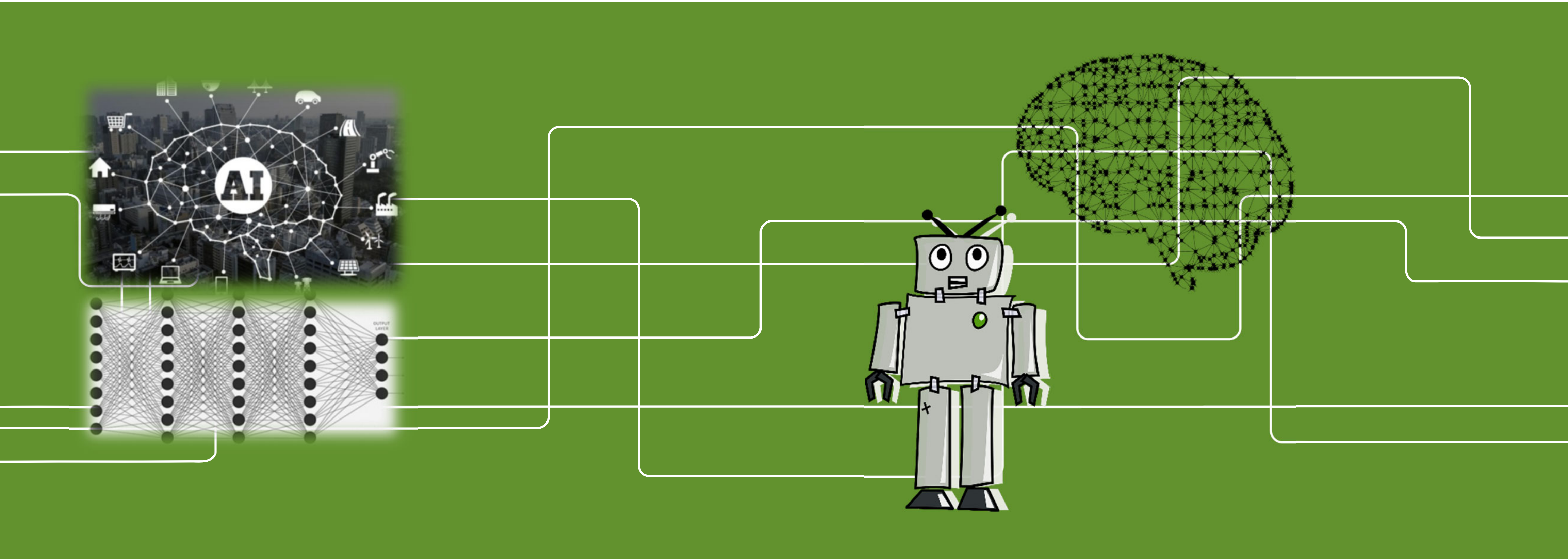
Introduction to AI and Machine Learning

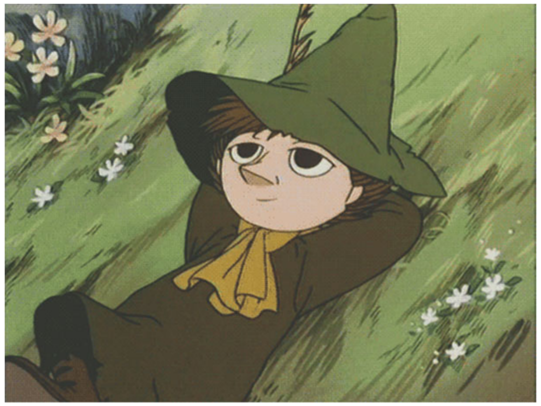
Pawel Herman

Department of Computational Science and Technology
School of Computer Science and Communication
KTH Royal Institute of Technology

PDC Summer School, August 2019

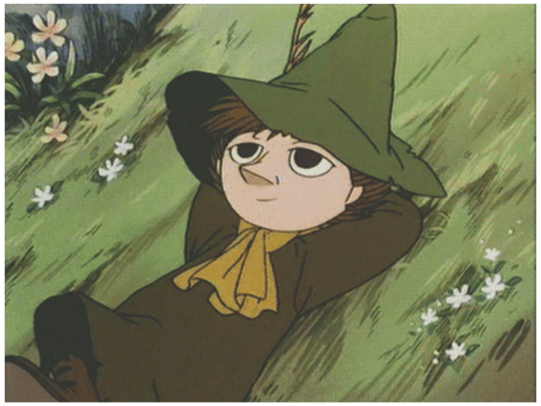


Why bother with AI and Machine Learning (ML)?



Why bother with AI and Machine Learning (ML)?

- Good to be familiar with current trends and buzz words – public attention
- Follow the currently “hot” scientific area
- Intriguing computational paradigm
- It may be useful
- It may be profitable
- ……



Main reasons for the AI/ML hype

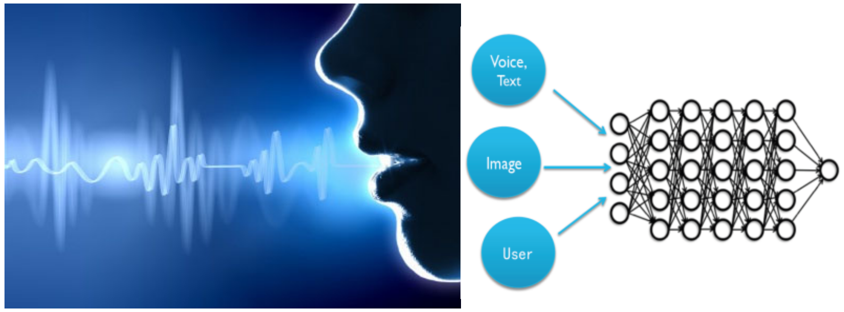
- Impressive scope of applications

Autonomous vehicles, agents, robotics

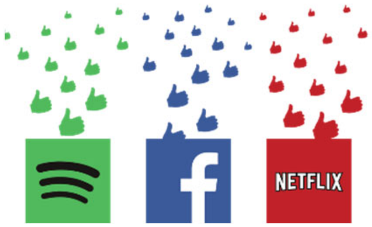
Medicine, healthcare

IBM WATSON

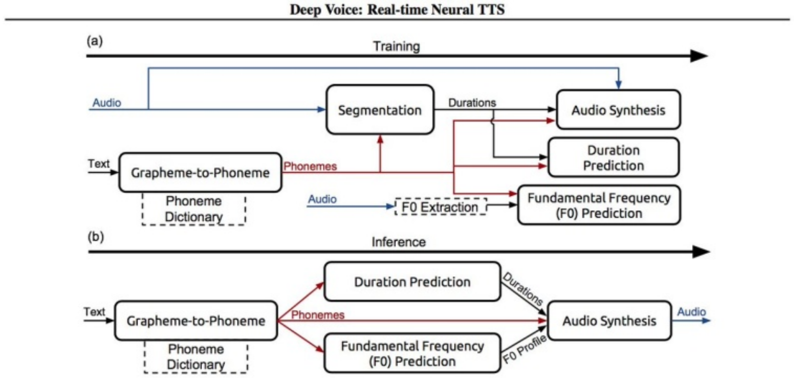
Automated speech recognition



Recommendation systems



Machine translation + text to speech transformation



Main reasons for the AI/ML hype

- Impressive scope of applications

Tesla

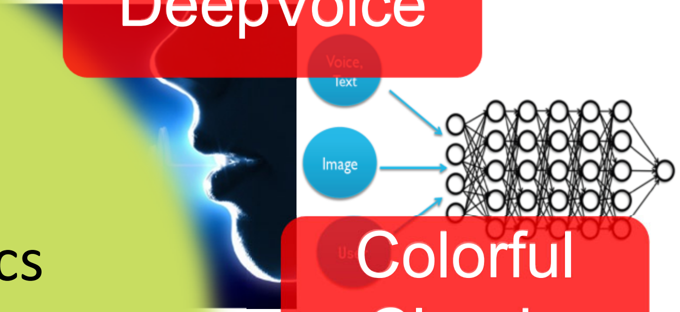
Autonomous vehicles, agents, robotics

- Image recognition
- Time series prediction
- Autonomous planning, logistics
- Decision making
- Spam filtering
- Financial applications
- Recommendation systems

Netflix

Automated speech recognition

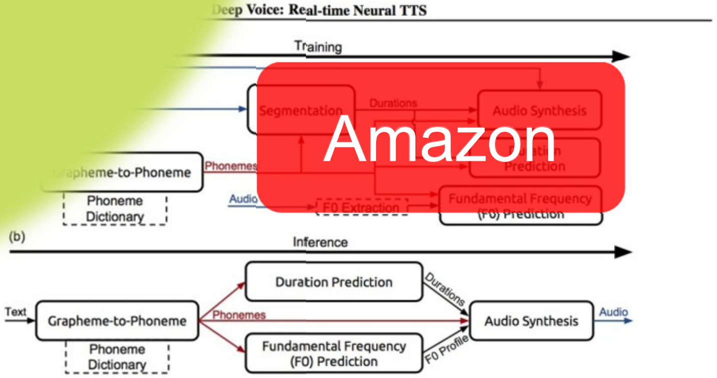
DeepVoice



Colorful Clouds

ation + text to speech
nsformation

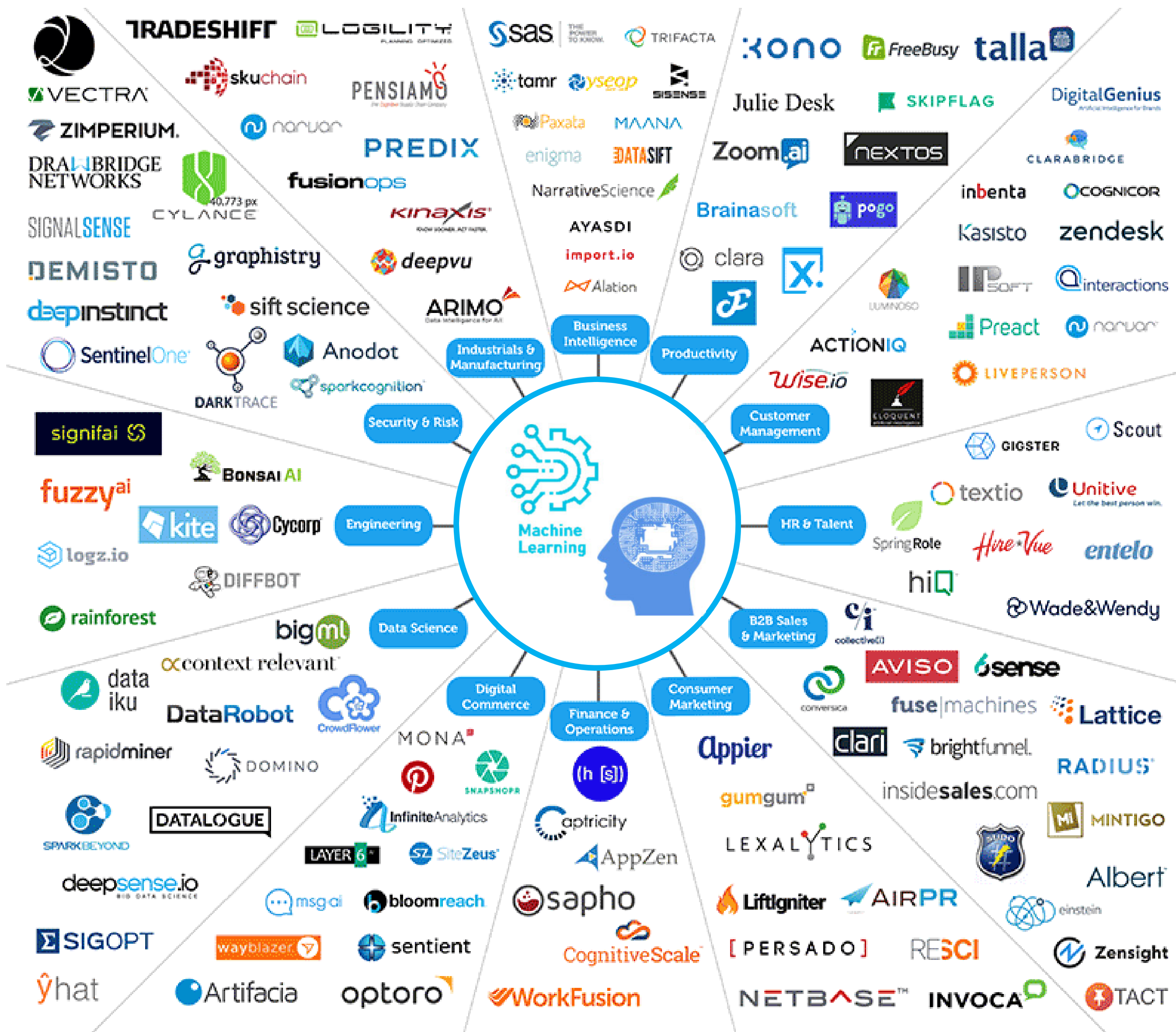
Amazon



Main reasons for the AI/ML hype

- Impressive scope of applications and intriguingly good performance
- A huge opportunity for industry to improve products, services at a relatively low cost – large profit margin

A fraction of the current AI/ML enterprise landscape



Main reasons for the AI/ML hype

- Impressive scope of applications and intriguingly good performance
- A huge opportunity for industry to improve products, services at a relatively low cost – large profit margin
- Renewed scientific interest following Deep Learning success

Main reasons for the AI/ML hype

- Impressive scope of applications and intriguingly good performance
- A huge opportunity for industry to improve products, services at a relatively low cost – large profit margin
- Renewed scientific interest following Deep Learning success
- Remarkable implications for other fields

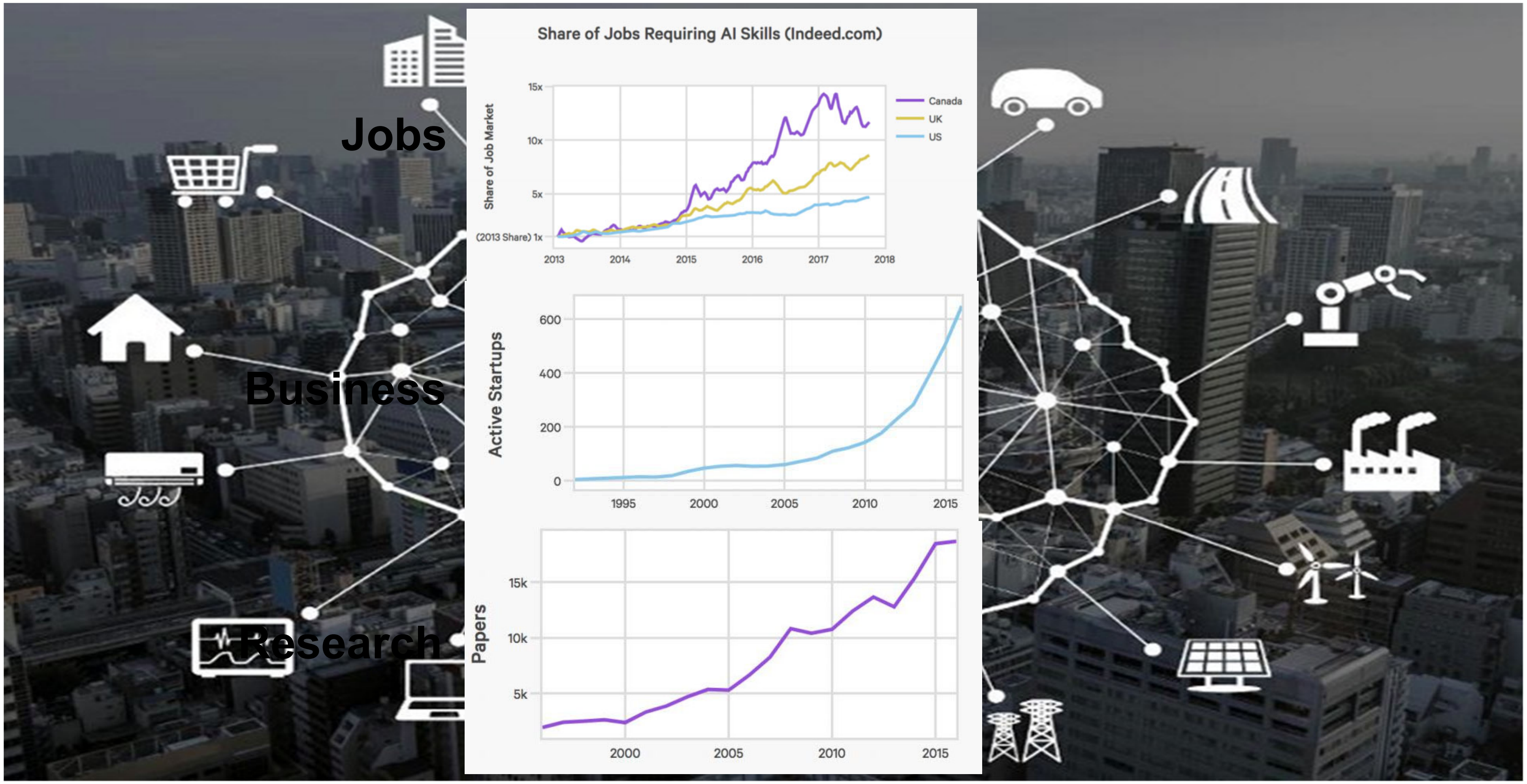
LIFE SCIENCES ENGINEERING PHYSICS ECONOMY

Main reasons for the AI/ML hype

- Impressive scope of applications and intriguingly good performance
- A huge opportunity for industry to improve products, services at a relatively low cost – large profit margin
- Renewed scientific interest following Deep Learning success
- Remarkable implications for other fields
- Good timing: data availability, growing compute power and the overall quest for automatization and optimisation

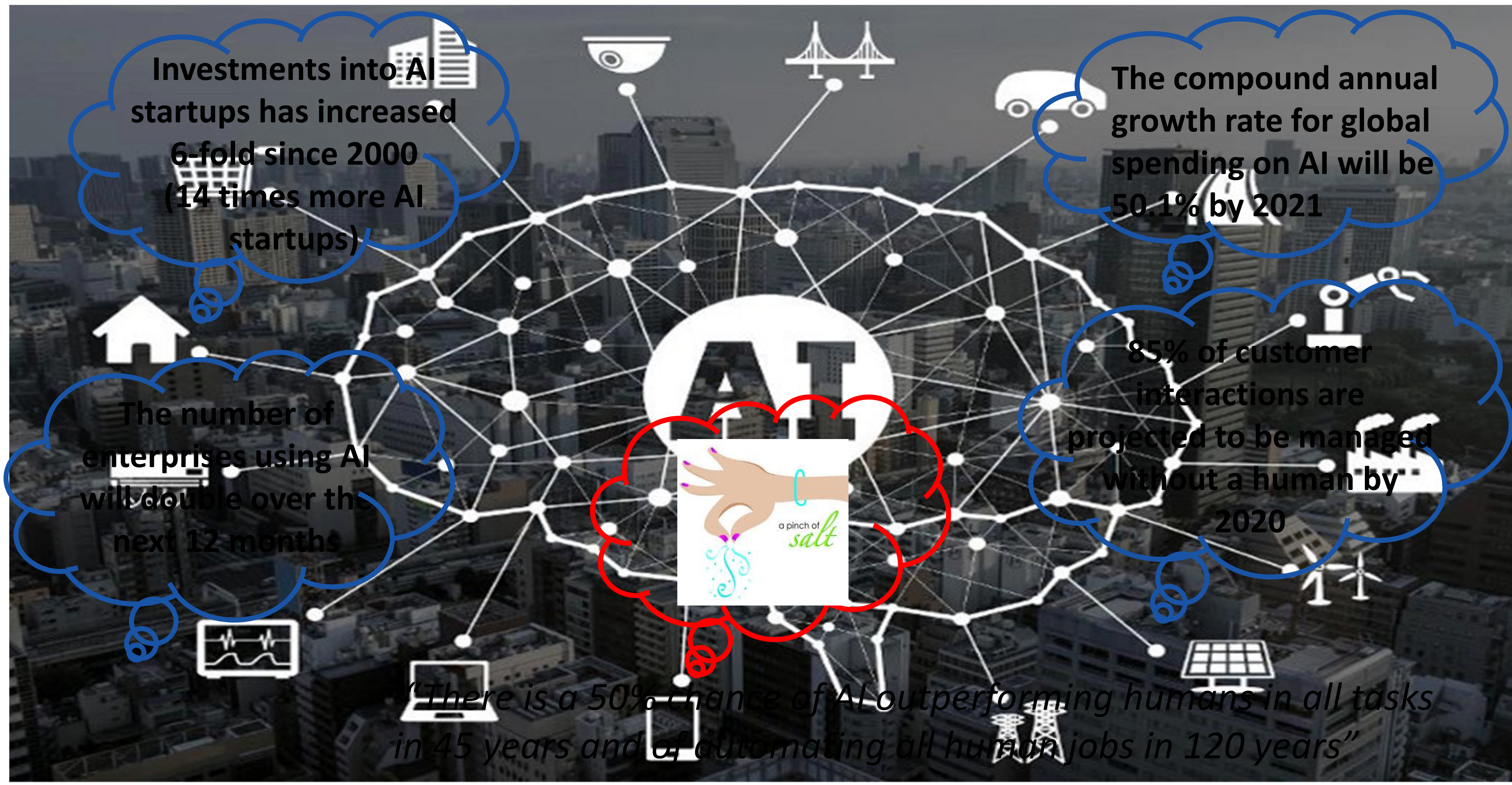
Main reasons for the AI/ML hype

➢ Sheer **impact!**



Main reasons for the AI/ML hype

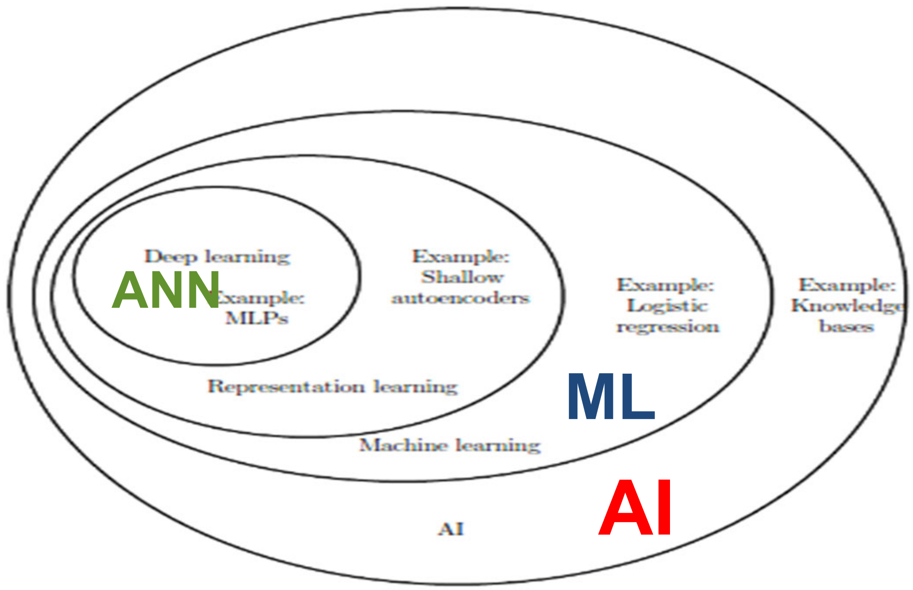
- Some facts and brave statements



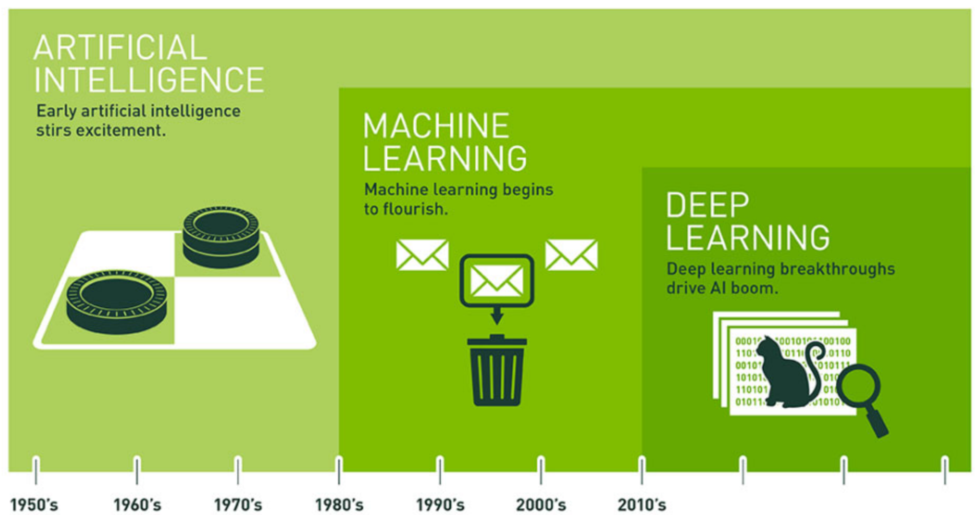
Main reasons for the AI/ML hype

- Impressive scope of applications and intriguingly good performance
- A huge opportunity for industry to improve products, services at a relatively low cost – large profit margin
- Renewed scientific interest following Deep Learning success
- Remarkable implications for other fields
- Good timing: data availability and growing compute power
- Vicious circle of expectations

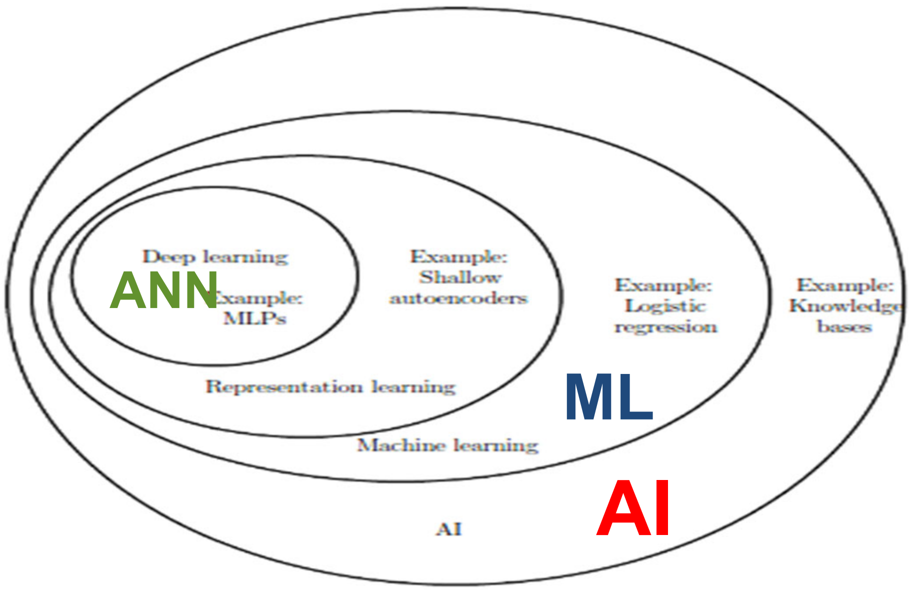
AI, ML, deep learning – what is the difference?



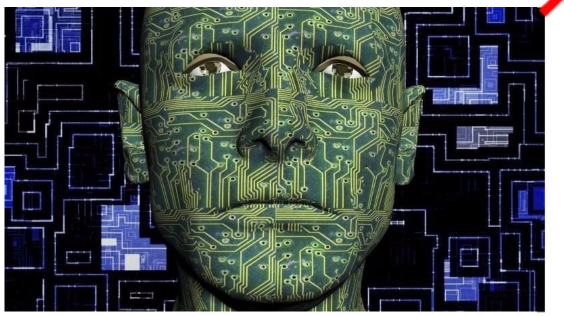
Goodfellow et al., 2016



AI, ML, deep learning – what is the difference?

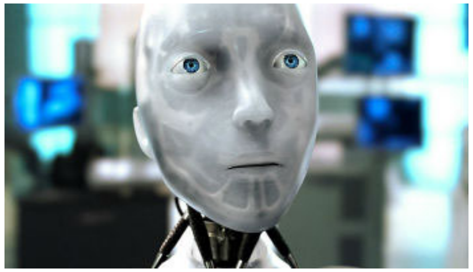


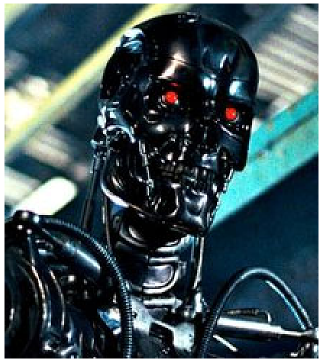
Goodfellow et al., 2016

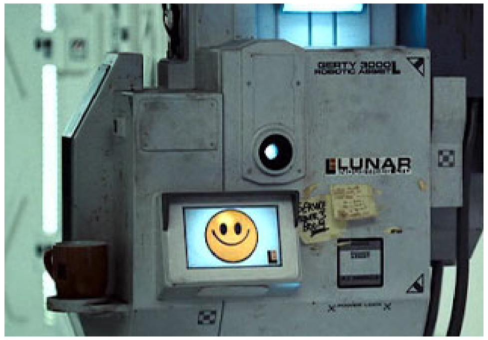


From General AI to Narrow AI

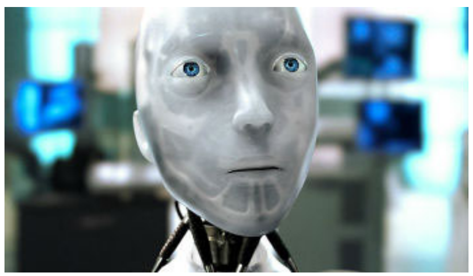
Artificial Intelligence – science fiction?







Artificial Intelligence – science fiction?



Central role of **intelligence**!

*"The science and engineering of making **intelligent** machines, especially **intelligent** computer programs".* John McCarthy

Defining AI

- Lack of precise and broadly accepted definition has likely facilitated the unrestrained growth at a very impressive speed
- Although the AI technology developments are fast, they are incremental in nature – gradual improvement

Defining AI

- Lack of precise and broadly accepted definition has likely facilitated the unrestrained growth at a very impressive speed
- Although the AI technology developments are fast, they are incremental in nature – gradual improvement
- How to measure intelligence? What are the key criteria?
  - complex phenomenon
  - interdisciplinary nature of the field

Interdisciplinary nature of AI as science

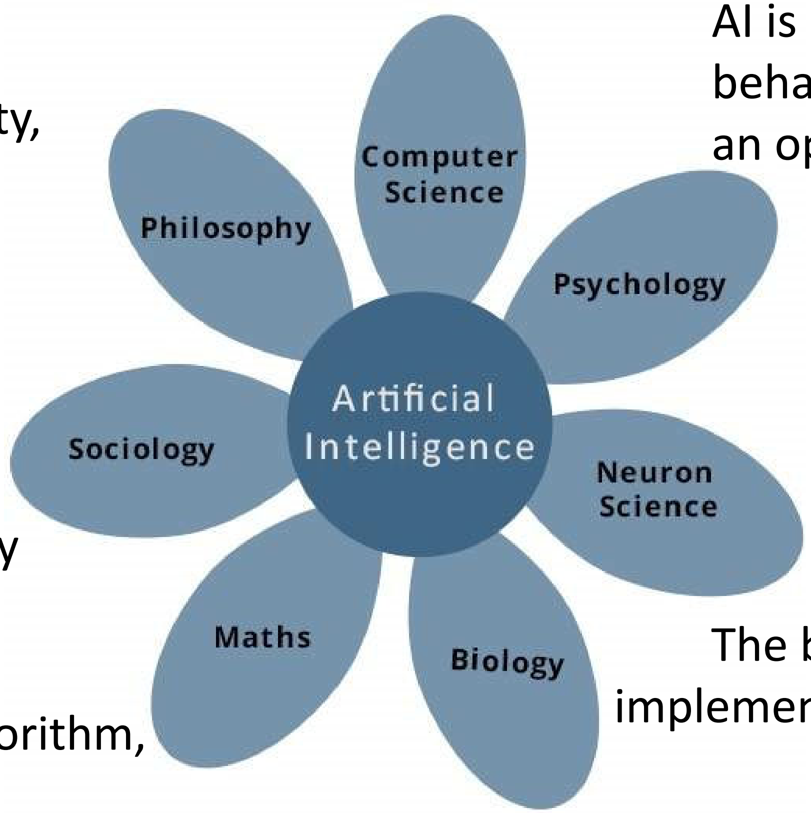
AI as a branch of computer science attempting to build machines capable of intelligent behaviour

Logic, rationality, reasoning

What is an intelligent behaviour in a group?

Economics – decision theory

Formal representations, algorithm, probability



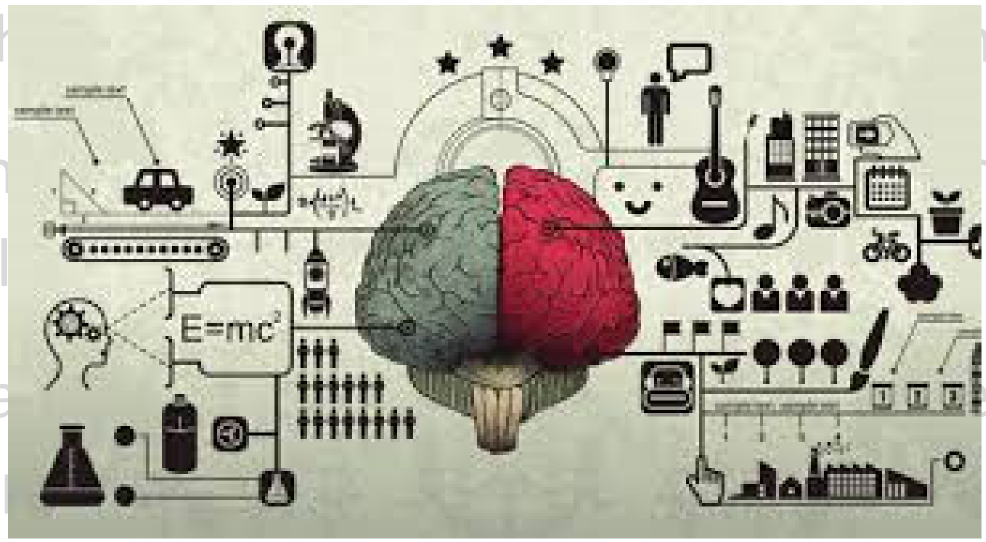
AI is helping us to understand human behaviour by recreating it and giving an opportunity to enhance it.

Linguistics –
How is language related to thought? Grammar, syntax.

The brain as an inspiration for the implementation framework for intelligence

Defining AI

- Lack of precise and broadly accepted definition has likely facilitated th



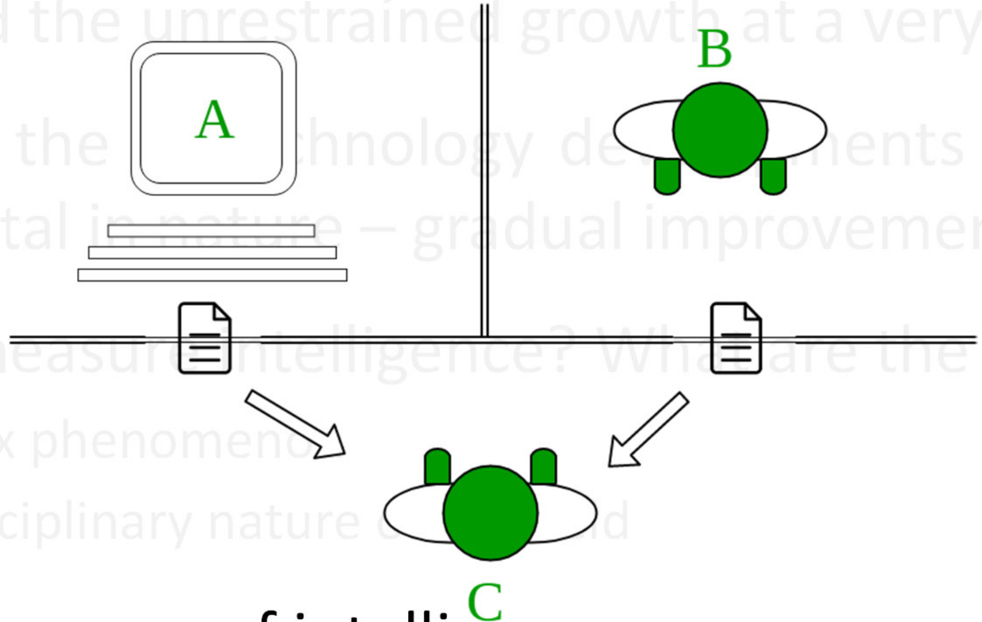
npressive speed

- Although th e fast, they are incremental
- How to mea y criteria?
  - complex pl
  - interdisciplinary nature of the field

- Human measure of intelligence
  - Unprecedented versatility, reasoning, planning capabilities, perception, understanding and generation of languages, creativity etc.
  - reference for benchmarking (still, is it only a set of abilities?)

Defining AI

Can machine A deceive human C that it communicates in a human-like way, just like B?



- Human measure of intelligence
  - Unprecedented versatility, reasoning, planning capabilities, perception, understanding and generation of languages, creativity etc.
  - reference for benchmarking (still, is it only a set of abilities?)
  - Turing test

How can AI be interpreted in the context of human intelligence?

Human approach to problem solving: based on abstract thought, high-level deliberate reasoning and pattern recognition

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think … *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning …" (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI …is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Russel and Norvig, 2010

How can AI be interpreted in the context of human intelligence?

Human approach to problem solving: based on abstract thought, high-level deliberate reasoning and pattern recognition

Thinking Humanly "The exciting new effort to make computers think ... *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ..." (Bellman, 1978)	**Thinking Rationally** "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	**Acting Rationally** "Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI ...is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Russel and Norvig, 2010

How can AI be interpreted in the context of human intelligence?

Central role of intelligence

*"The science and engineering of making **intelligent** machines, especially **intelligent** computer programs".* John McCarthy

Human approach to problem solving: based on abstract thought, high-level deliberate reasoning and pattern recognition

AI is a solution that appears to be intelligent and can often exceed the performance of humans. It is a broad description of any device that mimics human or intellectual functions (…)"

The continuing quest for machine intelligence

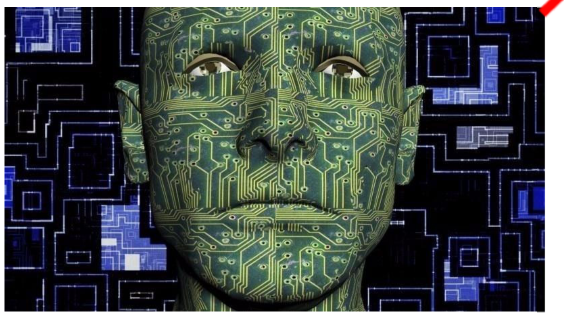
- ✓ Intelligence builds on powerful information processing, computations
- ✓ Ultimately, we desire to improve technology and there is an inherent need for continuous optimization

Are we on the course for intelligent machines?

The continuing quest for machine intelligence

- ✓ Intelligence builds on powerful information processing, computations
- ✓ Ultimately, we desire to improve technology and there is an inherent need for continuous optimization

Are we on the course for intelligent machines?



From General AI to Narrow AI

A brief historical overview

the gestation period

1940-1955

Turing test for intelligence

McCulloch & Pitts’ model of the brain (Boolean circuit)

Hebbian learning

A brief historical overview

the birth of AI as
a branch of CS

the gestation
period

1956

1940-1955

AI as *"the science and engineering of making intelligent machines"* (John McCarthy)

The core idea of duplicating human faculties such as creativity, self-improvement, and language use.

A brief historical overview

the gestation period

1940-1955

the birth of AI as a branch of CS

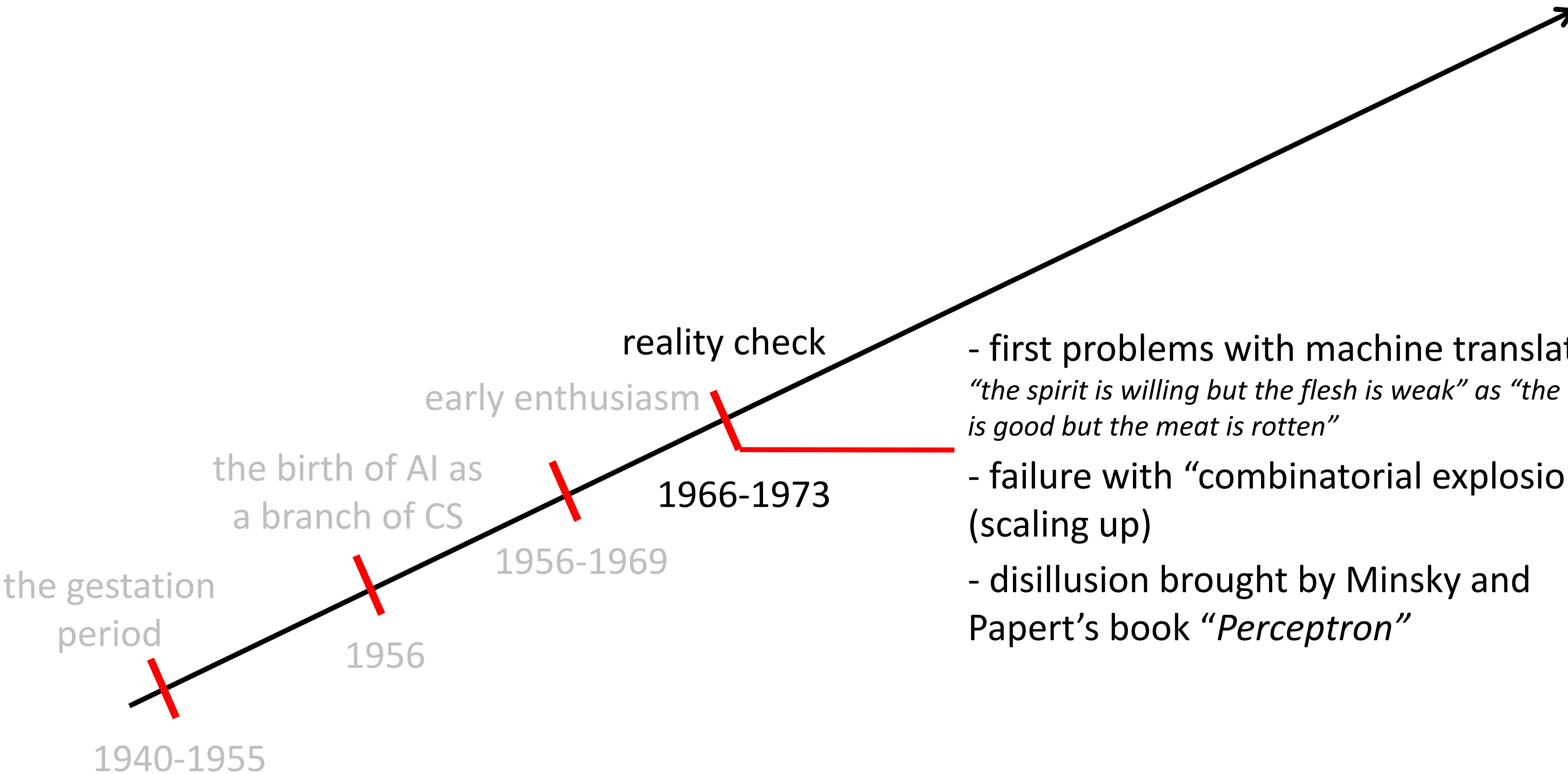
1956

early enthusiasm

1956-1969

- AI programming language **LISP**
- symbolic computing
- Logic Theorist, Geometry Theorem Prover, Advice Taker
- ”Programs with Common Sense” (no need to reprogram)
- perceptrons (simple neural networks)
- IBM computing infrastructure
- Minsky’s illusion that “machines will be capable, within 20 years, of doing any work human can do”

A brief historical overview

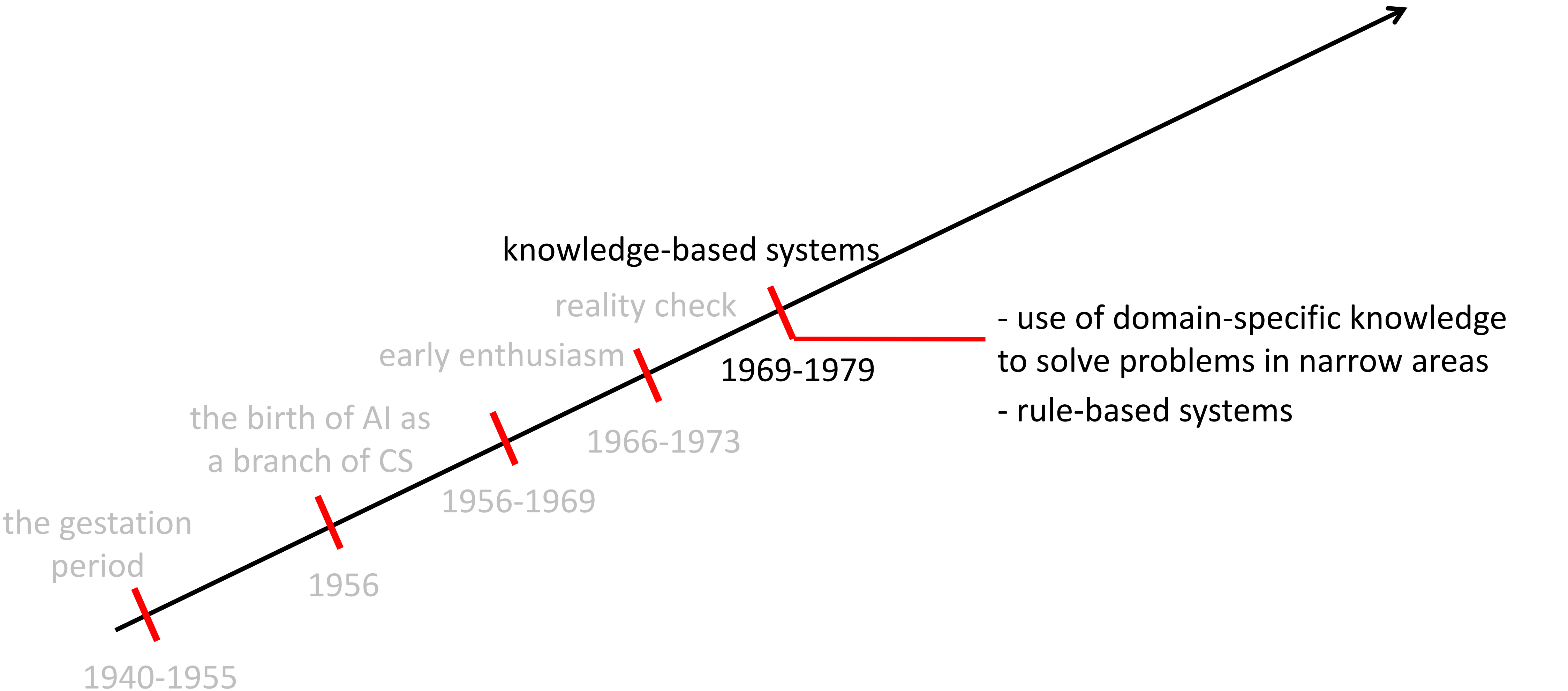


- first problems with machine translation
"the spirit is willing but the flesh is weak" as "the vodka is good but the meat is rotten"

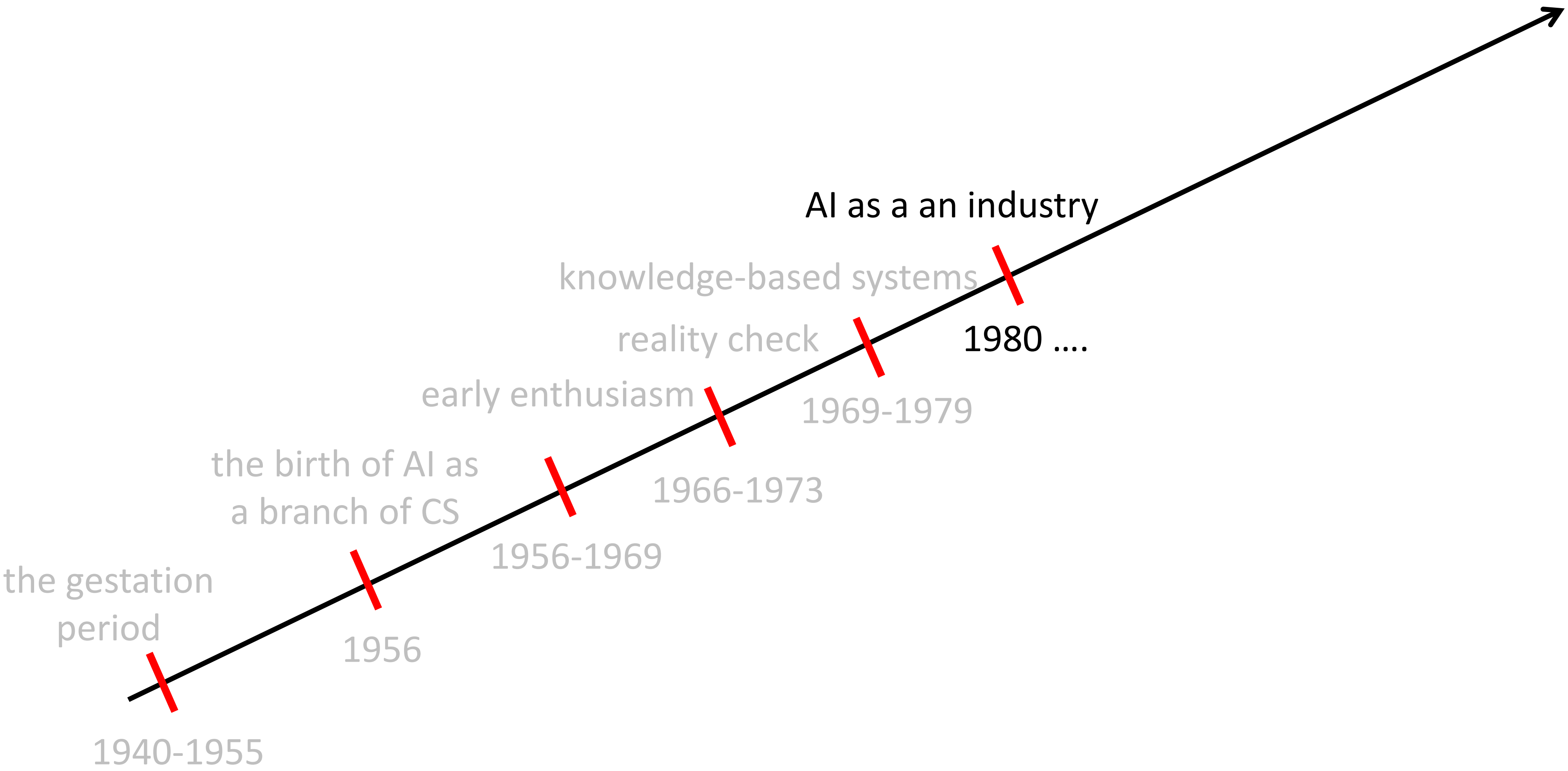
- failure with "combinatorial explosion" (scaling up)

- disillusion brought by Minsky and Papert's book "*Perceptron*"

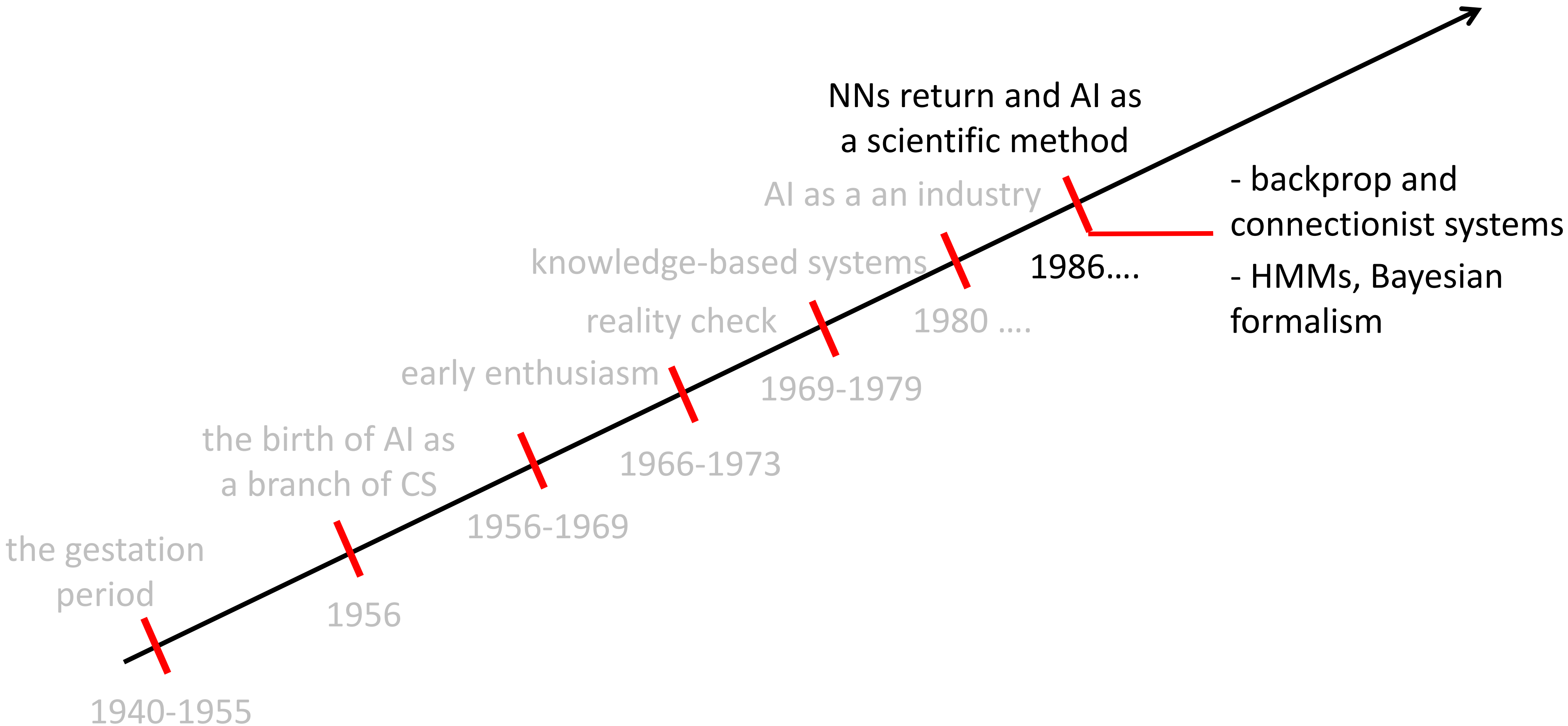
A brief historical overview



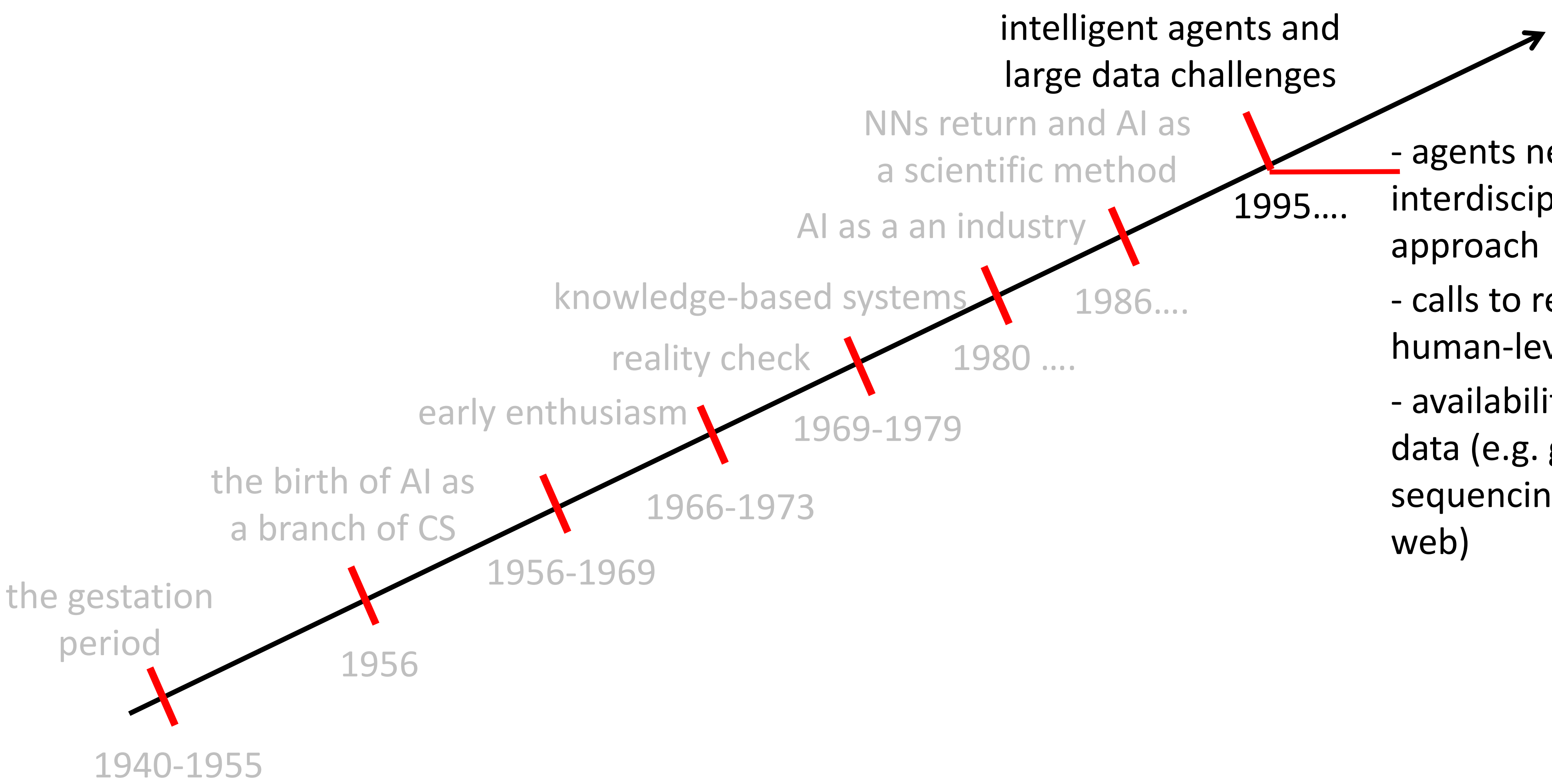
A brief historical overview



A brief historical overview



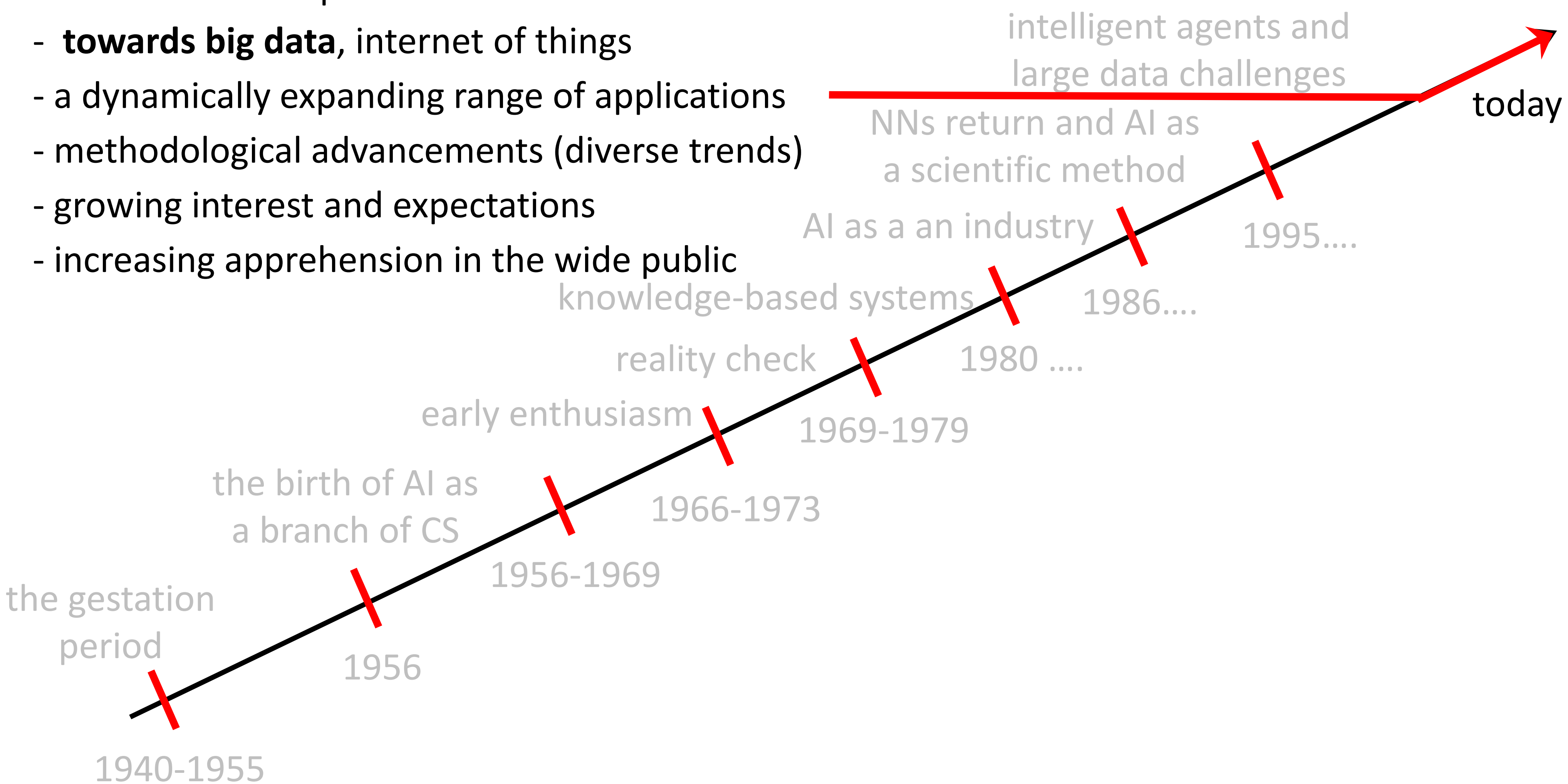
A brief historical overview



- agents need an interdisciplinary approach
- calls to return to human-level AI
- availability of large data (e.g. genome sequencing, test, web)

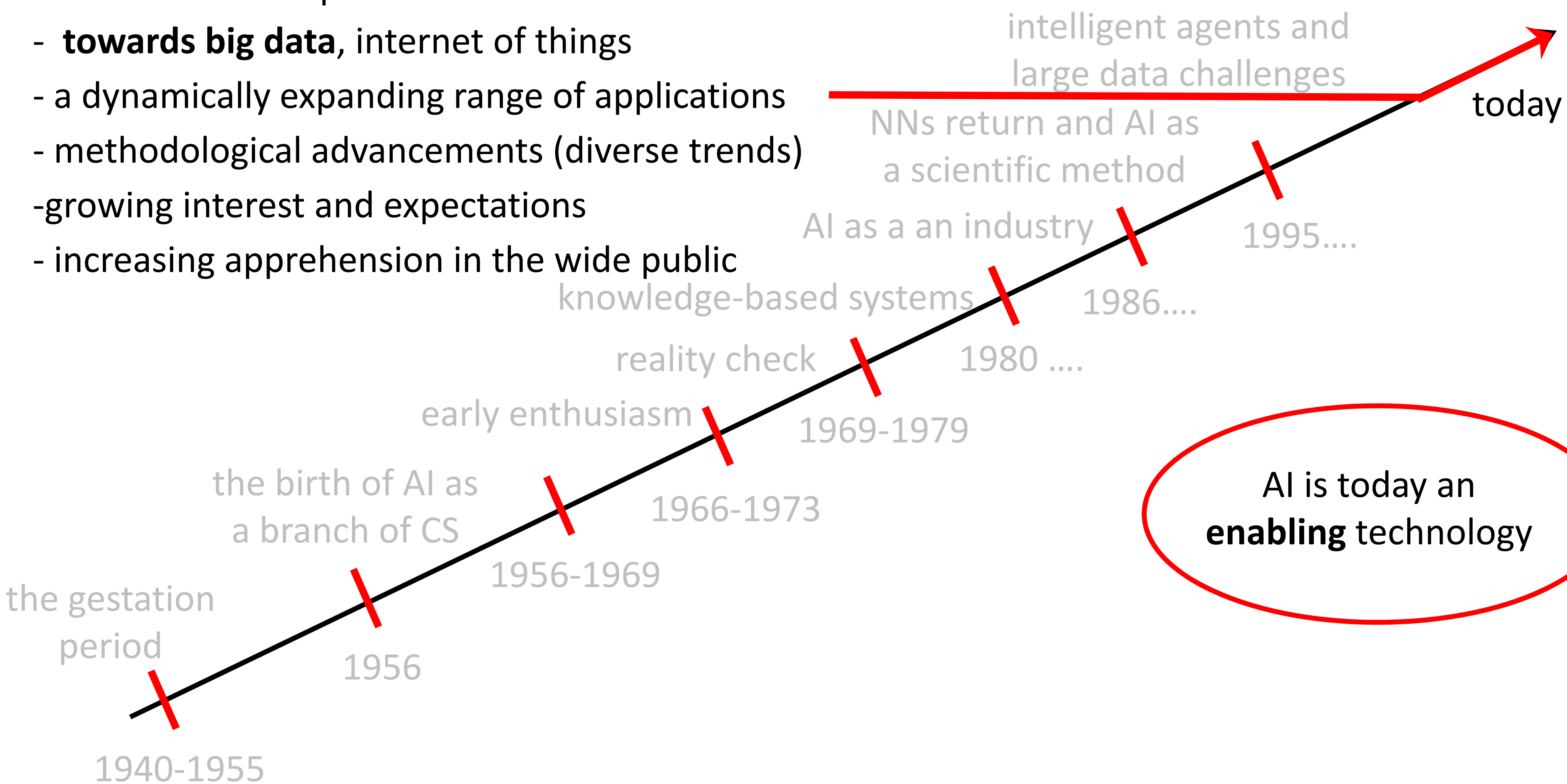
A brief historical overview

- **more powerful computational resources**, various hardware platforms
- **towards big data**, internet of things
- a dynamically expanding range of applications
- methodological advancements (diverse trends)
- growing interest and expectations
- increasing apprehension in the wide public



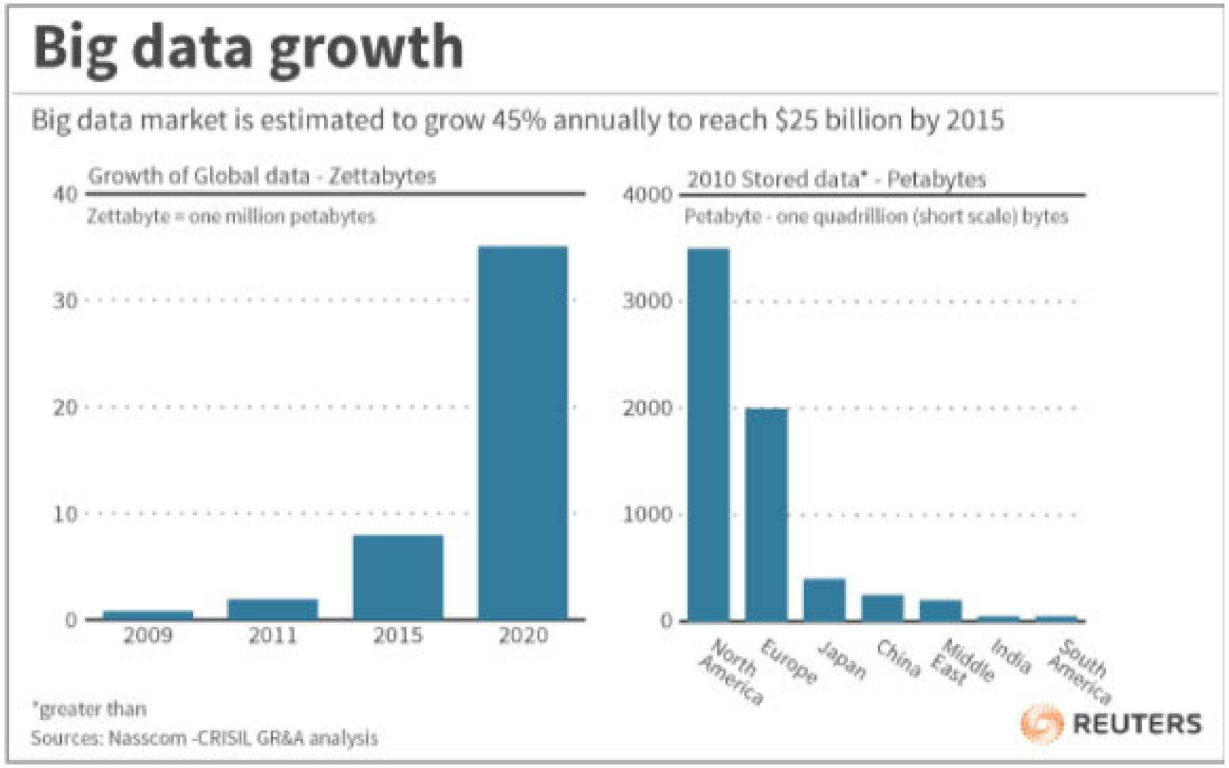
A brief historical overview

- **more powerful computational resources**, various hardware platforms
- **towards big data**, internet of things
- a dynamically expanding range of applications
- methodological advancements (diverse trends)
- growing interest and expectations
- increasing apprehension in the wide public

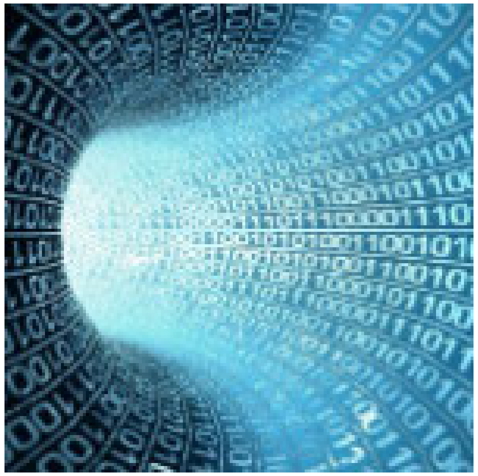


AI is today an **enabling** technology

The age of Big Data

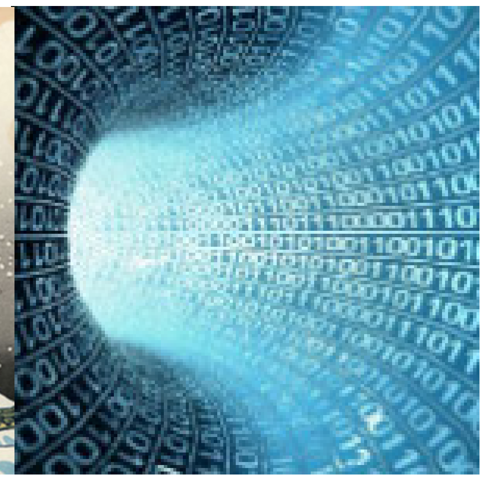


(From http://www.theopenstrategist.com/2012/10/big-data-growthchart.html)



The age of Big Data

on the brink of *"data revolution"*?

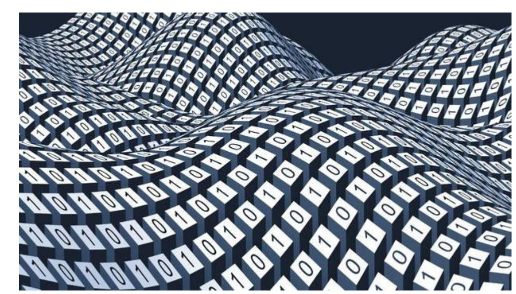
"Drowning in information but starved for knowledge."

John Naisbitt

"With the explosion of data generation, getting optimal solutions to data driven problems is increasingly becoming a challenge"

A. Kumar Kar

The predictive power of Big Data paves the way for *a new approach to understanding the world and making decisions*



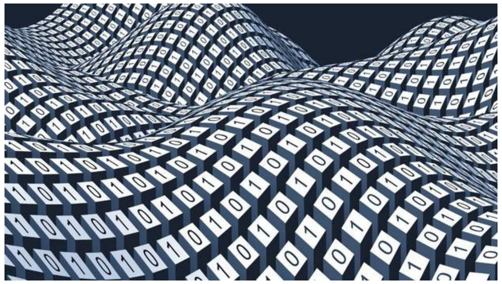
The age of Big Data

on the brink of *“data revolution”*?

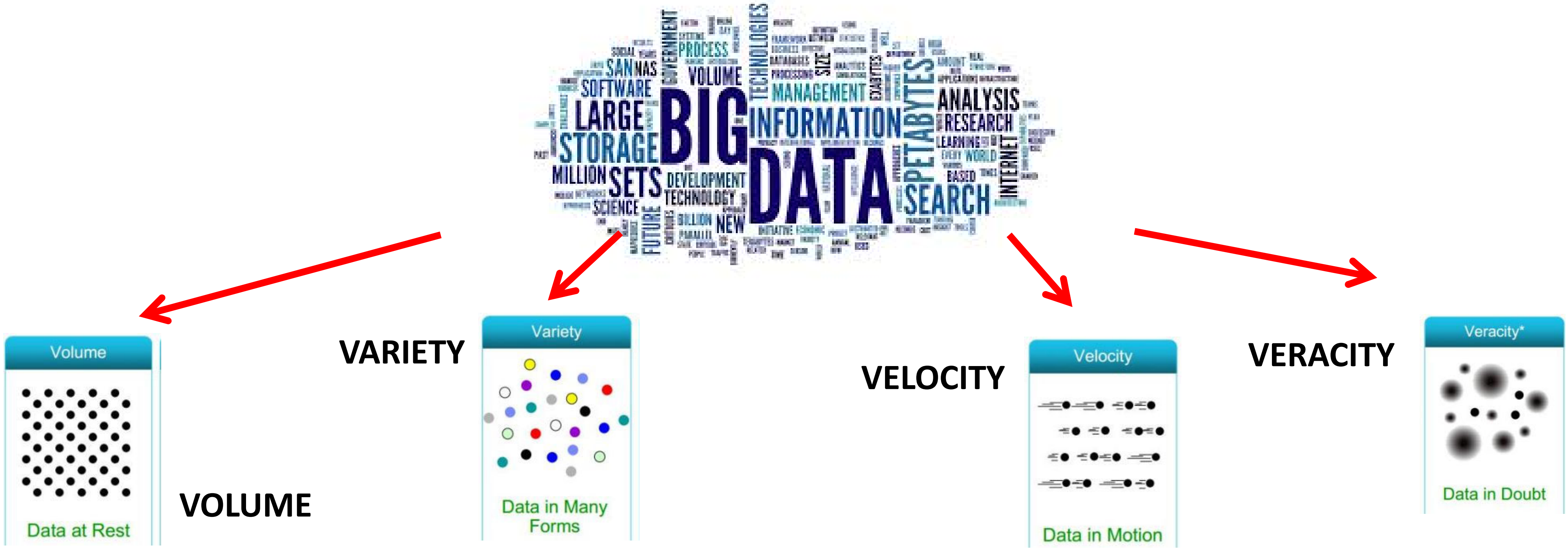
“Drowning in information but starved for knowledge.”

John Naisbitt

The key challenge to devise insightful and scalable computational methods to *summarise*, *describe* and *understand* **large-scale unstructured and multimodal data**.



General characteristics of Big Data, *four Vs*



Data Science: Data meets Machine Learning

PDC Summer School, Introduction to ML

Pawel Herman, Aug 30 2019, KTH

What is Machine Learning?

"The scientific discipline that explores the construction and study of algorithms that can ***learn from data****."*
(*wikipedia.com*)

"A computer's way of ***learning from examples****."*
(*businessinsider.com*)

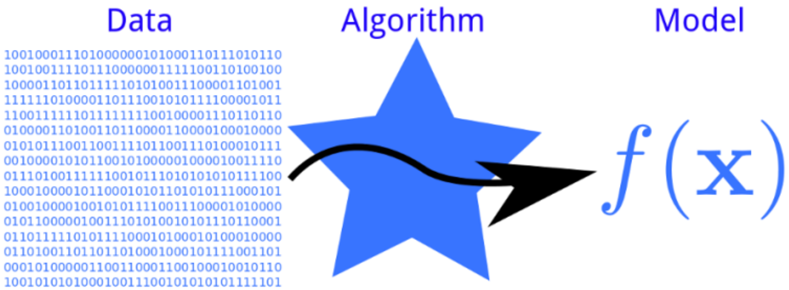
"The science of getting computers to ***act without being explicitly programmed****."*
(*Arthur Lee Samuel*)

"The field of machine learning is concerned with the question of how to construct computer programs that ***automatically improve with experience****."*
(*Tom Mitchell*)

"A computer program is said to ***learn from experience*** *E with respect to some class of* ***tasks*** *T and performance* ***measure*** *P, if its performance at tasks in T, as measured by P, improves with experience E."*
(*Tom Mitchell*)

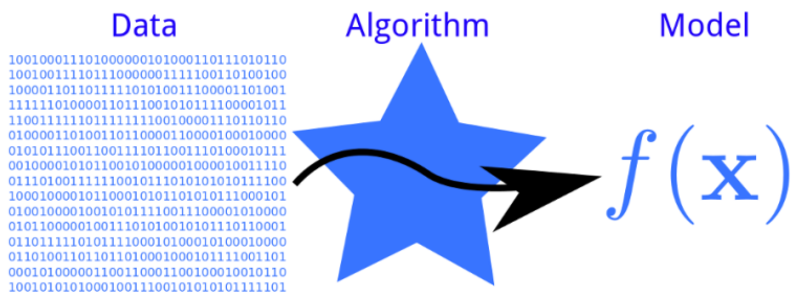
Machine Learning as a data driven approach

Learning from data



Machine Learning as a data driven approach

Learning from data



Classical approach to solving problems with computer algorithms

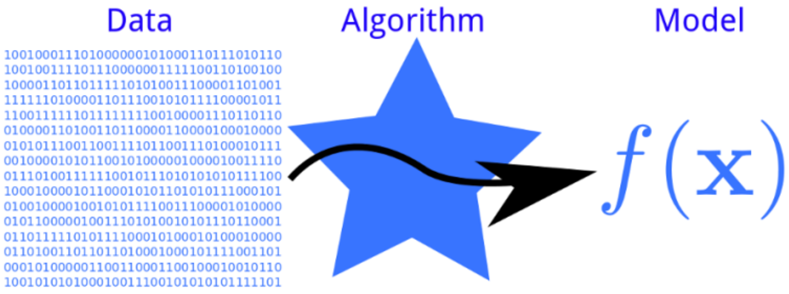
Data → Computer → Output
Program, Algorithm → Computer

Machine Learning approach – data driven approach

Data + Corresponding Output → Computer → Program

Machine Learning as a data driven approach

Learning from data



Capability to *acquire* and *integrate* the *knowledge* automatically from **data** and improve performance by learning

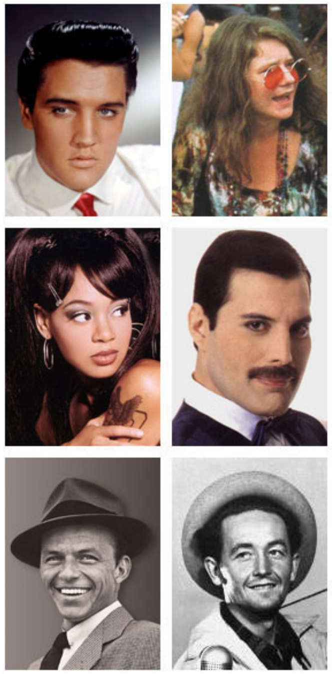
Scope for continuous self-improvement, adaptation to changing environment

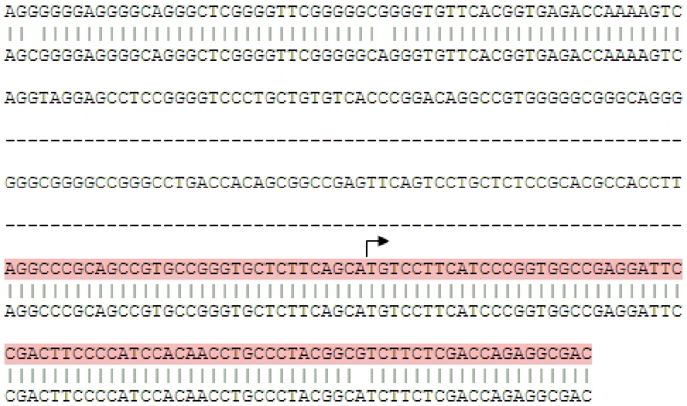
Discovery of patterns in complex data – data mining

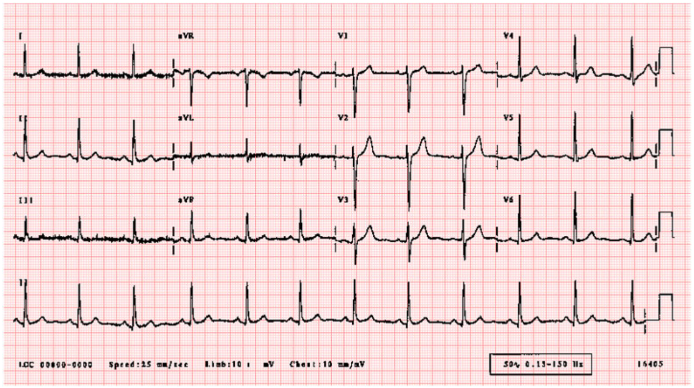
Generating predictions based on previous experience (data driven)

The art of searching for patterns

The problem of searching for patterns in data is a fundamental one and has a long and successful history.







The art of searching for patterns

The problem of searching for patterns in data is a fundamental one and has a long and successful history.

- the ubiquity of patterns in the surrounding world
- it could be an object, process, event etc. and is described by attributes (features)
- the importance of finding and categorising objects

ML for pattern recognition

- the ubiquity of patterns in the surrounding world
- it could be an object, process, event etc. and is described by attributes (features)
- the importance of finding and categorising objects

Examples of real-world problems involving *Pattern Recognition*

Identify risk factors for cancer based on clinical, demographic variables

Diagnose dementia based on electrophys. data

Predict the price of a stock on the basis of company performance measures and economic data

Predict short-term electricity demand and wind power production capabilities

ML for pattern recognition

- the ubiquity of patterns in the surrounding world
- it could be an object, process, event etc. and is described by attributes (features)
- the importance of finding and categorising objects

Pattern recognition amounts to grouping objects and often assigning categories (classes) to the objects.

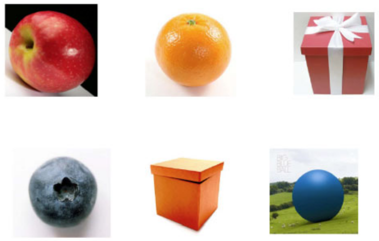
clustering

discrete classes - ***classification***

continuous labels - ***regression***

General taxonomy of ML problems

unsupervised learning

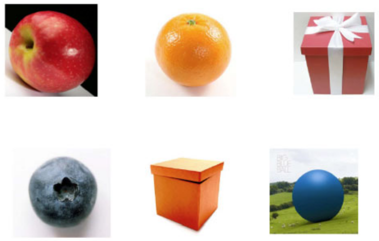


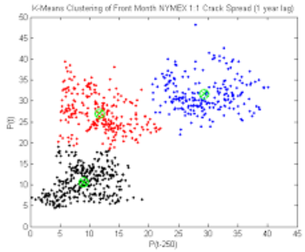
- grouping objects based on the similarity of their attributes
- clustering
- anomaly detection
- dimensionality reduction

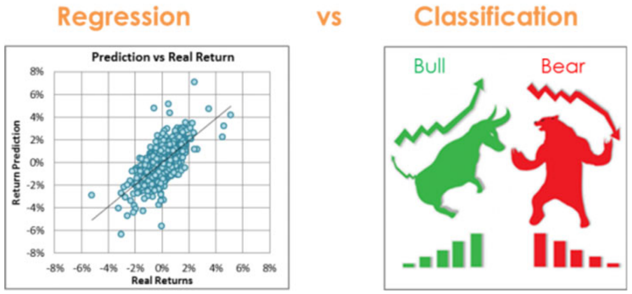
General taxonomy of ML problems

unsupervised learning

supervised learning



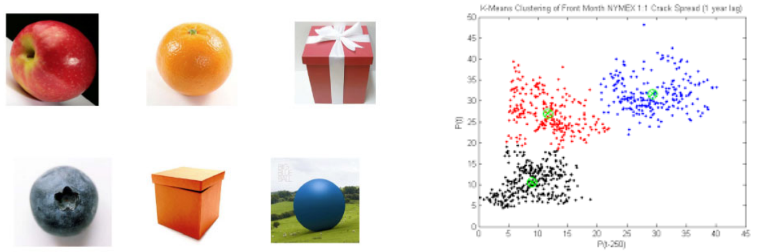




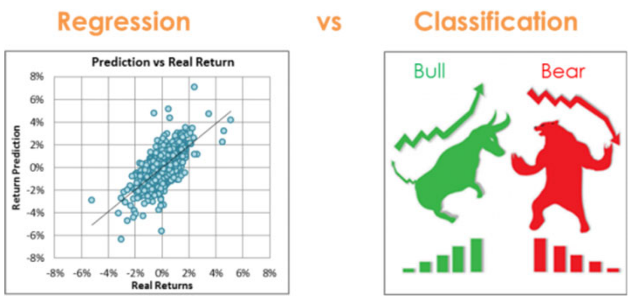
- mapping object attributes to other descriptors
(mapping one data part onto another)
- classification (discrete labels)
- regression (continuous target/output variables)

General taxonomy of ML problems

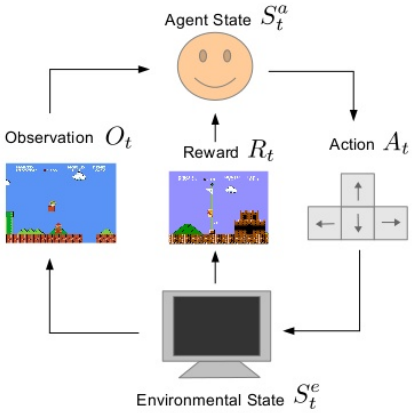
unsupervised learning



supervised learning



reinforcement learning



- no supervisor
- only delayed feedback: reward or penalty
- agent's actions affect the environment

Assumptions for ML process

1. Pattern exists.
2. We have no underlying mathematical model / explicit problem formulation.
3. There is data.

Basic premise of learning:

To uncover an underlying process using a set of observations

A general notion of a learning problem

Unknown function *f*

$$f : \mathbf{X} \rightarrow T$$

$$T = f(\mathbf{X}) + \varepsilon$$

Instance space – data

$$(\mathbf{X}, T) : \{(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2), ..., (\mathbf{x}_N, t_N),$$

A general notion of a learning problem

Unknown function *f* $f : \mathbf{X} \to T$

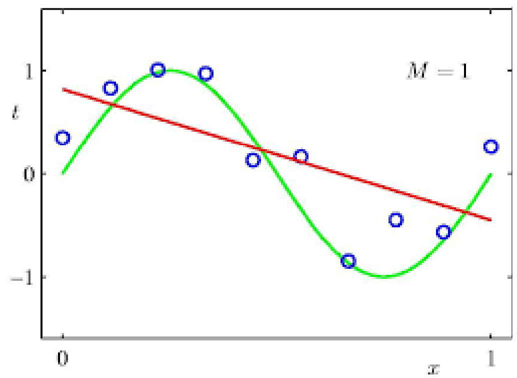
$$T = f(\mathbf{X}) + \varepsilon$$

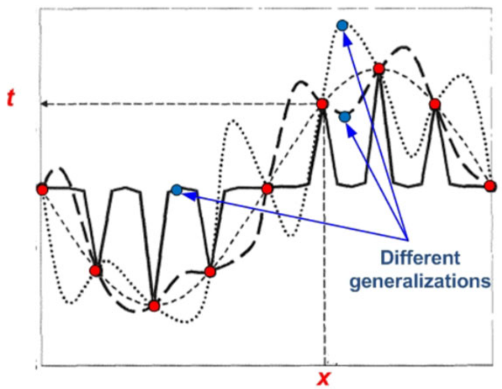
Instance space – data

$$(\mathbf{X}, T) : \{(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2), \ldots, (\mathbf{x}_N, t_N),$$

Hypothesis space

$$\mathcal{H} = \{f : \mathbf{X} \to T\}$$





A general notion of a learning problem

Unknown function *f*

$$f : \mathbf{X} \to T$$

$$T = f(\mathbf{X}) + \varepsilon$$

Instance space – data

$$(\mathbf{X}, T) : \{(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2), \ldots, (\mathbf{x}_N, t_N),$$

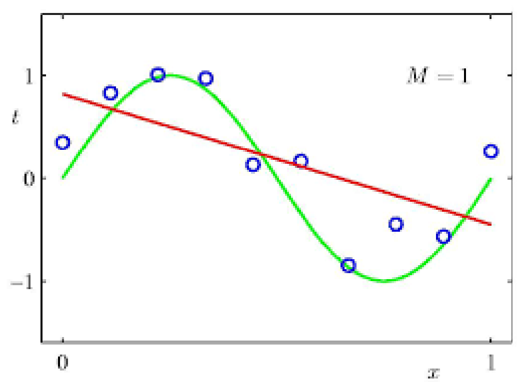
Learning

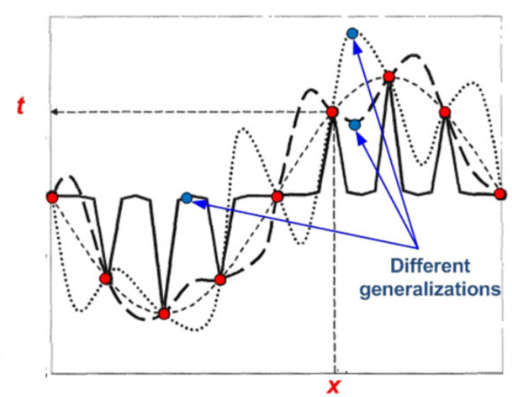
$$g = f$$

$$g \in \mathcal{H}$$

Hypothesis space

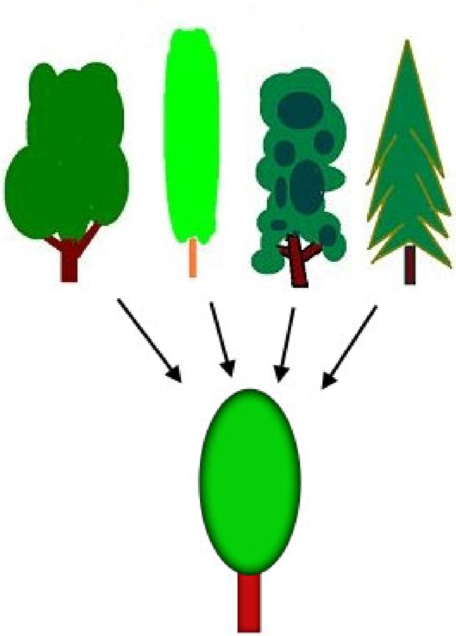
$$\mathcal{H} = \{f : \mathbf{X} \to T\}$$

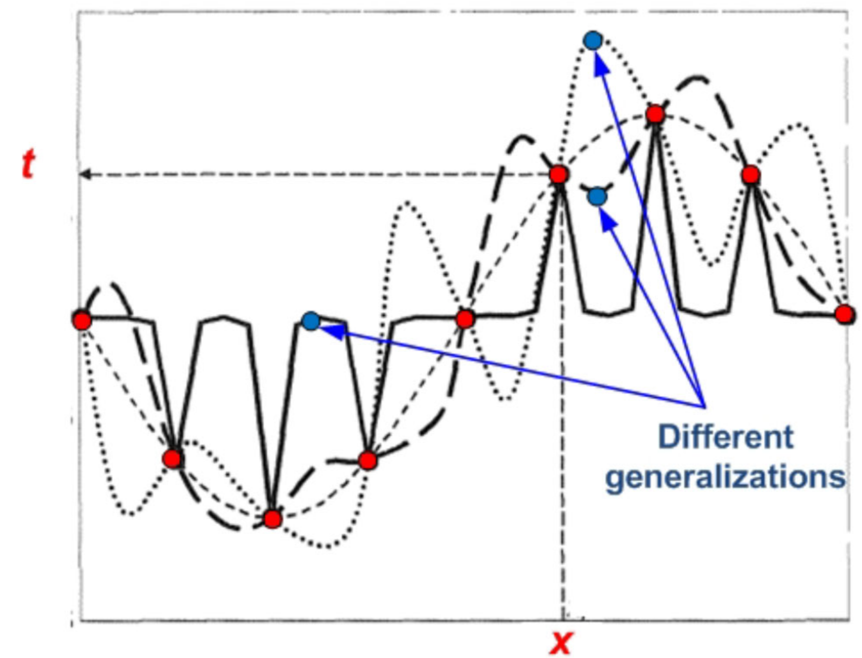




Generalisation

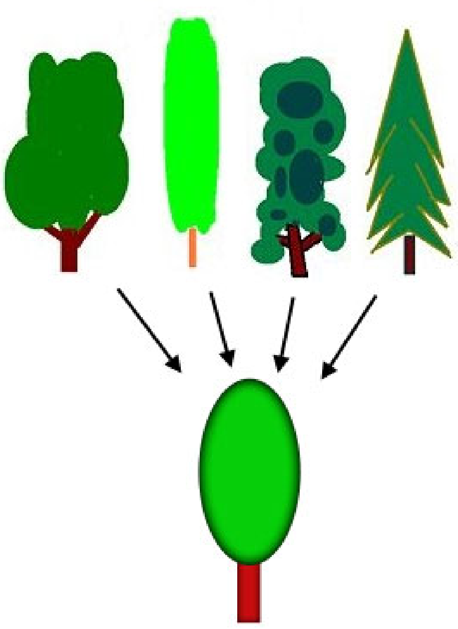
Capability to apply knowledge to *a new situation*, make a reliable <u>prediction</u> for *unseen* input

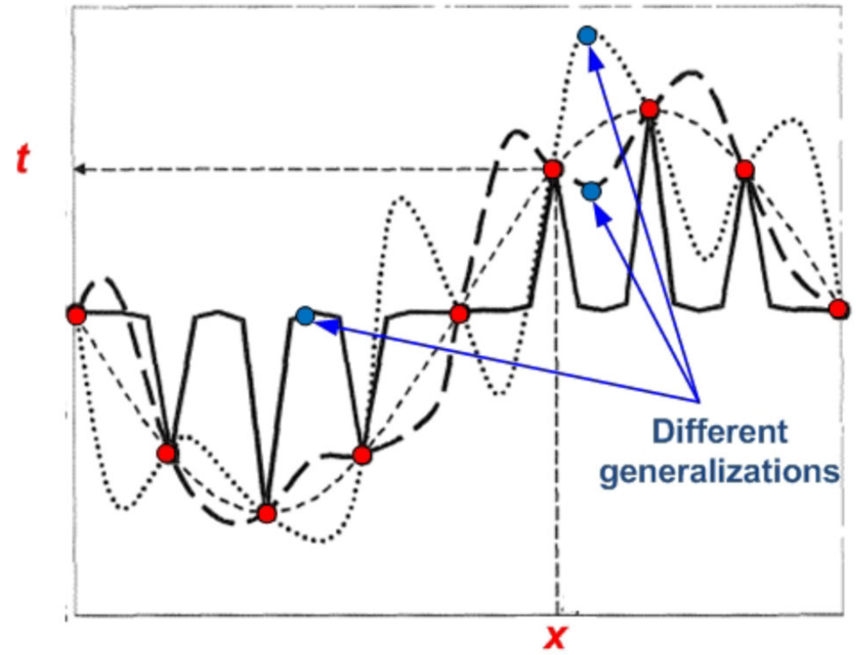




Generalisation

Capability to apply knowledge to *a new situation*, make a reliable <u>prediction</u> for *unseen* input





The risk for underfitting or overfitting!

How do we assess the generalisation?

The effectiveness of NN model $F(\mathbf{x}, \mathbf{w})$ can be defined as an estimator of the regression $f = \mathbb{E}[t \mid \mathbf{x}]$ for D :

$$\mathbb{E}_D\left[\left(F(\mathbf{x},\mathbf{w}) - f(\mathbf{x})\right)^2\right]$$

How do we assess the generalisation?

The effectiveness of NN model $F(\mathbf{x},\mathbf{w})$ can be defined as an estimator of the regression $f = \mathbb{E}[t \mid \mathbf{x}]$ for D :

$$\mathbb{E}_D\left[\left(F(\mathbf{x},\mathbf{w}) - f(\mathbf{x})\right)^2\right]$$

All possible realisations of data samples

How do we assess the generalisation?

The effectiveness of NN model $F(\mathbf{x},\mathbf{w})$ can be defined as an estimator of the regression $f = \mathbb{E}[t \mid \mathbf{x}]$ for D :

$$\mathbb{E}_D\left[\left(F(\mathbf{x},\mathbf{w}) - f(\mathbf{x})\right)^2\right]$$

But we do not have access to all data, *D!*

How do we assess the generalisation?

Maximum likelihood approach

- We rely on empirical error as an estimate of the generalization error – the problem of a data sample

$$R_{\mathrm{emp}}[F] = \frac{1}{N}\sum_{i=1}^{N}(t_i - F(\mathbf{x}_i, \mathbf{w}))^2$$

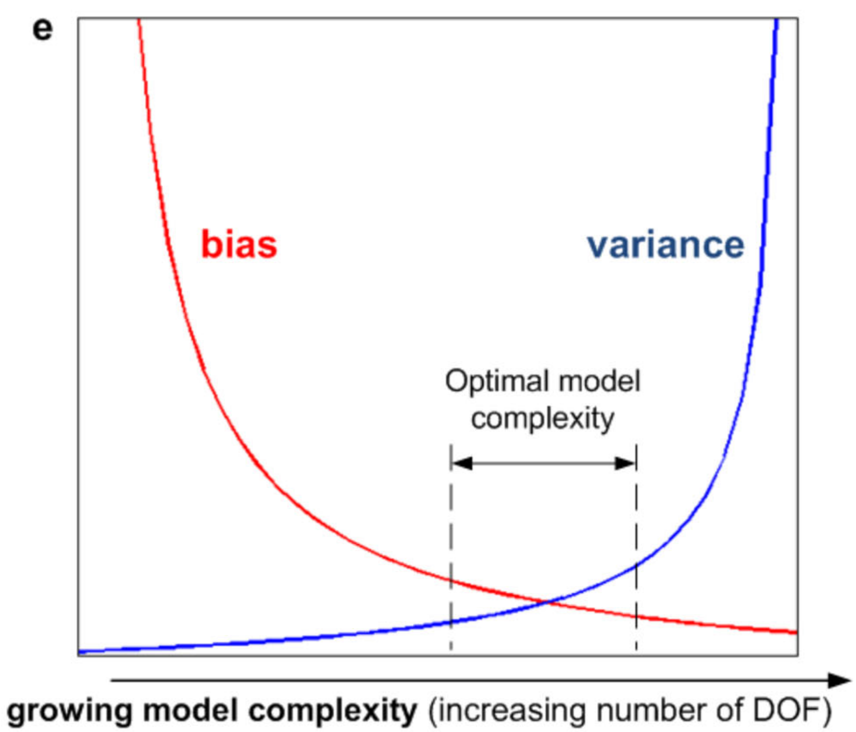
How do we assess the generalisation?

Maximum likelihood approach

- We rely on empirical error as an estimate of the generalization error – the problem of a data sample
- Bias-variance dilemma
  - Model selection
  - Occam’s razor

$$\mathbb{E}_D\left[\left(F(\mathbf{x},\mathbf{w})-f(\mathbf{x})\right)^2\right]=$$

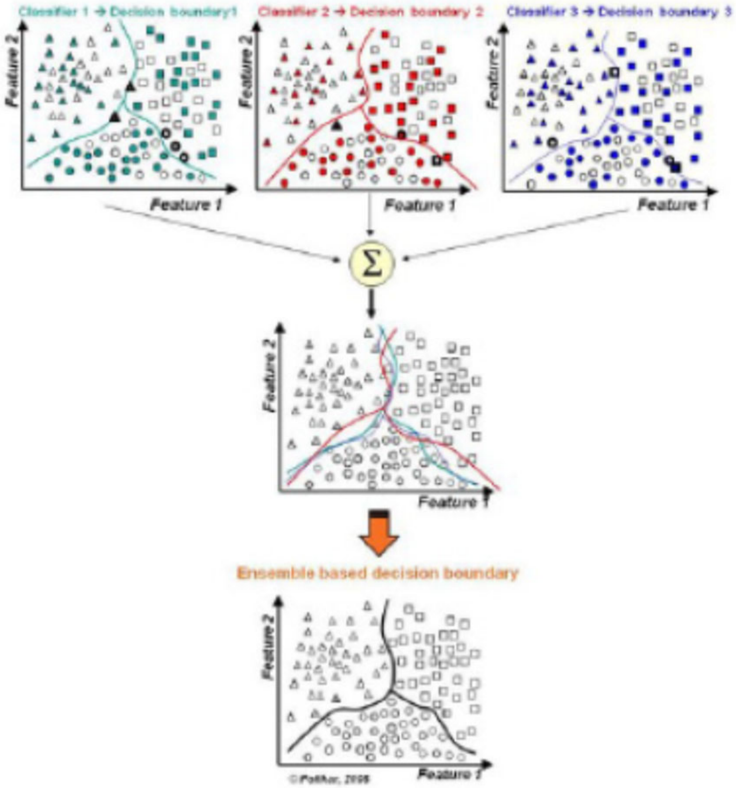
$$\left(\mathbb{E}_D[F(\mathbf{x},\mathbf{w})]-f(\mathbf{x})\right)^2+\mathbb{E}_D\left[\left(F(\mathbf{x},\mathbf{w})-\mathbb{E}_D[F(\mathbf{x},\mathbf{w})]\right)^2\right]$$



How do we assess the generalisation?

Maximum likelihood approach

- We rely on empirical error as an estimate of the generalization error – the problem of a data sample
- Bias-variance dilemma
  - Model selection
  - Occam's razor
  - Ensemble learning to reduce variance



How do we assess the generalisation?

Maximum likelihood approach

- We rely on empirical error as an estimate of the generalization error – the problem of a data sample
- Bias-variance dilemma
- Out-of-sample estimate of generalisation – simulation of unseen scenarios (out-of-sample vs in-sample performance)
- Data sampling techniques to improve an estimate of the generalization error
    - Data snooping
    - Sampling bias

How do we assess the generalisation?

Maximum likelihood approach

- We rely on empirical error as an estimate of the generalization error – the problem of a data sample
- Bias-variance dilemma
- Out-of-sample estimate of generalization – simulation of unseen scenarios (out-of-sample vs in-sample performance)
- Data sampling techniques to improve an estimate of the generalization error

How can we get away from max-likelihood?

Probabilistic perspective in a Bayesian framework

Philosophy of **Bayesian approach**

- Uncertainty is ubiquitous – describe all model components with probabilistic objects (*distributions, not point estimates*)
- Apply Bayesian machinery to propagate uncertainty

$$p(\mathbf{w} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \mathbf{w})\, p(\mathbf{w})}{p(\mathcal{D})}$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

Probabilistic perspective in a Bayesian framework

Philosophy of **Bayesian approach**

- Uncertainty is ubiquitous – describe all model components with probabilistic objects (*distributions, not point estimates*)
- Apply Bayesian machinery to propagate uncertainty
- Combine uncertain knowledge with data to reduce uncertainty (based on evidence from observations)
- Two levels of inference:
  parameter estimation and model selection

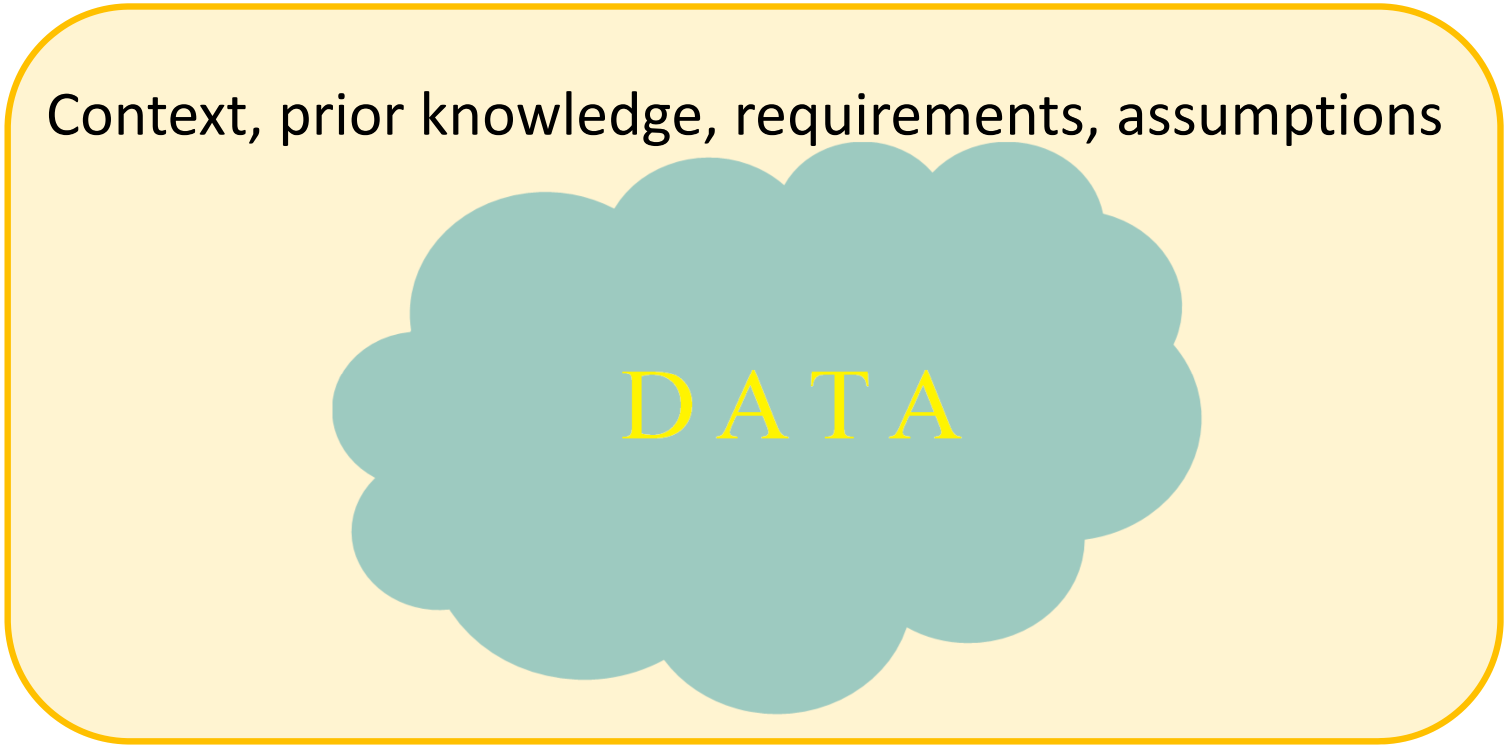
Machine Learning process –
from pattern recognition to predictions

1. **Identification of the application domain (what is the prior knowledge?)**

Context, prior knowledge, requirements, assumptions –
PROBLEM UNDERSTANDING

Machine Learning process –
from pattern recognition to predictions

1. Identification of the application domain (what is the prior knowledge?)
2. **Selection of datasets.**



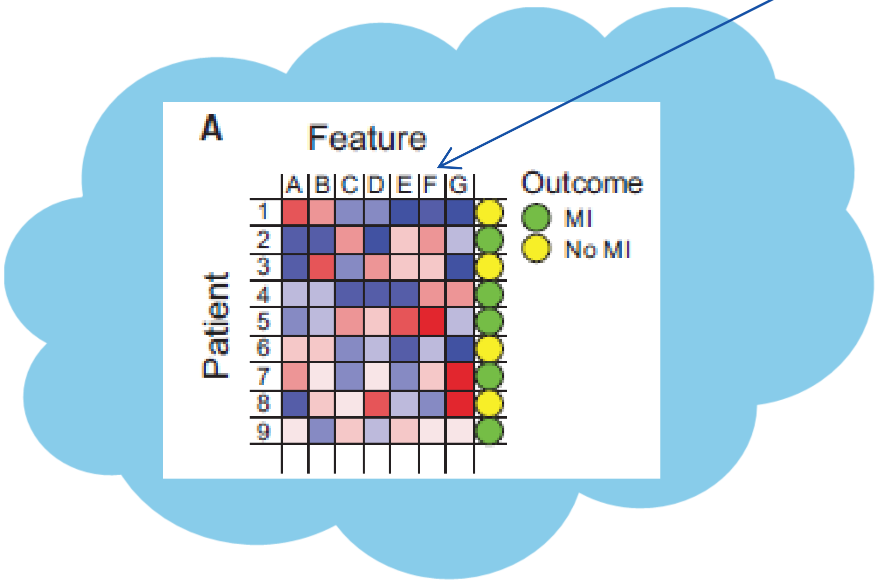
Machine Learning process –
from pattern recognition to predictions

1. Identification of the application domain (what is the prior knowledge?)
2. Selection of datasets.
3. **Data cleaning – preprocessing.**
   removing outliers, reducing noise etc.

Machine Learning process –
from pattern recognition to predictions

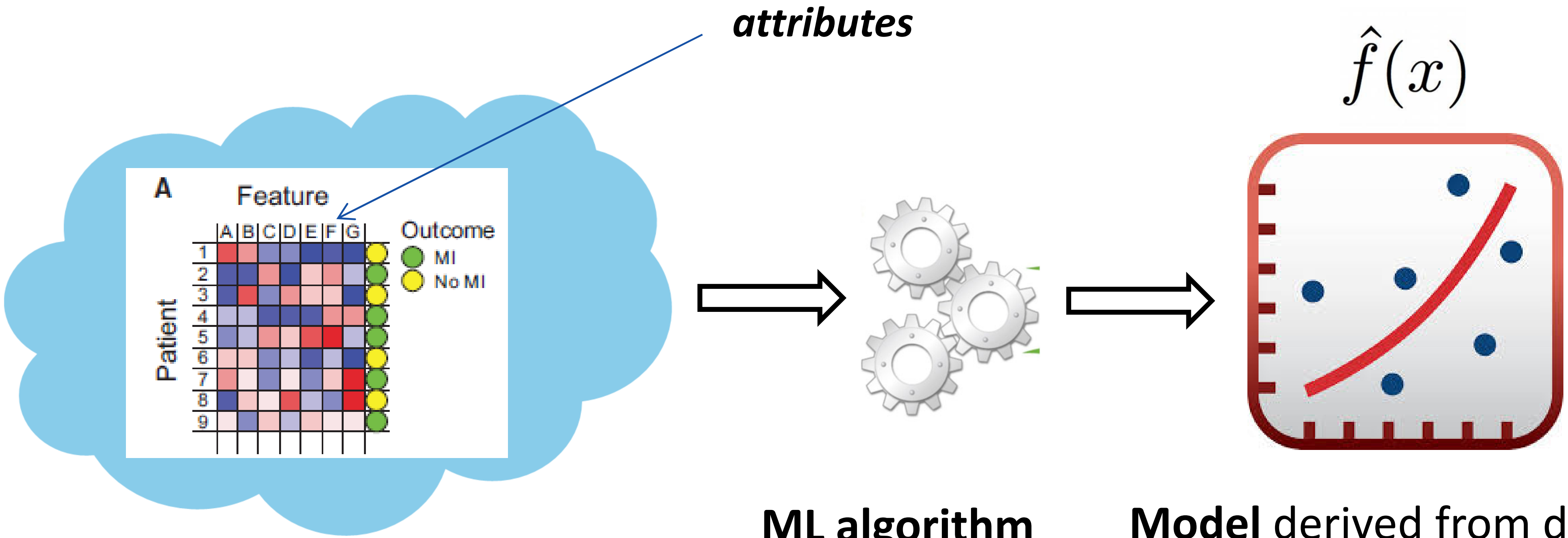
4. **Data reduction, feature selection** (e.g. subselection of EHR attributes)

attributes
e.g. blood pressure, glucose level, age



Machine Learning process –
from pattern recognition to predictions

5. **Building a data model,** extracting **patterns** and matching them with the related **outcomes (*historical data*)** using a selected **ML algorithm**

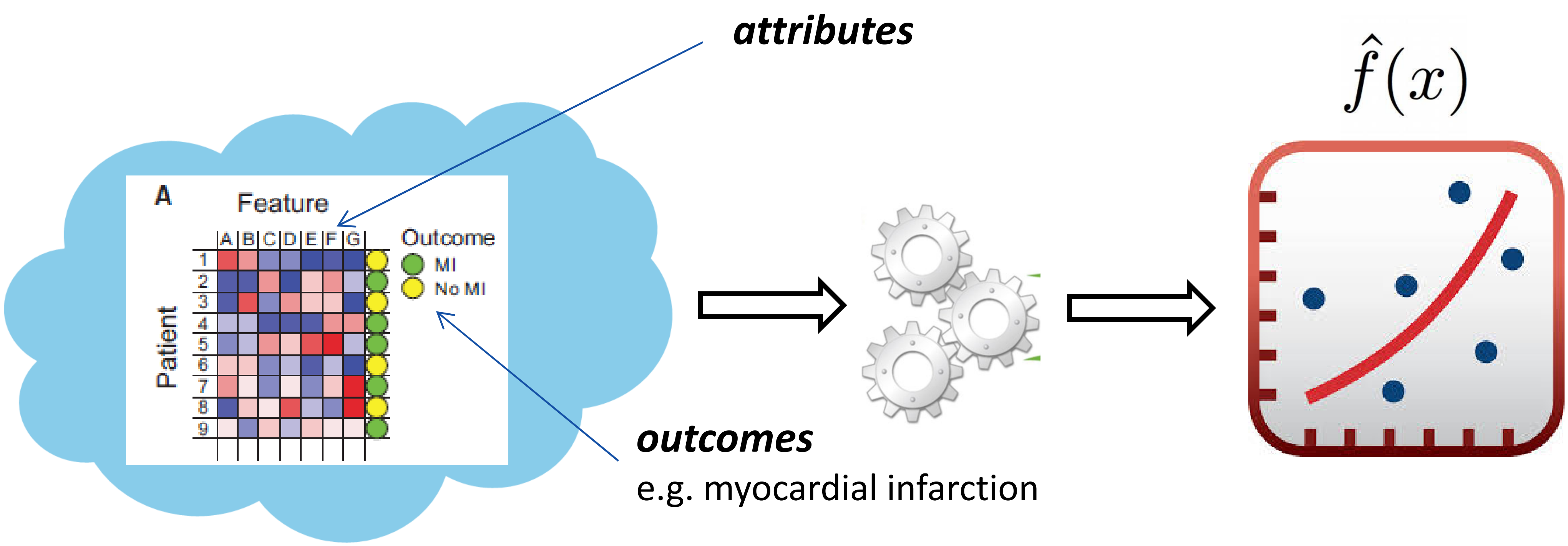


ML algorithm

Model derived from data using the **ML algorithm**

Machine Learning process – *from pattern recognition to predictions*

5. **Building a data model,** extracting **patterns** and matching them with the related **outcomes (*historical data*)** using a selected **ML algorithm**

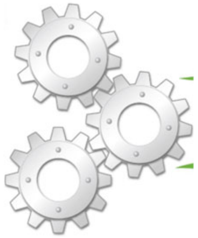


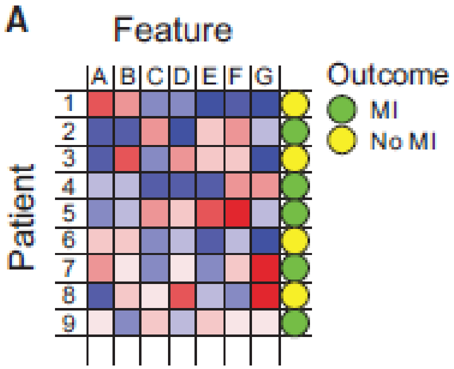
attributes (e.g. in EHR record) $\longrightarrow$ outcomes / labels (e.g., disease state)

Machine Learning process –
from pattern recognition to predictions

Supervised vs unsupervised learning

ML algorithm

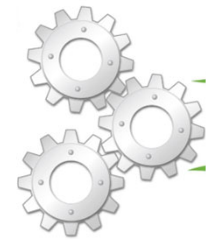




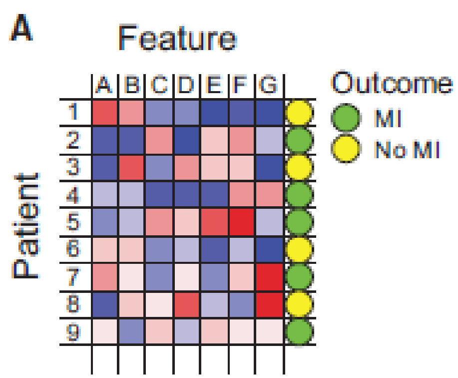
Data (attributes AND outcomes?)

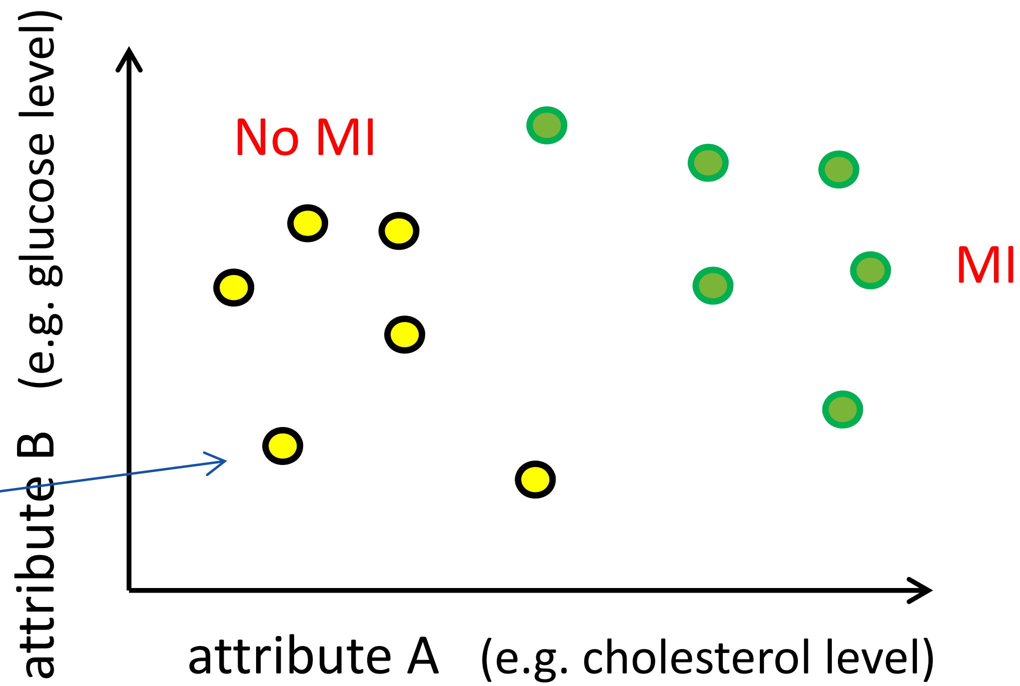
Machine Learning process –
from pattern recognition to predictions

Supervised vs unsupervised learning



Supervised approach taking into account outcomes

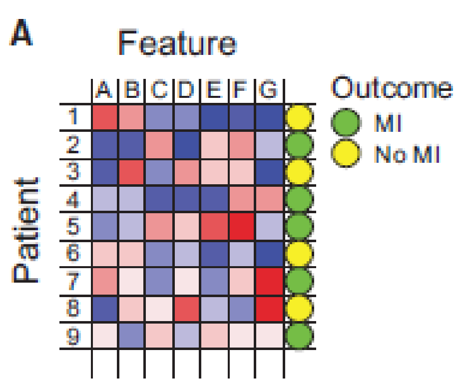


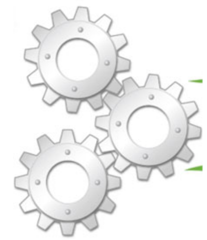


individual patients
(their attributes and disease outcomes)

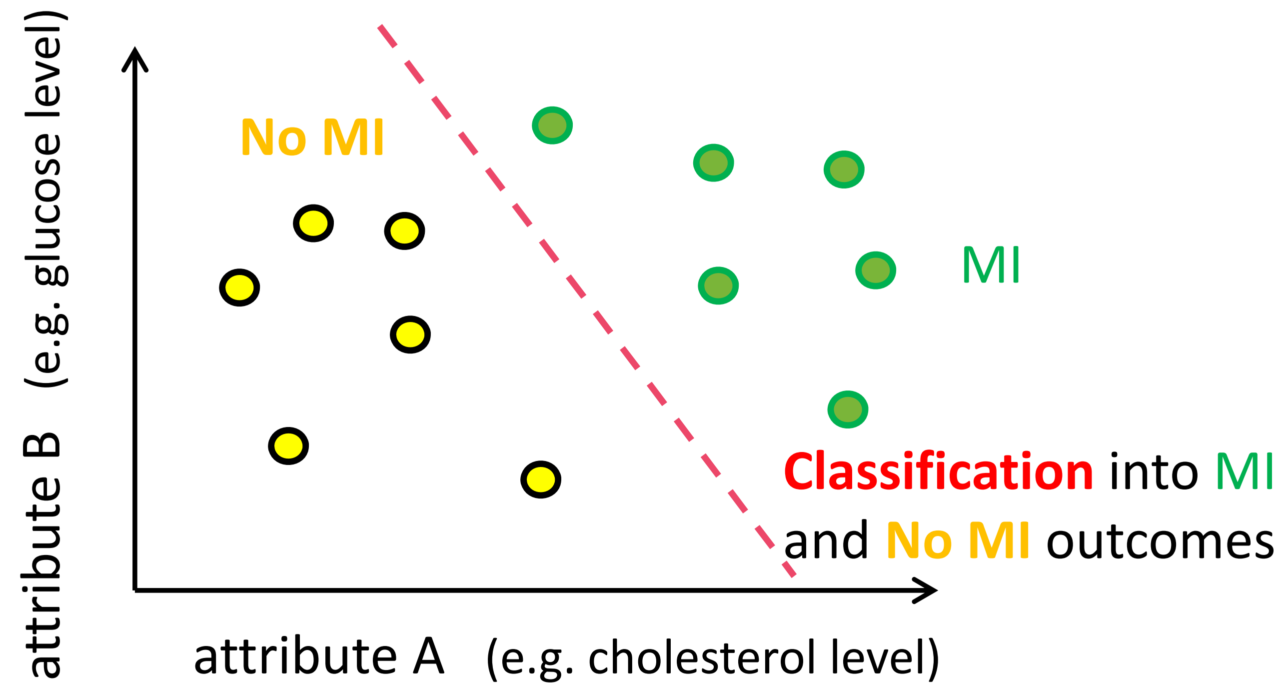
Machine Learning process –
from pattern recognition to predictions

Supervised vs unsupervised learning



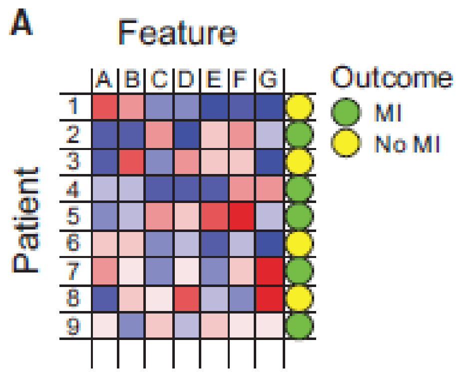


Supervised approach taking into account outcomes

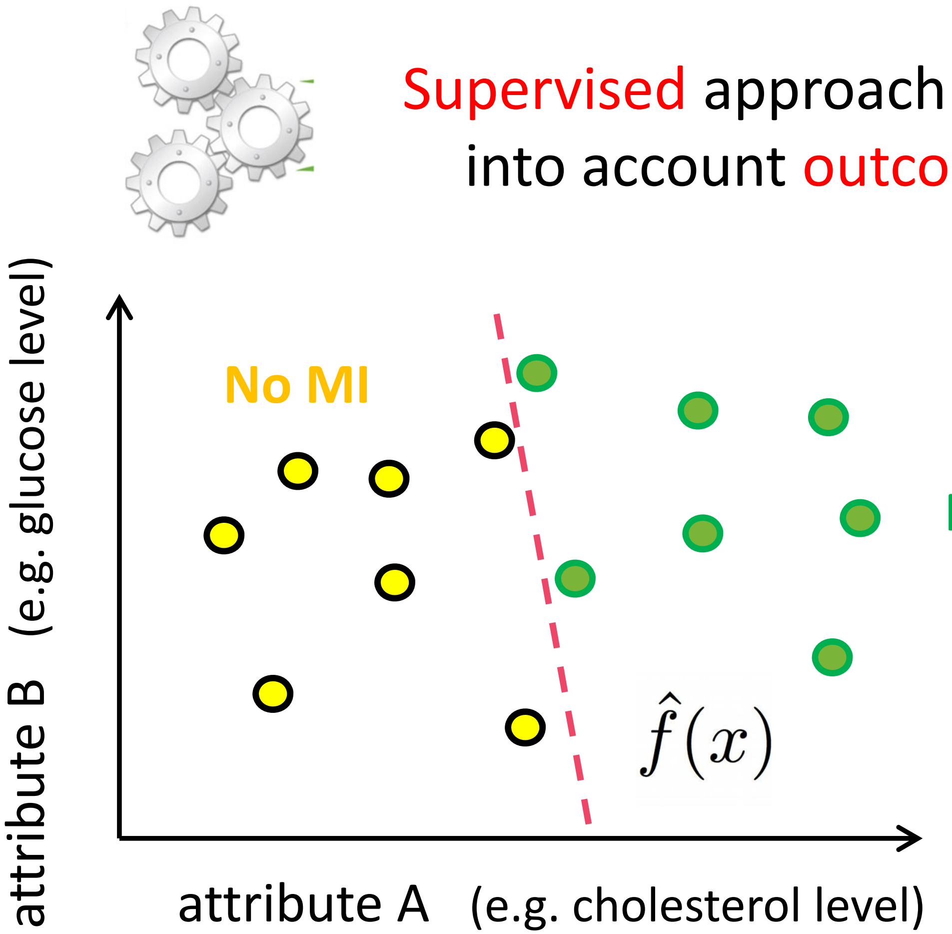


Machine Learning process –
from pattern recognition to predictions

Supervised vs unsupervised learning

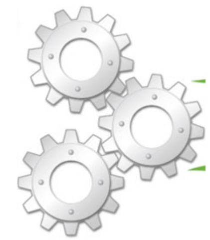


Supervised approach taking into account outcomes

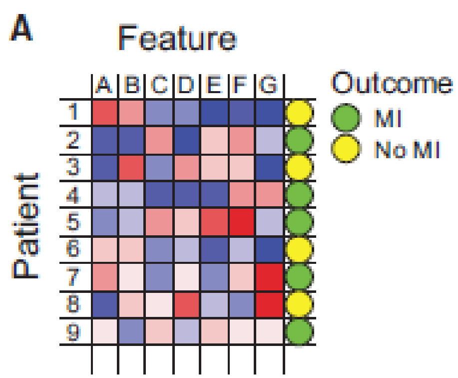


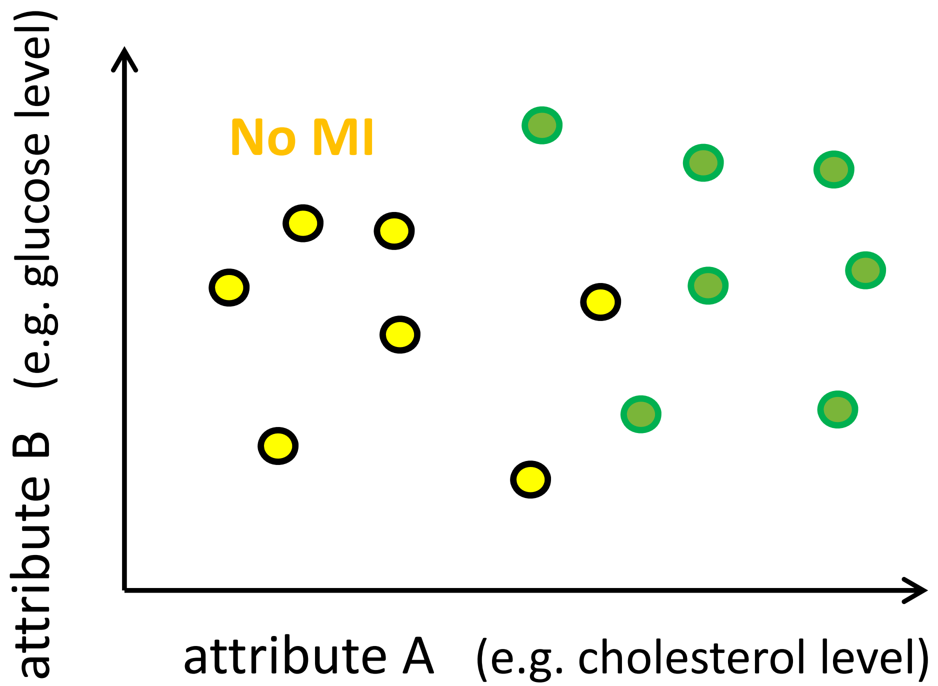
Machine Learning process – *from pattern recognition to predictions*

Supervised vs unsupervised learning



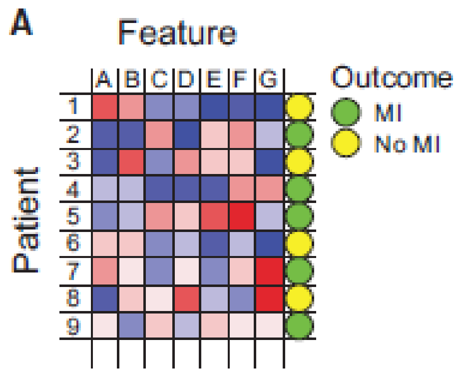
Supervised approach taking into account outcomes



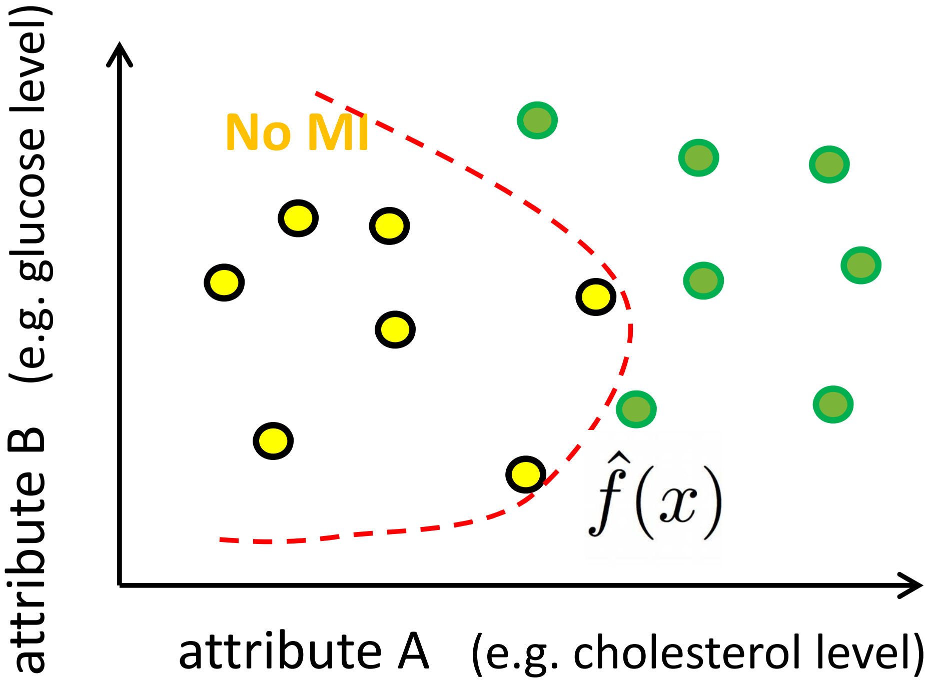


Machine Learning process –
from pattern recognition to predictions

Supervised vs unsupervised learning



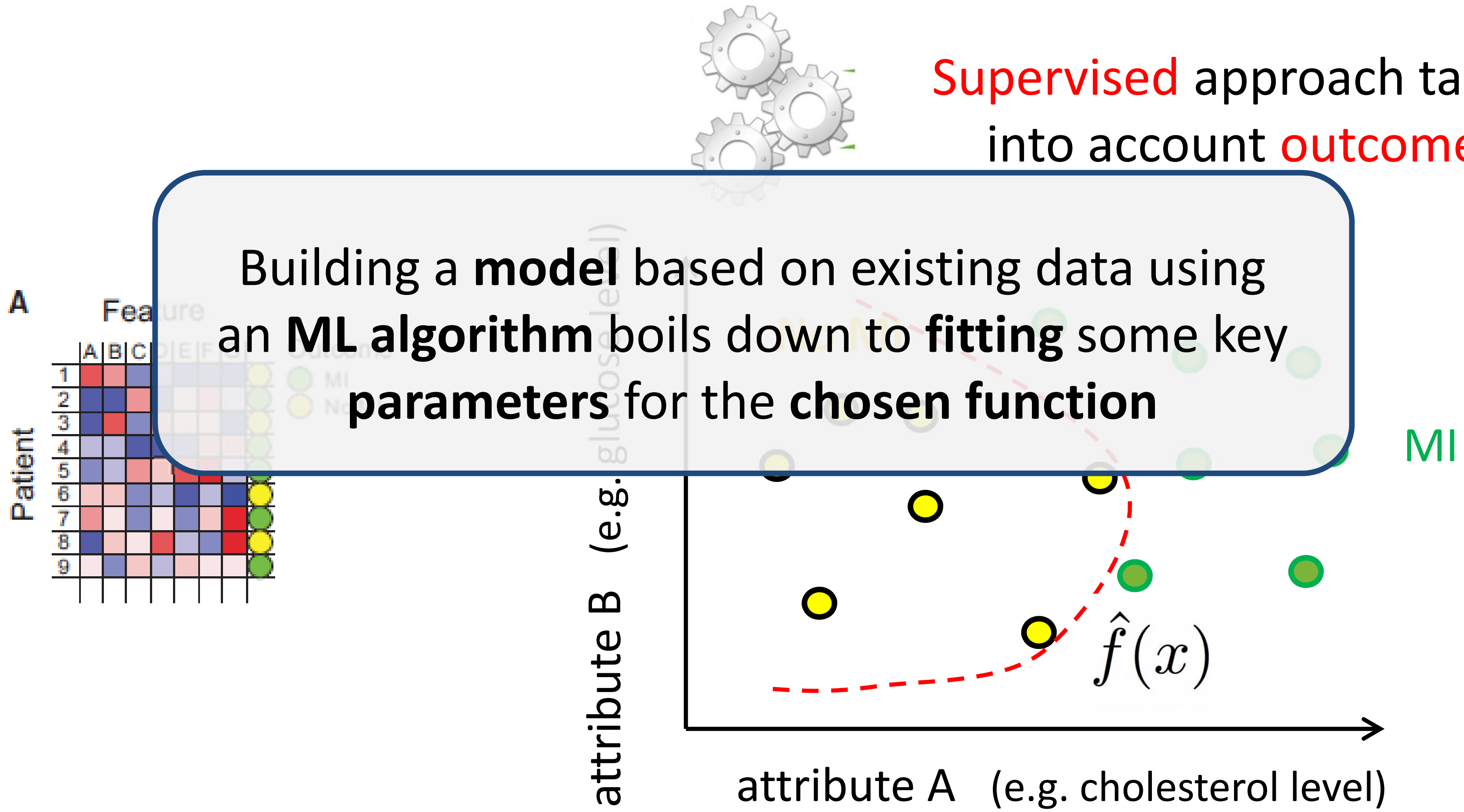
Supervised approach taking into account outcomes



Machine Learning process –
from pattern recognition to predictions

Supervised vs unsupervised learning

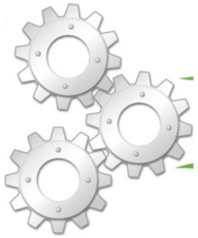
Supervised approach taking into account outcomes



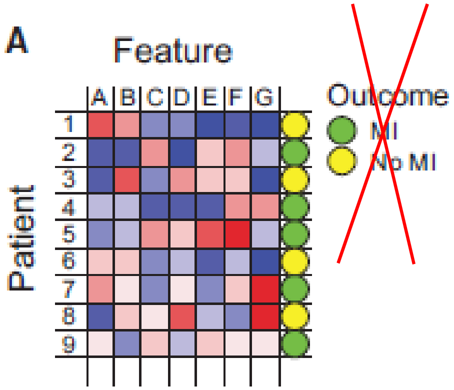
Building a **model** based on existing data using an **ML algorithm** boils down to **fitting** some key **parameters** for the chosen function

Machine Learning process –
from pattern recognition to predictions

Supervised vs **unsupervised learning**

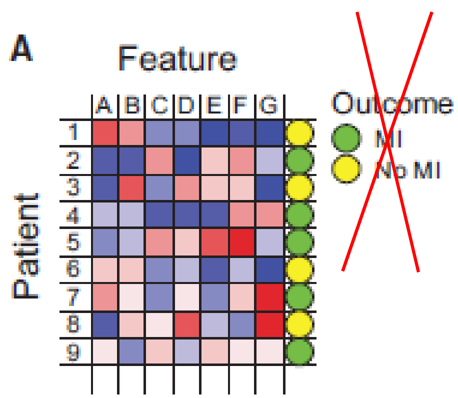


Unsupervised approach does not rely on outcomes

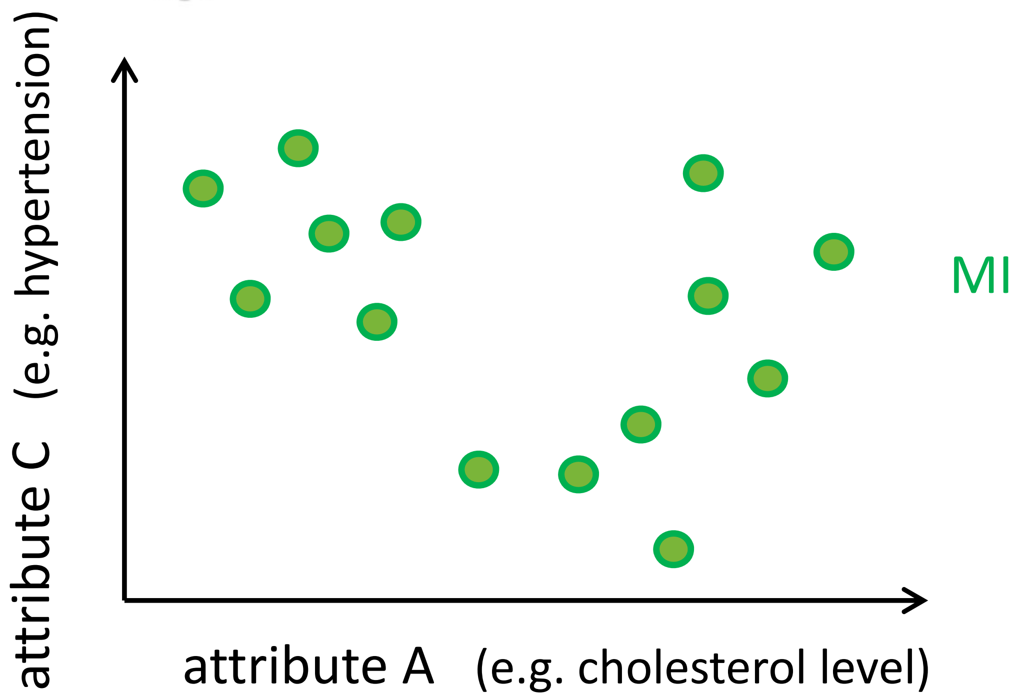


Machine Learning process –
from pattern recognition to predictions

Supervised vs **unsupervised learning**

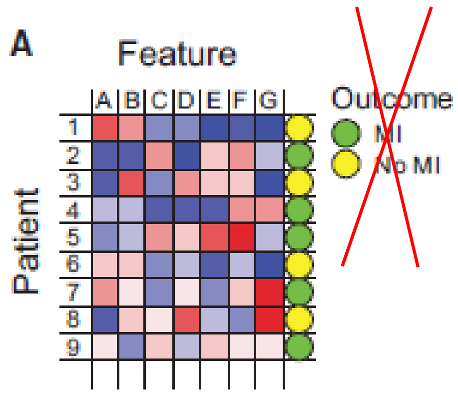


Unsupervised approach does not rely on outcomes

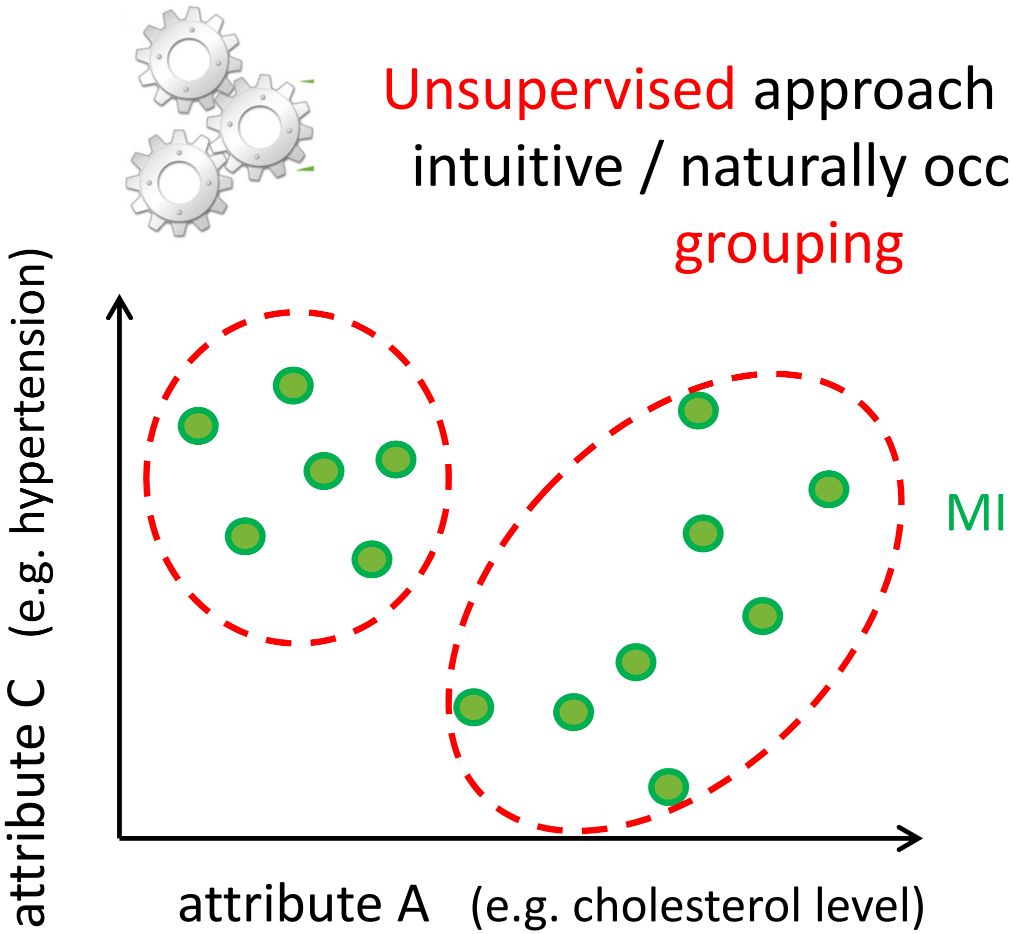


Machine Learning process –
from pattern recognition to predictions

Supervised vs **unsupervised learning**



Unsupervised approach implies intuitive / naturally occurring grouping



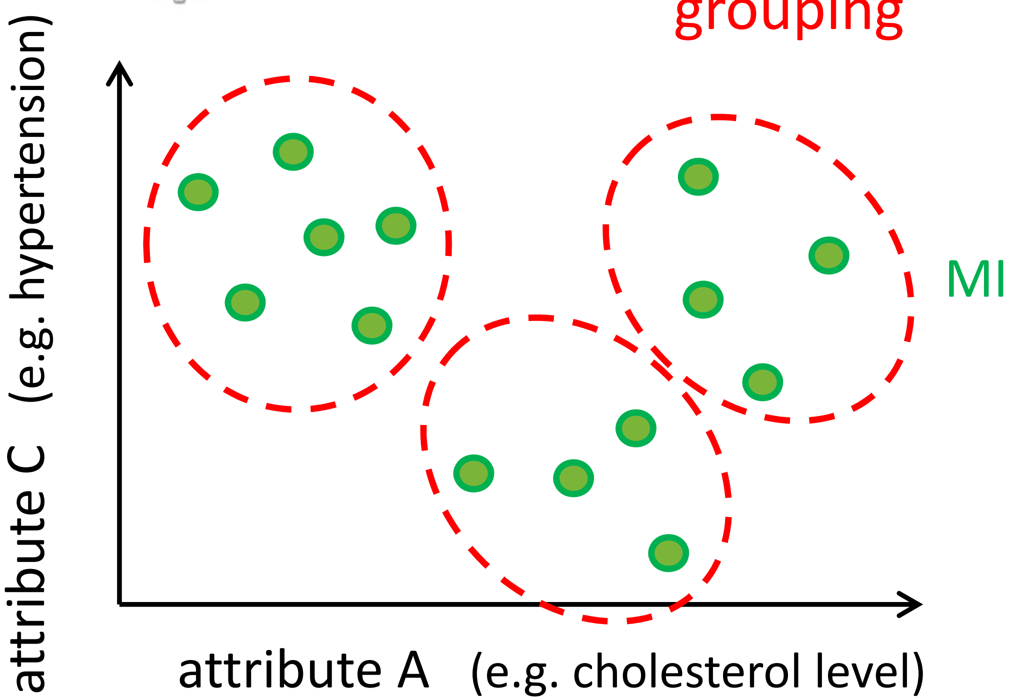
Machine Learning process –
from pattern recognition to predictions

Supervised vs **unsupervised learning**

Opportunities in precision medicine

- redefine disease according to pathophysiologic mechanisms
- heterogeneous disease phenotypes (e.g cardiac diseases)
- capacity for personalized medicine
- possibility to combine with supervised approaches to make more specific predictions
- Large amounts of data without annotations / associated outcomes - labels

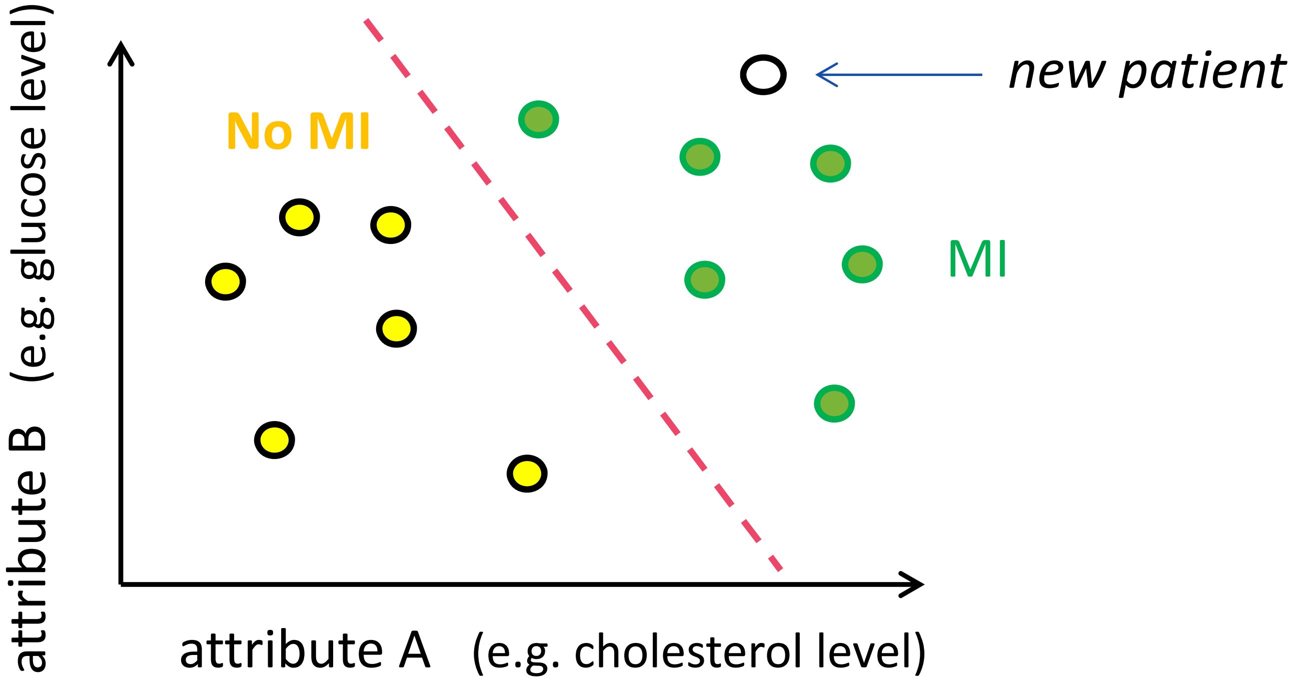
Unsupervised approach implies intuitive / naturally occurring grouping



Machine Learning process –
from pattern recognition to predictions

6. Making predictions for unseen data, e.g. new patients (their attributes)

$\hat{f}(x)$



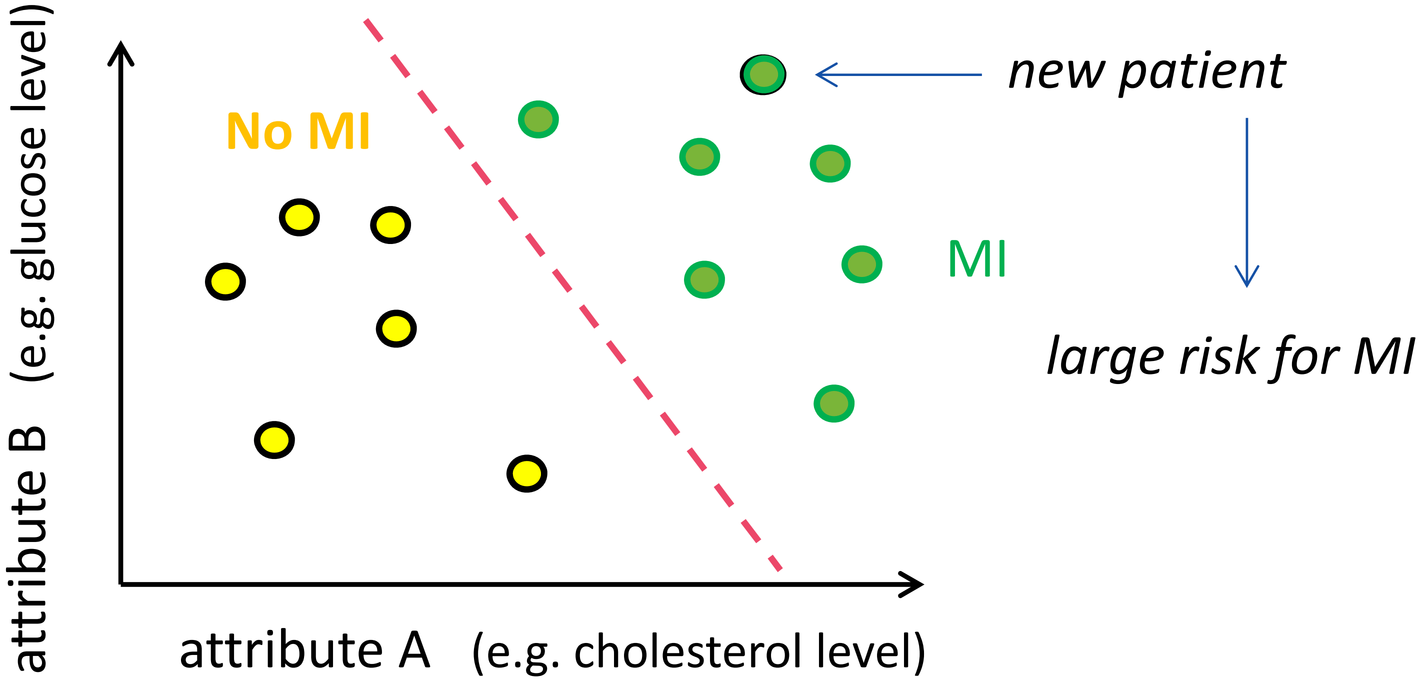
supervised learning context

Machine Learning process –
from pattern recognition to predictions

6. Making predictions for unseen data, e.g. new patients (their attributes)

$$\hat{f}(x)$$

supervised learning context



new data
(e.g. new patient’s EHR record)

$\longrightarrow$

predicted outcome
(e.g., disease state)

Machine Learning process –
from pattern recognition to predictions

6. Making predictions for unseen data, e.g. new patients (their attributes)

$\hat{f}(x)$

unsupervised learning context

new data
(e.g. new patient's EHR record)

- finding similar patients / data points
- recommending similar therapy, intervention
- detecting anomalies, for example
  - in EHR records
  - administrative, technical data

Machine Learning process –
from pattern recognition to predictions

6. Evaluation and interpretation of the ML model

$\hat{f}(x)$ 

- the need for evaluation on new datasets
  - performance measures (temptation to solely rely on ML performance measures)
  - generalization power (trade-off with the complexity, selection of model classes is very important: the simplest possible but not too simple)
- interpretability largely depends on the ML algorithm and the selected class of models

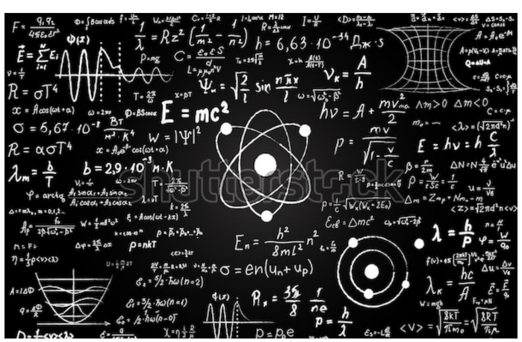
Key design decisions in a nutshell

1. Problem formulation
   - *What type of ML problem is it?*
2. Data representation/encoding scheme
   - *How do we map raw input to input space? What kind of feature are we using and how do we encode them?*
3. Loss function, evaluation metric definition
   - *What is a measure of success?*
4. Hypothesis space and learning algorithm selection
   - *What is a suitable hypothesis space? What are assumptions about the underlying model?*
   - *What are best suited learning algorithms?*

ML comes in many flavours

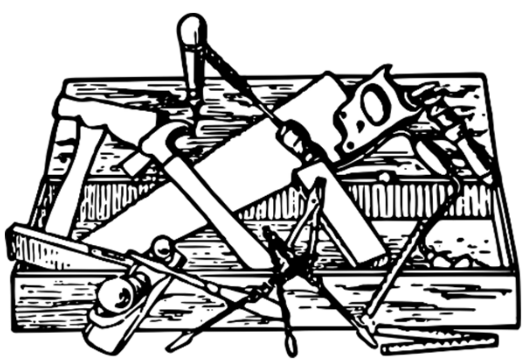
Theoretical perspective

- **statistical** learning theory
- **computational** learning theory
- optimization perspective
- inductive learning theory

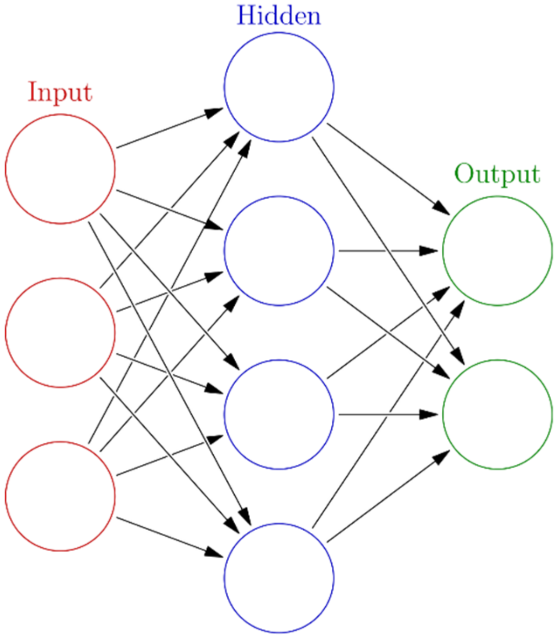


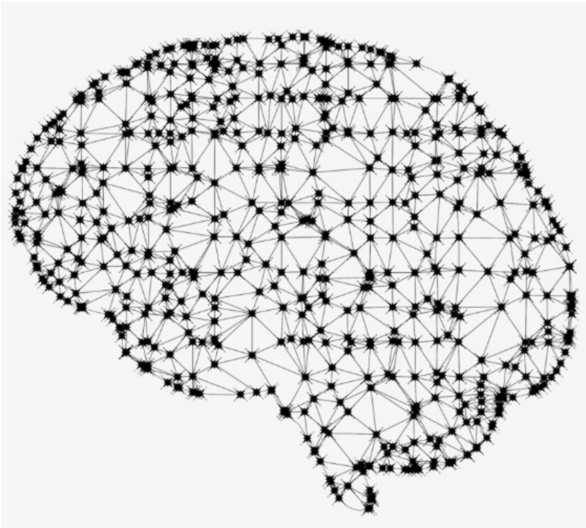
Toolbox – zoo of ML methods

- logistic regression
- linear discriminant analysis
- Bayesian inference machine
- support vector machines
- rule-based learning, fuzzy logic
- decision trees
- k-means
- evolutionary opt.
- Q-learning
- neural networks etc. etc.

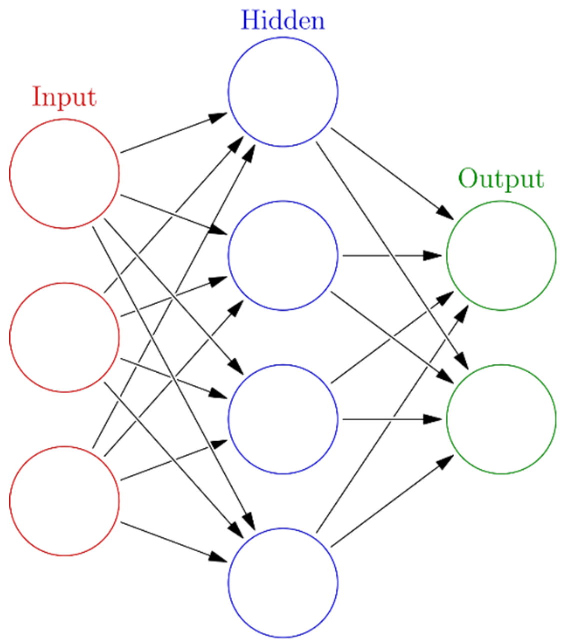


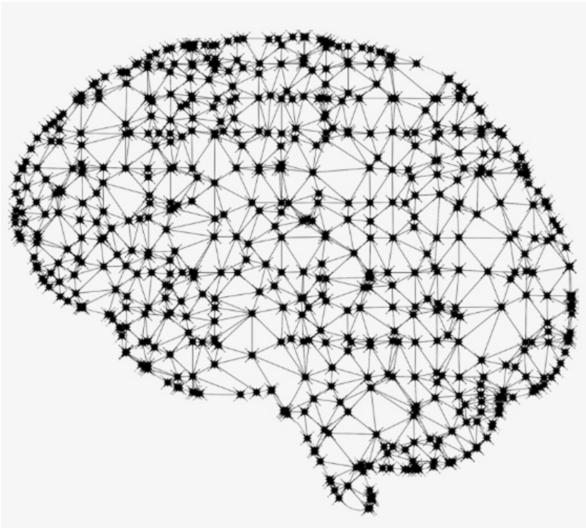
Artificial Neural Networks (ANNs)

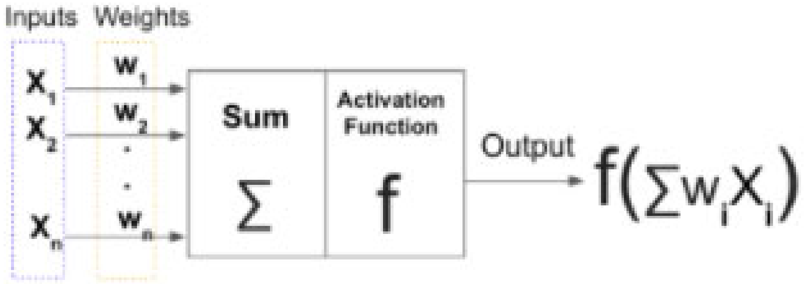




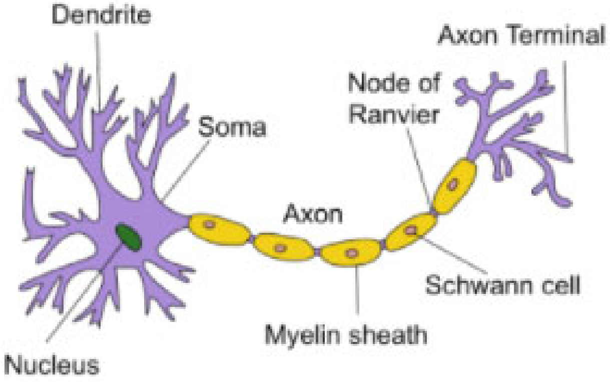
Artificial Neural Networks (ANNs)







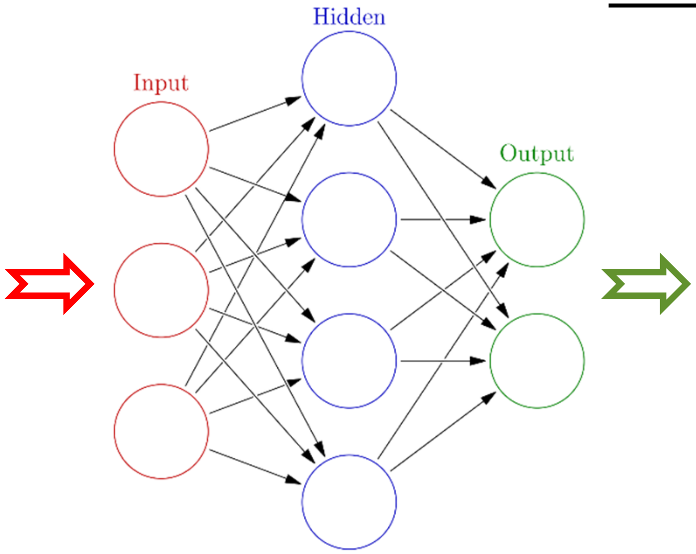
Structure of artificial neuron

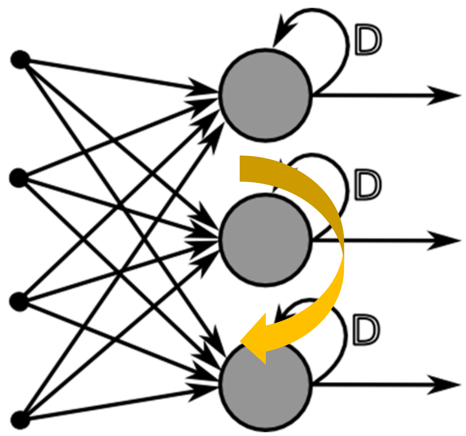


Structure of a typical neuron
(source: Wikipedia)

Artificial Neural Networks (ANNs)

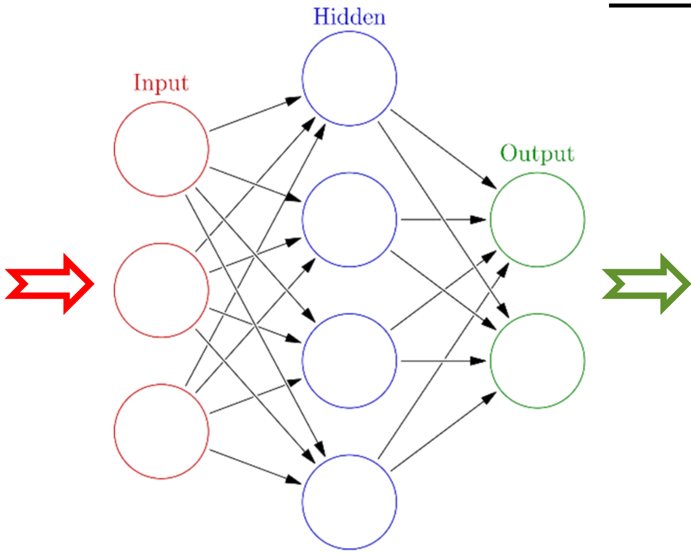
<u>Network architectures</u>

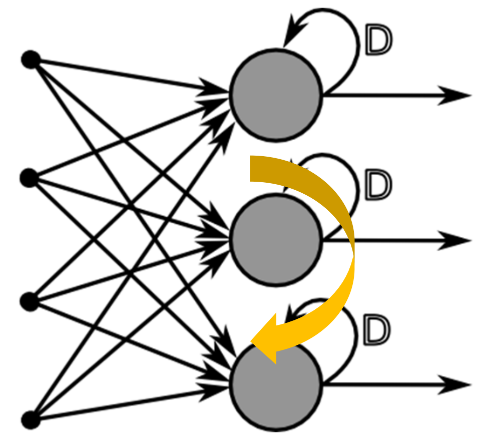




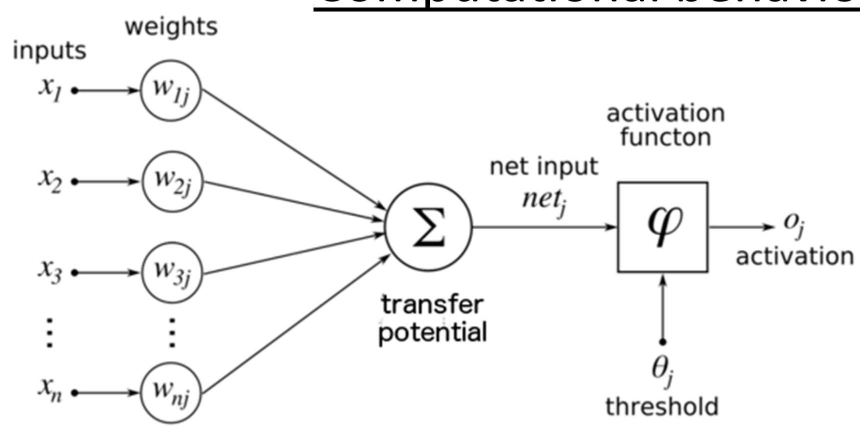
Artificial Neural Networks (ANNs)

Network architectures



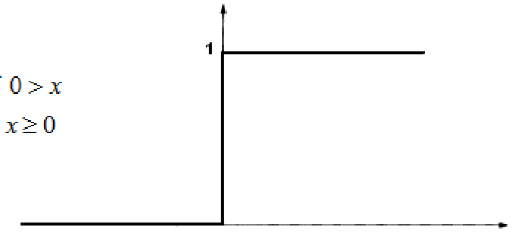


Computational behaviour of network units

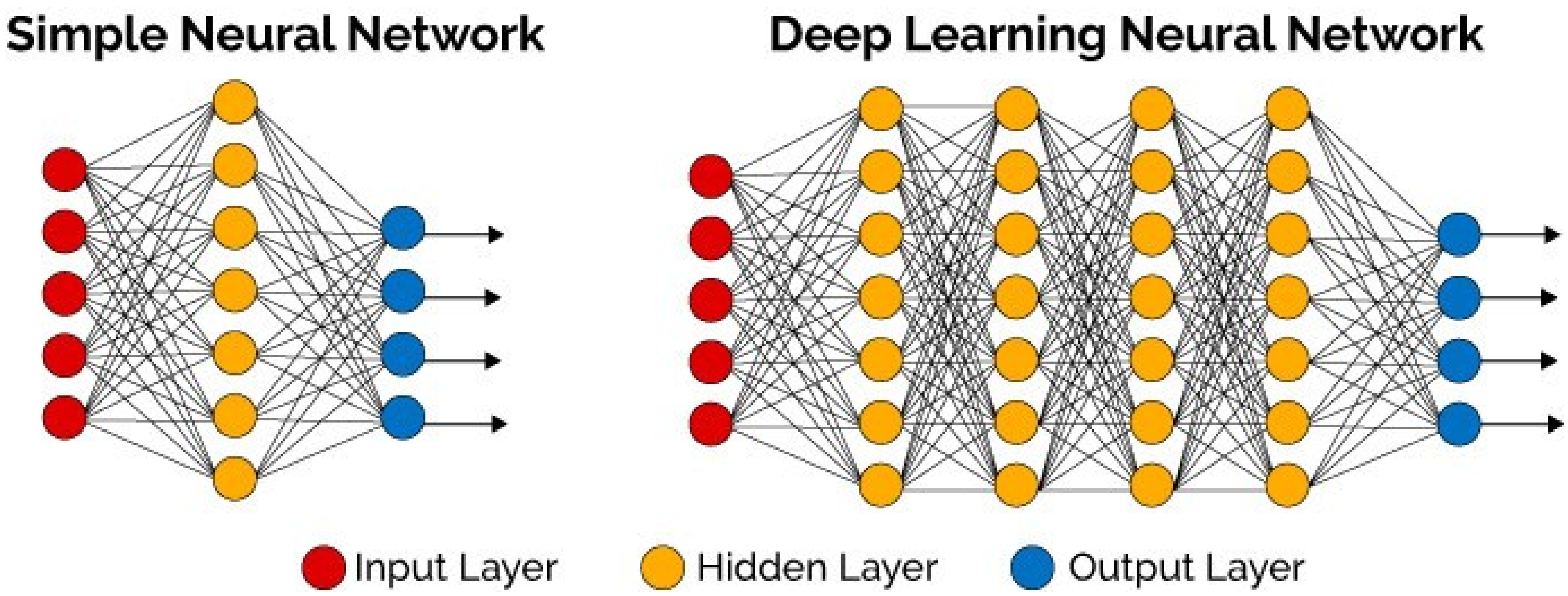


Unit step (threshold)

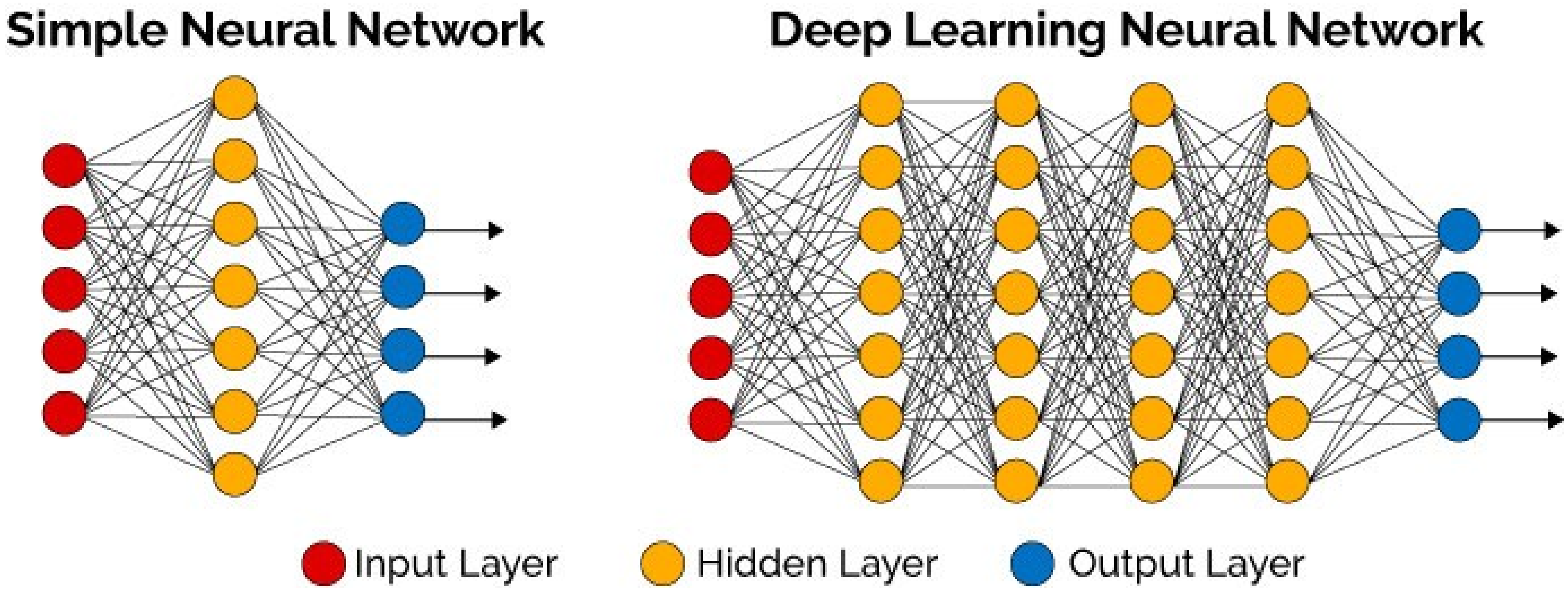
$$f(x) = \begin{cases} 0 \text{ if } 0 > x \\ 1 \text{ if } x \geq 0 \end{cases}$$



From ANNs to Deep Learning (DL)



From ANNs to Deep Learning (DL)



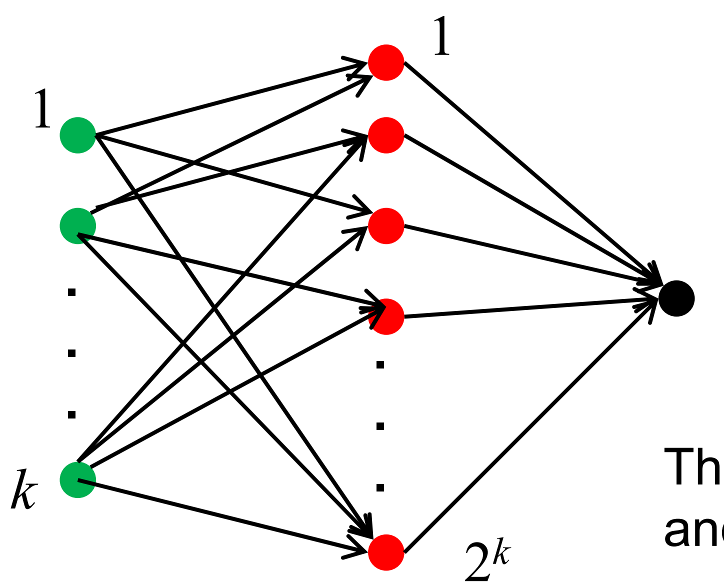
The multitude of hidden layers makes the learning problem more challenging but also allows for improved functionality and robustness.

Motivation for deep structures – why go deep?

Expressive power and compactness of models

(*expressibility* and *efficiency*)

- enhances generalisation, especially with limited training examples
- less degrees of freedom when handling complexity and nonlinearity – exponential gain

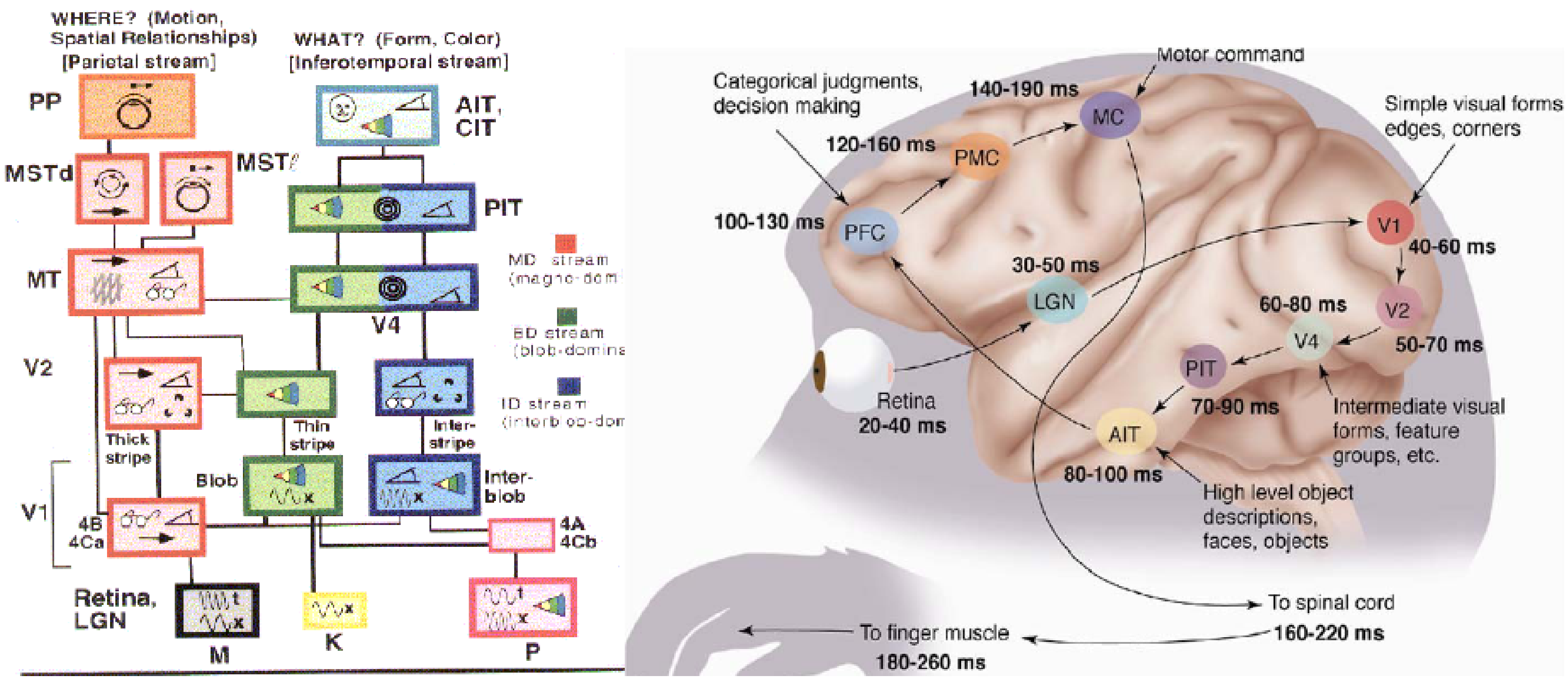


Shallow structure may need exponential size of hidden layer(s)

The universal approximation theorem and approximation costs.

Motivation for deep structures – why go deep?

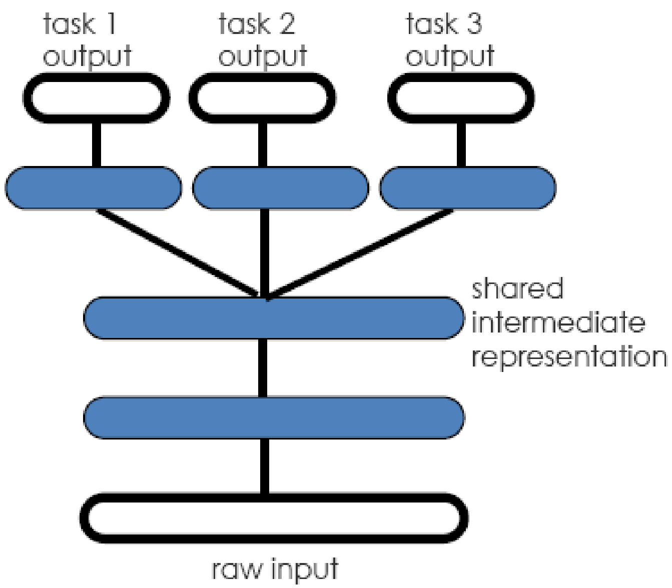
Inspirations from hierarchical brain organisation

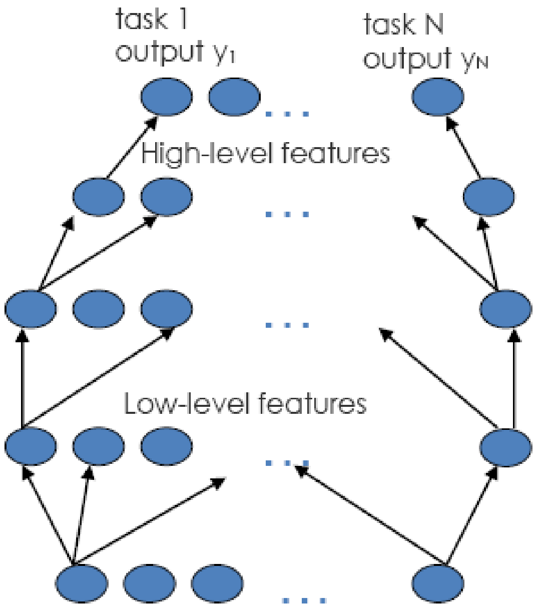


LeCun & Ranzato, 2013

Motivation for deep structures – why go deep?

Multiple levels of representations facilitate transfer and multi-task learning (hierarchy of representations, non-local generalisation).



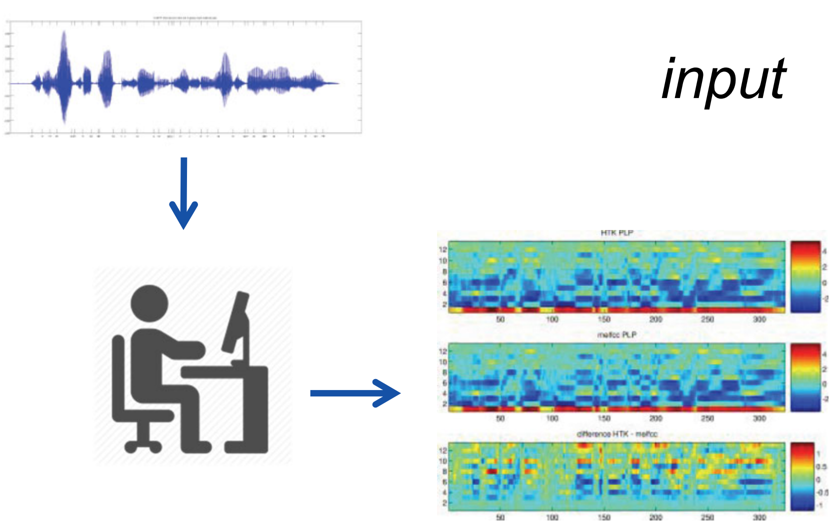


Lee, 2011

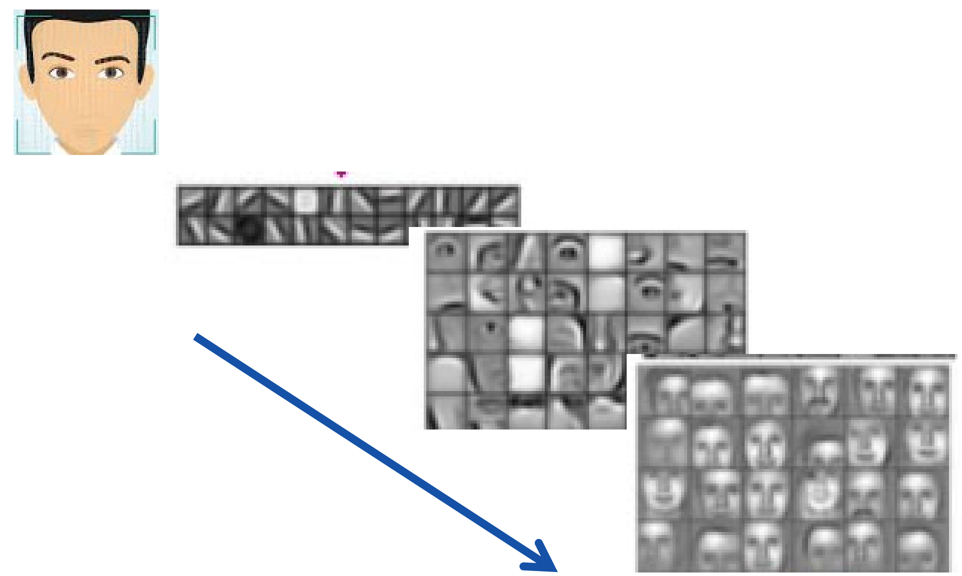
Representation learning in deep models

Hand-engineered features in a traditional pattern recognition approach

End-to-end networks with learned features spaces, data representations



input



features, representations

Generic characteristics of ML methodology

- generic approach with universal treatment of data
  - works well with high-dimensional and multi-modal data, problems with complex multivariate behaviours can be handled

  (problem if we have small amount of data though)
  - robust for noisy, corrupted data, even with missing values
  - can even work with partly unlabelled data (semi-supervised)
  - effectively deals with static and temporal incl. nonstationary data

Generic characteristics of ML methodology

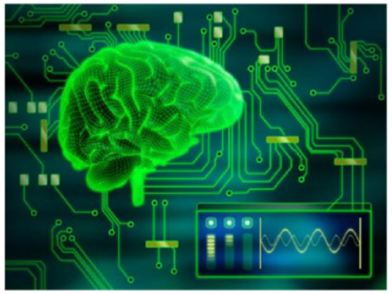
- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)

DATA

basic assumptions

without knowing mechanisms underlying data generation

PREDICTIVE MODELS, KNOWLEDGE



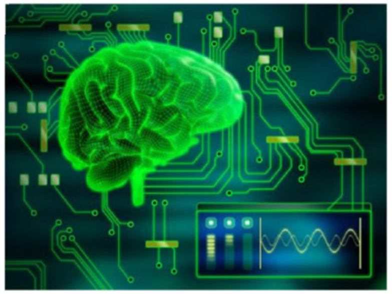
Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)

Robust LEARNING - parameter estimation methods

basic assumptions

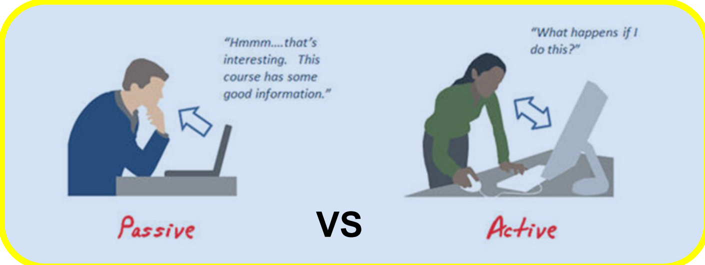
without knowing mechanisms underlying data generation



Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)

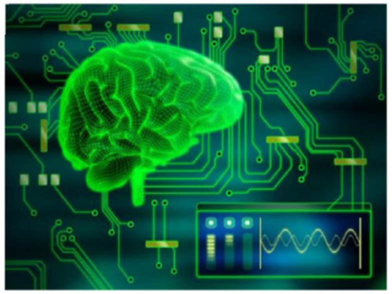
Robust LEARNING - parameter estimation methods



basic assumptions

⟹

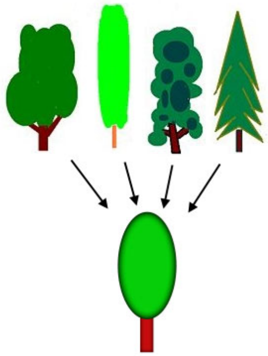
without knowing mechanisms underlying data generation



Generic characteristics of ML methodology

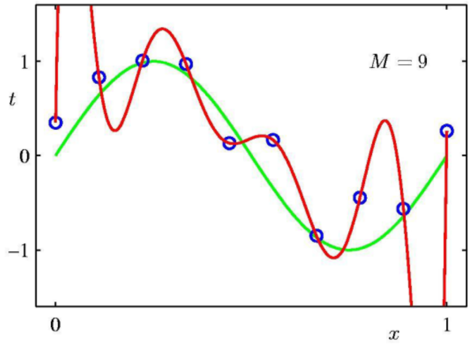
- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing

  - reliable parameter estimation/optimisation techniques
  - exploratory search for relevant / interesting patterns
  - potential for *good generalization* and consistent performance



Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
  - reliable parameter estimation/optimisation techniques
  - exploratory search for relevant / interesting patterns
  - potential for *good generalization* and consistent performance
  - mechanisms for *complexity calibration*



Generic characteristics of ML methodology

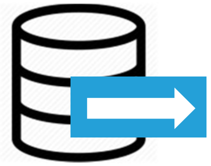
- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
  - reliable parameter estimation/optimisation techniques
  - exploratory search for relevant / interesting patterns
  - potential for *good generalization* and consistent performance
  - mechanisms for *complexity calibration*
  - *scalability*

Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
- different modes of accessing data
  - iterative batching

Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
- different modes of accessing data
  - iterative batching
  - one-pass



Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
- different modes of accessing data

  - iterative batching
  - one-pass
  - online

Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
- different modes of accessing data
- generative vs discriminative approach

Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
- different modes of accessing data
- generative vs discriminative approach
- so, when ML is particularly helpful?

Generic characteristics of ML methodology

- generic approach with universal treatment of data
- usually model-free approach with minimal assumptions
- data driven process (parameters estimated from data)
- effective data processing
- different modes of accessing data
- so, when ML is particularly helpful?

  - for tasks that are too <u>complex</u> to program or describe by rules
    (tasks performed by humans based on experience and tasks beyond our capabilities)

  - for tasks that require <u>adaptation</u>
    (where the behaviour should adapt to input data)

Some ML practicalities to be aware of

- all problems are different so approach has to be customized, *"no free lunch"*
- build as simple as possible but sufficiently complex models to learn from data

("Make things as **simple as possible**, **but not simpler**.", e.g. avoid linear models for nonlinear relationships but control the degrees of freedom)

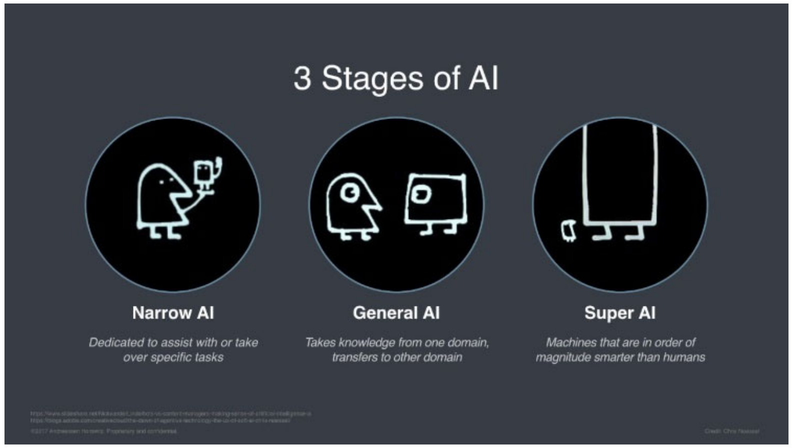
- be careful with inadequate amount of data for high-dimensional problems, i.e. "*curse of dimensionality*"
- choice of adequate performance criteria
- the need for proper validation of predictive models
- examine your data first, check outliers as they could be interesting
- try to incorporate as much domain specific knowledge as possible

Current limitations and challenges

- the need for large amounts of training data, good data and reliance upon lengthy batch training
- problems with nonstationary data, complex spatio-temporal patterns
- hard to handle interactive learning in real time
- scalability and compute time
- no free lunch: customised, "hand-crafted" solutions, no multi-tasking, no universal learner (meta ML?)
- poor interpretability (transparency vs robustness trade-off)
- accounting for correlations, not causation
- challenges involved in the integration of human expertise and ML
- limited suitability for high-level reasoning, abstraction or planning, intuition

Current limitations and challenges

Still in the era of narrow intelligence



Ongoing developments, trends

- data-driven analysis of behaviour, preferences, emotions to study how human make decisions (financial applications)
- towards autonomous cars, personalised medicine
- cloud-hosted intelligence, democratisation of ML
- increasing importance of **brain-like/inspired approaches** to AI, ML – towards general AI and for specific narrow applications
- **biological deep learning**
- contextual systems
- **unsupervised learning**
- meta- and automated ML systems
- growing the importance of ethical and moral issues

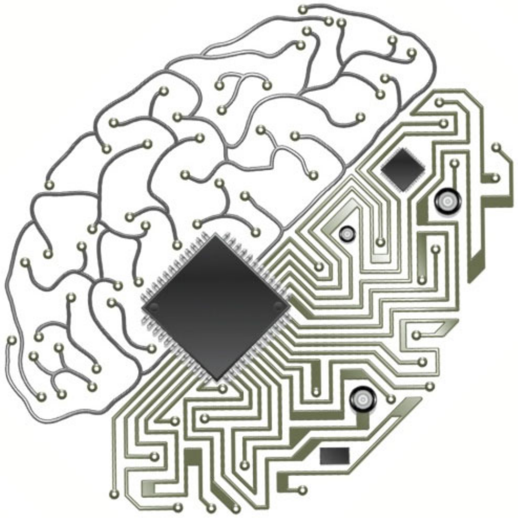
Neurocomputing, neuromorphic and brain-like approaches

”Brain machine”

analogue vs digital
(clock, temporal scales, etc.)

memory

embodiement



memory and information processing

self-organisation,
adaptation,
fault-tolerance

distributed computations with some modularity and hierarchy

hardware vs software
(wetware)

The inspirational brain

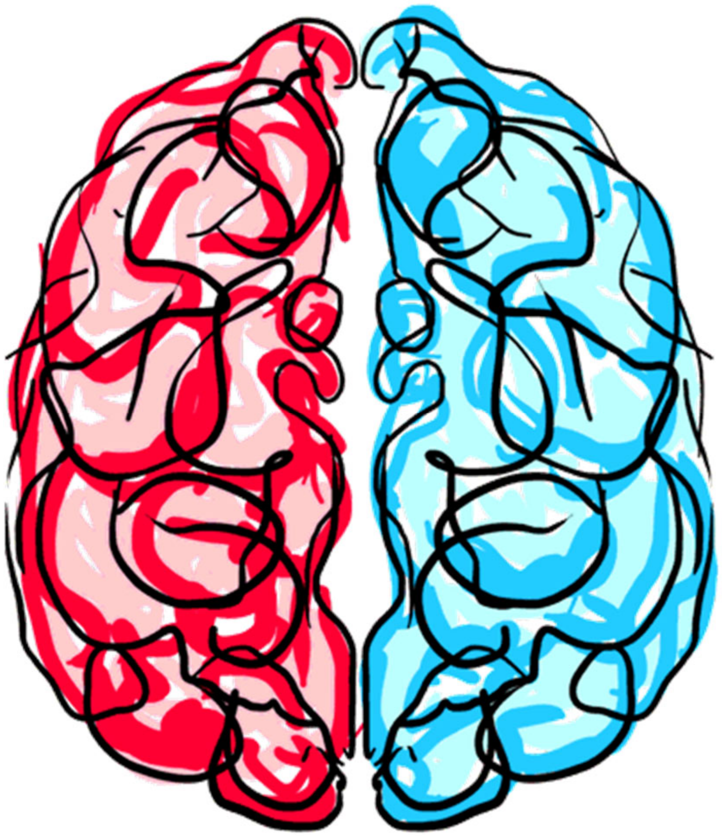
Why is there so much fuss about it?

Cognitive functionality, behaviour coordination

Plasticity, adaptivity, learning

Flexibility, multi-purposeness

Dealing with uncertainty, ambiguity, fragmentary information



Low energy consumption

Massively parallel

Fault-tolerance

Robust information processing

Simple facts about the human brain

Wetware

2% of the body weight (1.3-1.5 kg) but consumes 20% of the energy

Brain generates 15-30 W of electrical power

~86 bn neurons, each connecting to up to ~10 k neurons

Sparse activity, distributed representations (”memory is more activity than a place”)

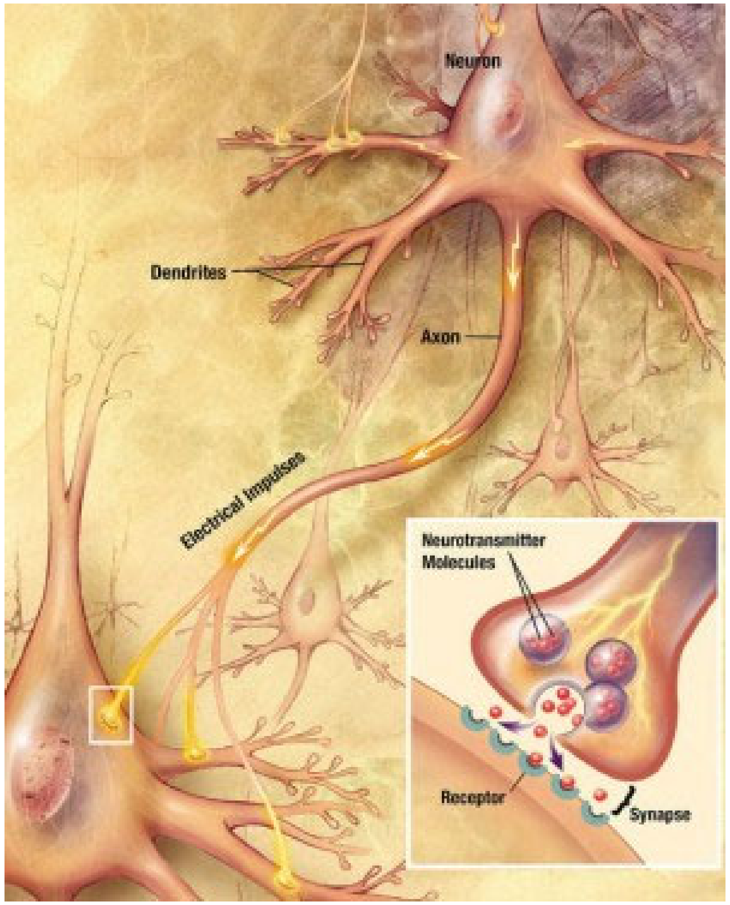
over 100 trillion synapses in total

Simple facts about the human brain

2% of the body weight (1.3-1.5 kg) but consumes 20% of the energy

~86 bn neurons, each connecting to up to ~10 k neurons

over 100 trillion synapses in total



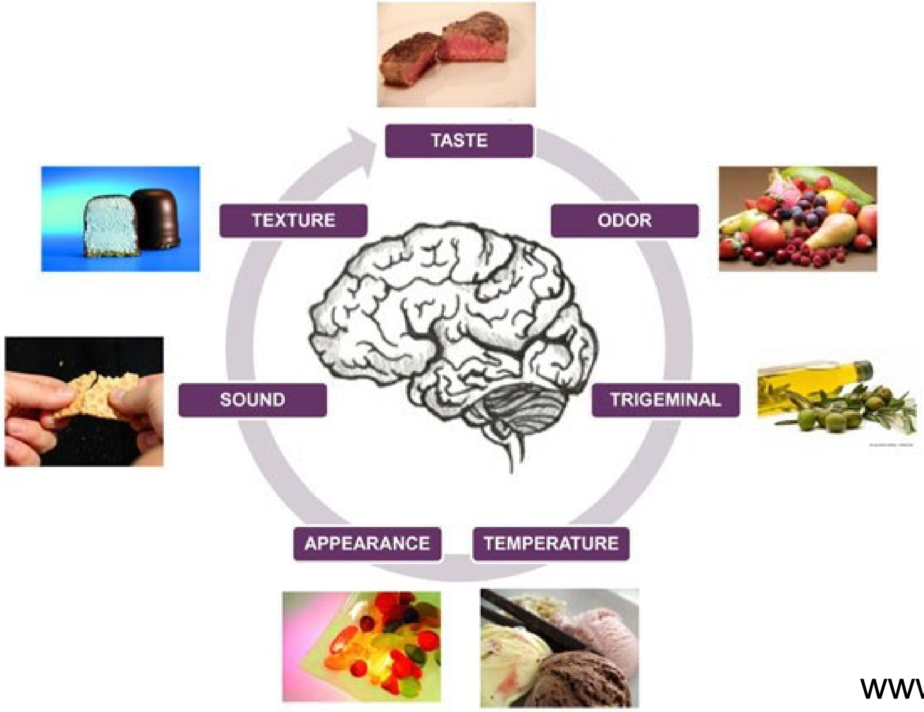
Information transmission using a combination of chemicals and electricity

Synapses handle communication (plasticity)

Amazing brain functionality

Function enables and can be directly manifested in our behaviour, cognition

Multi-modal perception, multi-sensory integration

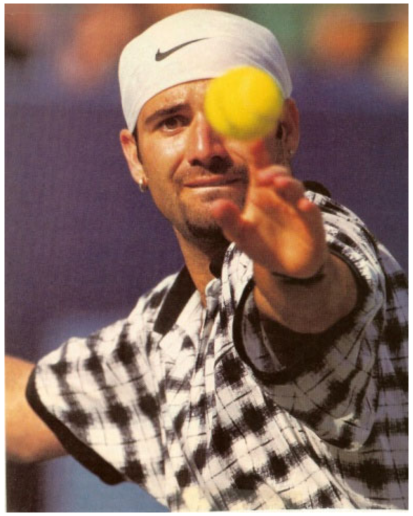


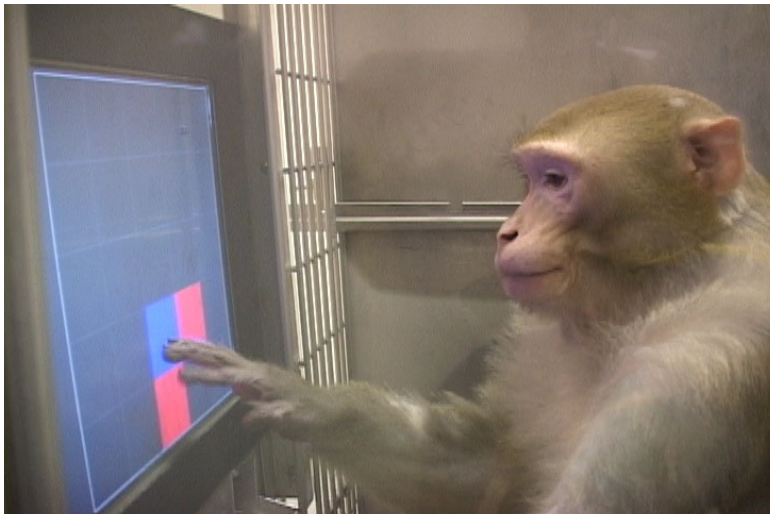
www.iw.fraunhofer.de

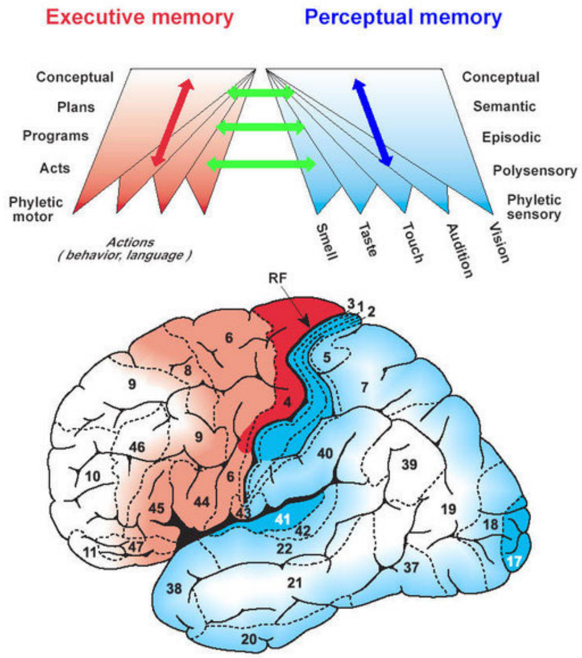
Amazing brain functionality

Function enables and can be directly manifested in our behaviour, cognition

Perception-action coupling and learning





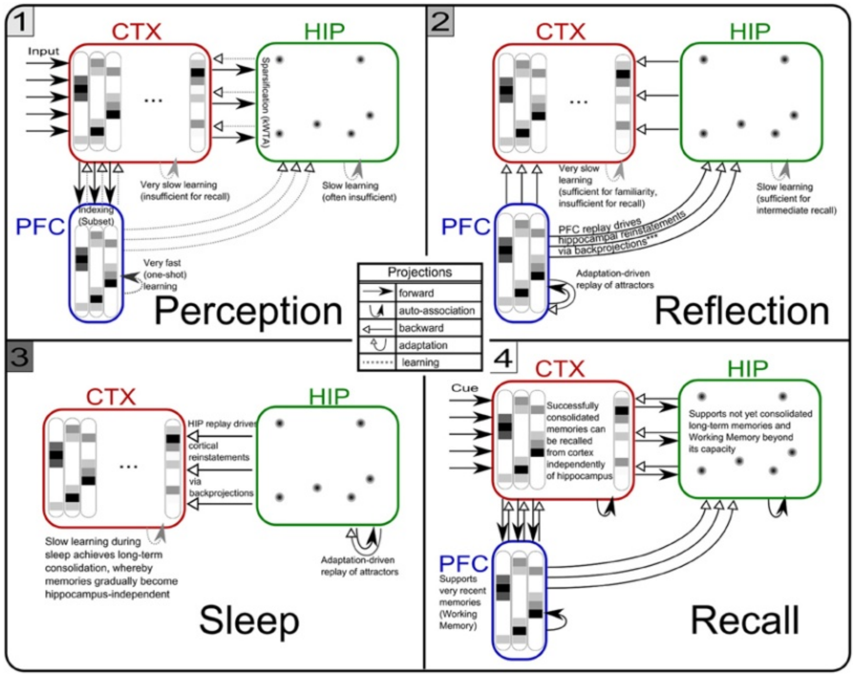


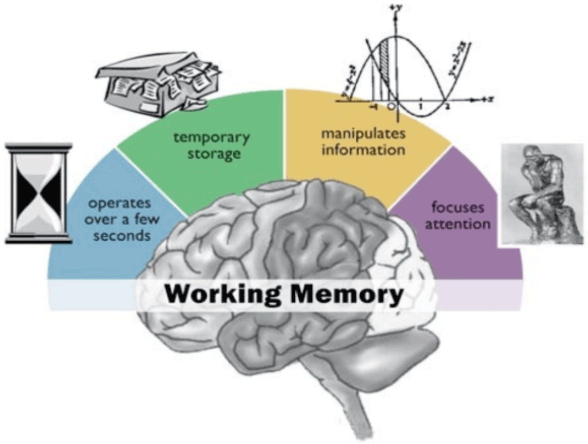
Fuster, JM (2004) Upper stages of the perception-action cycle. Trends in Cognitive Science 8:143-145

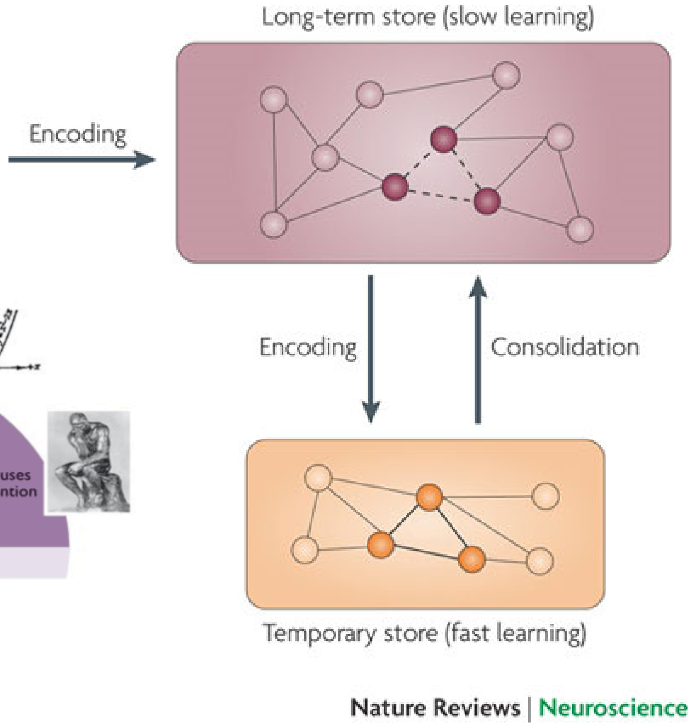
Amazing brain functionality

Function enables and can be directly manifested in our behaviour, cognition

Multi-scale memory and learning







Amazing brain functionality

- perception, sensation
- perception-action, decision making
- memory and learning
- motor behaviour, sensory-motor coupling
- emotions, arousal, motivational system
- etc

Amazing brain functionality

- perception, sensation
- perception-action, decision making
- memory and learning
- motor behaviour, sensory-motor coupling
- emotions, arousal, motivational system
- etc

- associative, content-addressable memory with storage, encoding, recall at different time scales
- graceful degradation, forgetting
- goal-directed behaviour
- (reinforecement) learning of states and actions
- decision making, reasoning
- generation of motor responses (motor behaviour)
- multi-modal object recognition, anomaly detection, prediction (time!)

Connectionist/network approaches

Activity

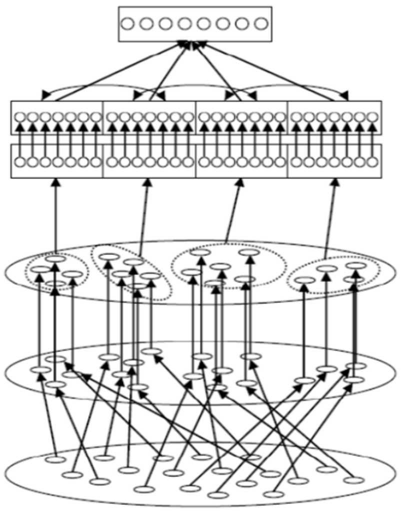
states of units – continuous vs
binary (spikes)
input – output transform

Representations

pattern of activation over the units
in the network,
sparse distributed representations

Learning (on-line)

adjustment of
connection weights



Processing

occurs through propagation of
activation signals

Global architectures

networks of networks, topography
importance of hierarchies !

Architecture

pattern of connectivity that
determine the nature of processing,
modularity

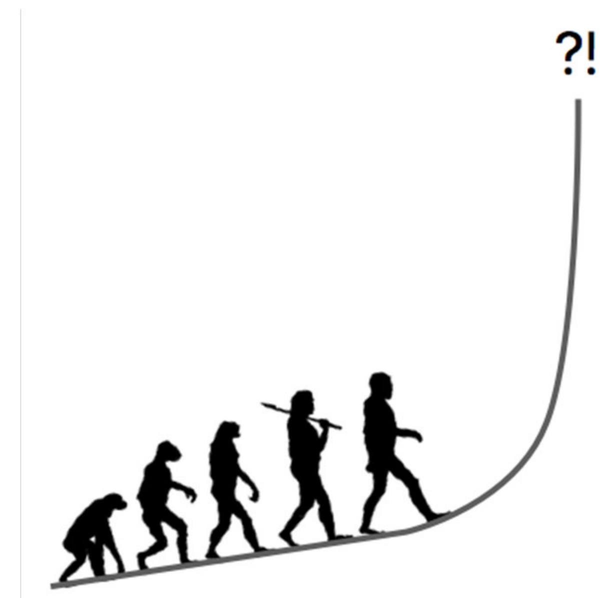
Future prospects – concerns beyond technological challenges (1)

- Like other technologies, AI can be used for good and bad purposes
  - there is still a lot of misunderstanding about what AI is and is not
  - there is a need for multi-aspect research (resources on research on the societal implications of AI technologies are scarce)
  - we need to prevent though from poorly informed regulation that would stifle innovation
  - education is a key factor
  - hardly any efforts to address the needs of low-resource communities
  - transparency problem
  - democratisation of AI (risk of widening inequality of opportunities)

Future prospects – concerns beyond technological challenges (2)

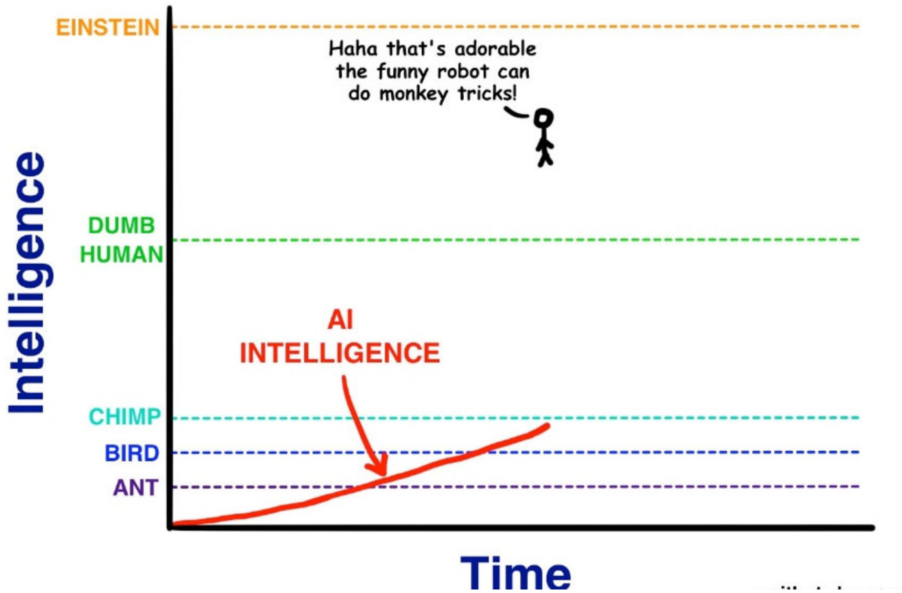
- Unintended consequences of using ML
  - overreliance on the capabilities of automation
  - “deskilling” in the longer run, i.e. reduction of the level of skill required to complete a task
  - data-dependent bias (AI should remove harmful biases)
- Concerns about human jobs
- Privacy and security for enabling data use and sharing
- Building public trust
- Ethical and legal issues, policies

Superintelligence on the horizon?



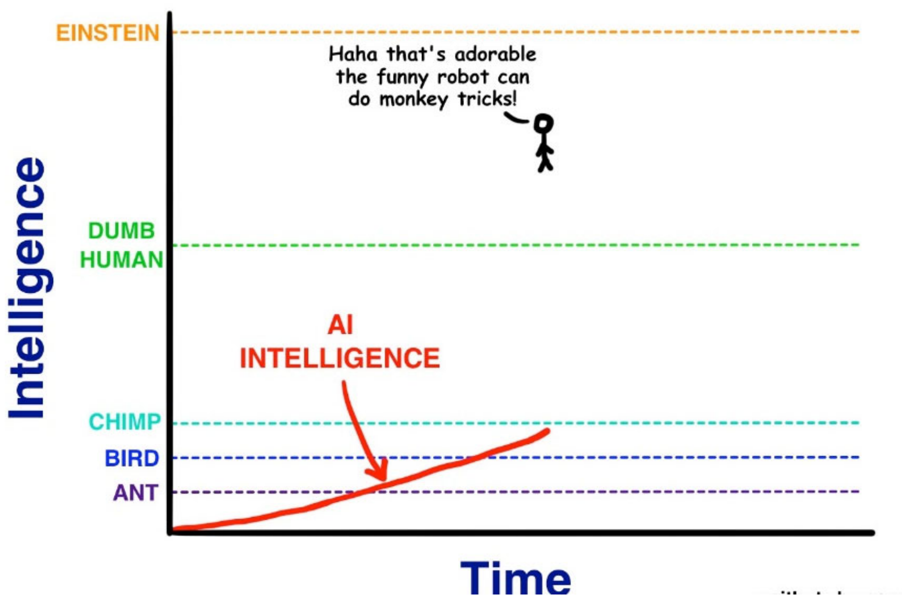
Superintelligence on the horizon?

Our distorted view

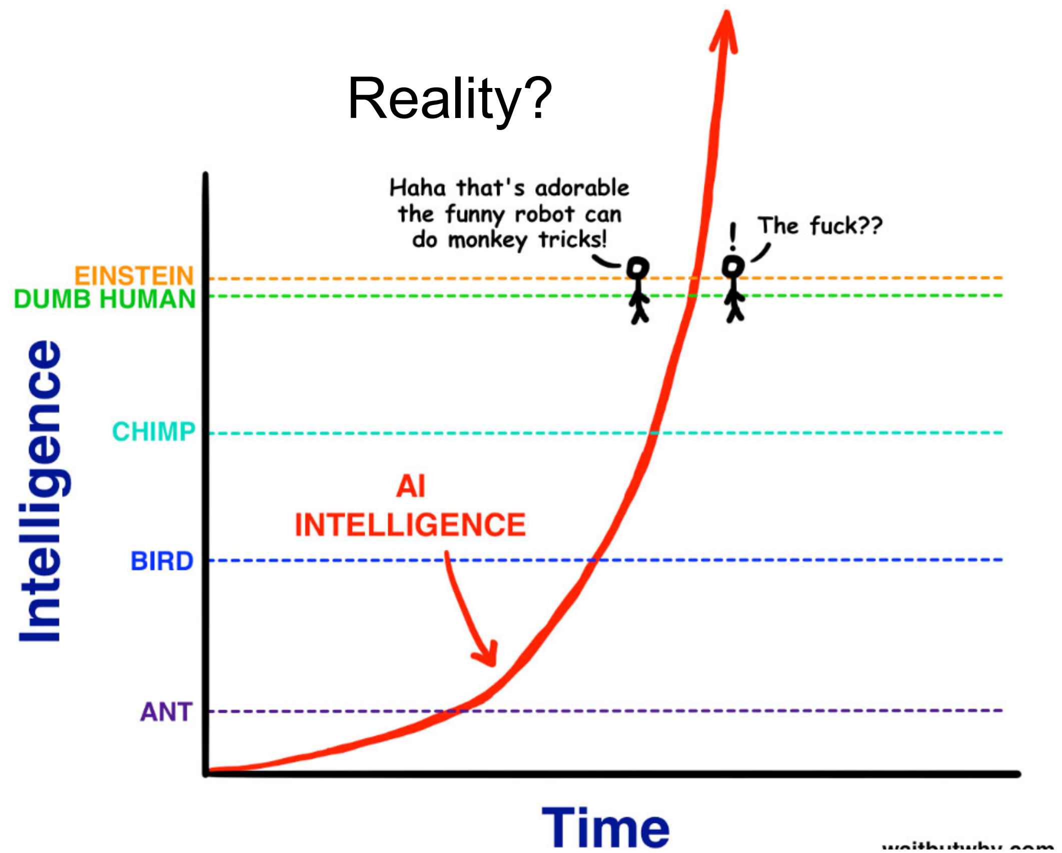


Superintelligence on the horizon?

Our distorted view

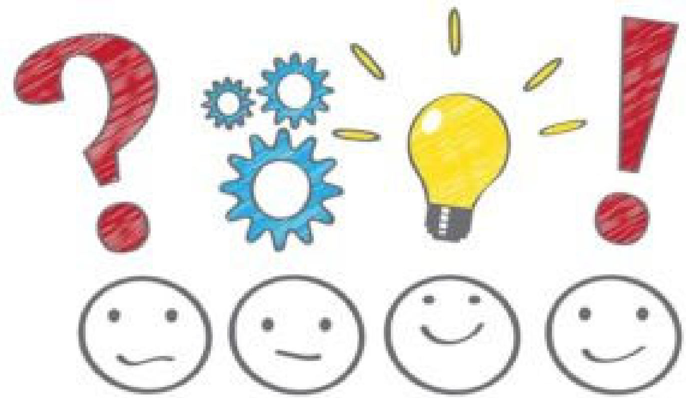


Reality?



Questions, discussion

Thank you for your attention!



Limitations, challenges and special issues

- Causality vs correlative evidence
- Unintended consequences of using ML
  - overreliance on the capabilities of automation
  - “deskilling” in the longer run, i.e. reduction of the level of skill required to complete a task
- Privacy and security for enabling data use and sharing
- Ethical issues
- Practicalities: computational power, data access etc.

Predictions

✓ The market for AI solutions will have annual growth exceeding 50% over the next 5 years

✓ By 2019 Internet of Things will provide growing amounts of data and applications providing automated assistance

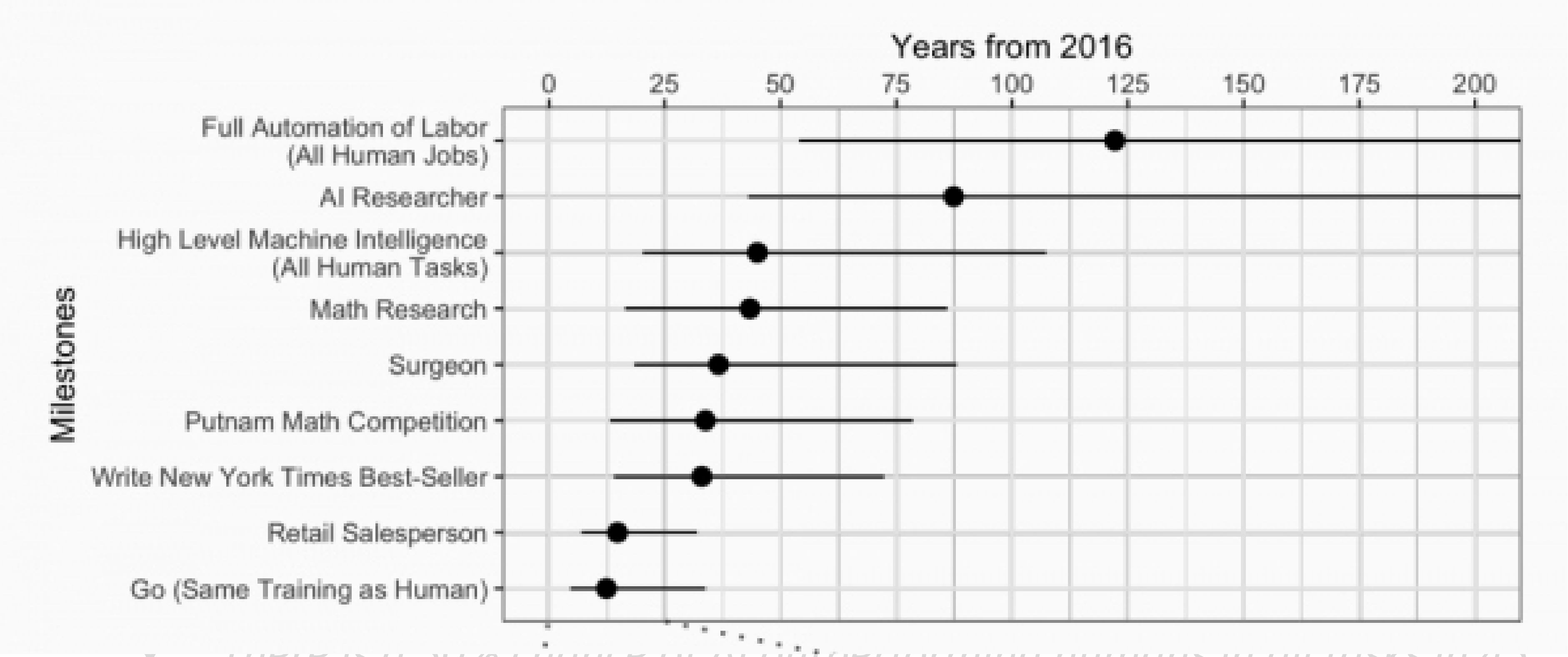
✓ AI will move towards unconstrained problem solving capabilities with increasing feasibility for general AI artefacts

✓ Computing architectures will be increasingly geared towards AI solutions (AI supercomputers, neuromorphic chips, etc.)

✓ *"There is a 50% chance of AI outperforming humans in all tasks in 45 years and of automating all human jobs in 120 years"*

Predictions

✓ The market for AI solutions will have annual growth exceeding 50%



✓ *“There is a 50% chance of AI outperforming humans in all tasks in 45 years and of automating all human jobs in 120 years”*