Course Overview

Trần Trung Kiên
ttkien@fit.hcmus.edu.vn

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Machine learning

An image → **f** → **Cat?**

Not machine learning

A number → **f** → **Prime?**

We KNOW f

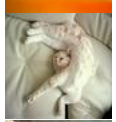
Do you think you can write down a function to recognize cats?

Image source: Fei-Fei Li's TED talk

Machine learning

An image → **f** → **Cat?**

We do NOT KNOW f,
but we have DATA
which can be used to approximate f

Input	Output
	Cat
	Cat
	Cat
	Not cat
	Not cat
⋮	⋮

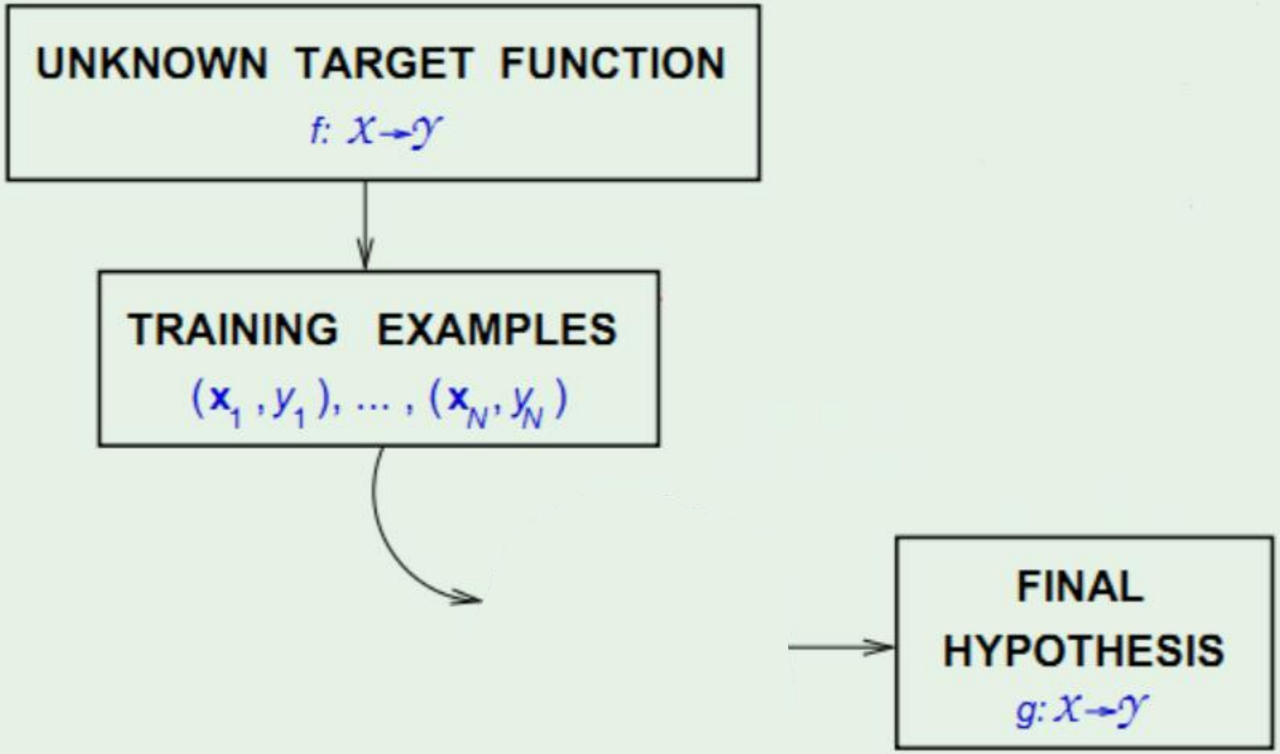
Not machine learning

A number → **f** → **Prime?**

We KNOW f

Conditions to apply machine learning

- A pattern exists
- We cannot pin it down mathematically
- We have data on it



We want: $g \approx f$

Image source: Yaser S. Abu-Mostafa, Learning from Data

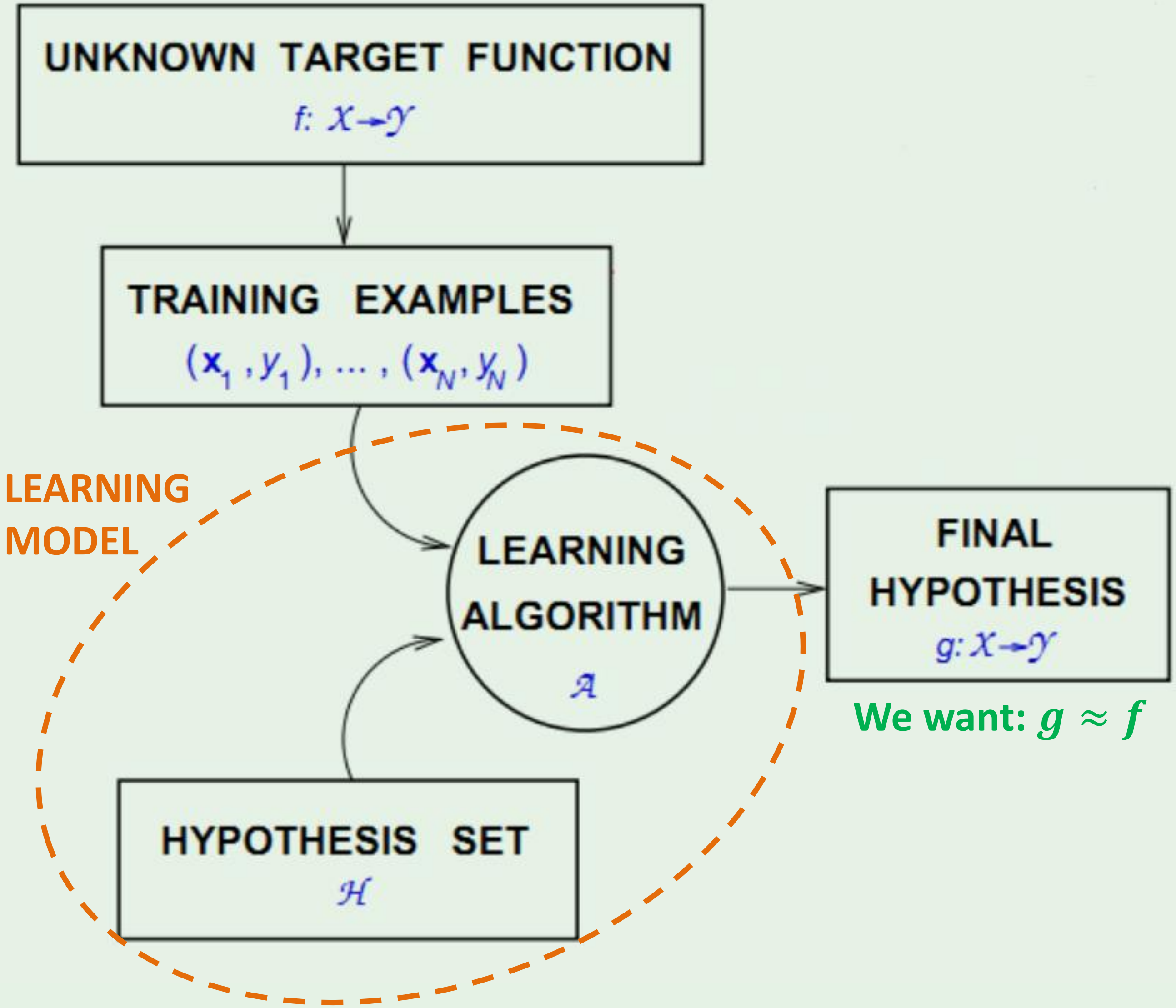


Image source: Yaser S. Abu-Mostafa, Learning from Data

Examples: machine learning or not?

- Speech recognition
- Determining the time it would take a falling object to hit the ground
- Predicting whether you will pass this course or not

ML can be used in a lot of fields

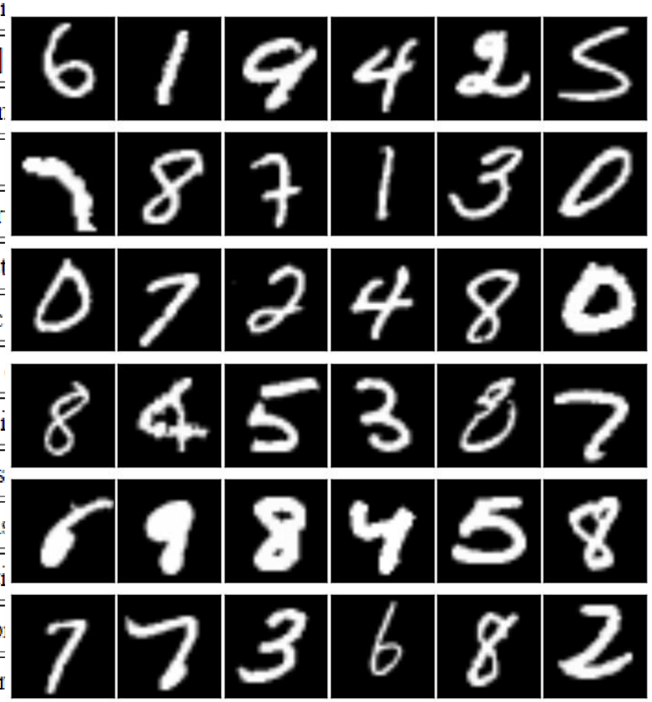
- Computer vision
- NLP
- Medicine
- Economics
- ...

Handwritten digit recognition

yann.lecun.com/exdb/mnist/

Convolutional net LeNet-4	none	1.1
Convolutional net LeNet-4 with K-NN instead of last layer	none	1.1
Convolutional net LeNet-4 with local learni		1.1
Convolutional net LeNet-5, [no distortions]		0.95
Convolutional net LeNet-5, [huge distortio		0.85
Convolutional net LeNet-5, [distortions]		0.8
Convolutional net Boosted LeNet-4, [distor		0.7
Trainable feature extractor + SVMs [no dist		0.83
Trainable feature extractor + SVMs [elastic		0.56
Trainable feature extractor + SVMs [affine		0.54
unsupervised sparse features + SVM, [no di		0.59
Convolutional net, cross-entropy [affine dis		0.6
Convolutional net, cross-entropy [elastic di		0.4
large conv. net, random features [no distorti		0.89
large conv. net, unsup features [no distortio		0.62
large conv. net, unsup pretraining [no distor		0.60
large conv. net, unsup pretraining [elastic di		0.39
large conv. net, unsup pretraining [no distortions]	none	0.53
large/deep conv. net, 1-20-40-60-80-100-120-120-10 [elastic distortions]	none	0.35
committee of 7 conv. net, 1-20-P-40-P-150-10 [elastic distortions]	width normalization	27 +-0.02
committee of 35 conv. net, 1-20-P-40-P-150-10 [elastic distortions]	width normalization	0.23

MNIST Samples



Better than human ☺

The ILSVRC-2012 competition on ImageNet

- The dataset has 1.2 million high-resolution training images.
- The classification task:
  - Get the "correct" class in your top 5 bets. There are 1000 classes.
- The localization task:
  - For each bet, put a box around the object. Your box must have at least 50% overlap with the correct box.
- Some of the best existing computer vision methods were tried on this dataset by leading computer vision groups from Oxford, INRIA, XRCE, …
  - Computer vision systems use complicated multi-stage systems.
  - The early stages are typically hand-tuned by optimizing a few parameters.

cheetah

bullet train

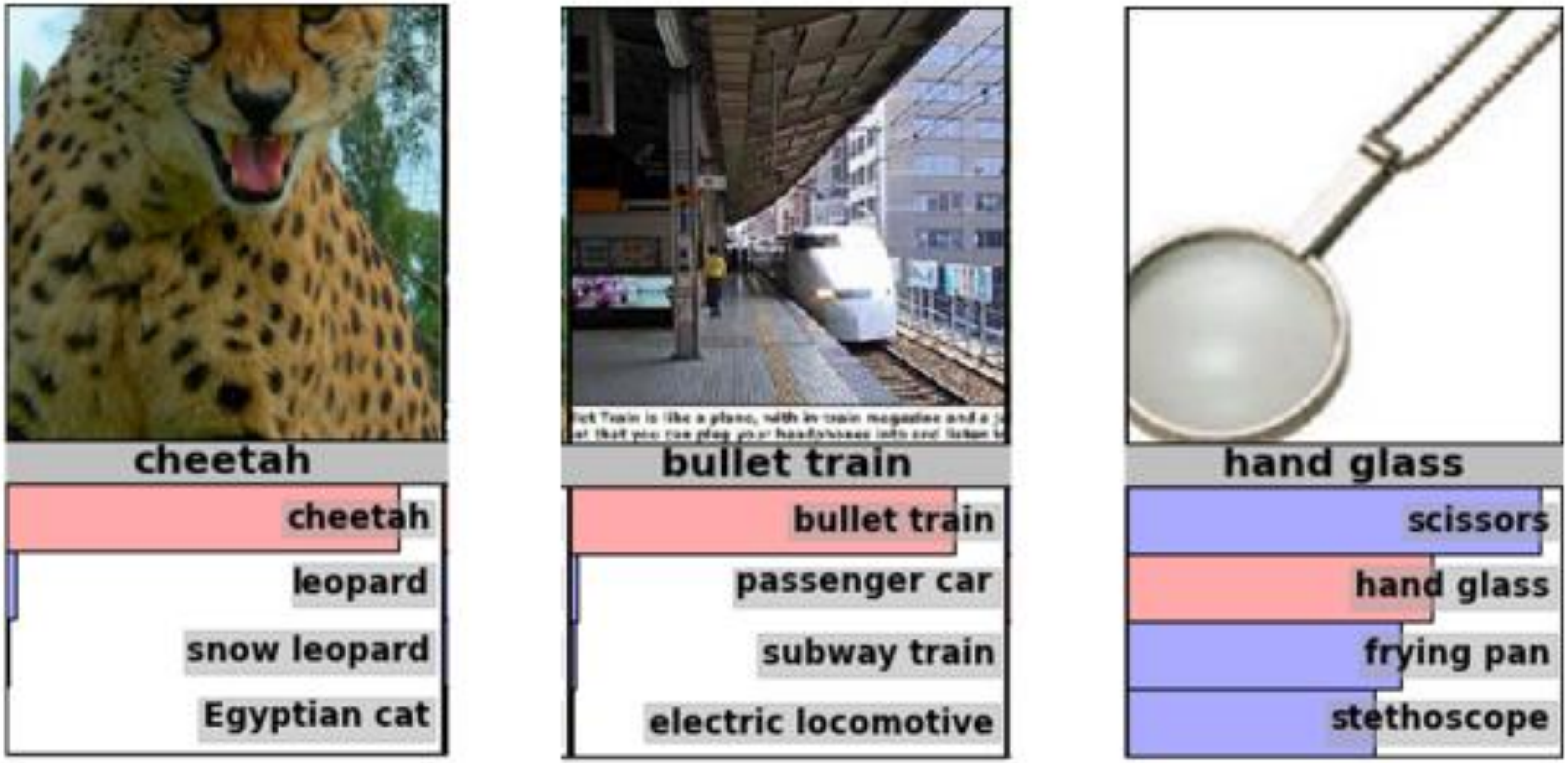
hand glass

- University of Toronto (Alex Krizhevsky) — 16.4% — 34.1%

Error rates on the ILSVRC-2012 competition

	classification	classification &localization
University of Tokyo	26.1%	53.6%
Oxford University Computer Vision Group	26.9%	50.0%
INRIA (French national research institute in CS) + XRCE (Xerox Research Center Europe)	27.0%	
University of Amsterdam	29.5%	

Examples from the test set (with the network's guesses)



Clarifai (ImageNet 2013 winner)

See how well computers can understand images

band
stage
concert
music
club
rock show

garden
kid
flower
woman
child
tree

dog
winter
snow

street
alley
old
mexico
italian

metal
key
door
old
lock
wood

climb
rock climber
mountain
sport

Source: https://www.clarifai.com/

Go beyond object recognition: image captioning

"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

"two young girls are playing with lego toy."

"girl in pink dress is jumping in air."

"black and white dog jumps over bar."

"young girl in pink shirt is swinging on swing."

Source: http://cs.stanford.edu/people/karpathy/deepimagesent/

Machine learning (ML) is hot nowadays

- Why?
  - Data get bigger and bigger
  - Computers get faster and faster
- If you want to be in academia, ML is a good place for you.
- If you want to be in industry, ML is a good place for you too.
  - In big IT companies like Google, Facebook, ..., ML is one of the most important skills.
  - In Viet Nam, recently some ML companies have emerged ☺.

Types of learning

- The type of learning we have talked about so far is called **supervised learning**

  data have the form: (input, correct output)

  It'll be the main focus in this course
- There are other types of learning out there
  - **Unsupervised learning**
    - (input)
  - **Reinforcement learning**
    - (input, **some** output, **grade for this output**)
  - …

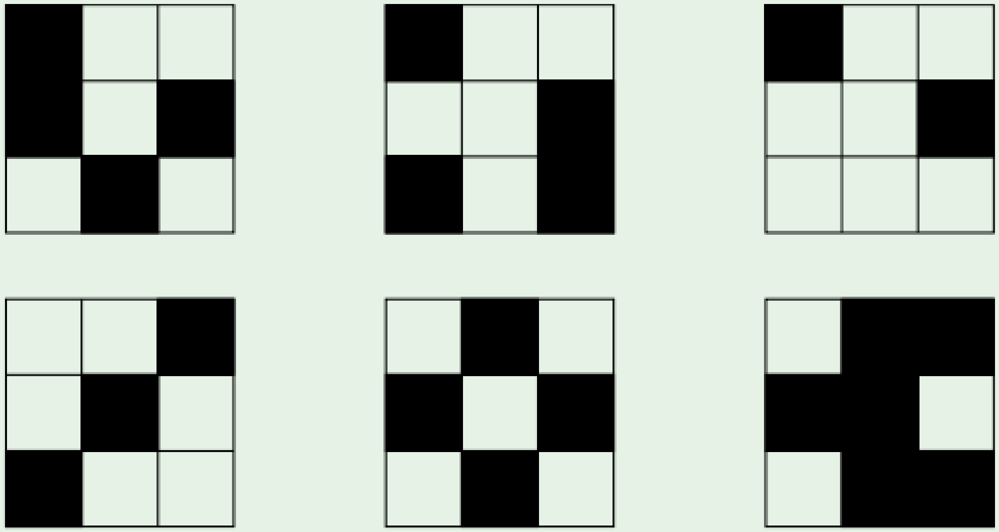
Course contents:

- What is machine learning?
- Can a machine learn?

After successful completion of this course, you will be able to:

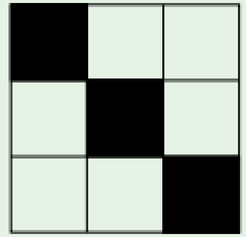
- Explain what machine learning is

Can a machine learn?



$$f = -1$$

$$f = +1$$



$$f = ?$$

Course contents:

- What is machine learning?
- Can a machine learn?
- How to learn?
- How to learn well?

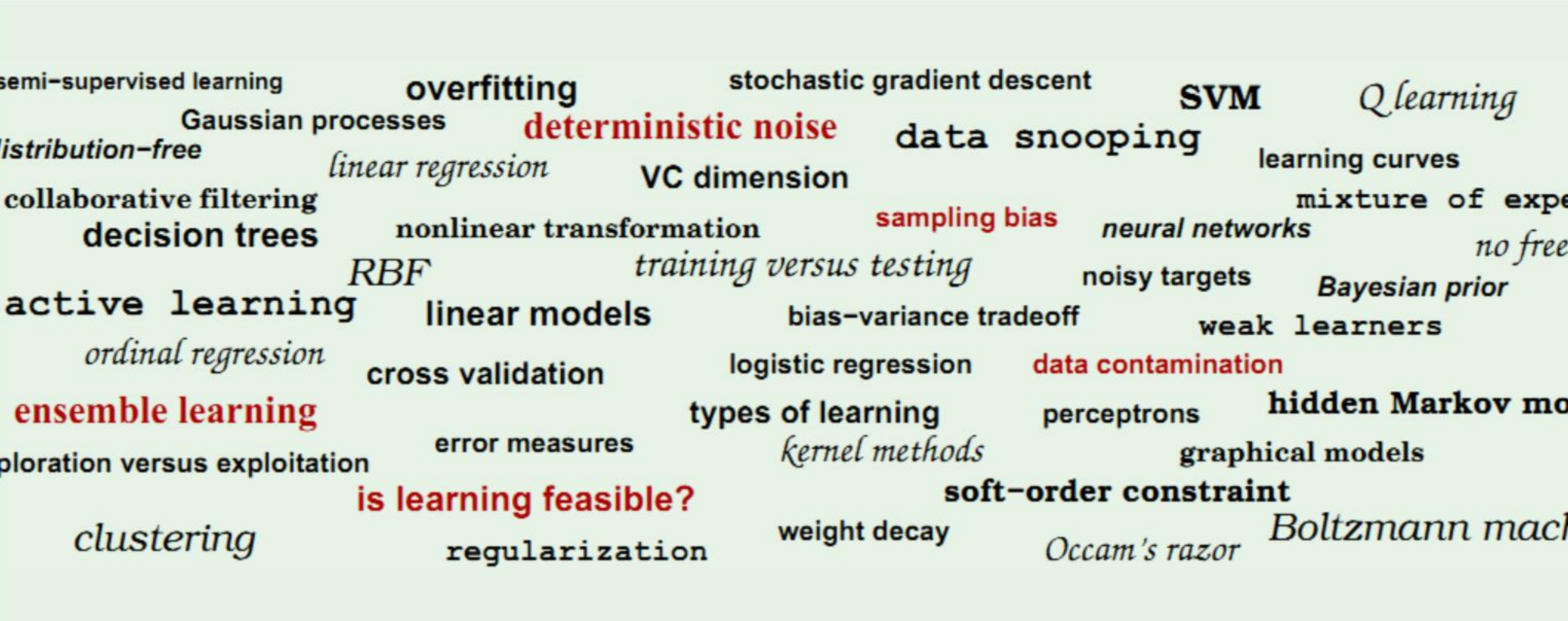
After successful completion of this course, you will be able to:

- Explain what machine learning is
- Explain why a machine can learn
- Explain learning models and implement them (in Python)
- Analyze results when applying a learning model to a dataset, and propose solutions to improve results
- Listen and read provided English materials

How do we teach and learn?

- We will follow the excellent online course taught by Caltech Professor Yaser Abu-Mostafa
  - Why excellent?

A ML jungle out there



Caltech course teaches **the foundations of ML**
so that you can go further in the future easily

Image source: Yaser S. Abu-Mostafa, Learning from Data

How do we teach and learn?

- We will follow [the excellent online course taught by Caltech Professor Yaser Abu-Mostafa]
- There are 18 video lectures and we have 15 weeks
  - Weeks 01 → 02: video lecture 01 + Python
  - Weeks 03 → 15: video lectures 02 → 14
  - You'll drink coffee

    and watch video lectures 15 → 16 at home
  - Video lectures 17 → 18 are optional

How do we teach and learn?

- **Before class**, you will watch the video lecture I require, and post to the forum at least one question
- **In class**, I will explain this video lecture and answer some questions
  - If you don't watch the video before, you may find it difficult to follow
- **After class**, you will pick at least one question and post your answer to the forum

How are students assessed?

- **Exercises (40% of the grade)**
  - 1 exercise / 2 video lectures
  - Total: 6 exercises / 12 video lectures (01→12)
  - Each exercise includes both theoretical and programming questions
  - Programming language: Python (+ NumPy)
    - It allows us to implement ML models very quickly
    - Jupyter Notebook is awesome (text + live code + latex)

How are students assessed?

- **Final project (50% of the grade)**
  - Train SVM (a state-of-the-art learning model) to classify images of hand-written digits
  - Group?
  - Present on the final exam day
- **Forum Q & A (10% of the grade)**
  - Easy to get this 10%
  - Penalize spam

How are students assessed?

- **Remember: the main goal is to learn, truly learn**
- You can discuss ideas with others, but your writing and code must be your own, based on your own understanding
- If you violate this rule, you will get 0 score for the course

Last slide

This course will be difficult ☹

But if you can enjoy difficulties, focus on your own learning, eliminate noise, ...

... you will be able to grasp the foundations of ML ☺