Machine learning tutorial

psychobiology summer school

June 26, 2017

Gal Chechik

Gonda Brain Research center, Bar Ilan University

Google research, Mountain View California

http://chechiklab.biu.ac.il/

Self-driving cars completed millions of miles

(fewer accidents than people)

Speech recognition in every phone

(that actually works)

deep networks learn new games

(and beat world champions)

Interest in machine learning is exploding

Media

The New York Times

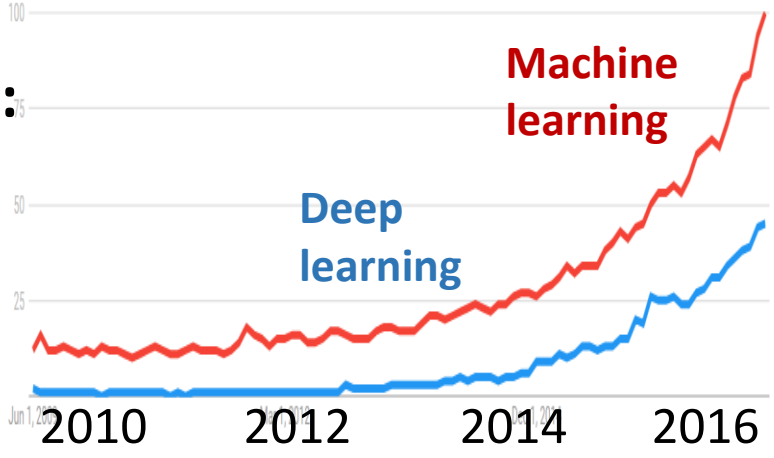
The Great A.I. Awakening

The Real Threat of Artificial Intelligence

Business:

Beware the Hype of **Artificial Intelligence**
Fortune - Jun 23, 2017
Artificial intelligence has made great strides in tl
generated much hype over its current capabilities

Google searches:



Science

nature.com > subjects > machine learning

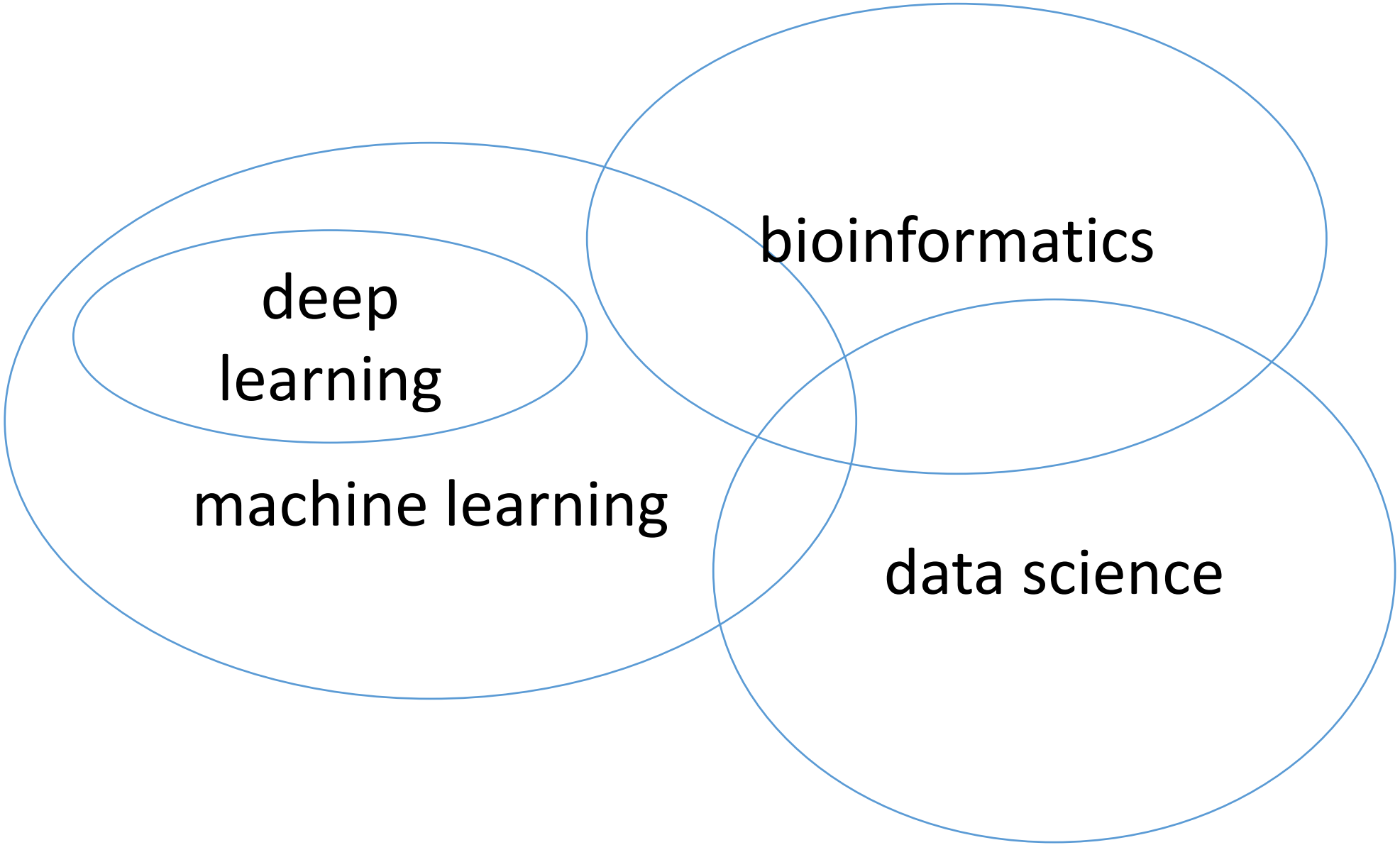
nature.com

Editorial | 10 February 2017

Auspicious machine learning

The accelerating diagnostic power of deep learning will soon empower physicians.

Many buzzwords:



This talk will cover basic concepts:

1. Data science
2. Machine learning
3. Deep learning
4. Recent advances

Part 1: Data science

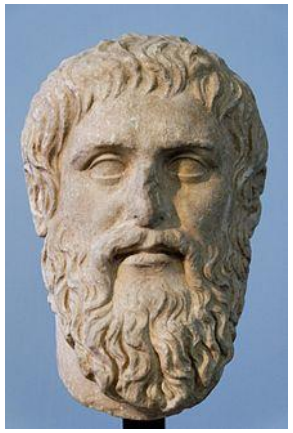
Data science and data-driven science

Isn’t all science “data driven”?

A fundamental question:

Epistemology: How can we gain knowledge?

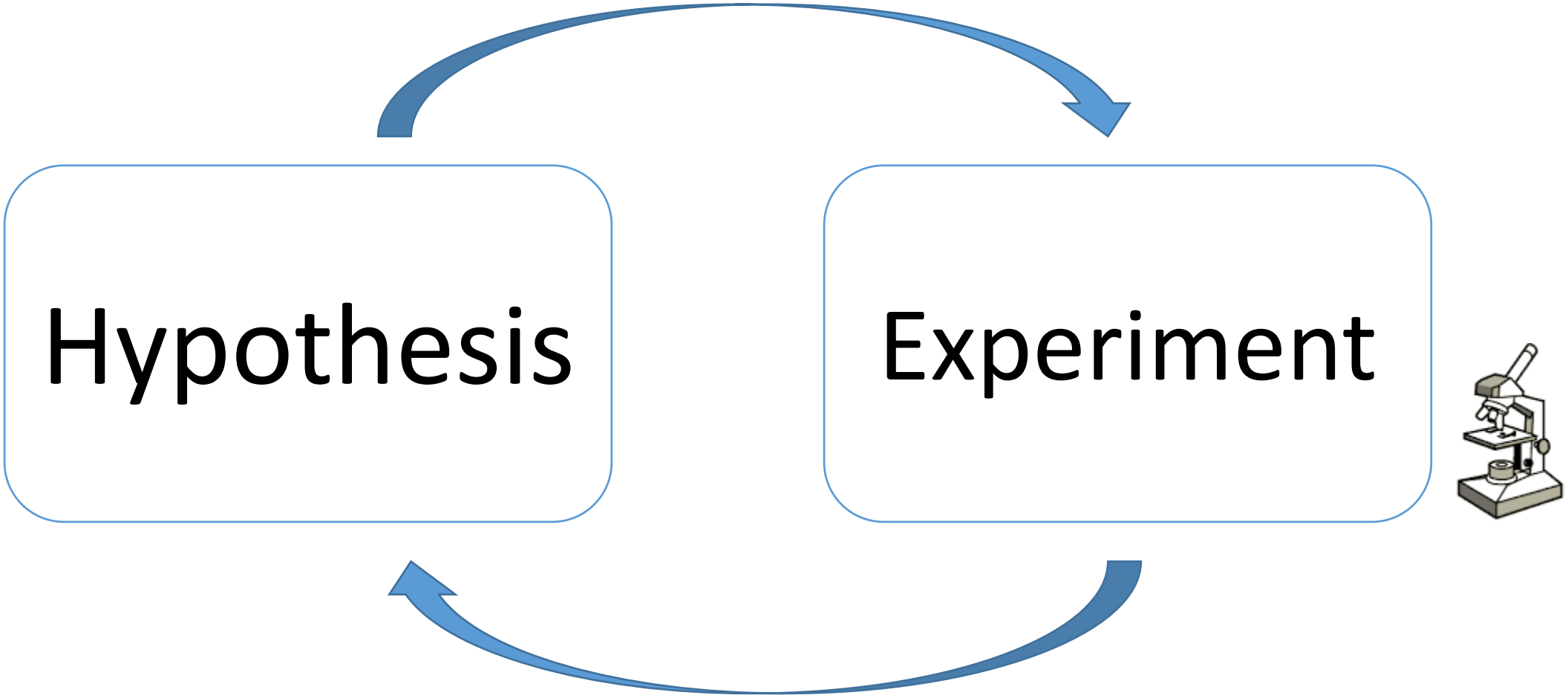
Plato



Descartes

Bacon

The scientific method

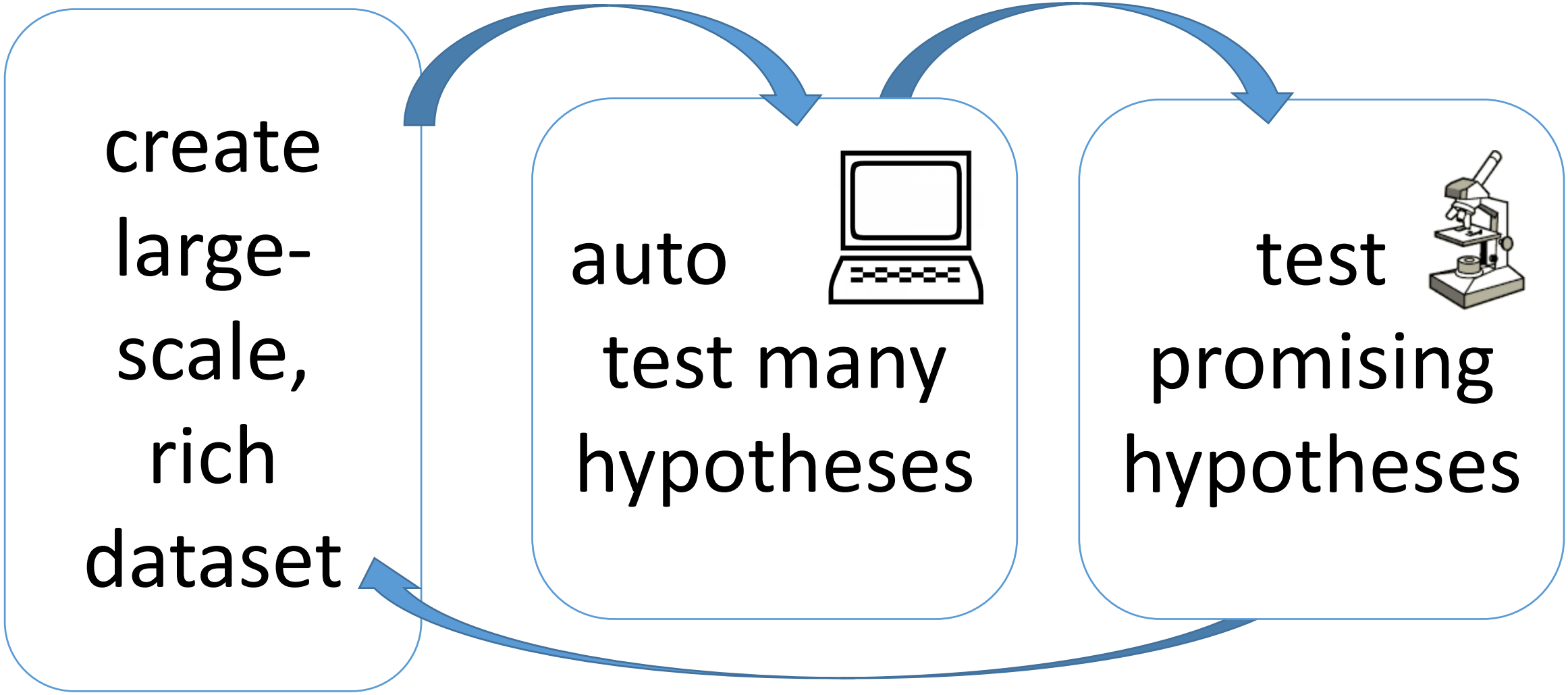


This approach is slow and inefficient

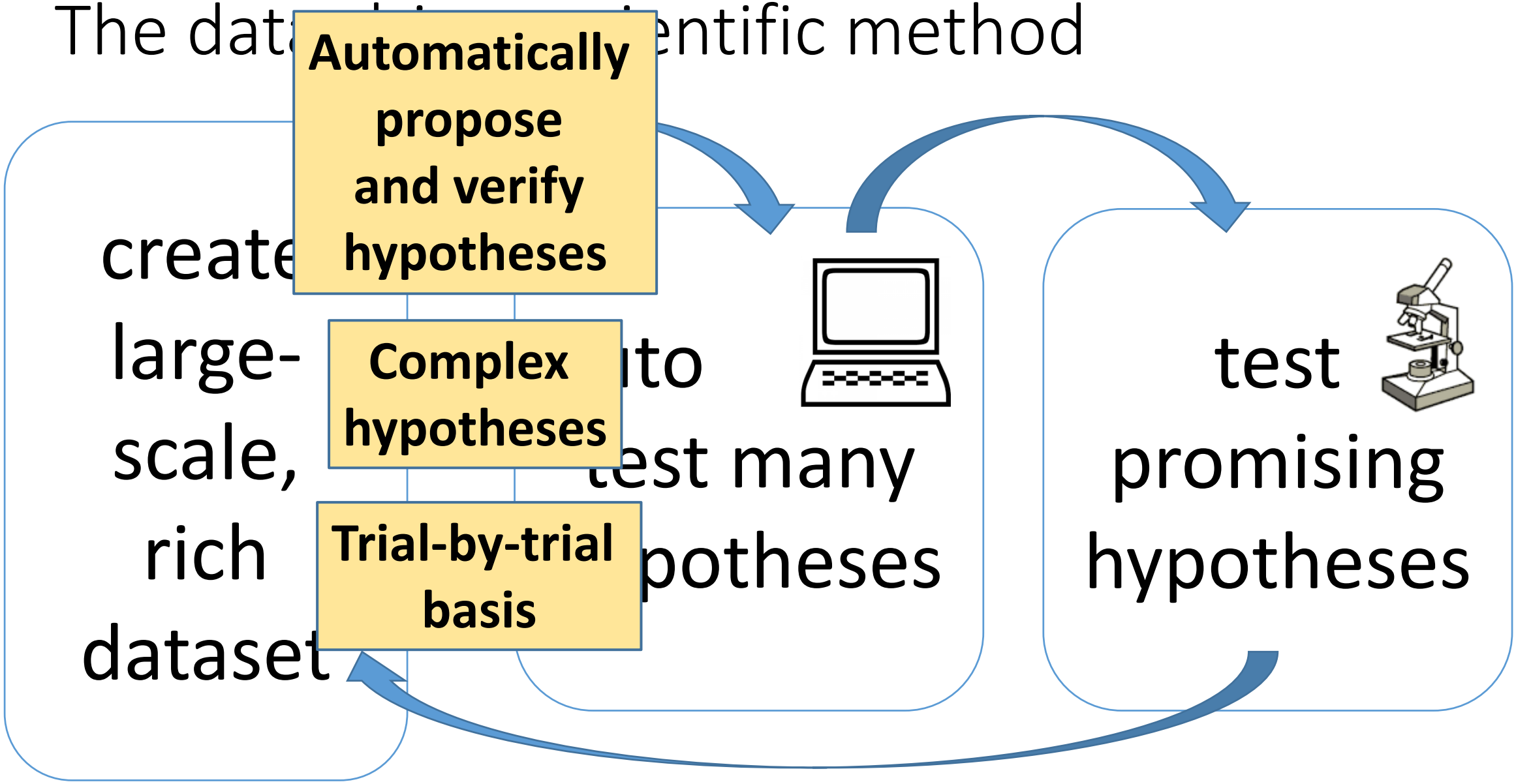
Experimental techniques become very complex. Only few (rich) experts can conduct experiments.

Sharing data, in addition to sharing results, is more efficient.

The data-driven scientific method



The data-driven scientific method



Part 2: Machine learning

Machine learning

Algorithms that learn from experience
Learn **rules from examples**
Learning vs inference

Example: A 2-class classifier

Classify one of two faces

Learning phase:

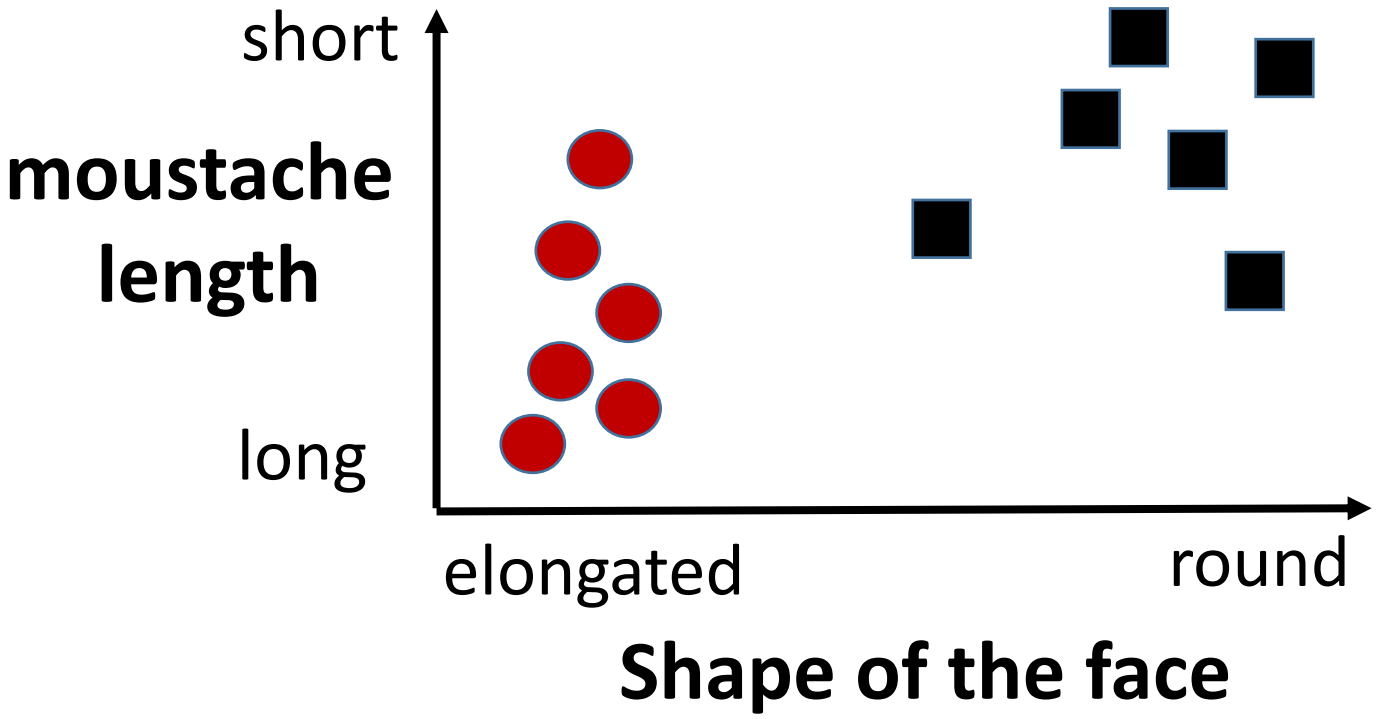
Example: A 2-class classifier

Classify one of two faces

Inference phase:

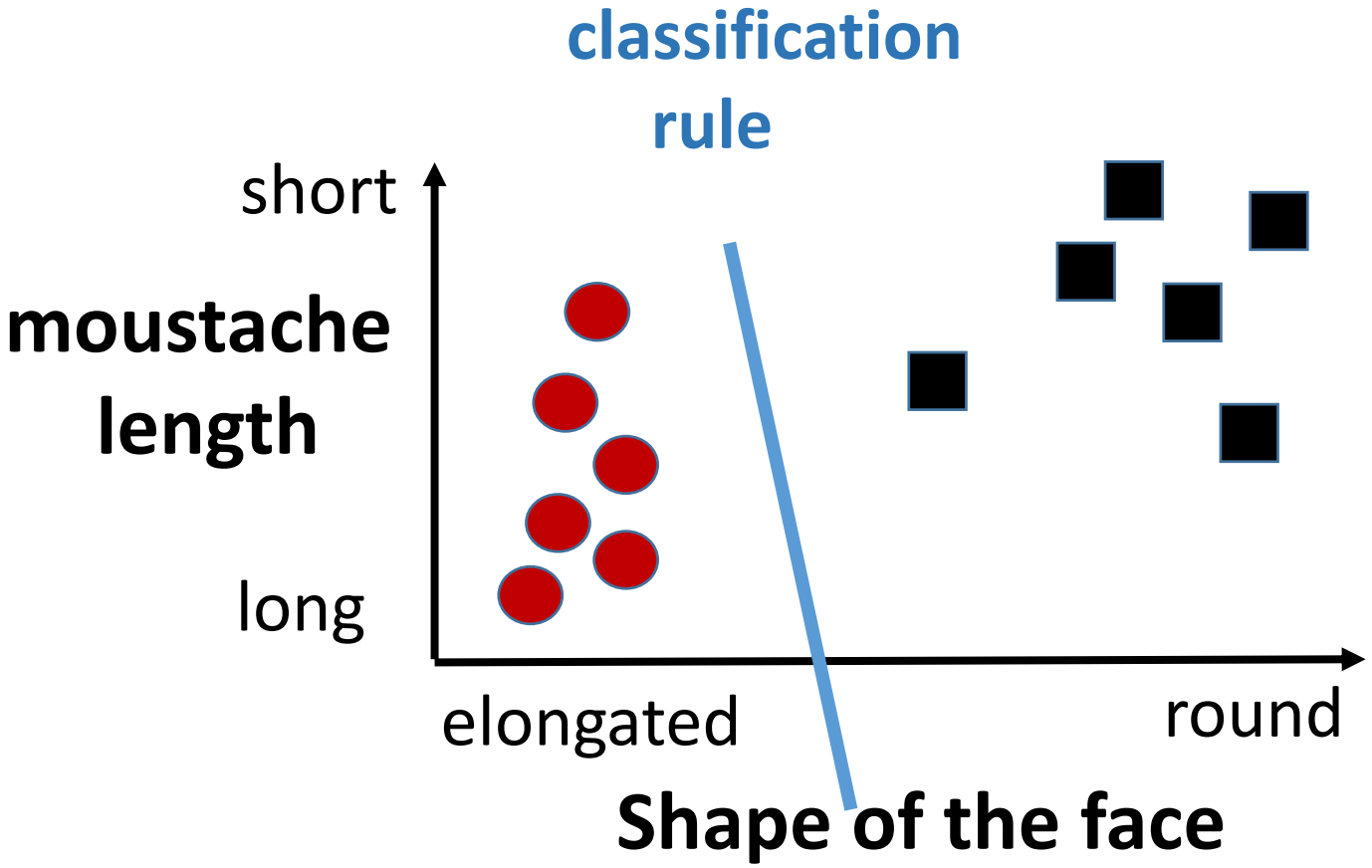
Example: A 2-class classifier

Classify one of two faces



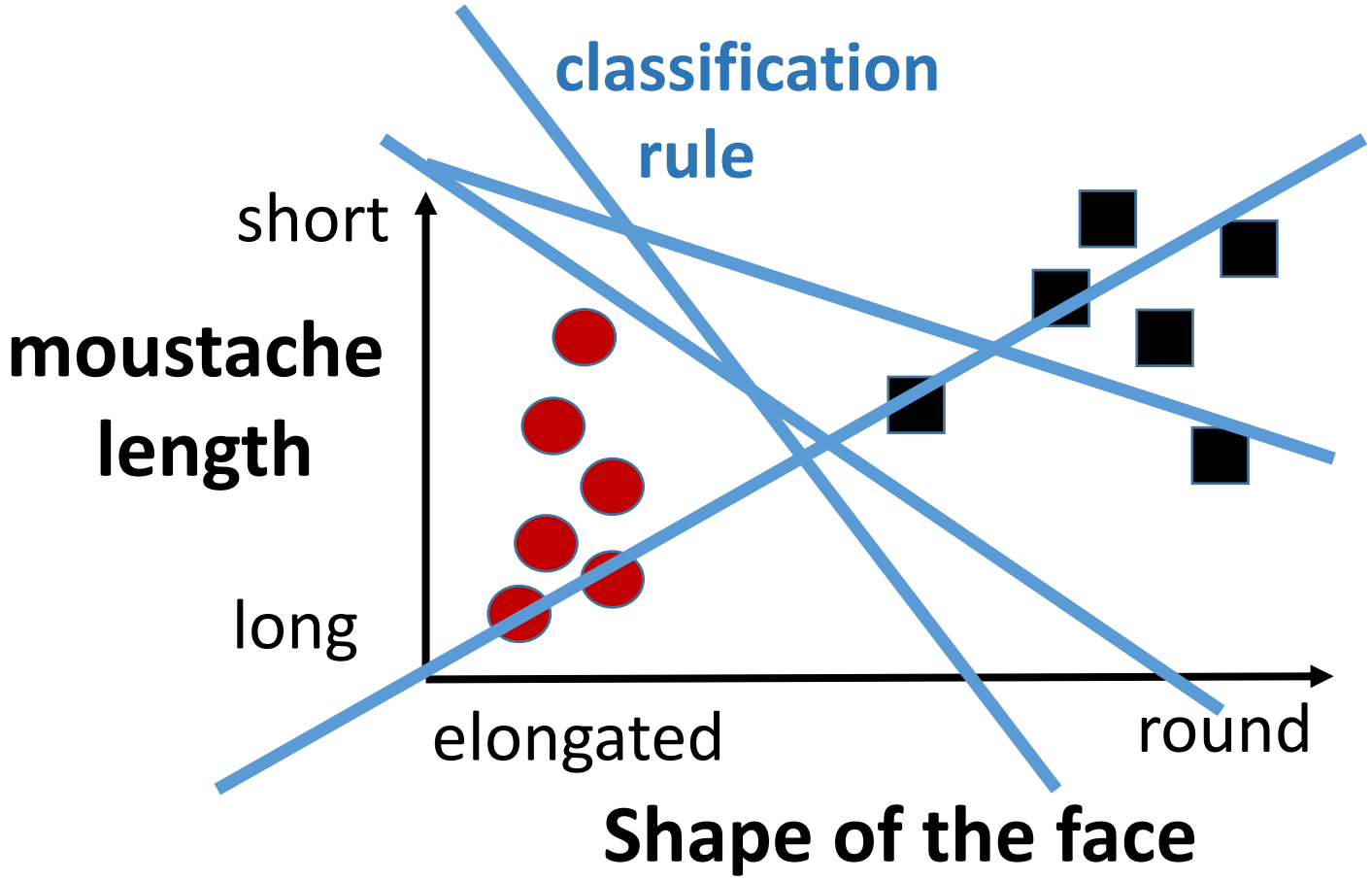
Example: A 2-class classifier

Classify one of two faces



Example: A 2-class classifier

Classify one of two faces

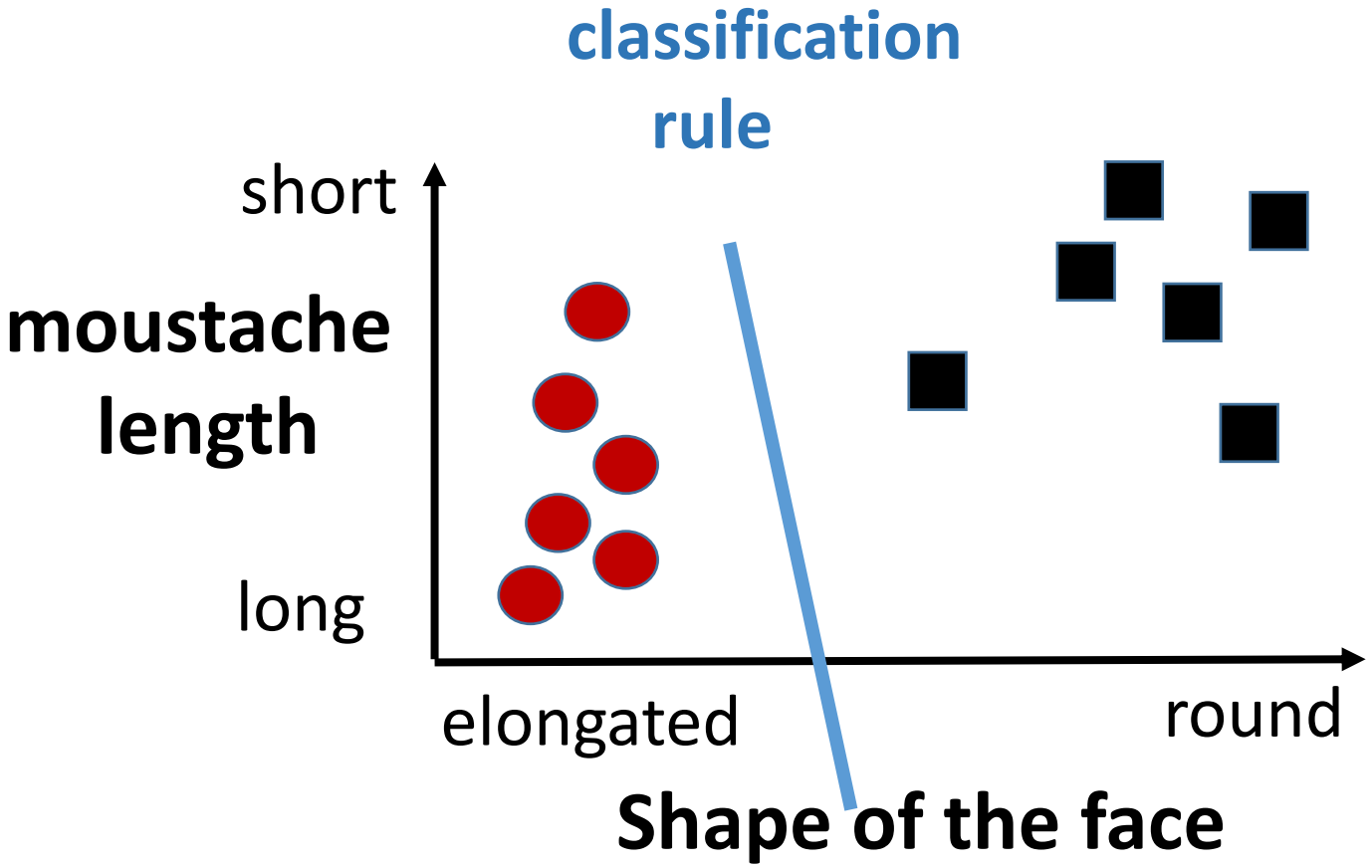


How to search over possible rules?

Stochastic gradient descent:
Iteratively take small steps that reduce errors
"one sample at a time"

Example: A 2-class classifier

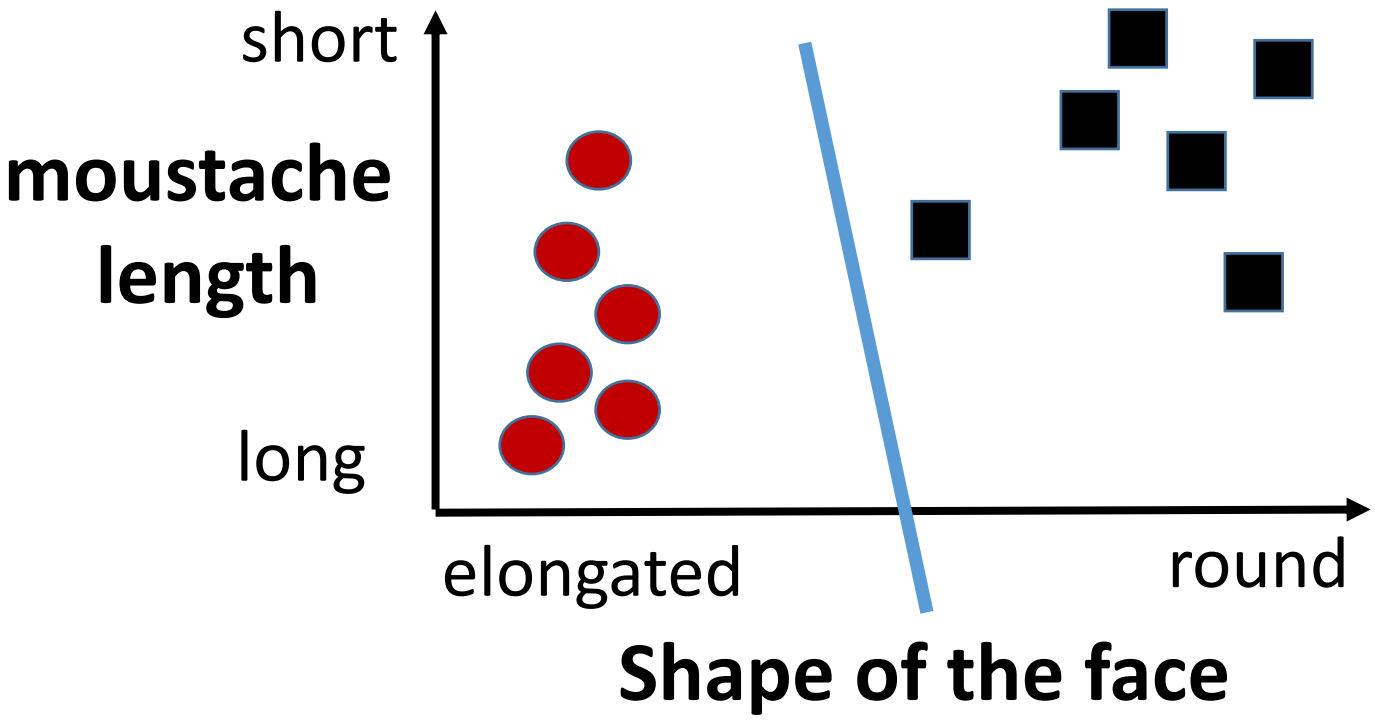
Classify one of two faces



Real life is more complex

- Many facial features and their combinations (high-dimensional representation)
- More philosophers

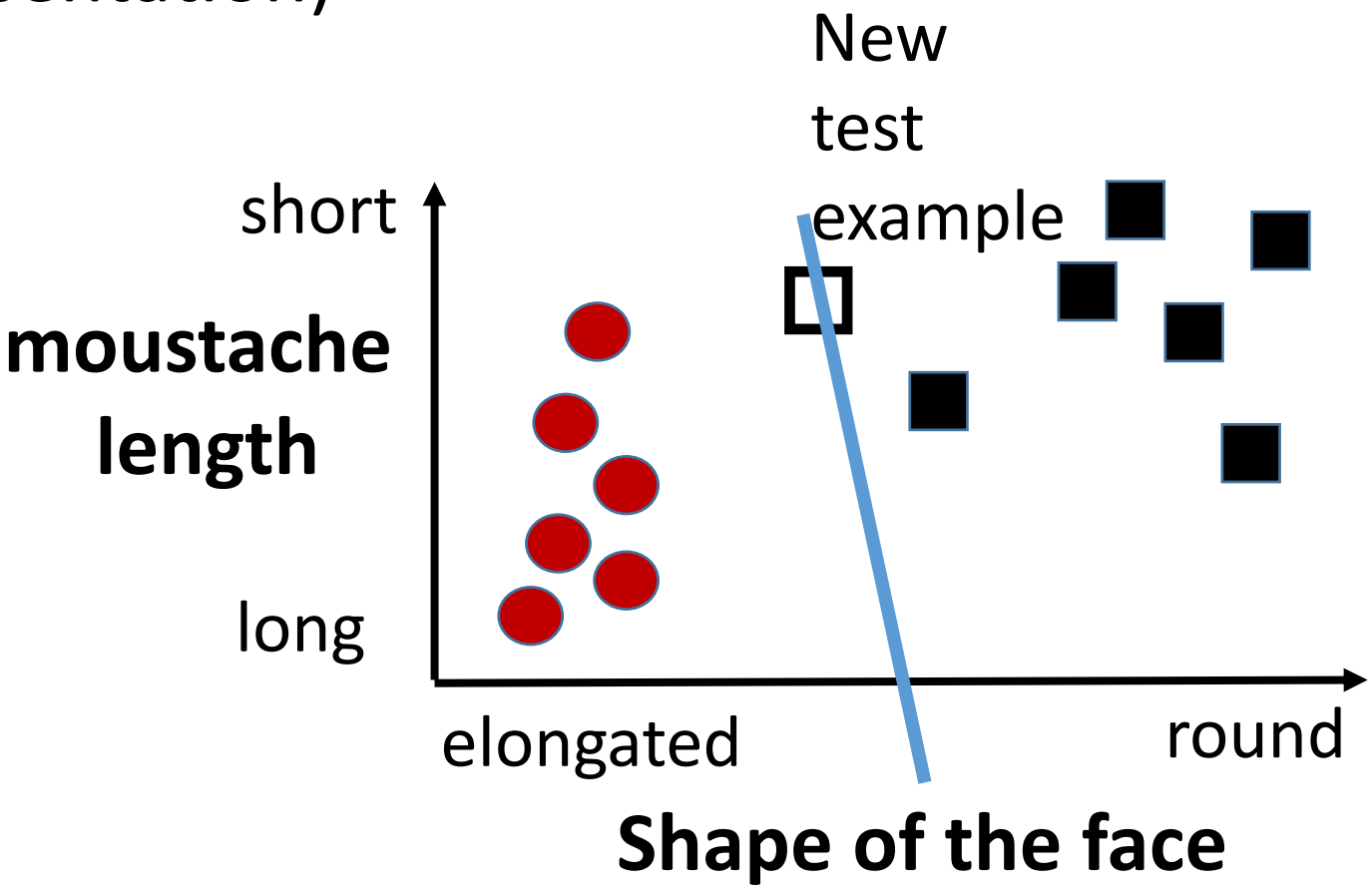
- We care about **generalization to new examples!**



Real life is more complex

- Many facial features and their combinations (high-dimensional representation)
- More philosophers

- We care about **generalization to new examples!**



An alternative approach: Programming a rule-based system

Ask experts to explicitly specify the rules that allow to classify.

Often doesn’t work well

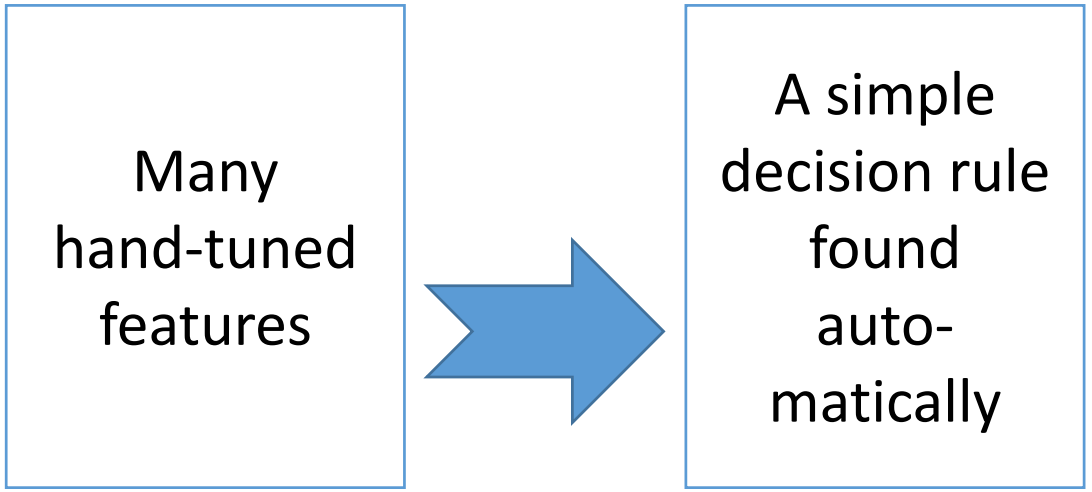
- People don’t actually know the rules
- Cannot formalize them explicitly
- Disagreement among experts
- Time consuming and inefficient

Machine learning is like data-driven programming

Instead of programming rules explicitly, provide enough examples, and let the computer search for the rule

A typical machine-learning system until 2012

1. People extract many (hand tuned) features
2. Machines find simple rules based on these complex features

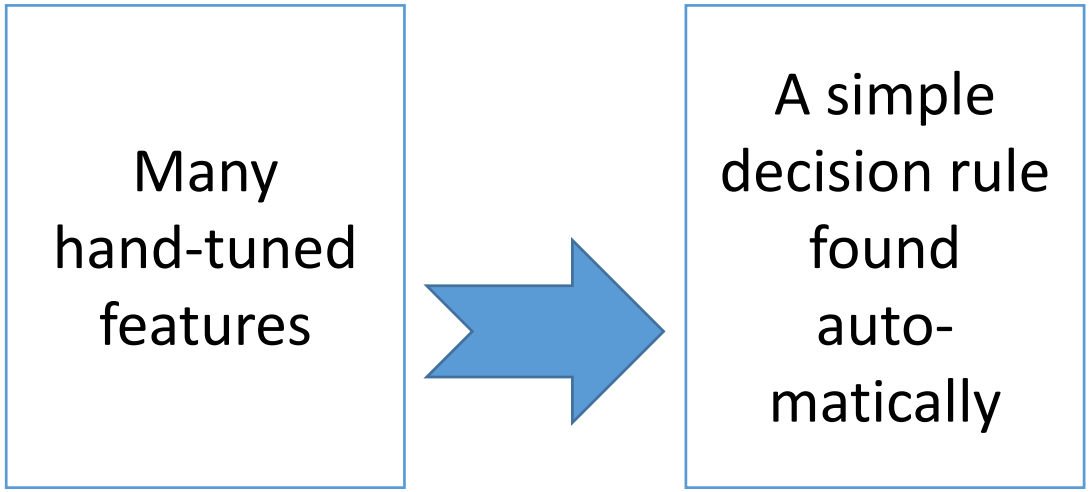


Linear separators are called "neurons". Because real neurons fire if a weighted sum of their inputs exceeds a threshold

Part 3: Deep learning

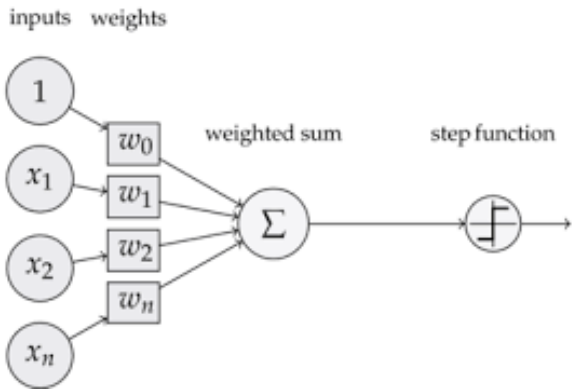
A typical machine-learning system until 2012

1. People extract many (hand tuned) features
2. Machines find simple rules based on these complex features



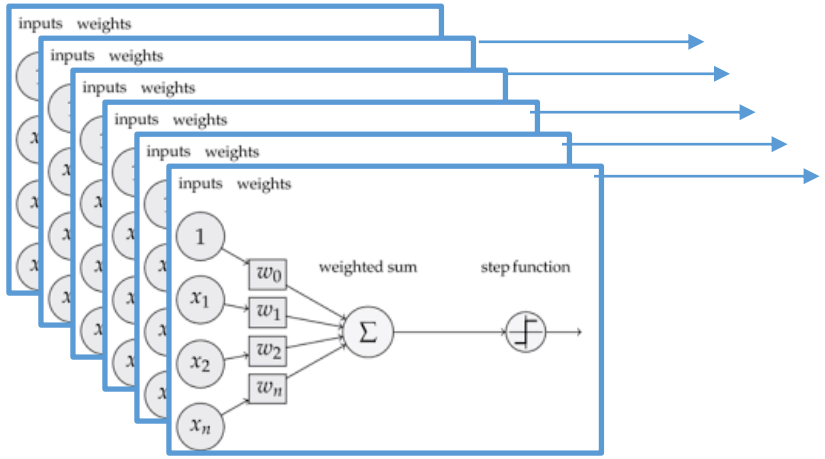
Linear separators are called "neurons". Because real neurons fire if a weighted sum of their inputs exceeds a threshold

Single neuron

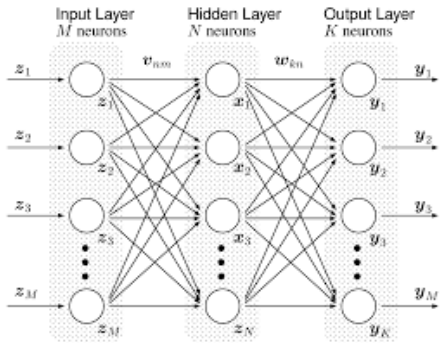


הפלט הוא סכום משוקלל
של הקלטים.
(או אפס אם מתחת לסף)

Layer of neurons

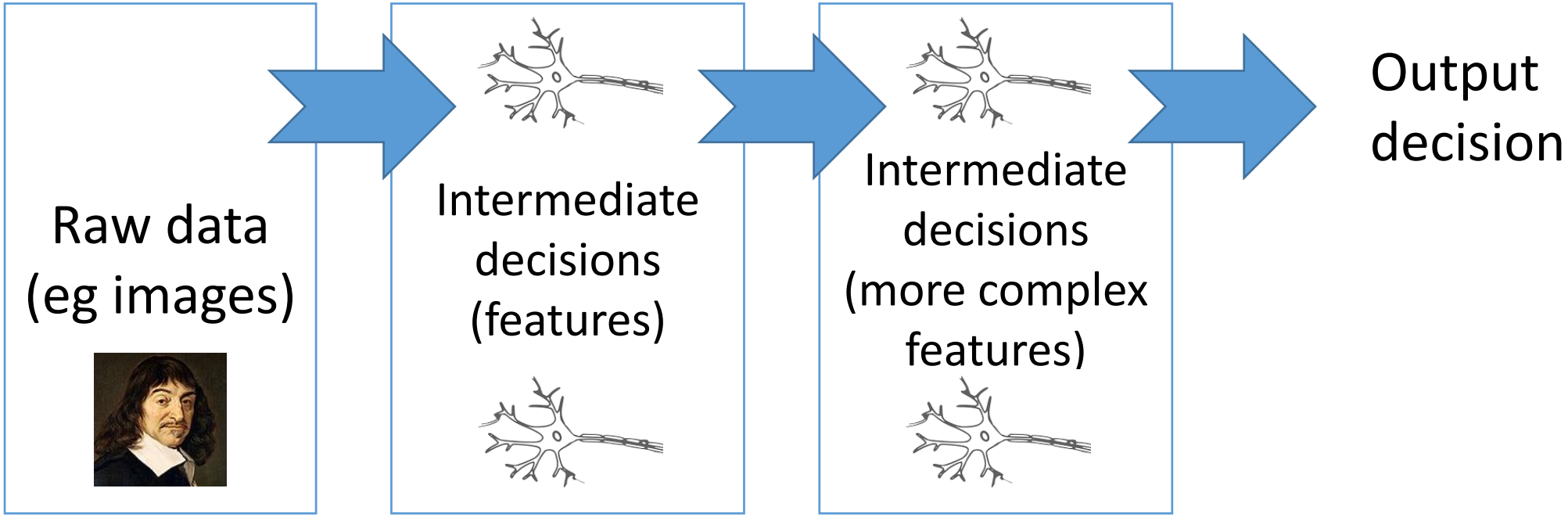


Multiple layers



A typical machine-learning system after 2012

Let the algorithm find the features



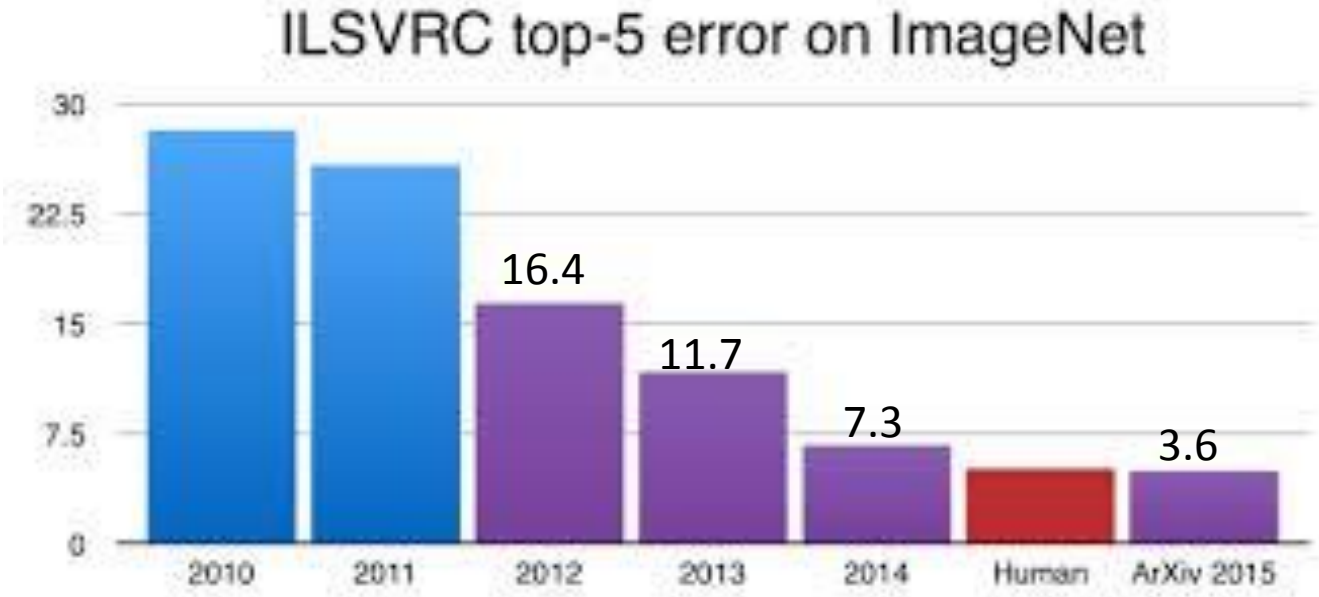
Called "**deep learning**" because has many layers

Deep learning succeed in numerous fields

Image-Net

Image-Net: 1M images, 1K classes

Super-human performance



Early approaches

Deep networks

Deep learning works well with massive data

More is better:
Deep networks require to tune many parameters; requires massive labeled dataset.

Adding “noise” helps:
Different than human inference:

Part 4: recent advances

Sequence to sequences: Machine translation

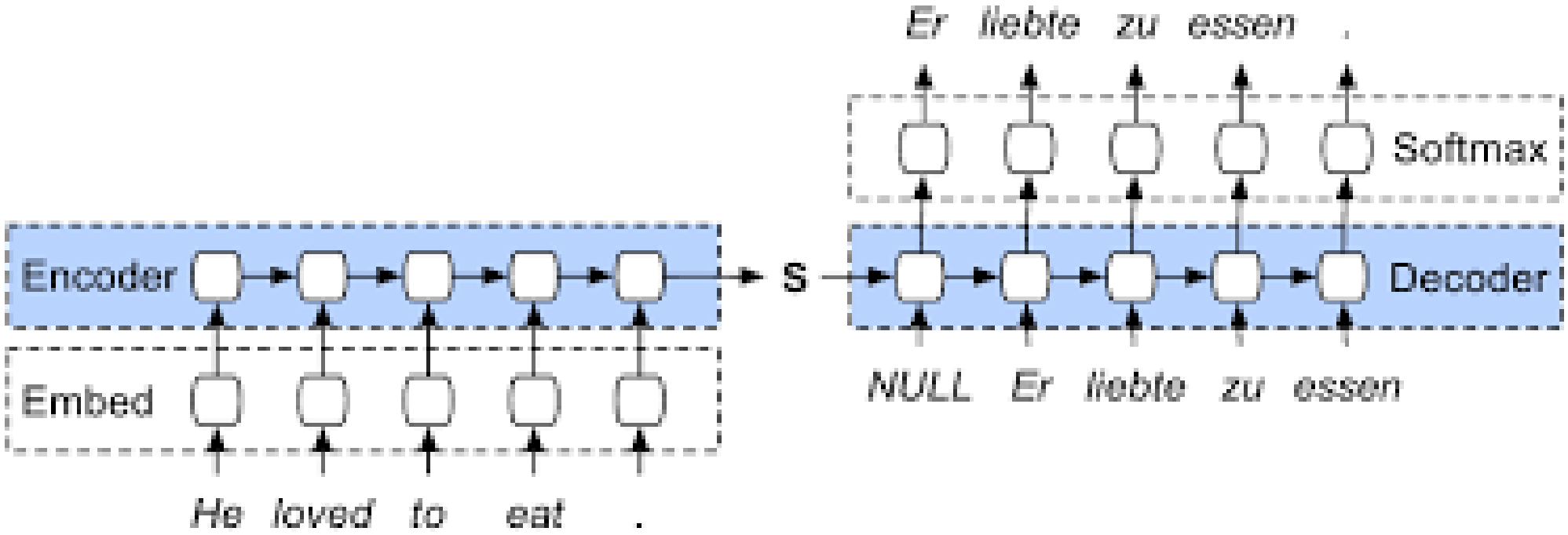


Image to natural language descriptions

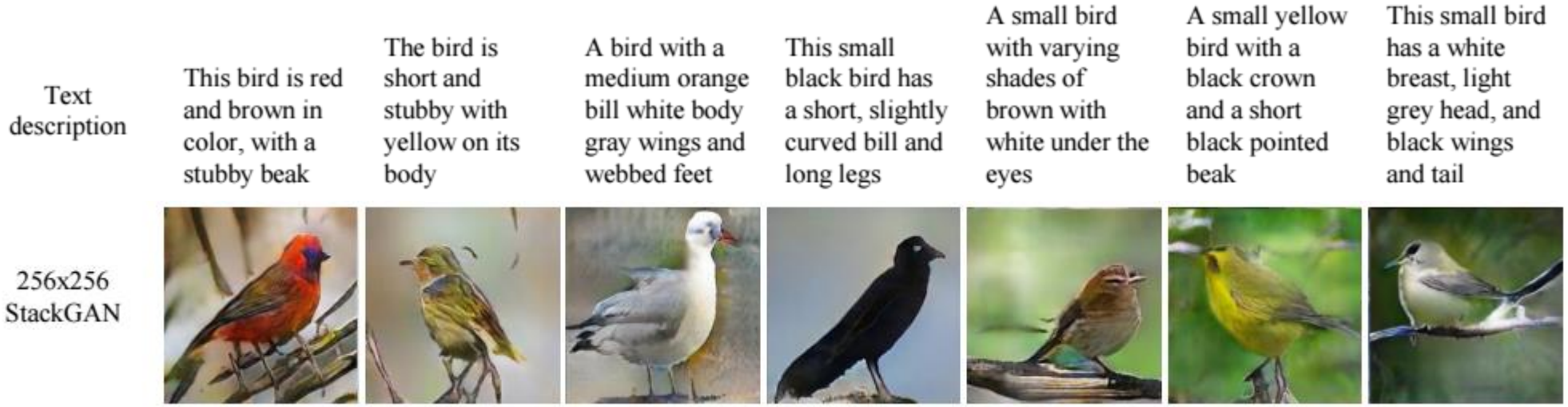
Language is a universal interface for knowledge:
perfect for describing complex scenes.

Generative models:

One learner generates data. Trying to trick a second classifier to think the image is natural

input

output



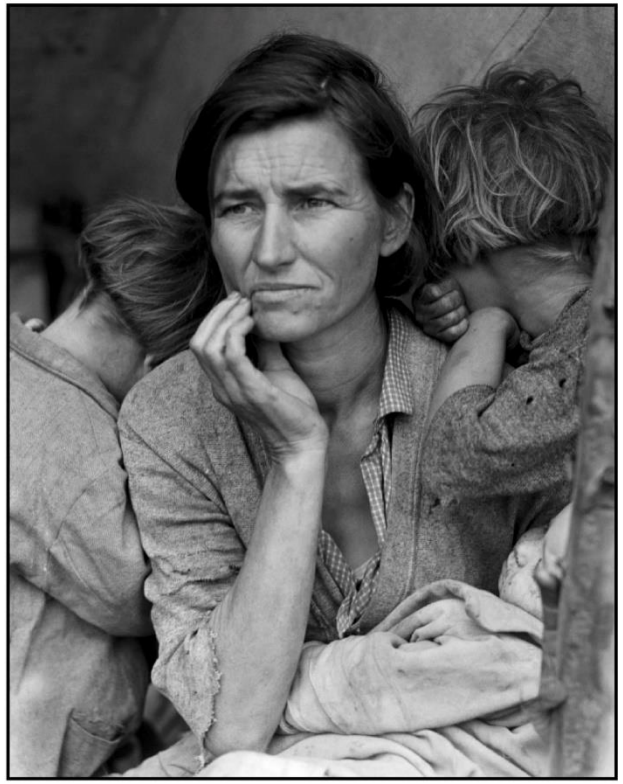
Pixel-to-pixel models

The output is another image

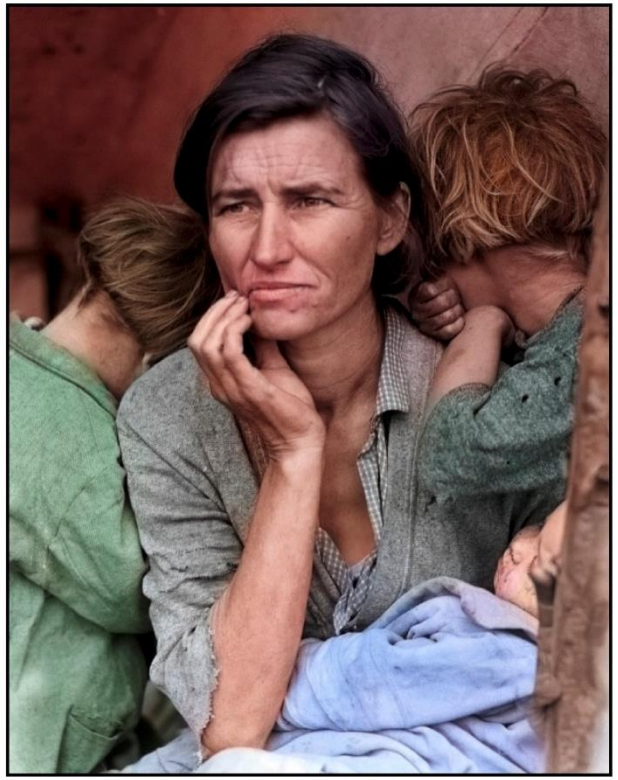
Learning phase:
Map given images to B/W; use as a training set.

Inference phase:
Turn new B/W images to color.

Input



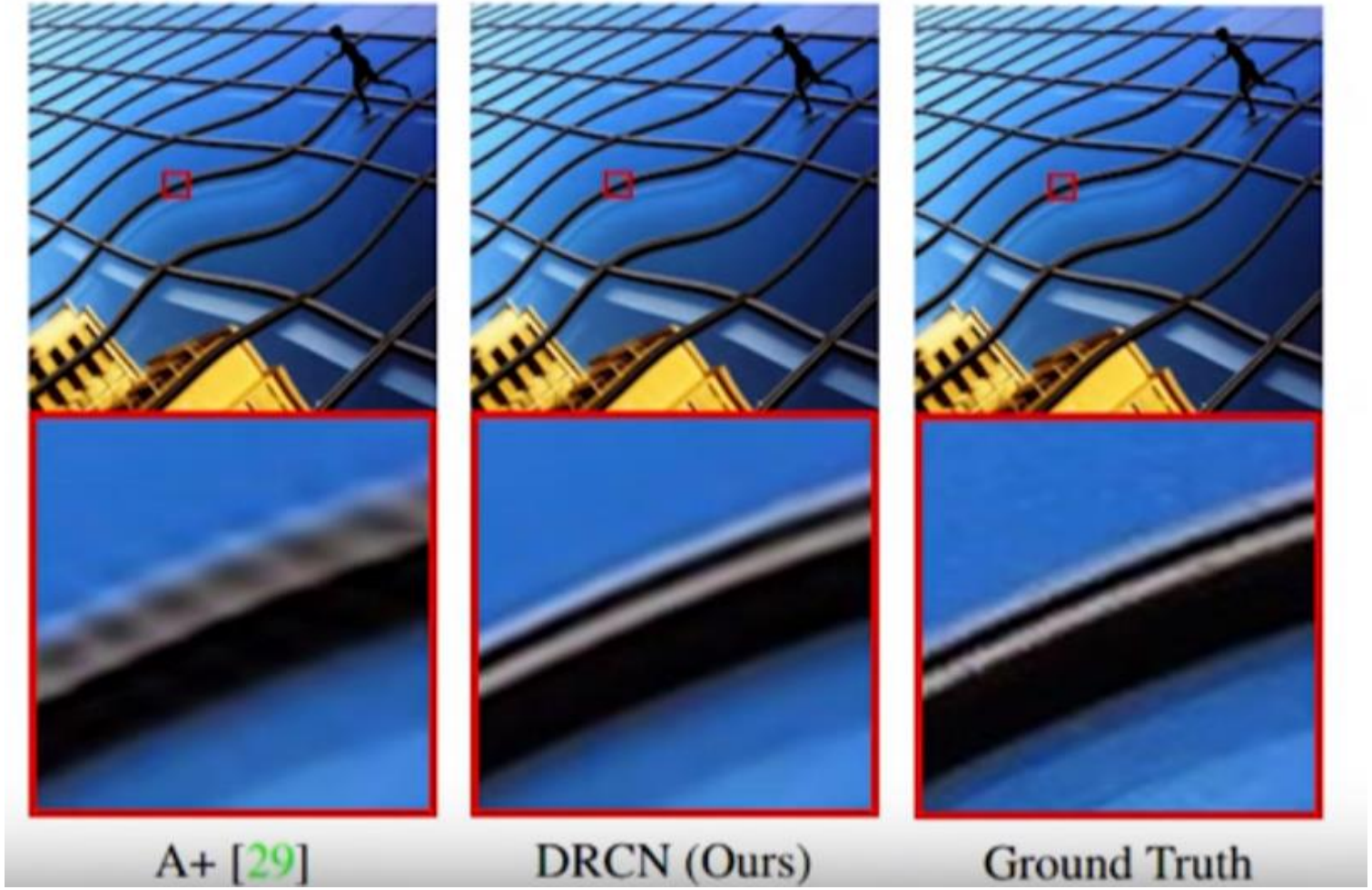
output



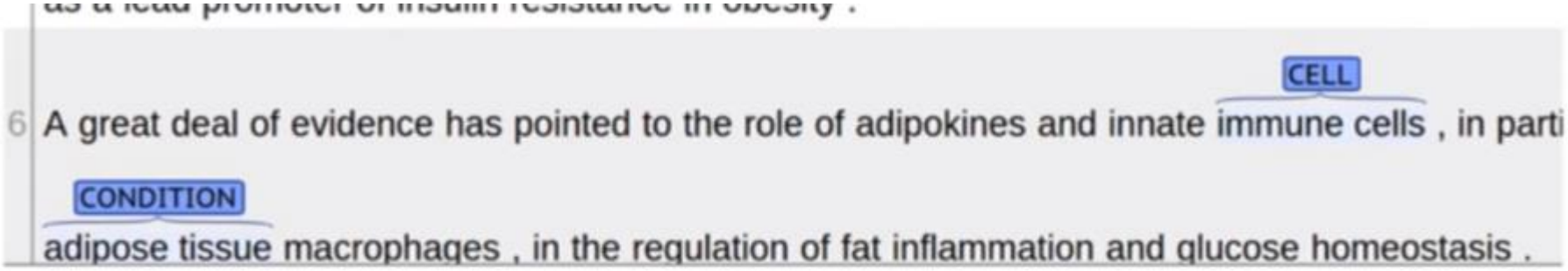
Transfer style

Input **output**

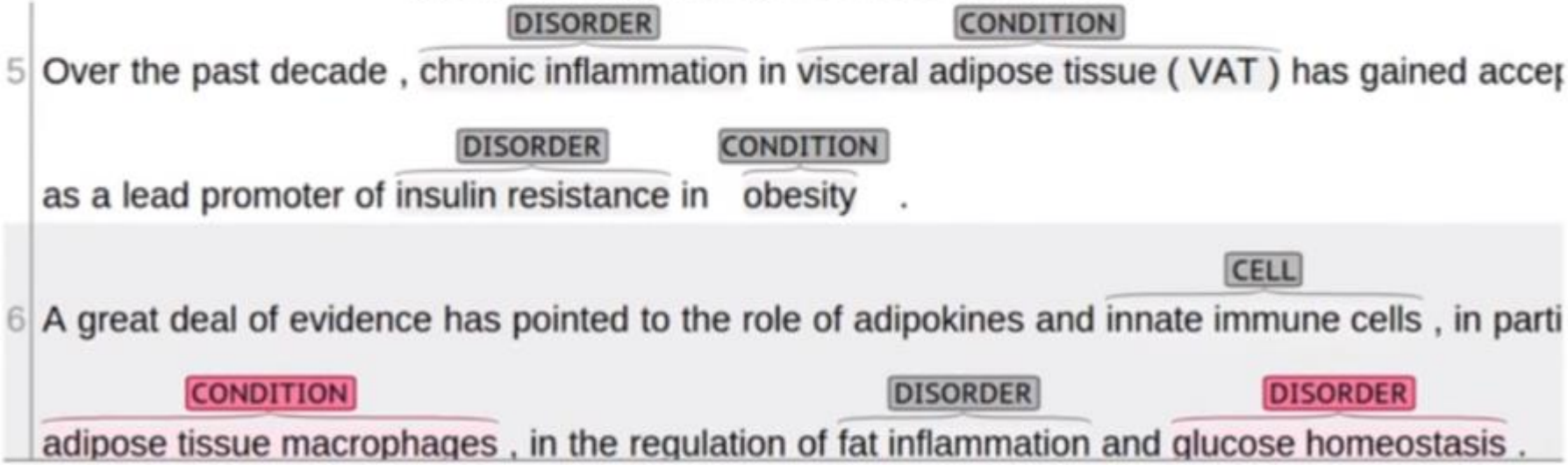
Super resolution (CSI in real life)



Analysis of scientific literature



(b) Automatic suggestions after 5 abstracts are annotated.



Justification: Example results

Tennessee Warbler

This is a **grey bird** with yellow on its wings and a **white eyebrow**

Mourning Warbler

Small **green green** and red bird with medium tarsus and short beak

Black and white Warbler

Part 6: ML can benefit psychobiology

- Diagnosis
- Prediction (side effects)
- Finding predictive bio-markers
- Automated analysis of behavior, video, images, etc.

Thank you

http://chechiklab.biu.ac.il/