

Introduction to Machine Learning

Zifan Jiang 01/06/2022

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



Zifan Jiang

PhD Student

Tel.: +41763376465

Raumbezeichnung: AND-2-18

→ jiang@cl.uzh.ch

→ [Webseite](#)

Phonetics & Speech Sciences

Digital Linguistics

Former Staff Members and Alumni

Student Assistants

Institutsleitung

Hi, my name is Zifan [ts̯ən̩: fæn̩] Jiang [tʃən̩]

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.: +41 76 337 6465

Raumbezeichnung: AND-2-18

→ jiang@cl.uzh.ch

→ Webseite

Hi, my name is Zifan [ts̯ʃ: fæn] Jiang [tʃæŋ]

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

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

Zifan Jiang

Team

Steven Moran

Tiana Danner

Marco Maiolini

Zifan Jiang

Carlos Silva

Publications

Open positions

News



Zifan Jiang

Scientific programmer

Top of the page

Print

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

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

What We Do ...

By leveraging data science and machine learning techniques

Machine translation
Natural language processing

Machine learning

Computational biology
Audio synthesis / reconstruction

Image classification

Human pose estimation

Evolution of phonetics and phonology (EVOPHON)

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

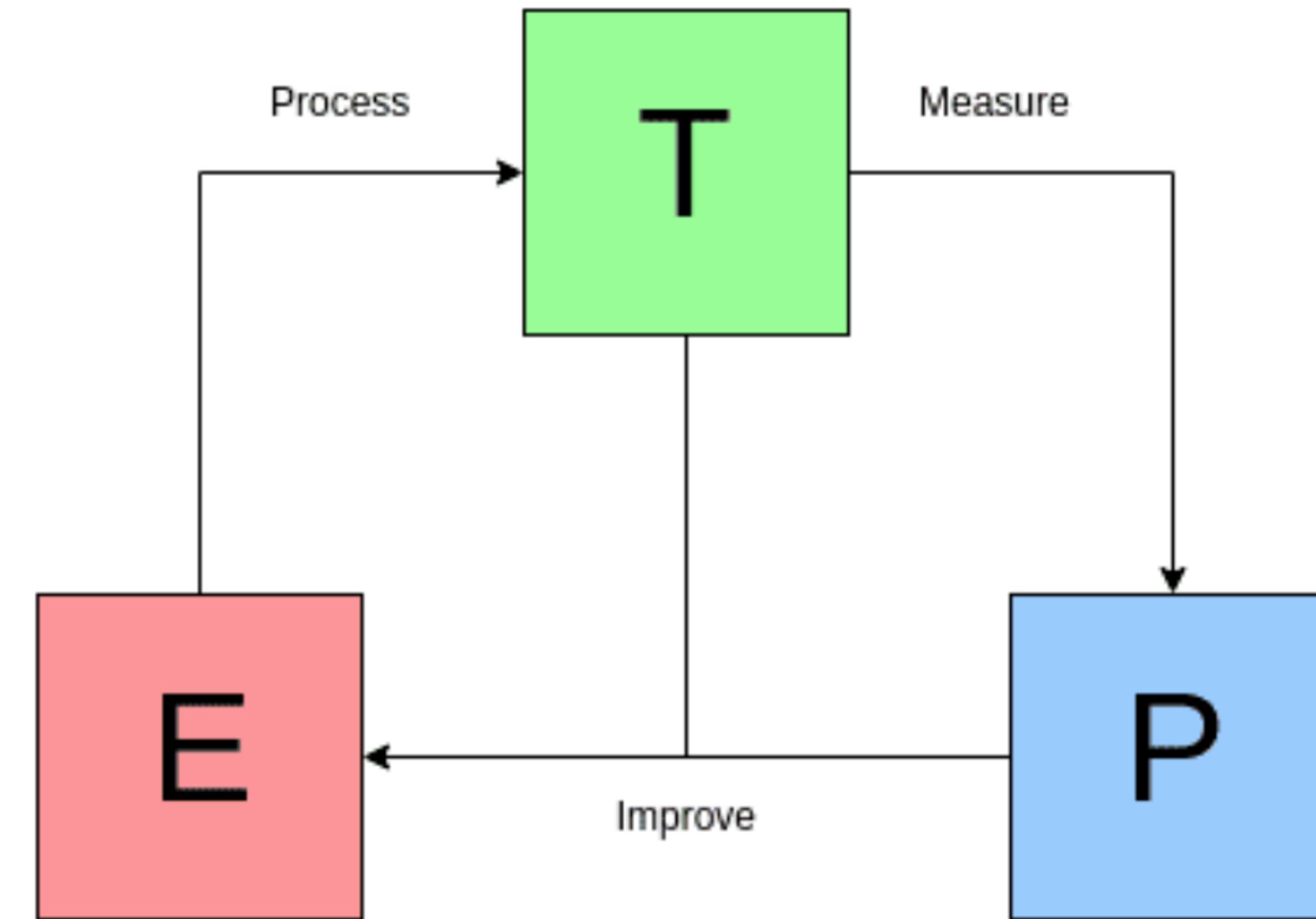
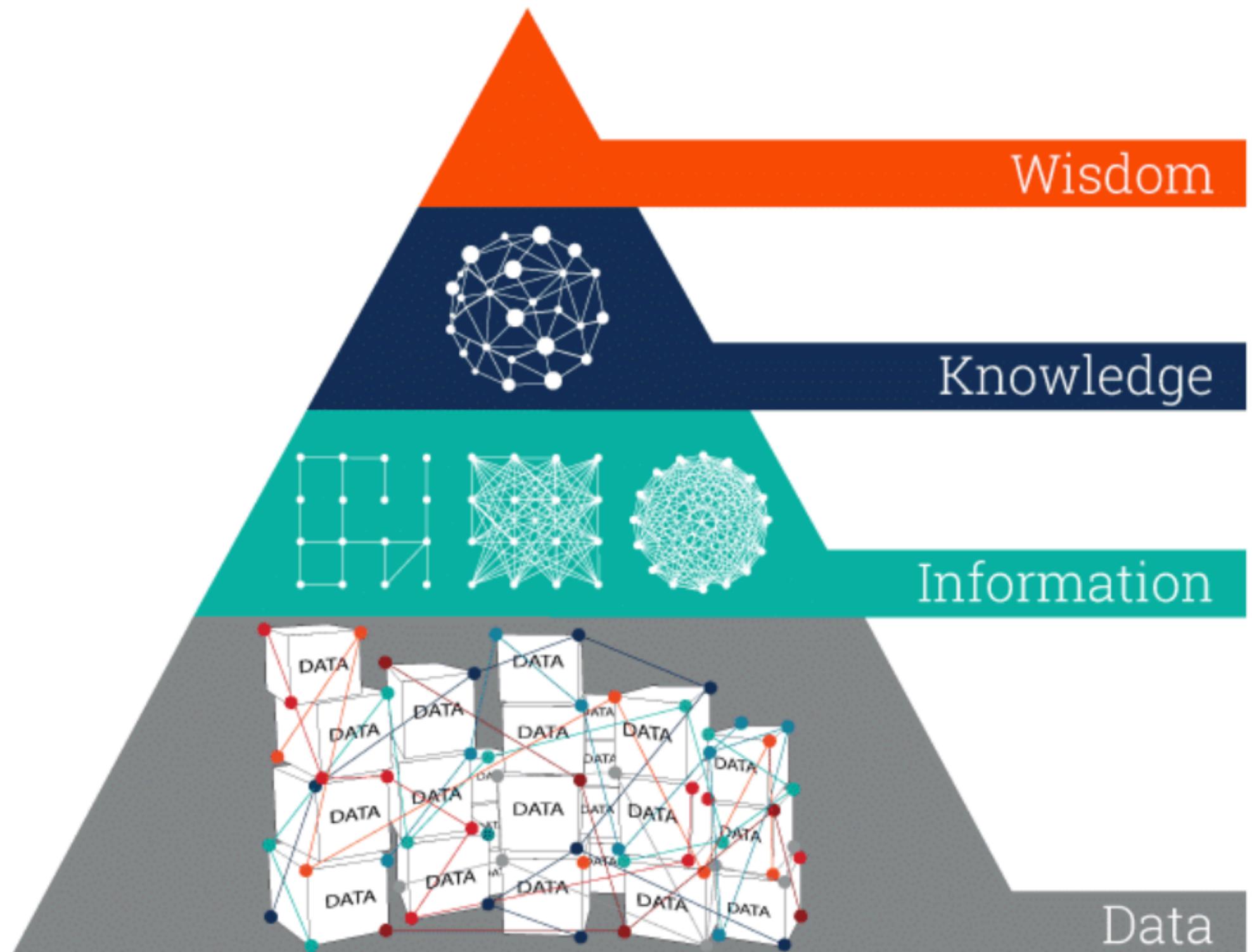


Figure 1. The Mitchell Paradigm, visualized.

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

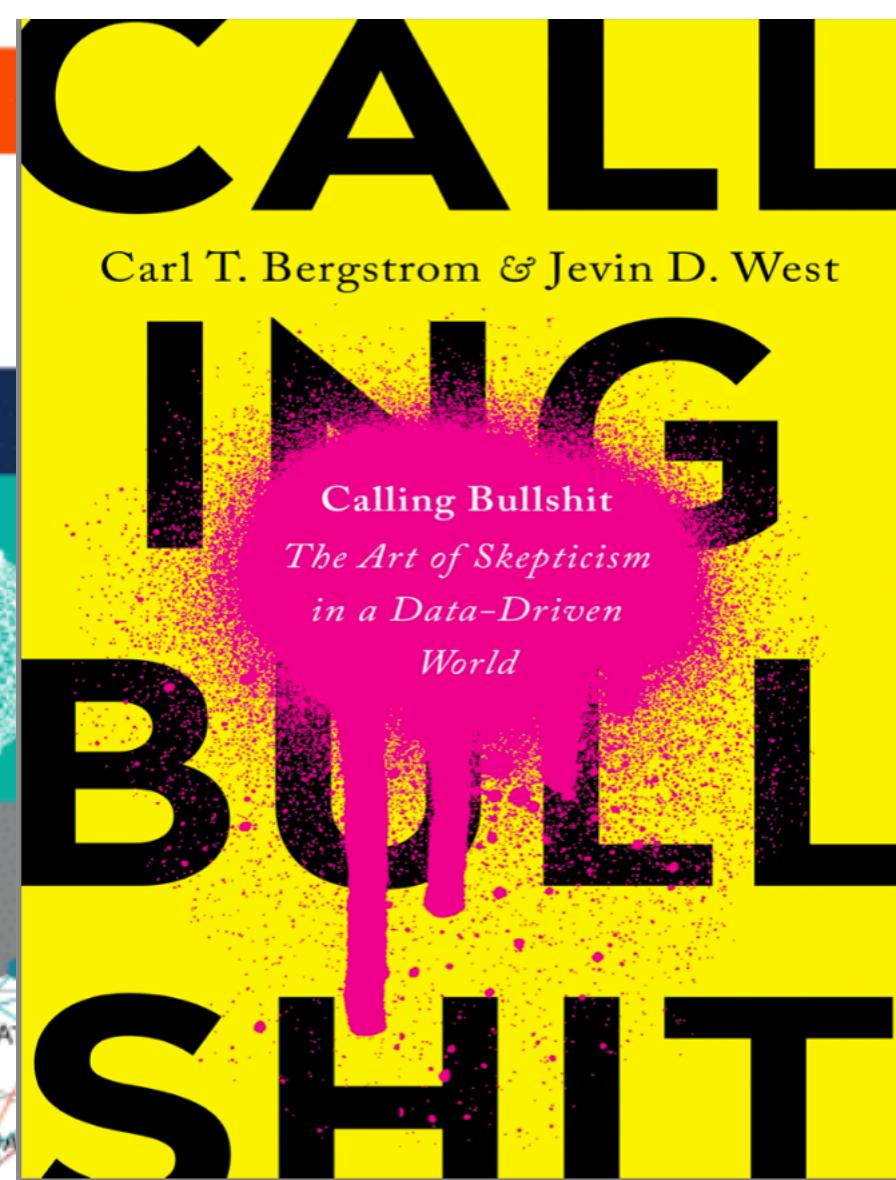
Machine Learning plays an important role in this pyramid



Data, Information, Knowledge, Wisdom (DIKW) Pyramid

What is Machine Learning: the Role

Machine Learning plays an important role in this pyramid



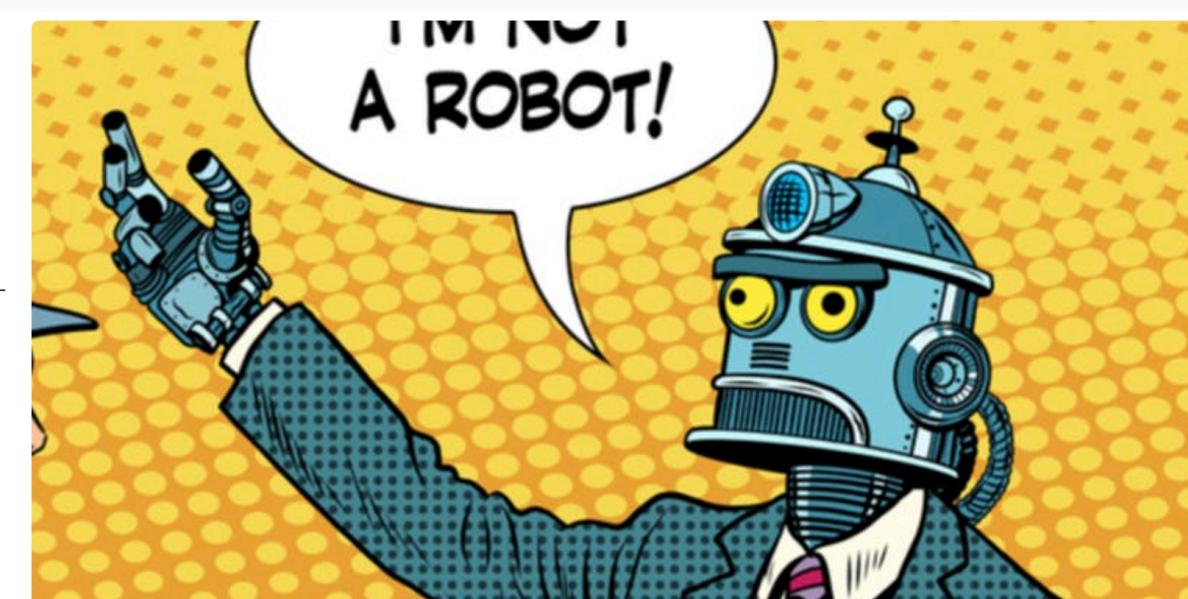
TECHNOLOGY
Welcome to the Next Level of Bullshit

The language algorithm GPT-3 continues our descent into a post-truth world.

BY RAPHAËL MILLIÈRE September 9, 2020



Share



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



Environment

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

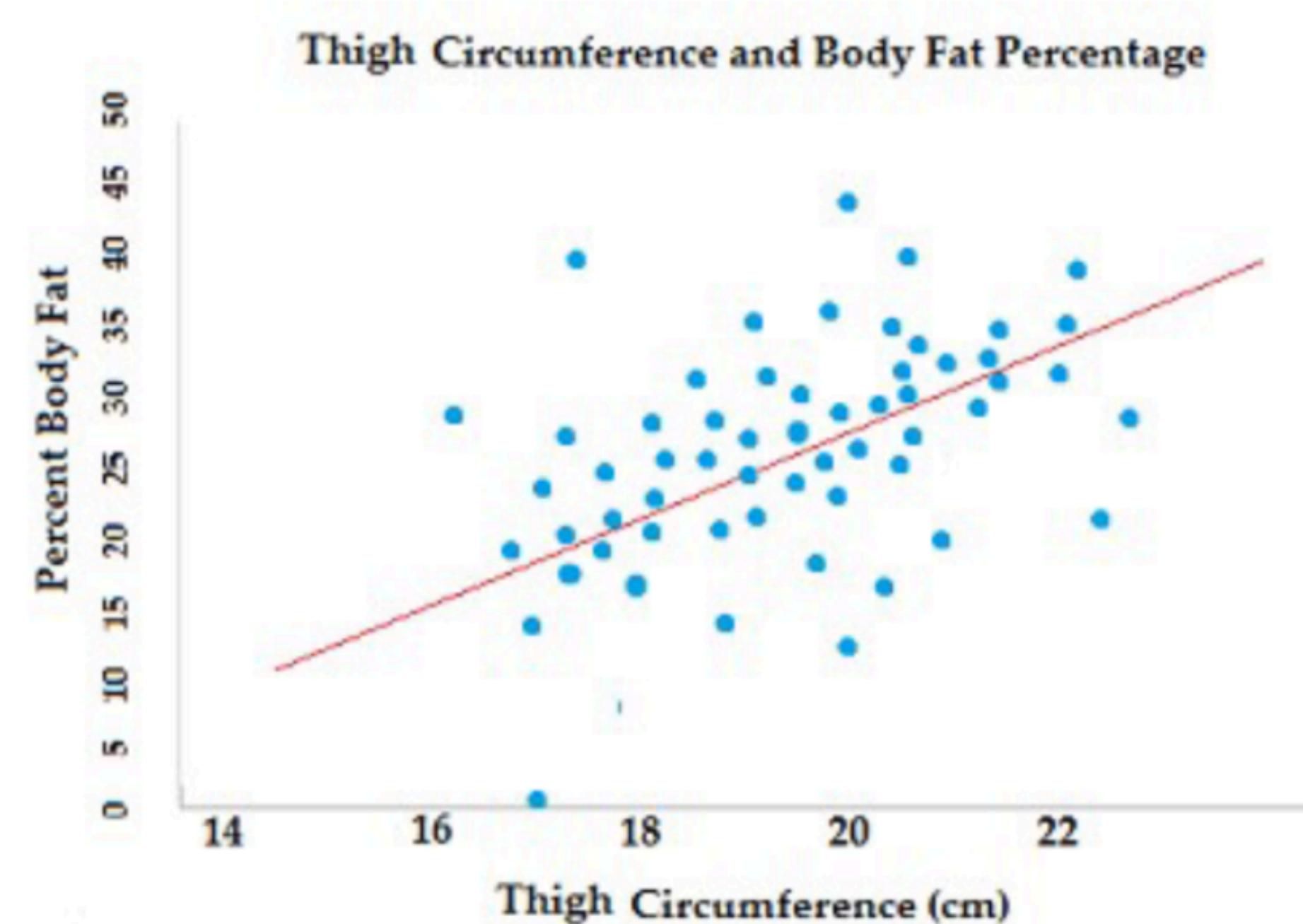


Figure from lecture [Linear models I](#)

Supervised Learning - Regression

Goal: Predict real valued labels

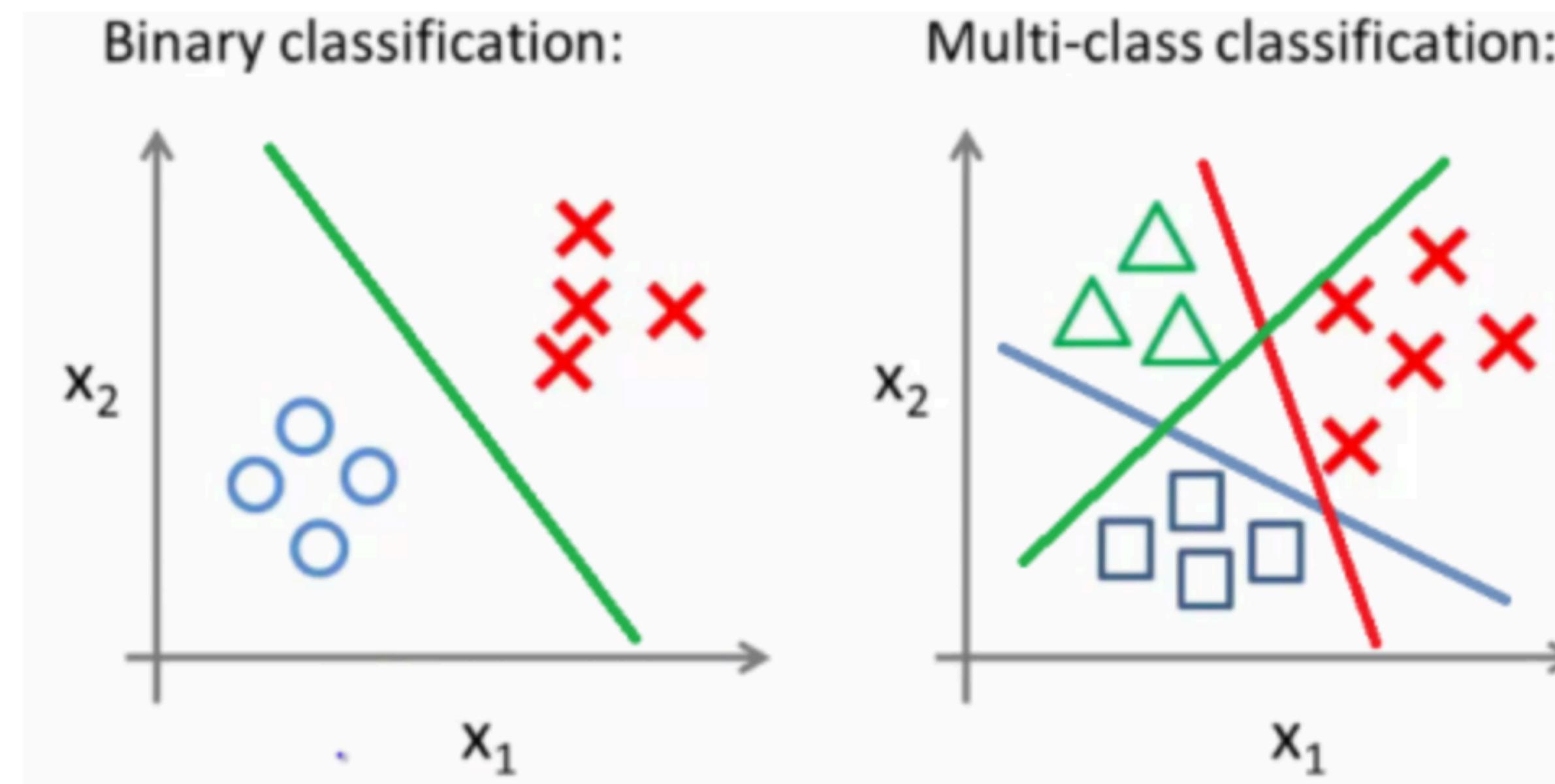
Summary

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>

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



Supervised Learning - Classification

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

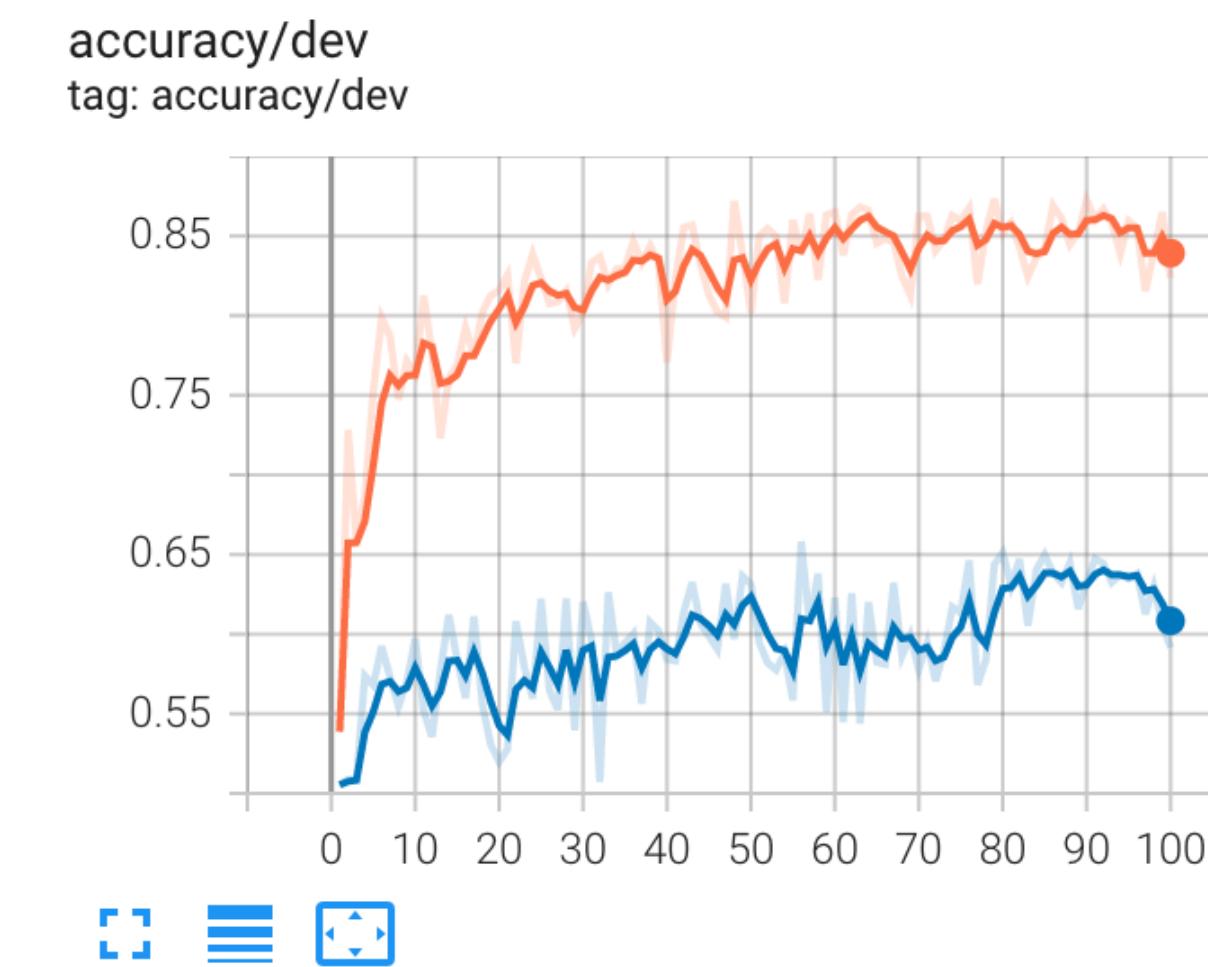
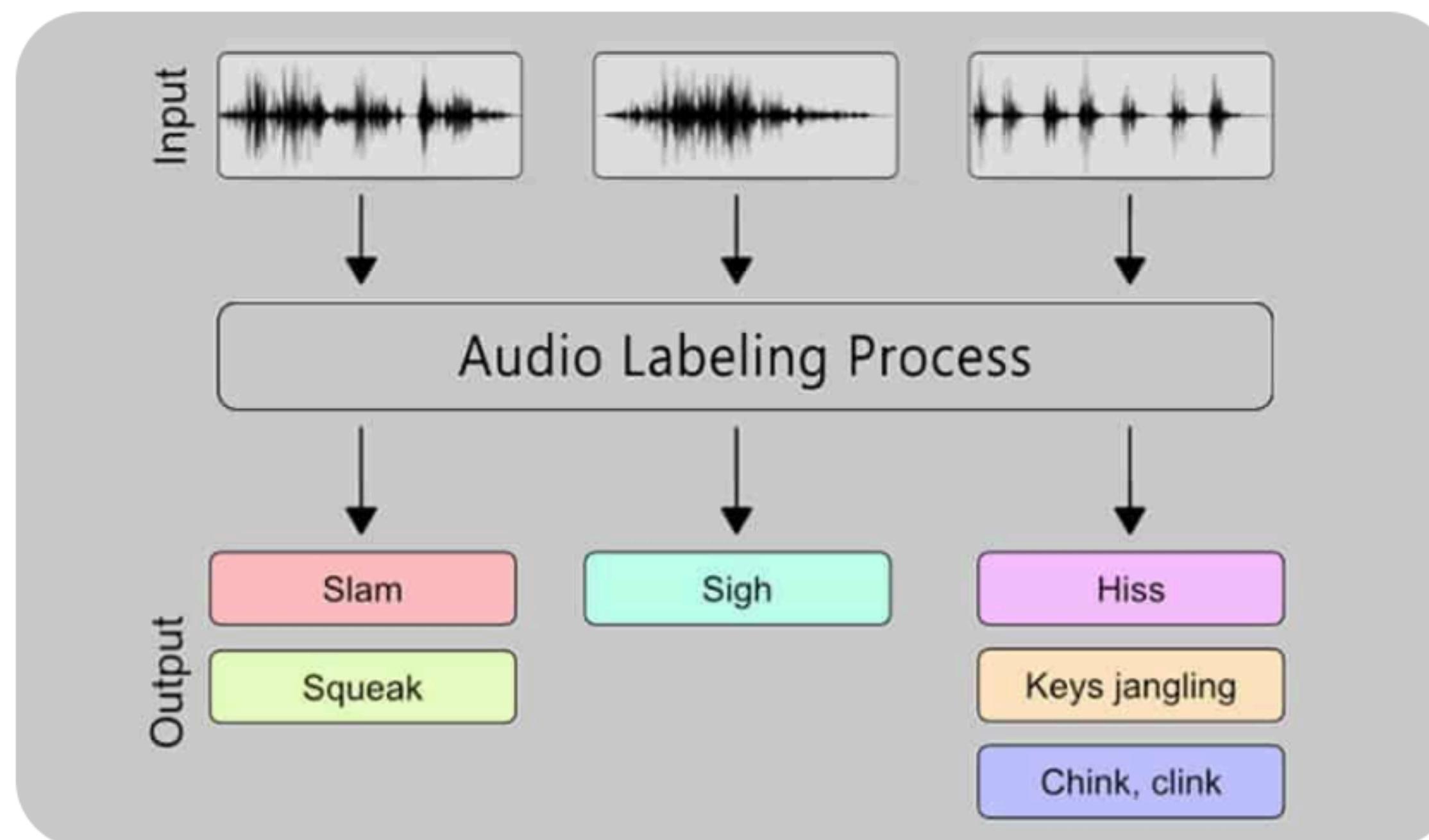


<https://www.freecodecamp.org/news/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d/>

Supervised Learning - Classification

Example: extract and classify monkey calls from raw audio

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

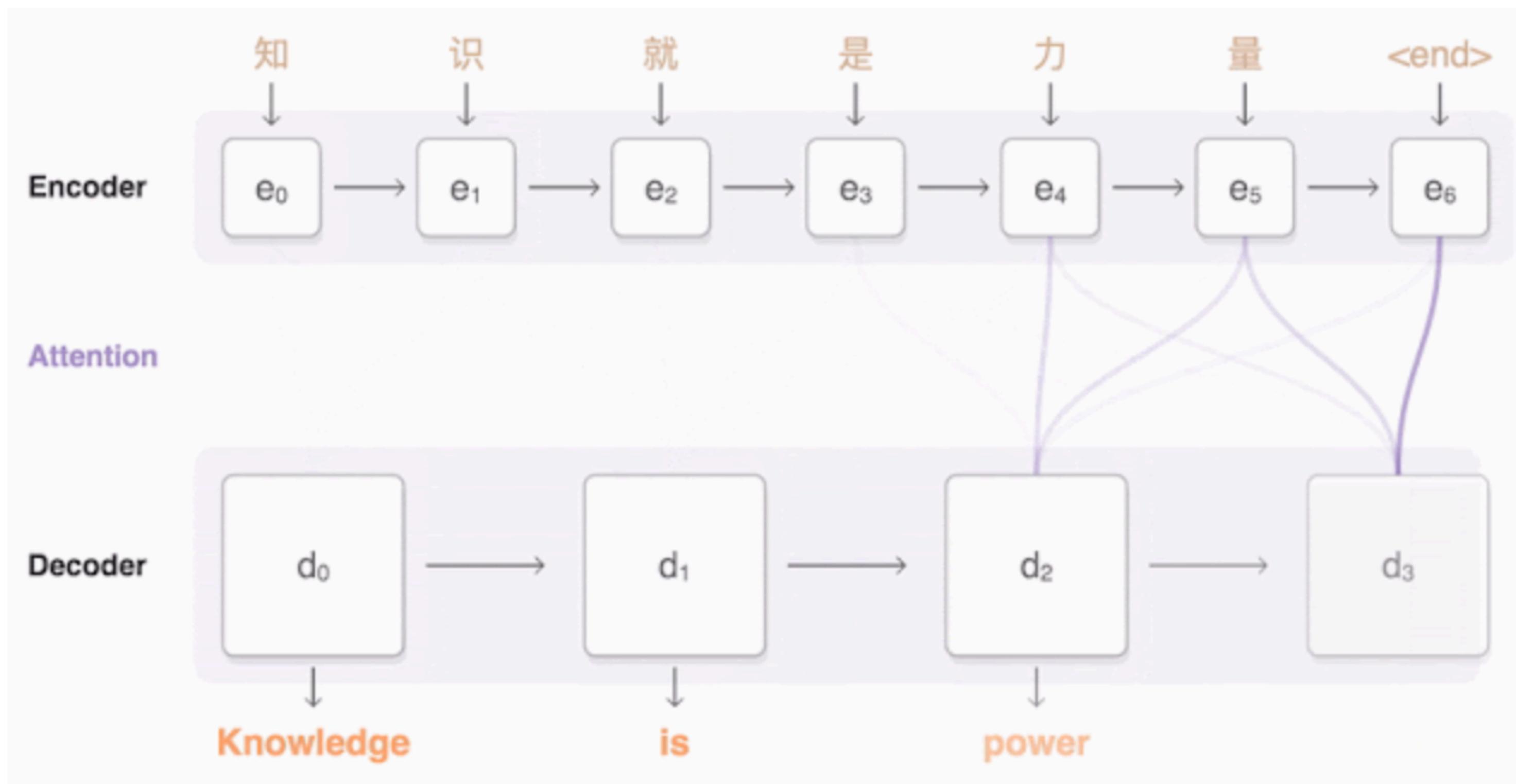


Statistics from our research

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

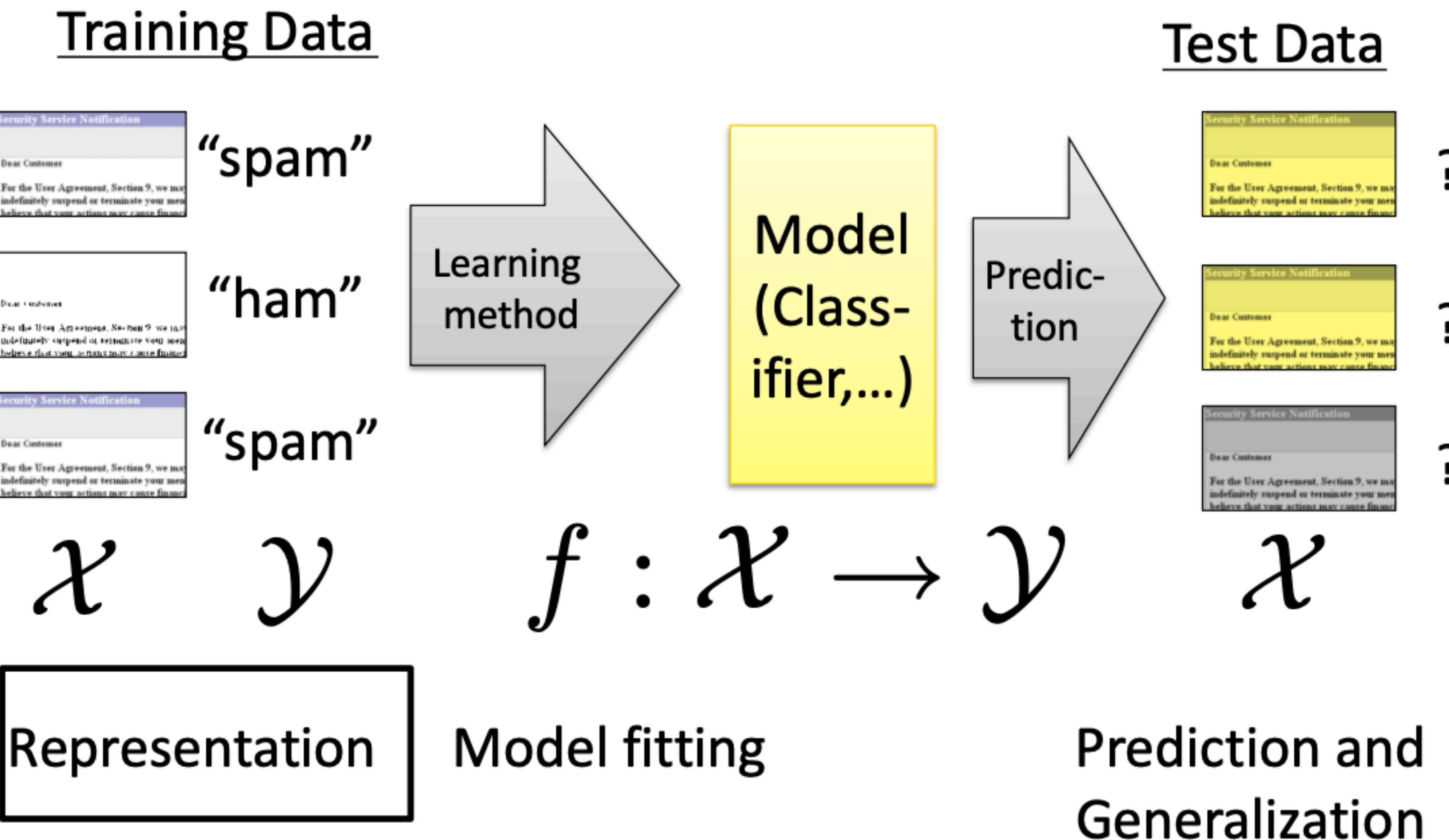


Figure from course <https://las.inf.ethz.ch/teaching/introml-s20>

Data Representation

How to represent data in a machine-understandable fashion?



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]



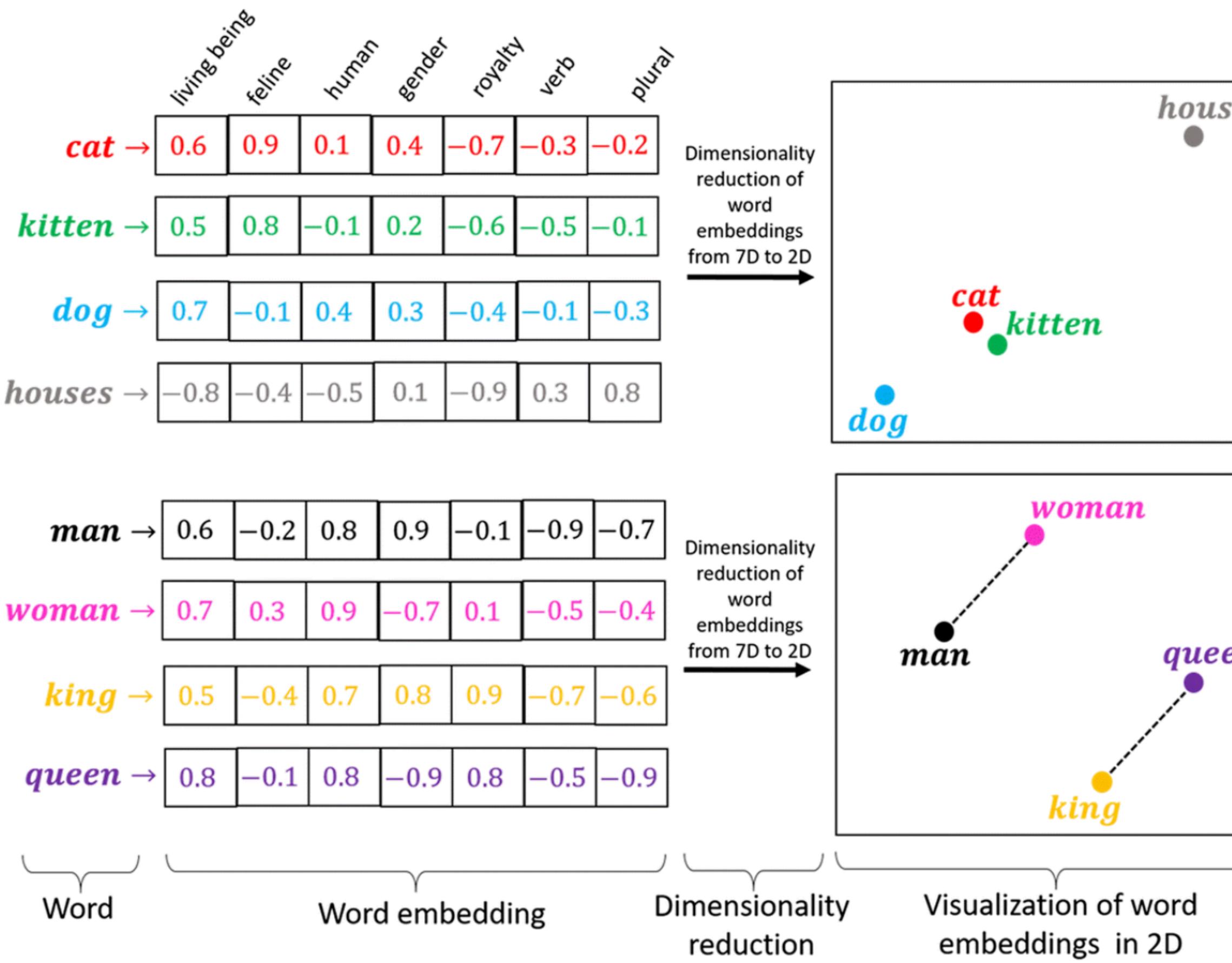
[.3 .01 .1 2.3 0 0 1.1 ...]

Vectors!

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

Data Representation - Text

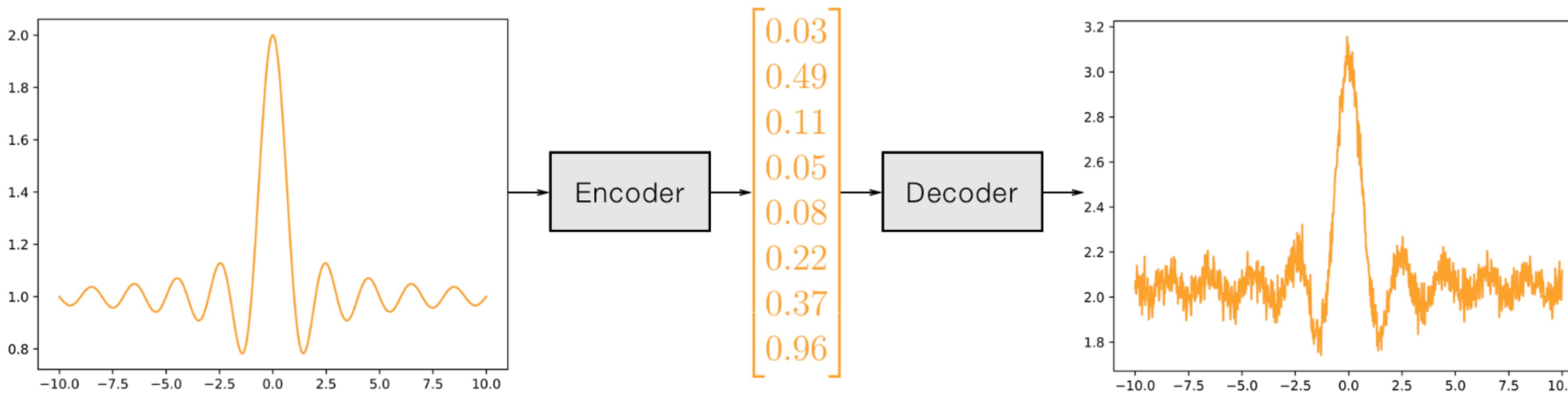
How to represent data in a machine-understandable fashion?



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

Data Representation - Audio

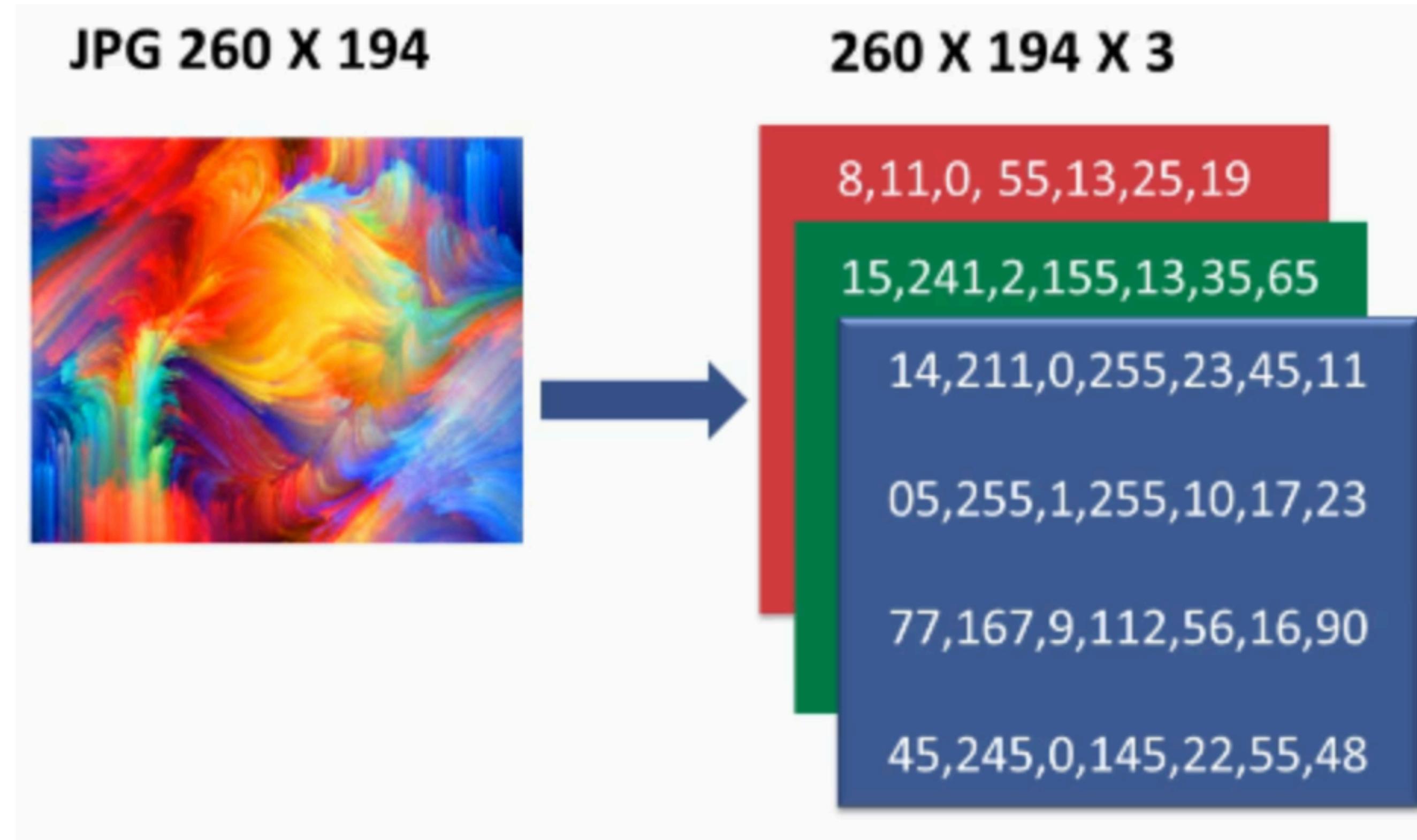
How to represent data in a machine-understandable fashion?



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

Data Representation - Image

How to represent data in a machine-understandable fashion?



Unsupervised Learning - Pipeline / Paradigm

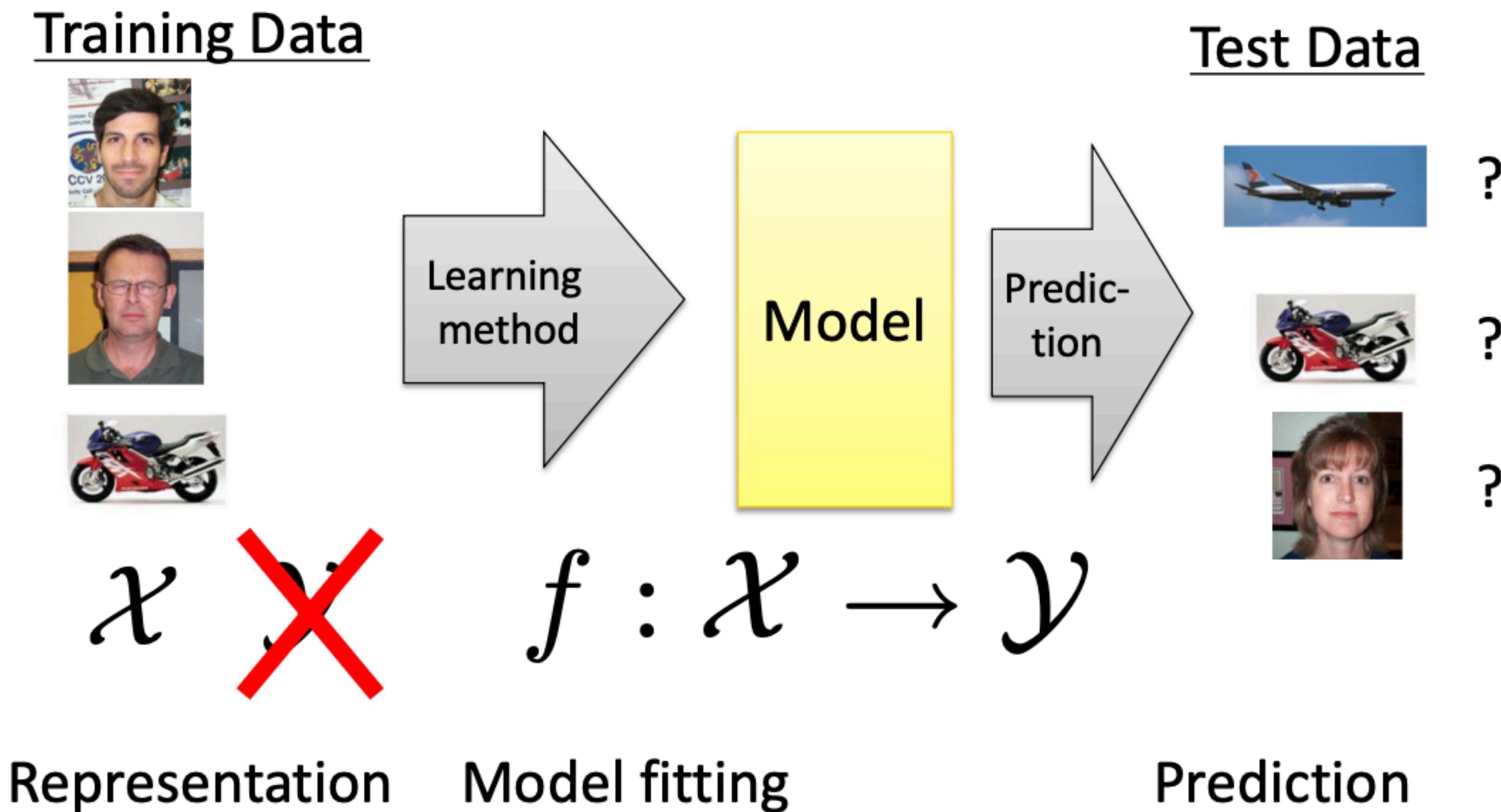


Figure from course <https://las.inf.ethz.ch/teaching/introml-s20>

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)

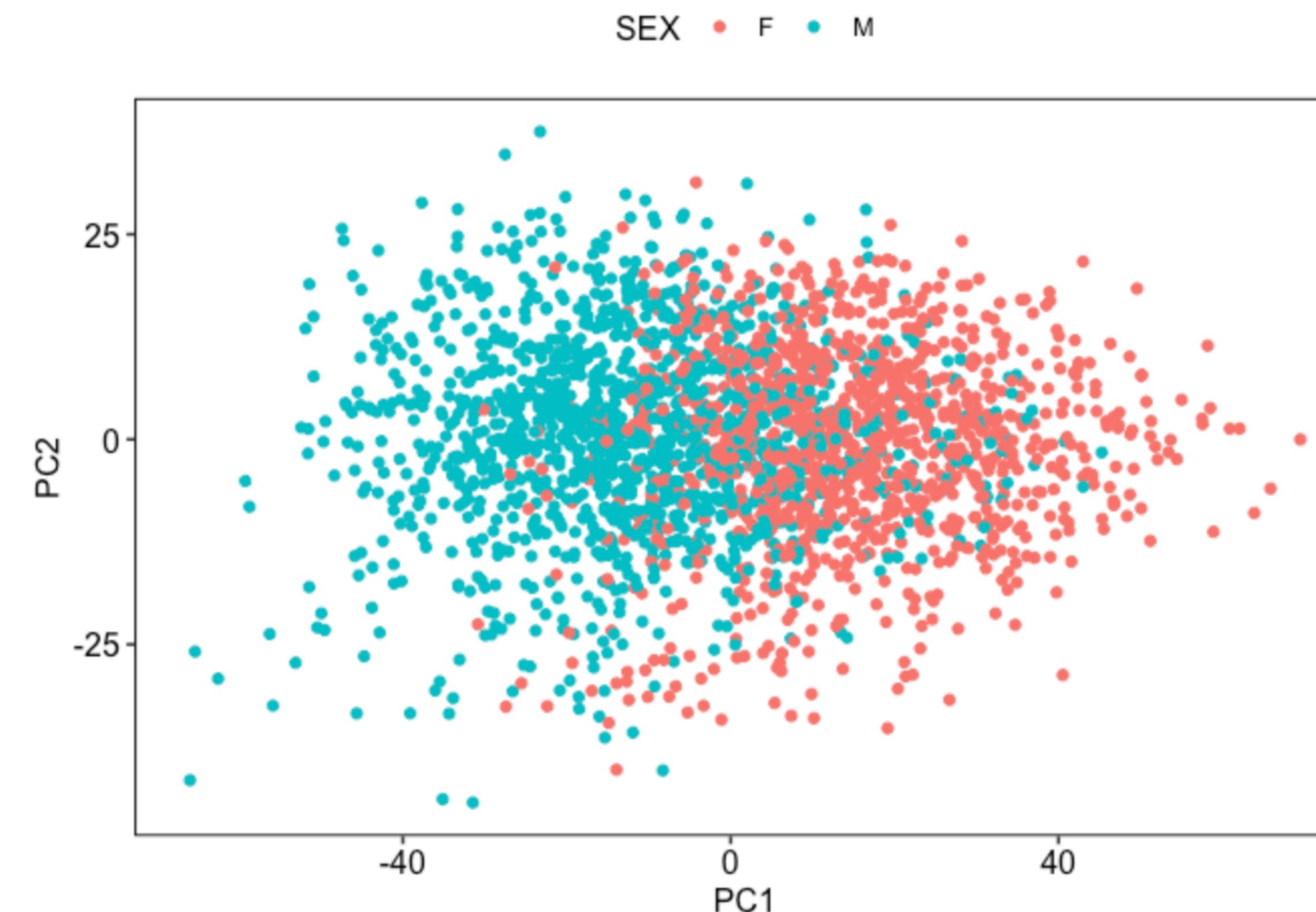
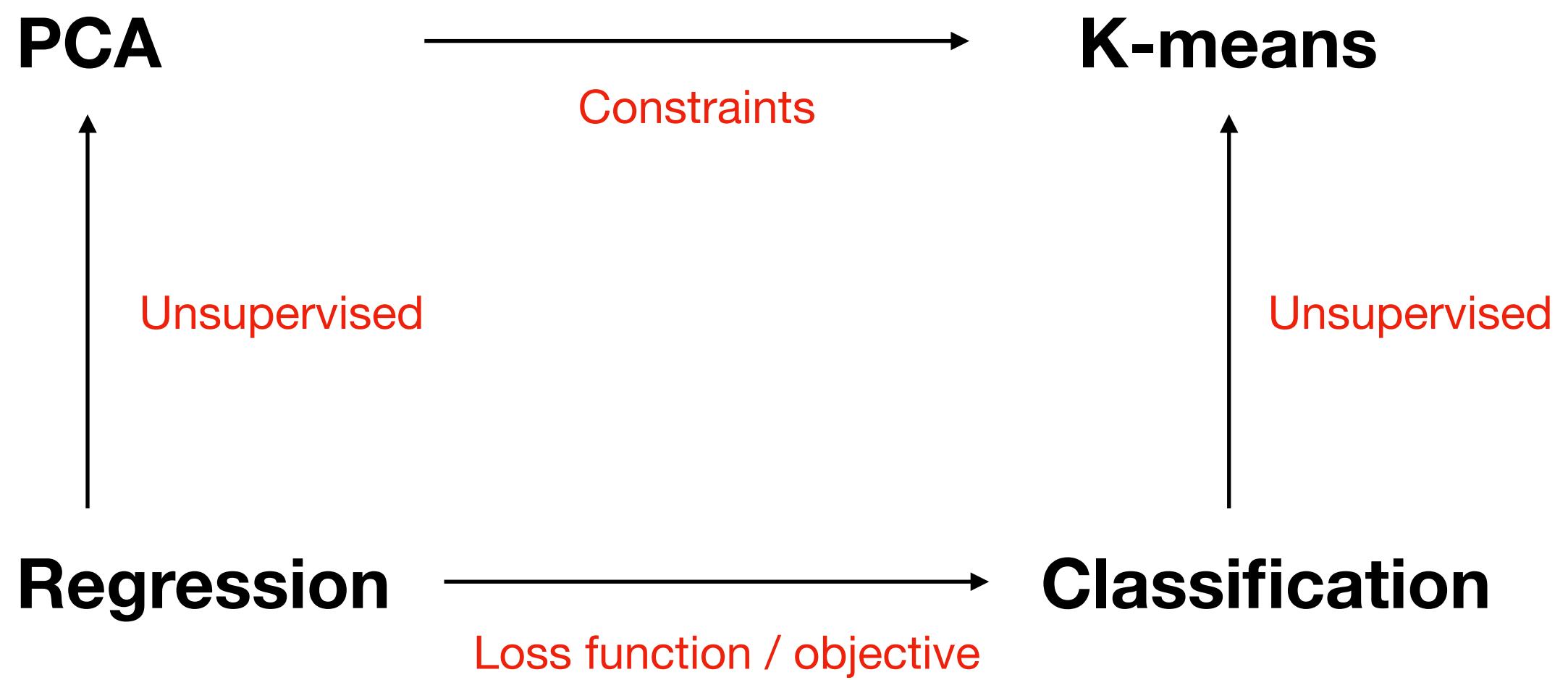


Figure from lecture Dimensionality reduction and clustering I

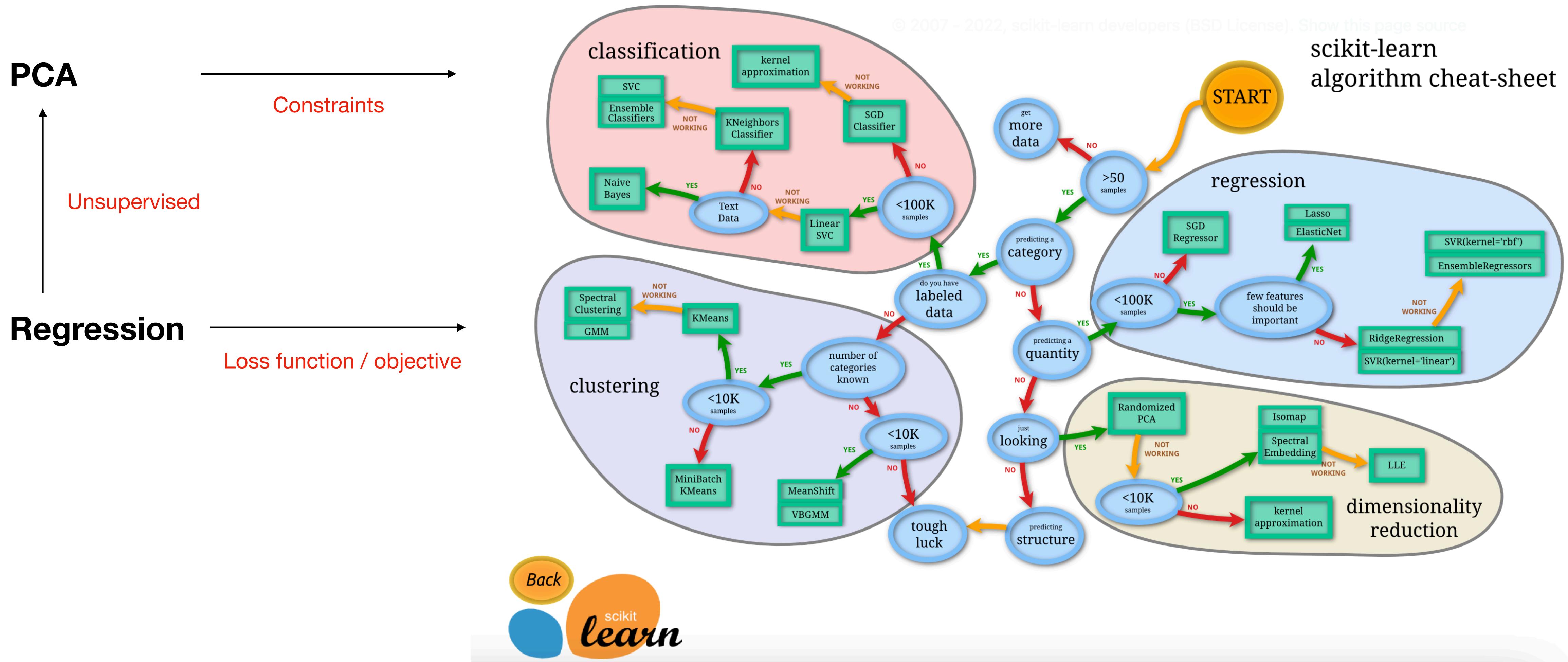
Unsupervised Learning - Comparison

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



Unsupervised Learning - Comparison

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



Unsupervised Learning - Autoencoders

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

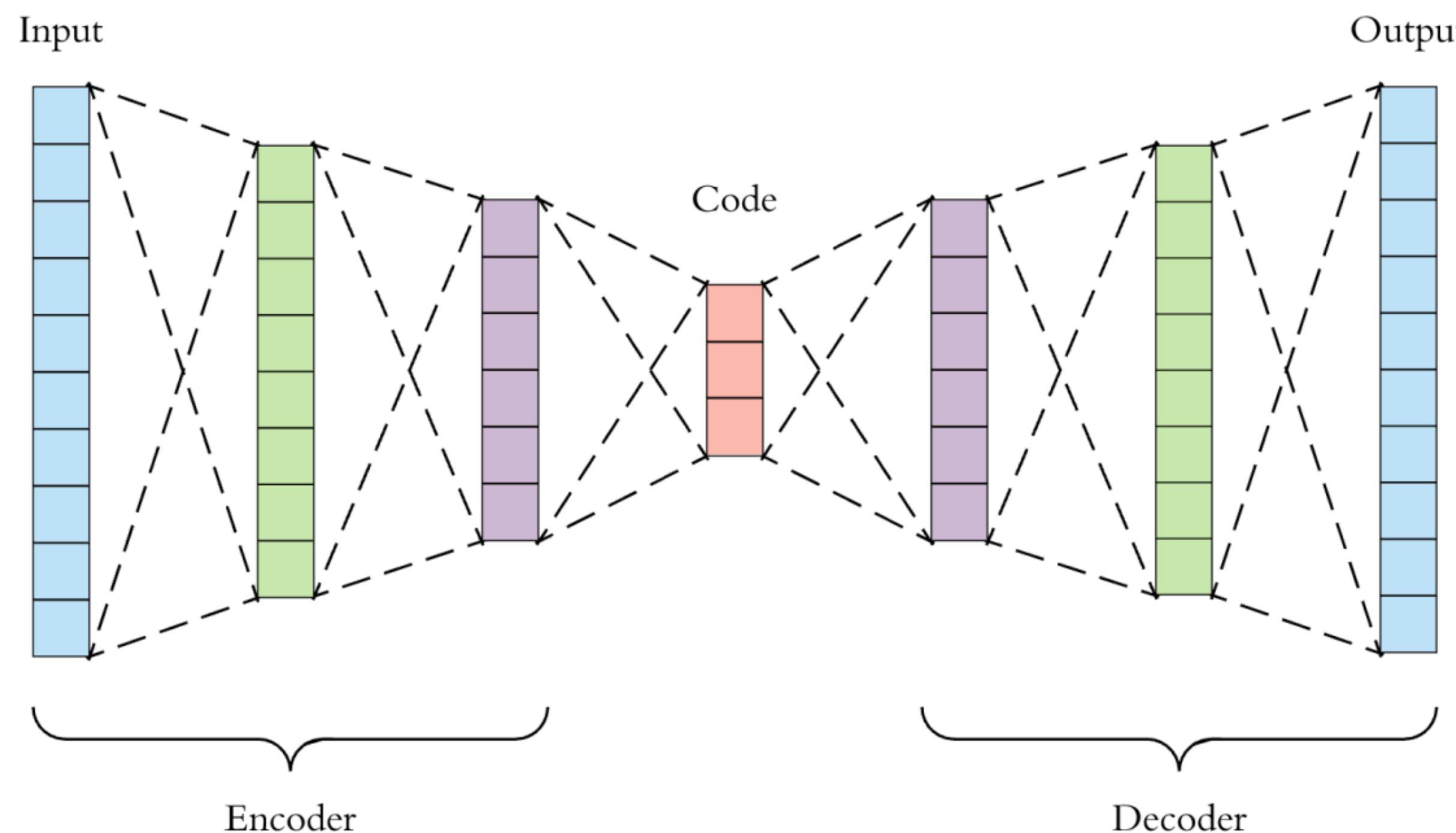
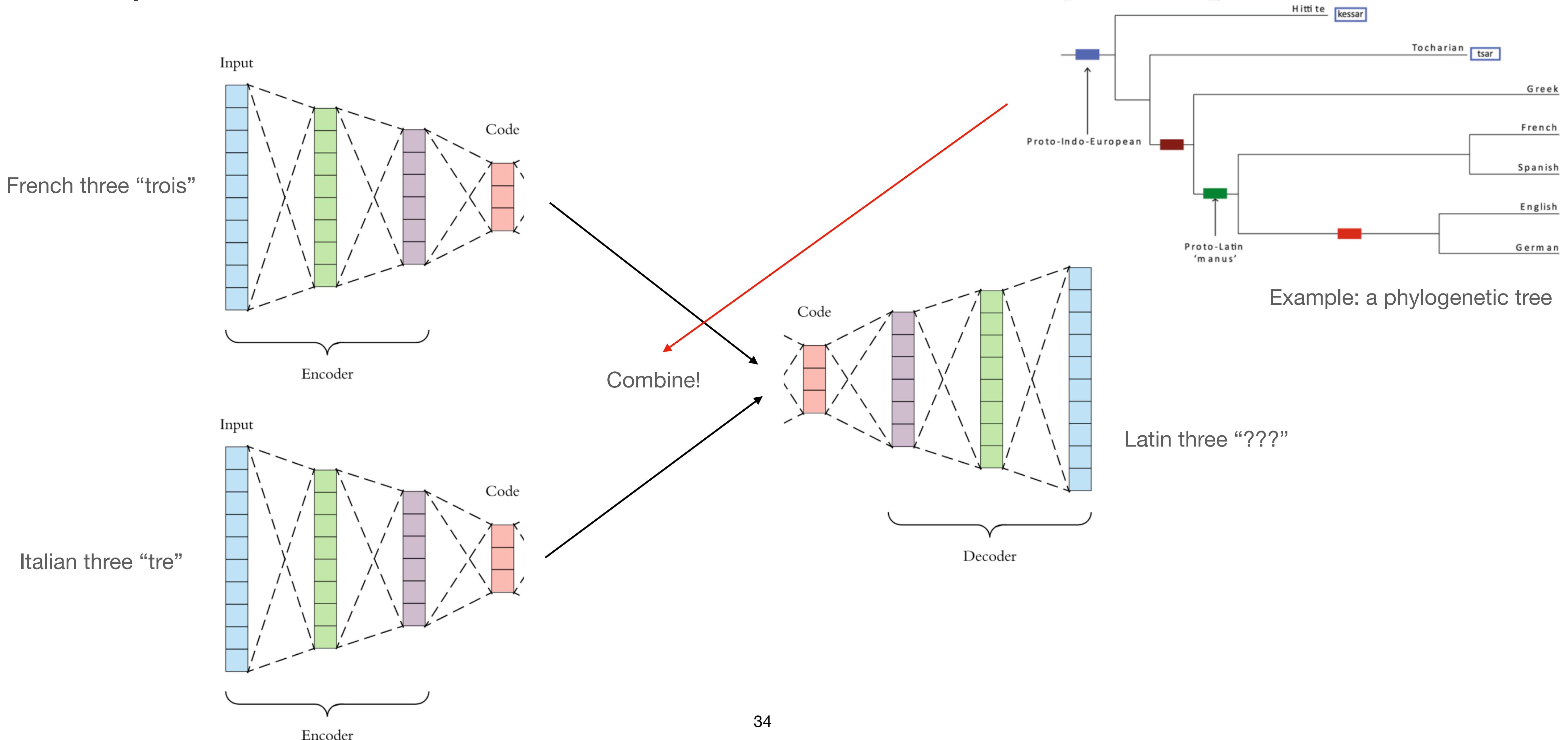


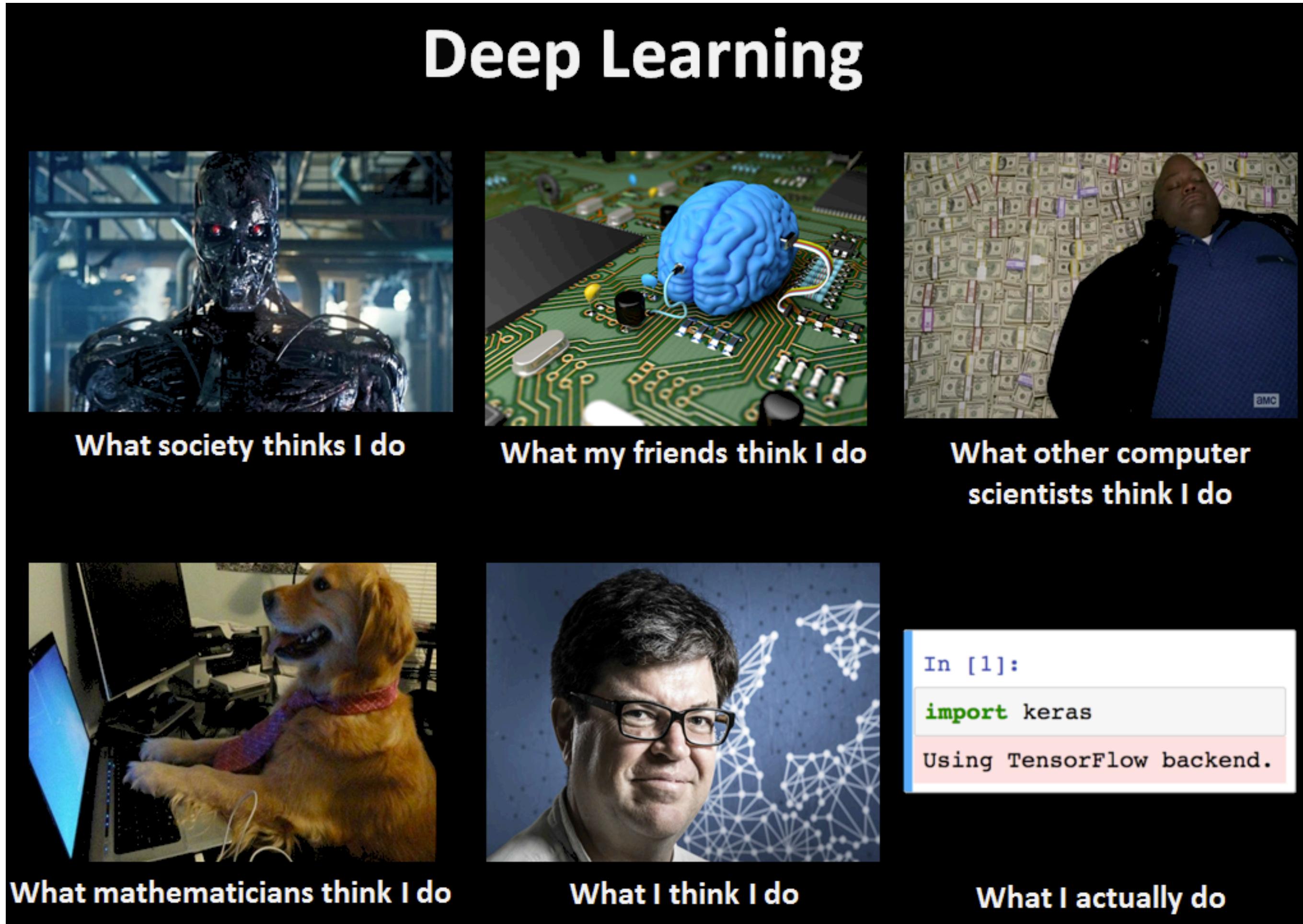
Figure from course <http://www.da.inf.ethz.ch/teaching/2022/CIL/>

Unsupervised Learning - Autoencoders

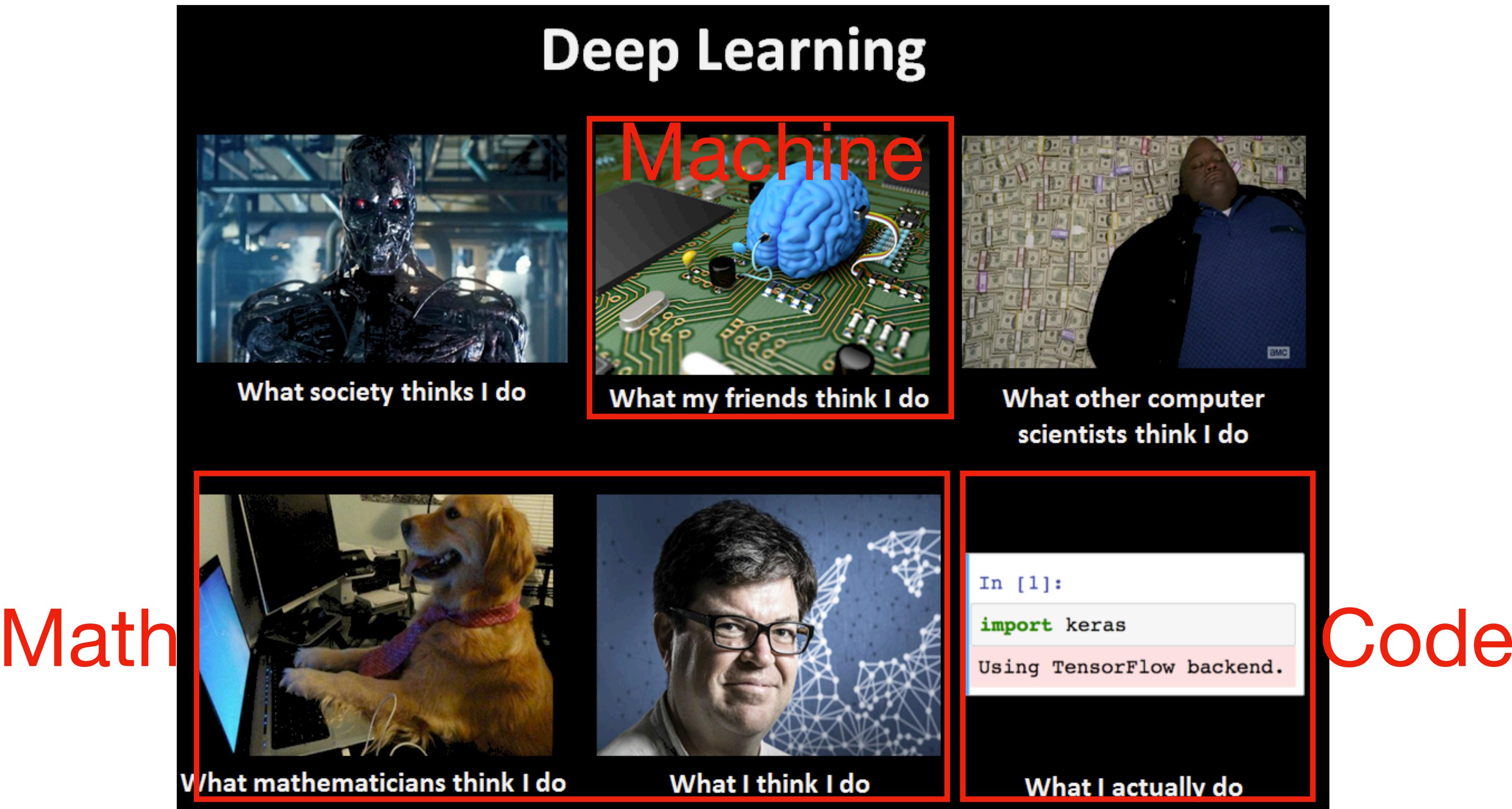
Example: reconstruct word sound from thousands of years ago ...



Deep Learning

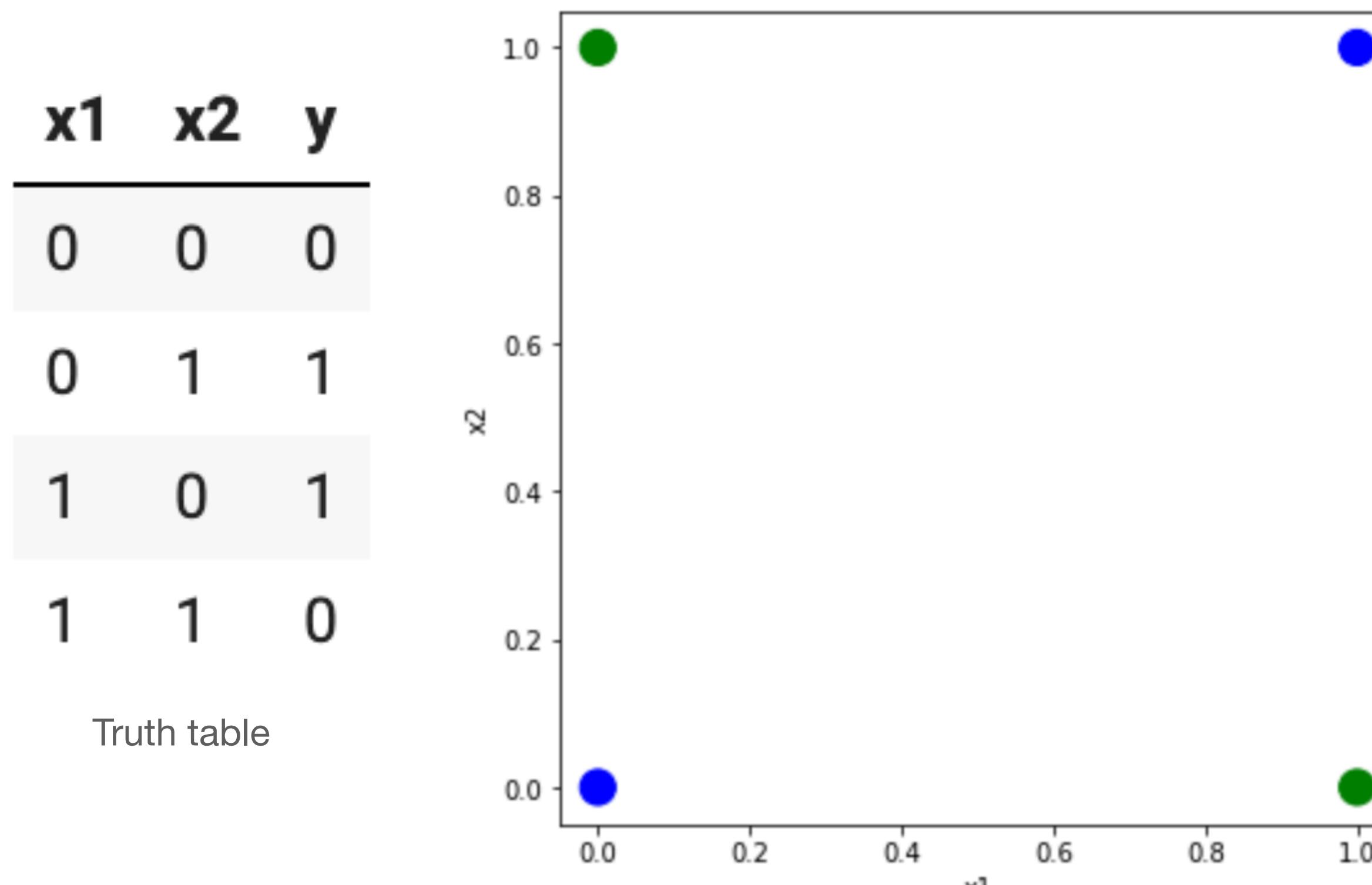


Deep Learning



Deep Learning - Power of Nonlinearity

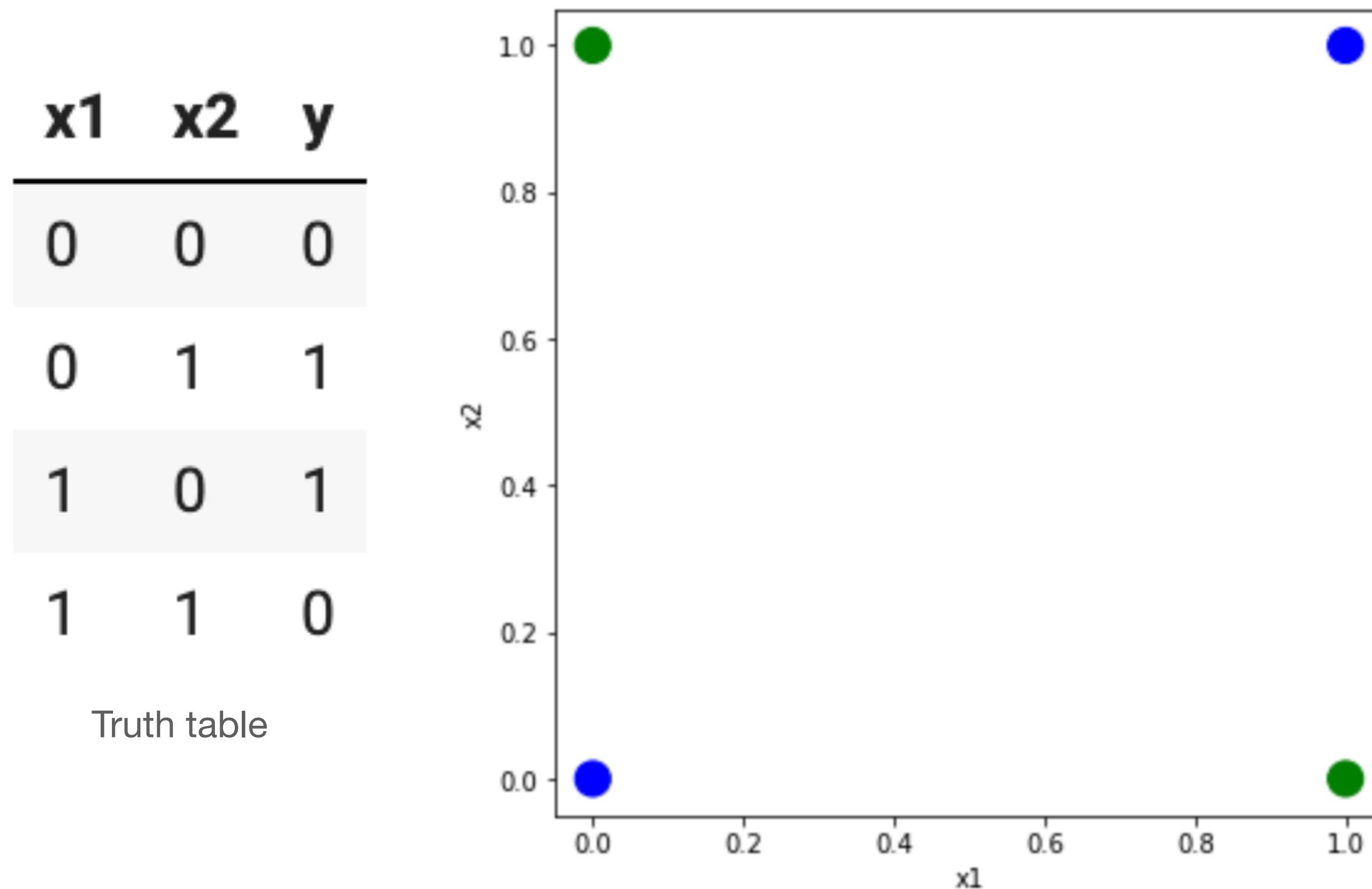
Example: XOR problem



Visualization - how to draw a decision line?

Deep Learning - Power of Nonlinearity

Example: XOR problem

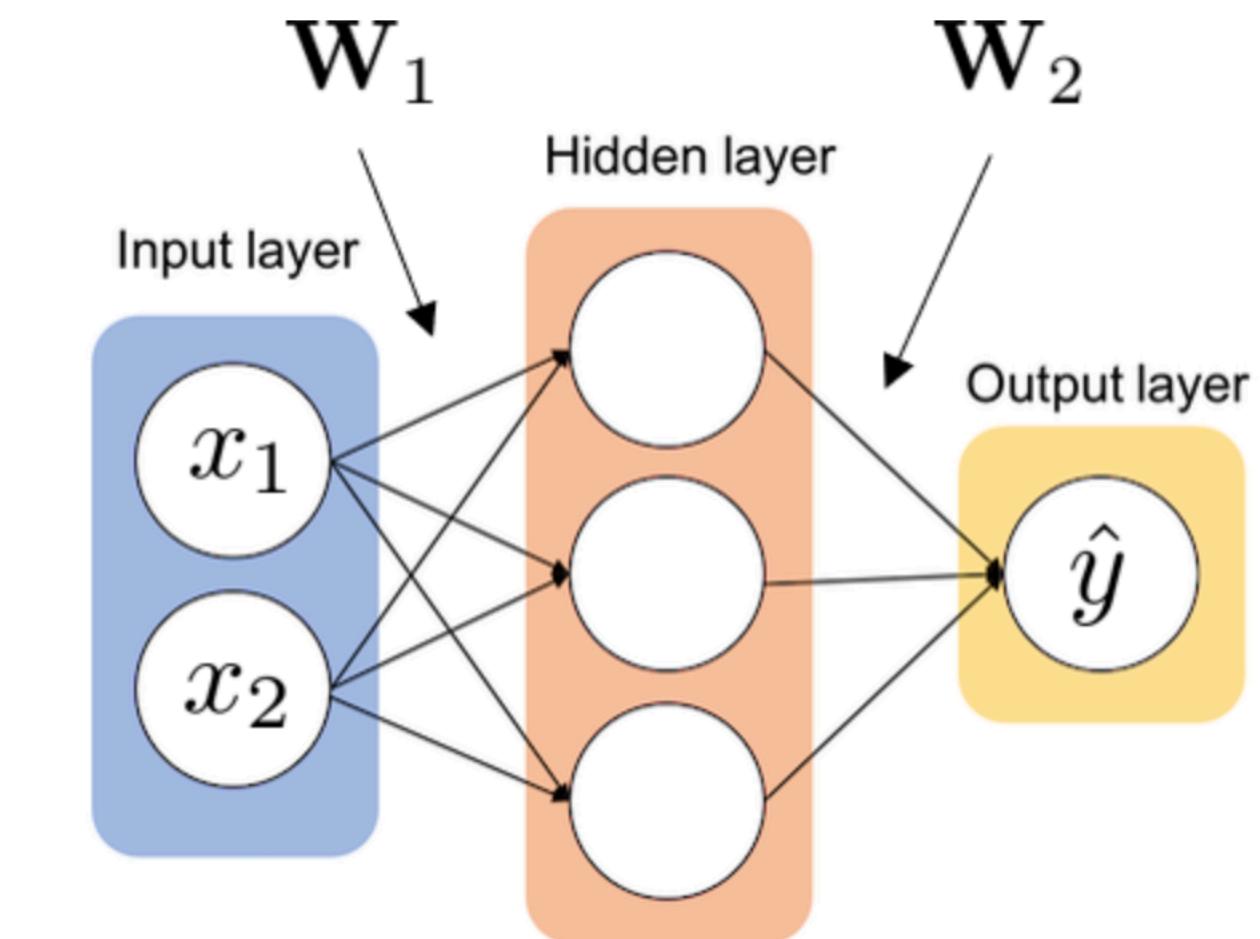
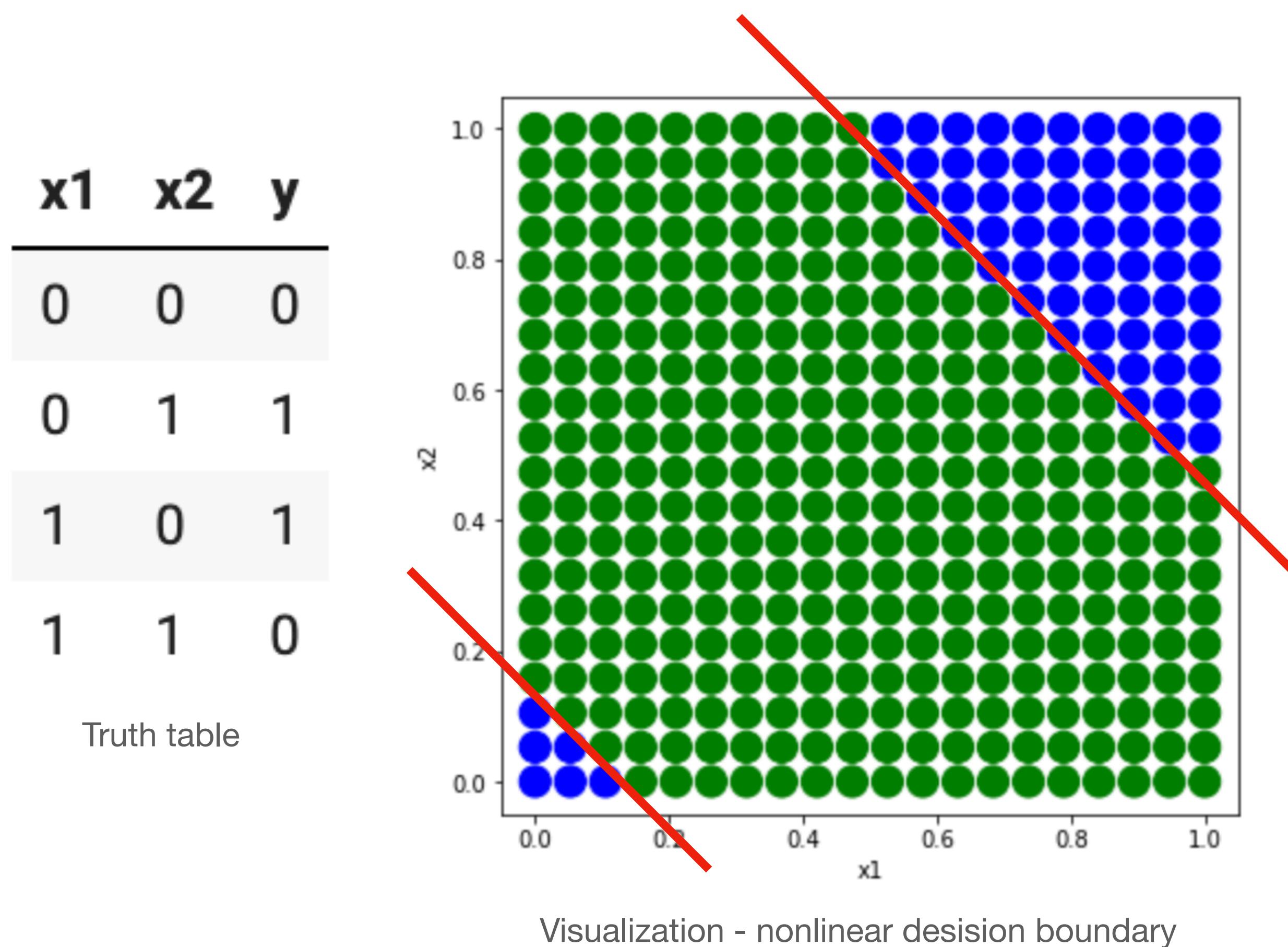


Visualization - how to draw a decision line?

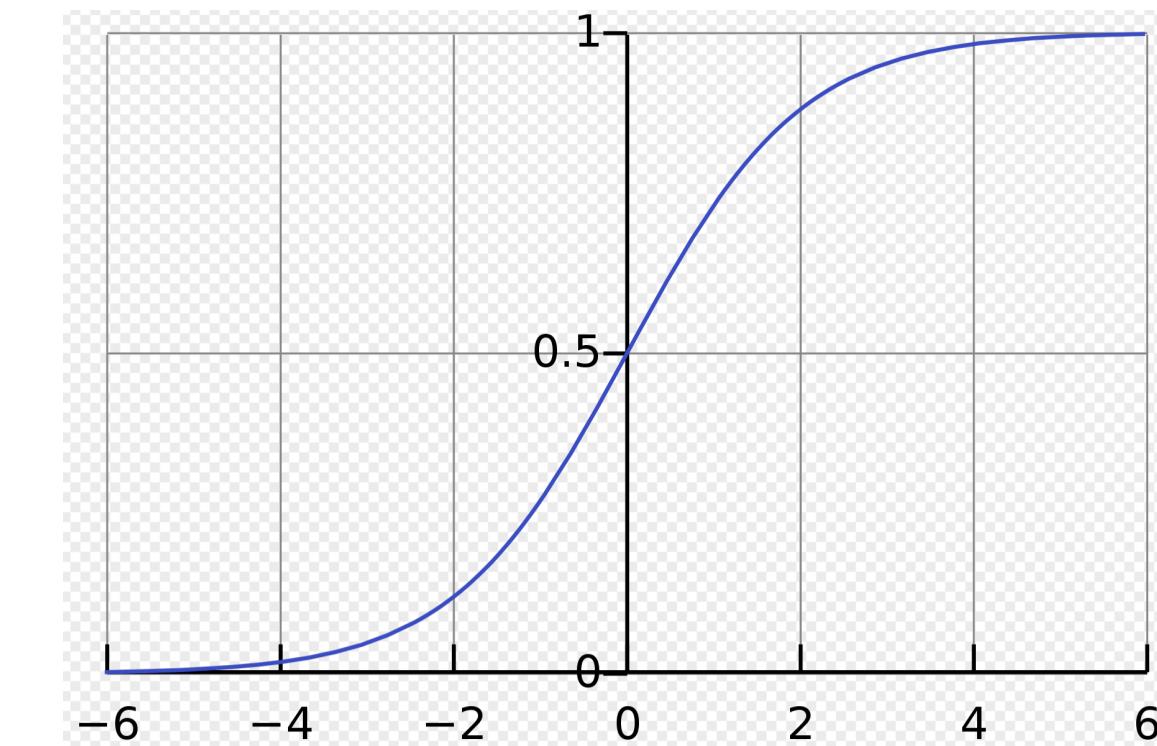


Deep Learning - Power of Nonlinearity

Example: XOR problem



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

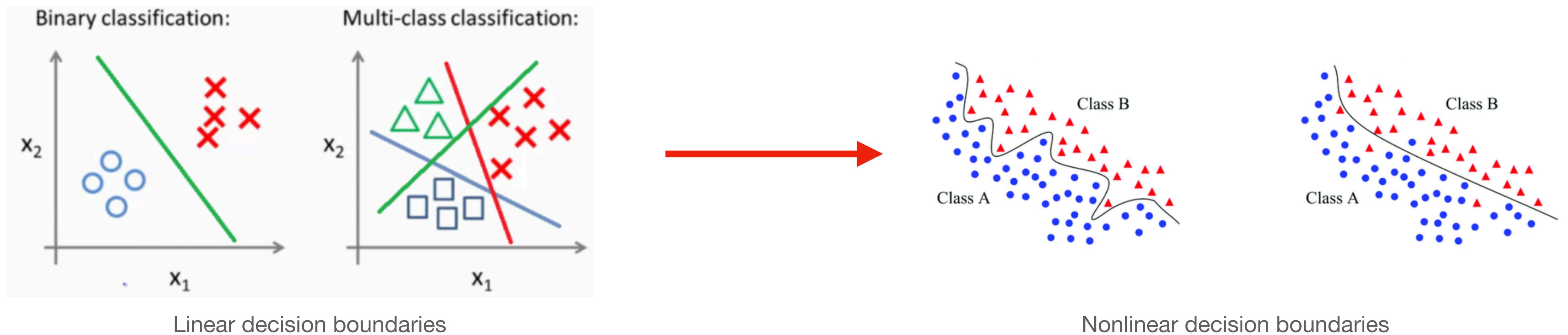


$$S(x) = \frac{1}{1 + e^{-x}}$$

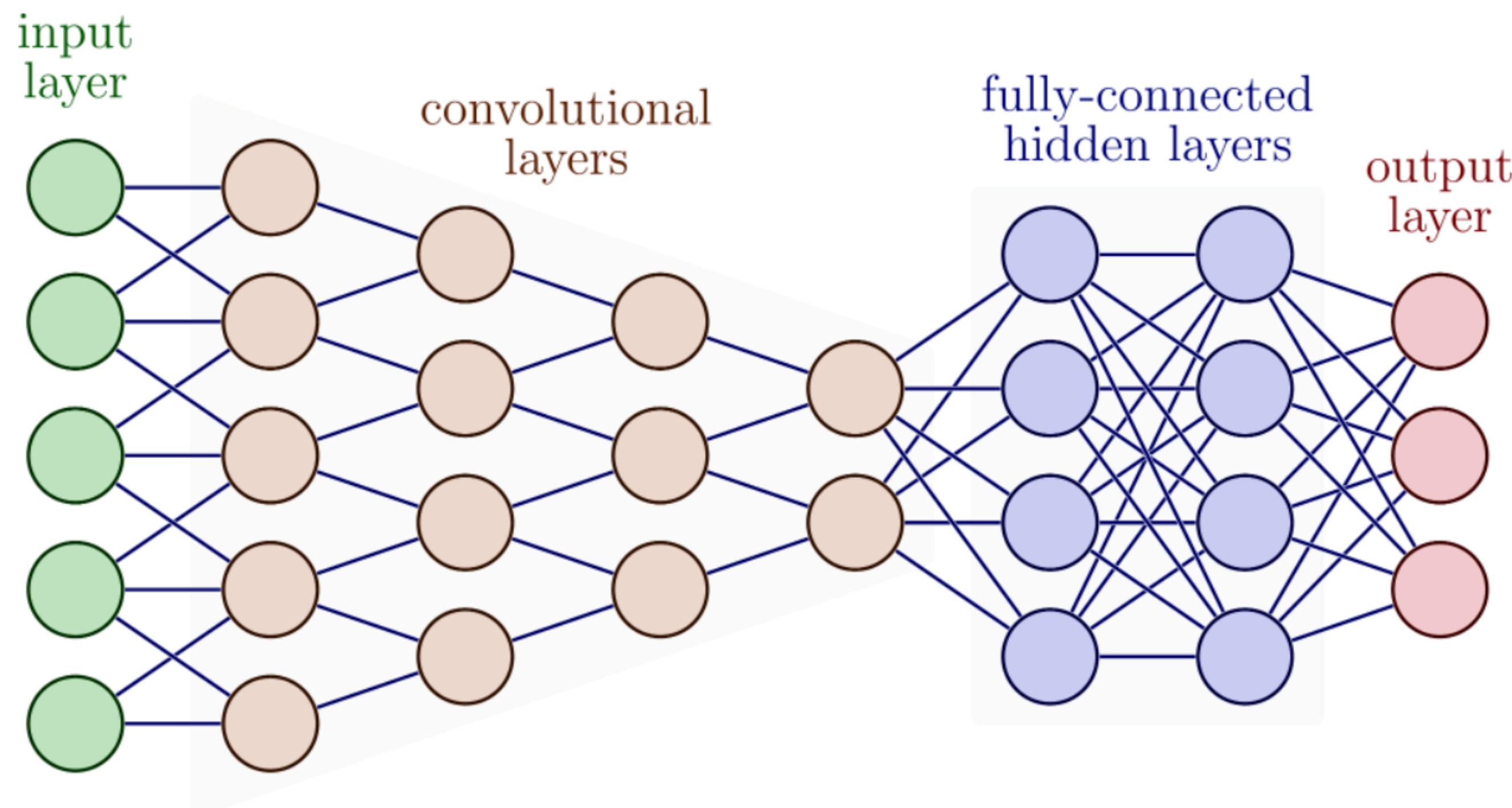
Solution: g is a sigmoid activation function

Deep Learning - Power of Nonlinearity

From linear to nonlinear decision boundaries ...



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



Write With Transformer

Get a modern neural network to
auto-complete your thoughts.

This web app, built by the Hugging Face team, is the official demo of the
`🤗/transformers` repository's text generation capabilities.

 Star 62,093

<https://transformer.huggingface.co/>

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