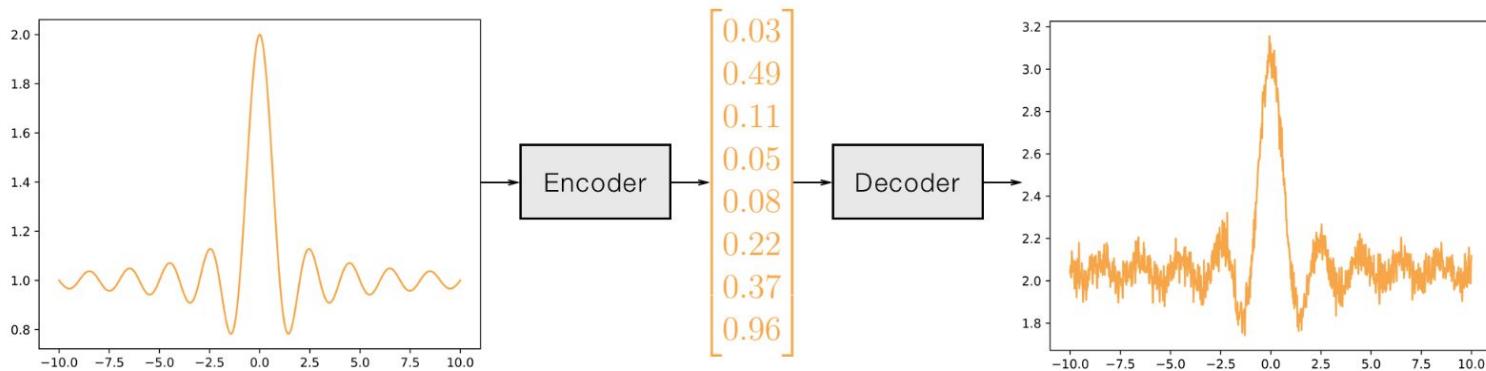
How to represent data in a machine-understandable fashion?



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

Data Representation - Audio

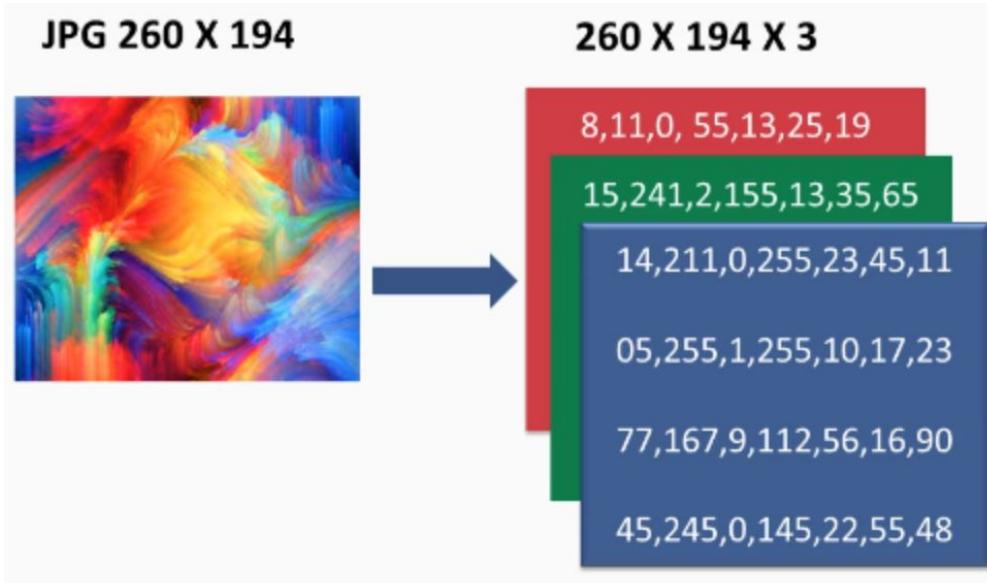
How to represent data in a machine-understandable fashion?



<https://github.com/shobrook/sequitur>

Data Representation - Image

How to represent data in a machine-understandable fashion?



Unsupervised Learning - Pipeline / Paradigm

Training Data



Test Data



\mathcal{X}

$$f : \mathcal{X} \rightarrow \mathcal{Y}$$

Representation

Model fitting

Prediction

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

Unsupervised Learning - Clustering

Unsupervised classification

Example: K-means clustering

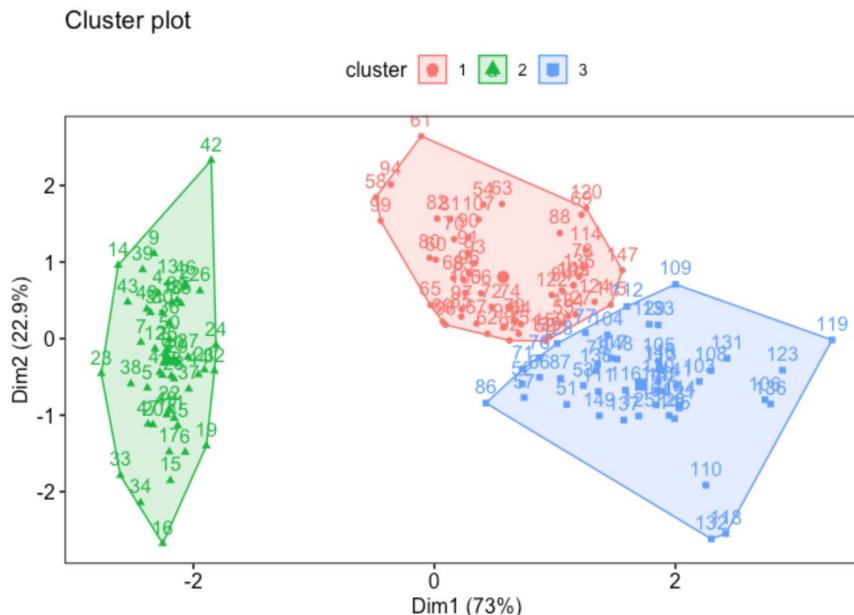


Figure from lecture [Dimensionality reduction and clustering II](#)

Unsupervised Learning - Dimension Reduction

Unsupervised regression

Example: principal component analysis (PCA)

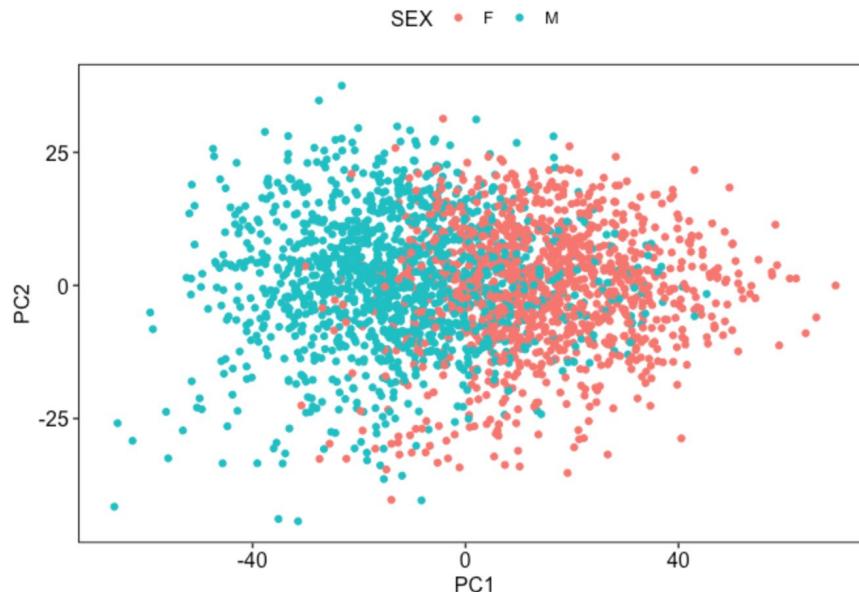
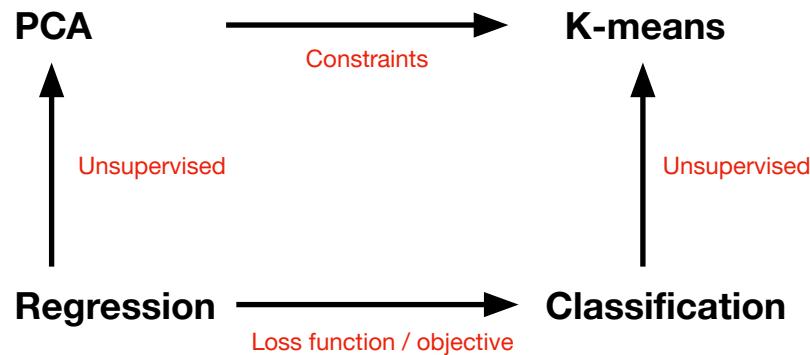


Figure from lecture [Dimensionality reduction and clustering I](#)

Unsupervised Learning - Comparison

Look at everything so far in a big picture ...



Unsupervised Learning - Comparison

Look at everything so far

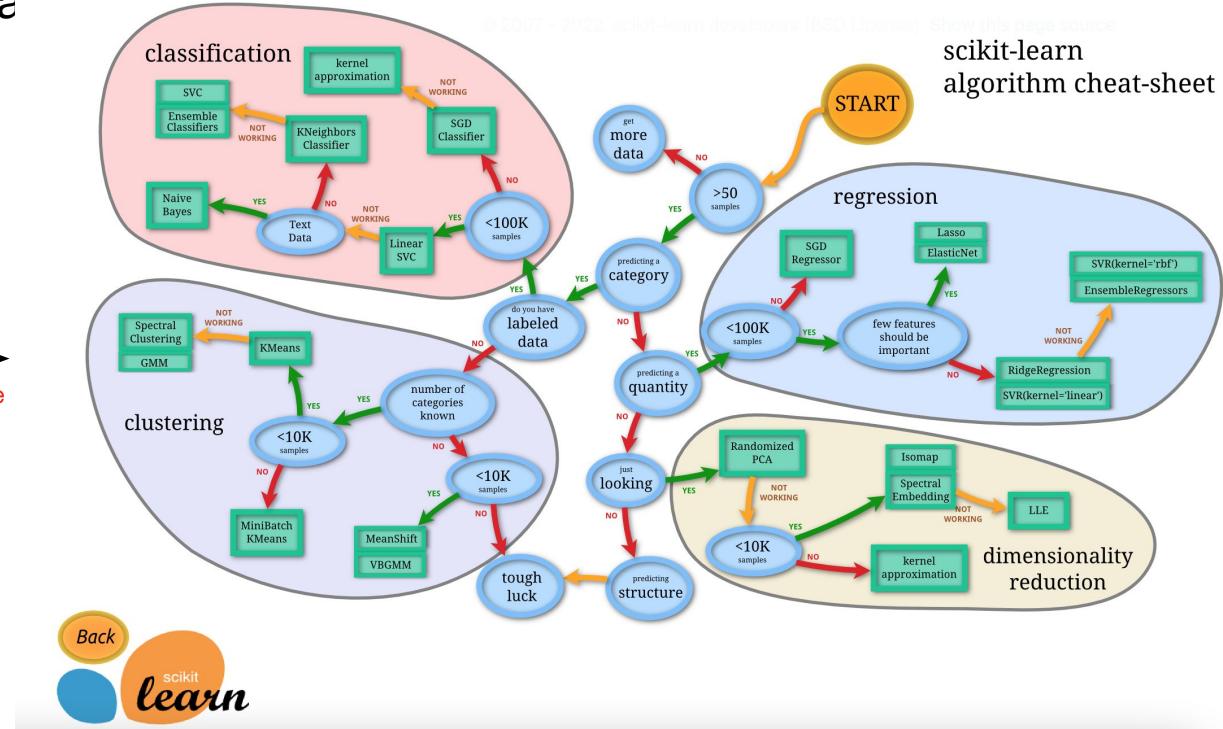
PC



A

Unsupervised

Regression



Unsupervised Learning - Autoencoders

First reduce the dimension, then reconstruct to the original ...

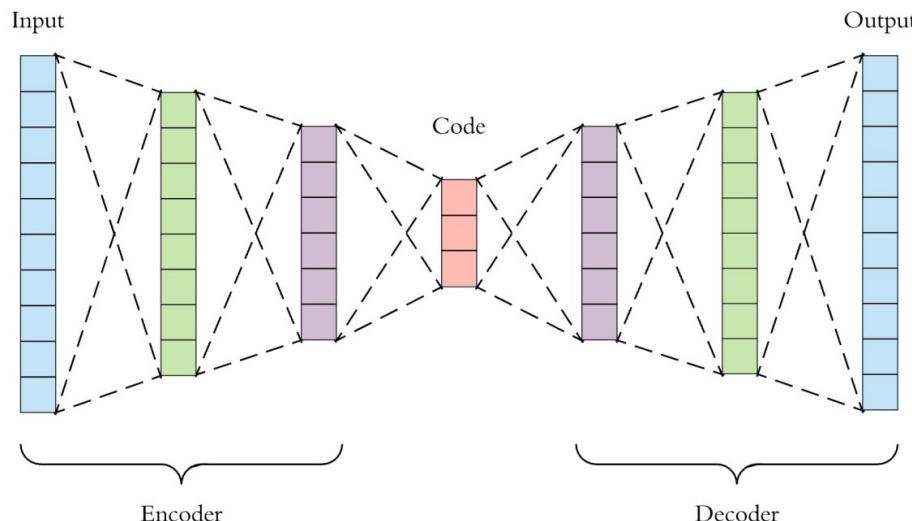
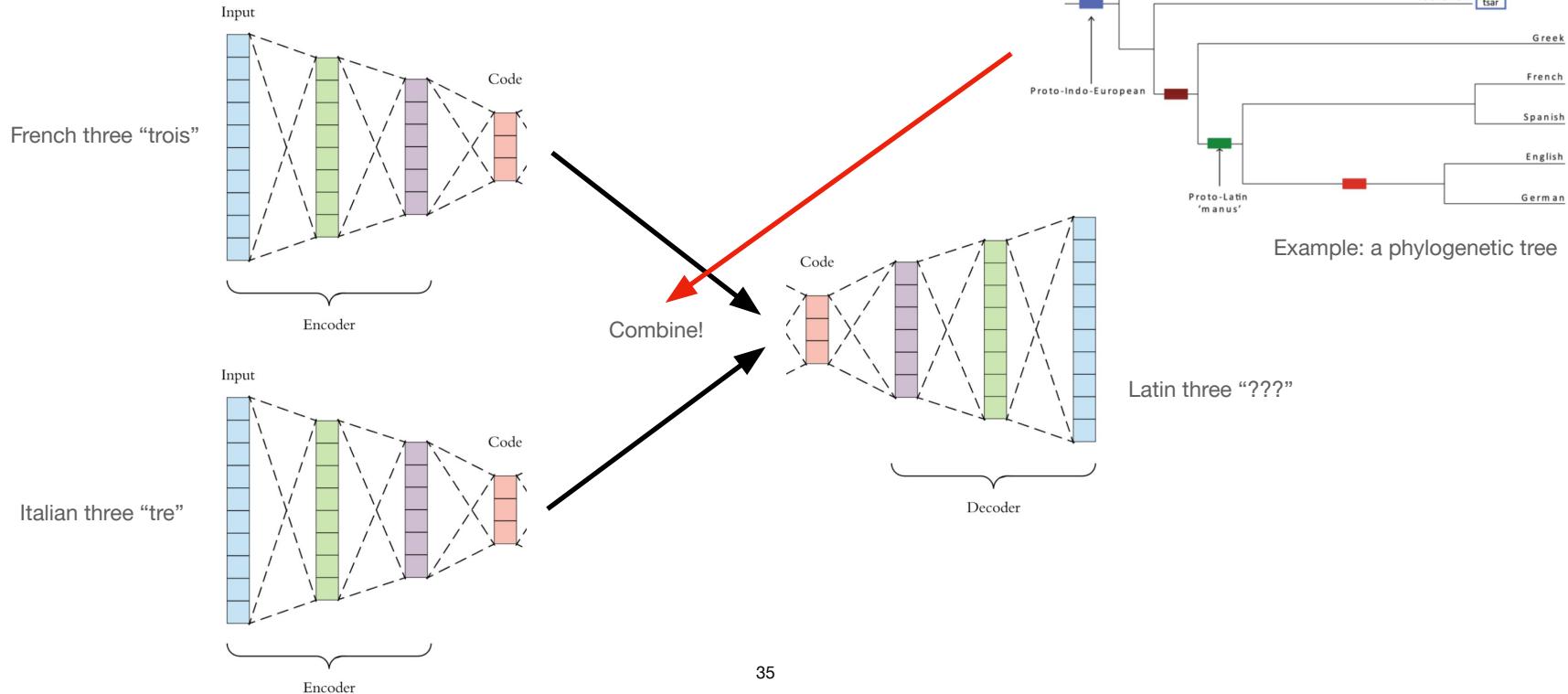


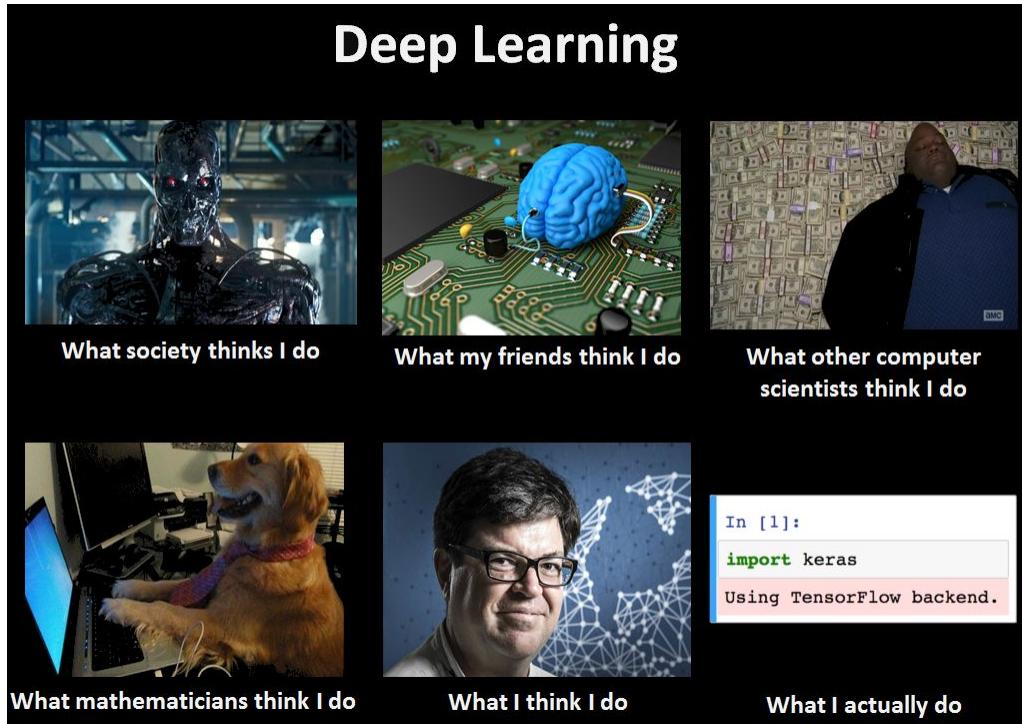
Figure from course
<http://www.da.inf.ethz.ch/teaching/2022/CIL/>

Unsupervised Learning - Autoencoders

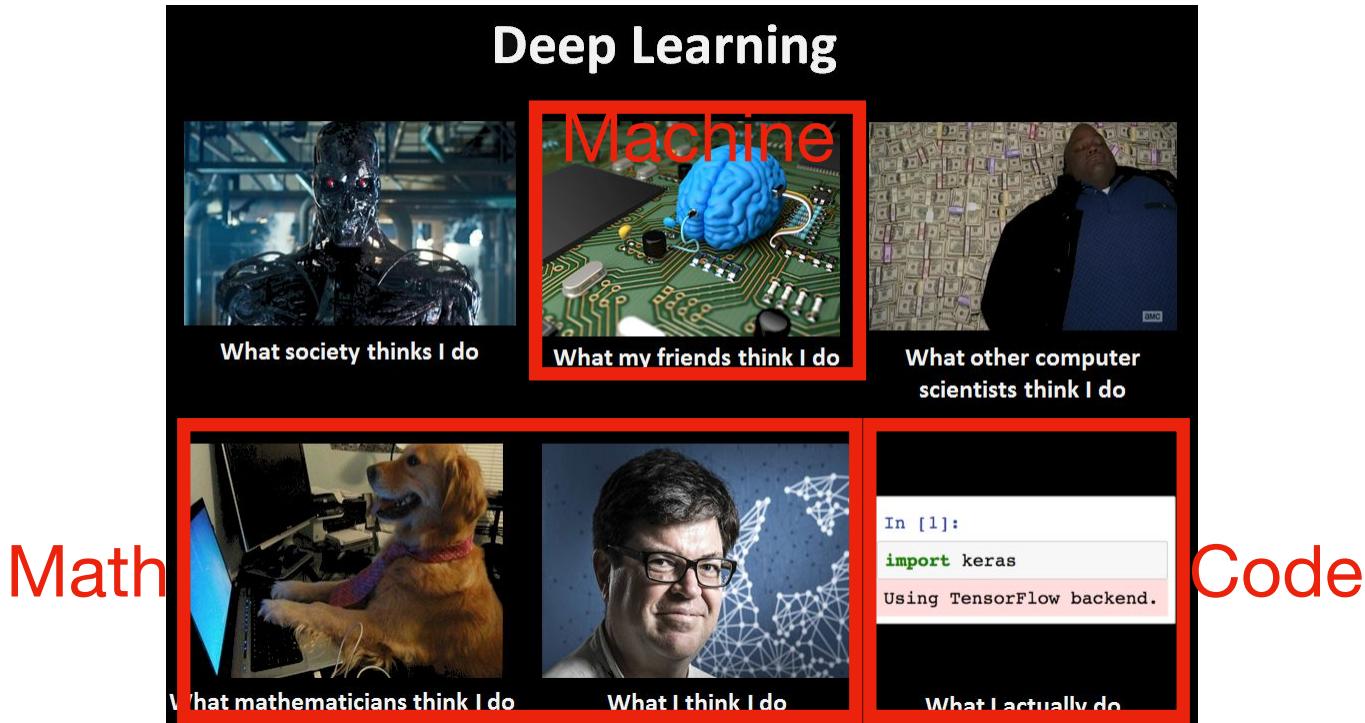
Example: reconstruct word sound from thousands c



Deep Learning



Deep Learning

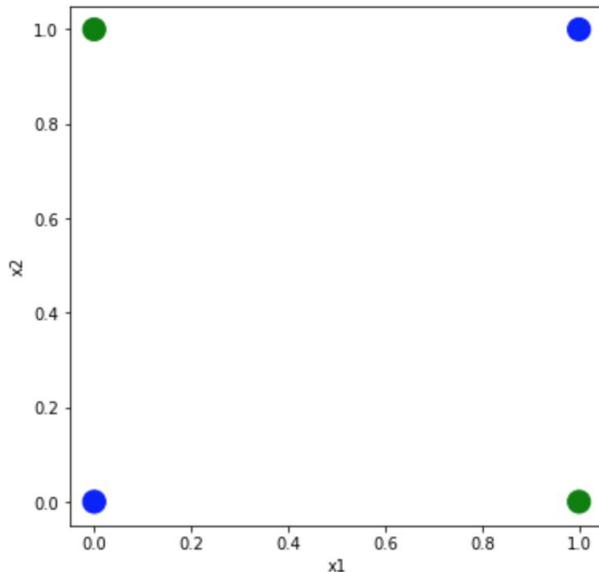


Deep Learning - Power of Nonlinearity

Example: XOR problem

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

Truth table



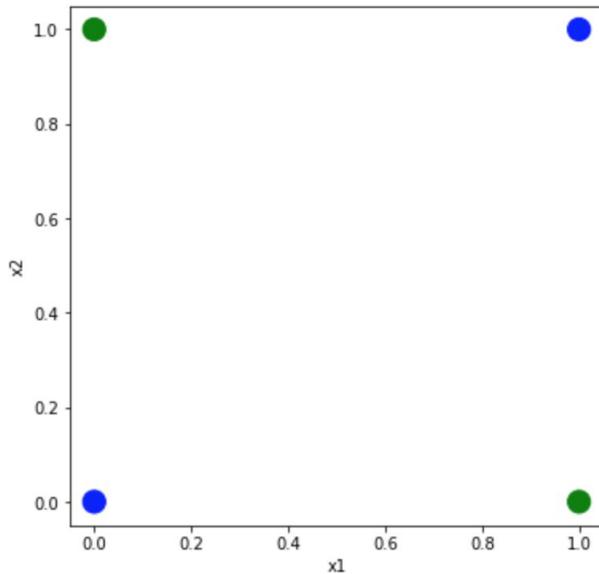
Visualization - how to draw a decision line?

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?

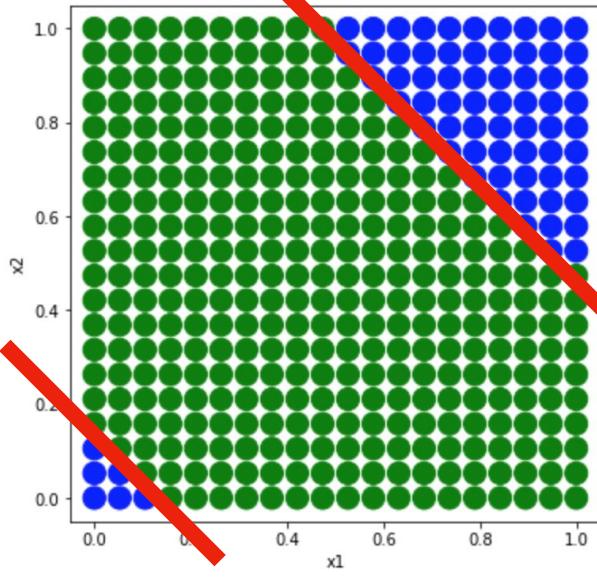


Deep Learning - Power of Nonlinearity

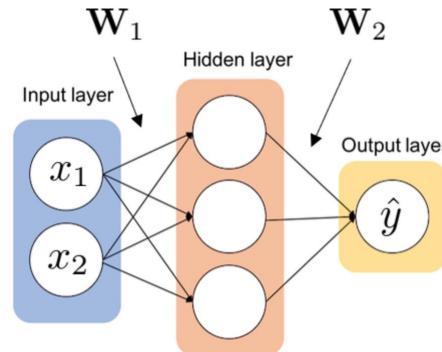
Example: XOR problem

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

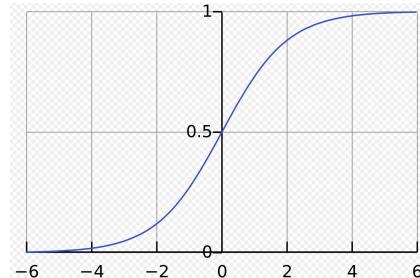
Truth table



Visualization - nonlinear desision boundary



$$\mathbf{a} = g(\mathbf{h}) = g(\mathbf{x}^T \mathbf{W}_1)$$
$$\hat{y} = \mathbf{a} \mathbf{W}_2$$



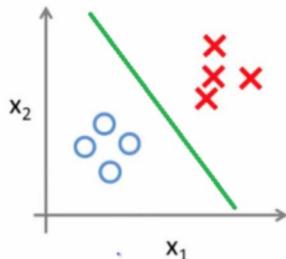
$$S(x) = \frac{1}{1 + e^{-x}} :$$

Solution: g is a sigmoid activation function

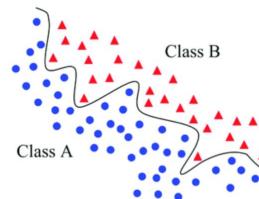
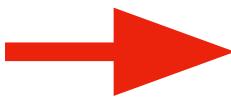
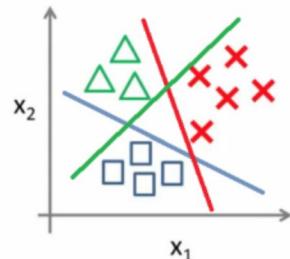
Deep Learning - Power of Nonlinearity

From linear to nonlinear decision boundaries ...

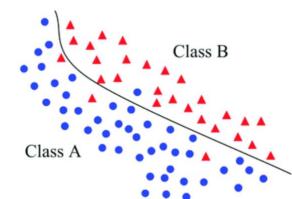
Binary classification:



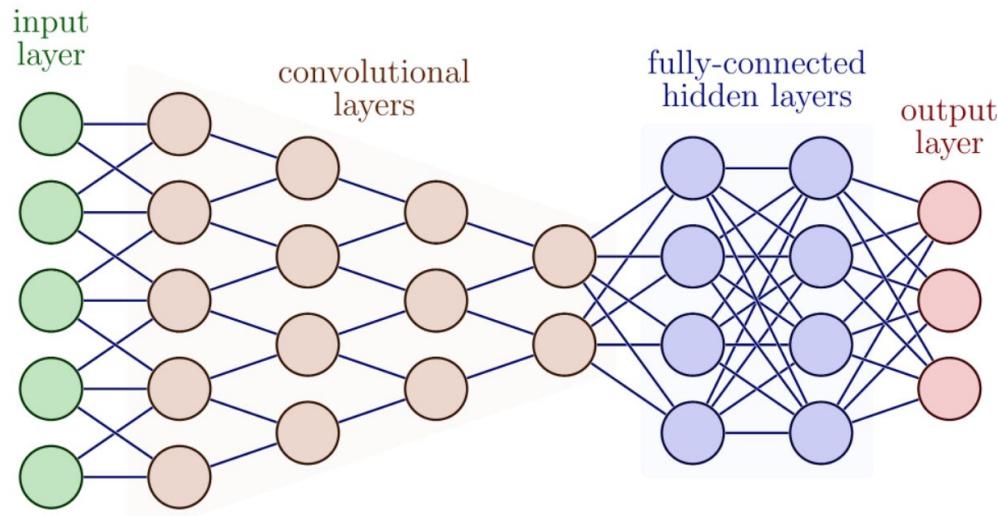
Multi-class classification:



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
→ 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

Z show me how to do linear regression in R

2 / 2 To perform linear regression in R, you will need to have the `lm` function from the base R package installed. This function fits a linear model to the input data and produces a model object that contains the coefficients, the residuals, and other information about the model.

Here is an example of how to use the `lm` function to fit a linear regression model to some data in R:

```
# load the base R package
library(base)

# define the input data
x <- c(1, 2, 3, 4, 5)
y <- c(2, 4, 5, 4, 5)

# fit a linear model to the data
fit <- lm(y ~ x)

# print the coefficients of the model
print(fit$coefficients)
```

The `lm` function will automatically fit the best possible linear model to the input data by finding the values of the coefficients that minimize the residuals (the differences between the predicted values and the observed values).

<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