POLITECNICO
MILANO 1863

Machine Learning

- Course Intro -

Matteo Matteucci, PhD (matteo.matteucci@polimi.it)
Artificial Intelligence and Robotics Laboratory
Politecnico di Milano

AIRLAB
ARTIFICIAL INTELLIGENCE AND ROBOTICS LAB

«Me, Myself, and I»

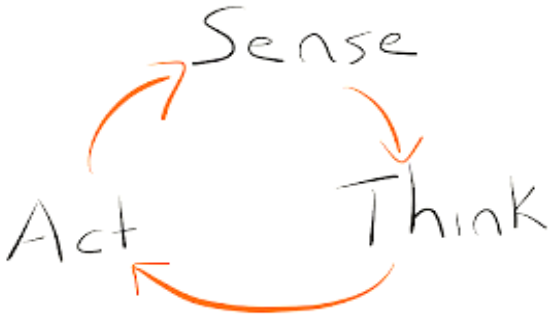
Matteo Matteucci, PhD
Full Professor
Dept. of Electronics, Information & Bioengineering
Politecnico di Milano
matteo.matteucci@polimi.it

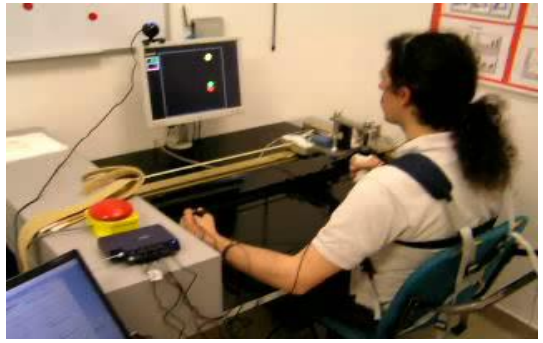
My research interests

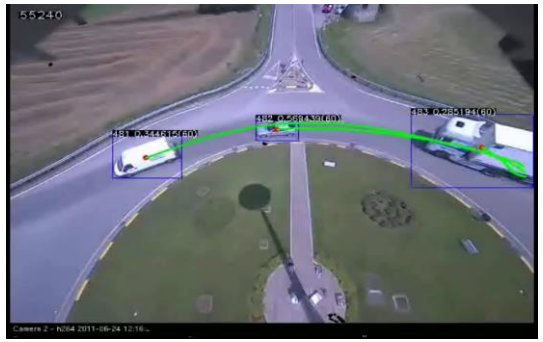
- Robotics & Autonomous Systems
- Machine Learning
- Pattern Recognition
- Computer Vision & Perception

Courses I teach

- Robotics (BS + MS)
- Cognitive Robotics (MS)
- Machine Learning (MS)
- Deep Learning (PhD)







Enable physical and software autonomous systems to perceive, plan, and act without human intervention in the real world

Machine Learning

Third B4CG course edition …

- Based on a previous four year edition course
- Based on a freely available book
- Practical exercises with python (new)

Lectures given by:

- Matteo Matteucci (Lecturer): matteo.matteucci@polimi.it
- Stefano Samele (Teaching Assistant): stefano.samele@polimi.it

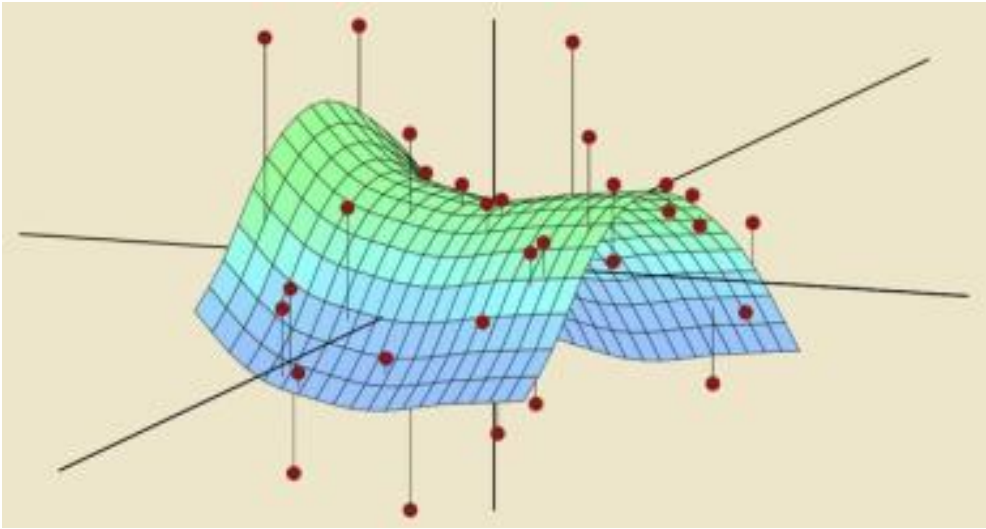
Material (and more) on the course website

https://chrome.deib.polimi.it/index.php?title=Machine_Learning_Bio

ML Course Program and Laboratories

Topics covered during the course

- Statistical Learning Theory
- Regression methods
- Classification methods
- Clustering methods



Practicals

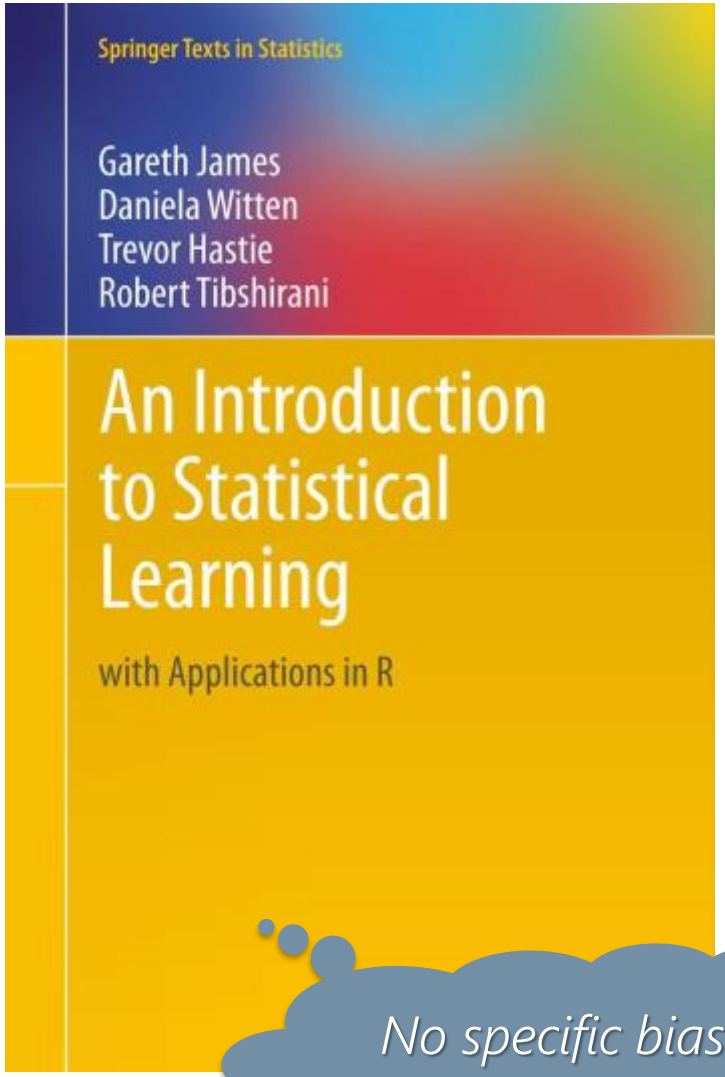
- Need to have a laptop/desktop (No need to run under linux)
- Need to install anaconda and python 3.7+
  https://docs.anaconda.com/anaconda/install/
- Need to do this in a couple of weeks!

Course Material

Book for this year edition

- Same authors of ESL (the best!), but easier!
- Practical perspective
- Labs and Exercises using the R language
- Available online as free pdf

www.statlearning.com



No specific bias toward biology …

Material from the teachers

- Slides from the book (except clustering)
- Jupiter notebooks (from teaching assistant)
- Online resources (you can suggest too)

Hypothesis behind Book & Course (1-2)

1. *Many statistical learning methods are relevant and useful in a wide range of academic and non-academic disciplines, beyond just the statistical sciences.*

   - Rather than attempting to consider every possible approach (an impossible task), we concentrate on presenting the methods that we believe are most widely applicable.

2. *Statistical learning should not be viewed as a series of black boxes.*

   - No single approach will perform well in all possible applications. We attempt to carefully describe the model, intuition, assumptions, and trade-offs behind each of the methods that we consider.

Hypothesis behind Book & Course (3-4)

3. *While it is important to know what job is performed by each cog, it is not necessary to have the skills to construct the machine in the box!*

   - We minimize discussion of technical details related to fitting procedures and theoretical properties. We assume the reader is comfortable with basic mathematical concepts, but we do not assume a graduate degree in mathematical science.

4. *We presume that the reader is interested in applying statistical learning methods to real-world problems.*

   - We devote some time to R computer labs. In each lab, we walk through a realistic application of the methods considered during the lectures.

Exam Rules for Machine Learning

The exam grade is composed by a written exam & homework

- WE - Written Exam is compulsory and up to 26/32 points
  - 2 Theoretical questions + 2 Practical Exercises (no code)
  - Covering the 4 topics previously introduced

Replaced by an oral exam in 2019/2020!

- HW – Homeworks, up to 6/32 points are NOT compulsory, but …
  - … they help you in understanding the course and prepare WE
  - … they get you in the "nitty gritty" of implementations
  - … they increase the final grade!
- Final GRADE combines the two
  - GRADE = 26/32+6/32 = 30 *cum Laude*

Questions?

POLITECNICO
MILANO 1863

Machine Learning

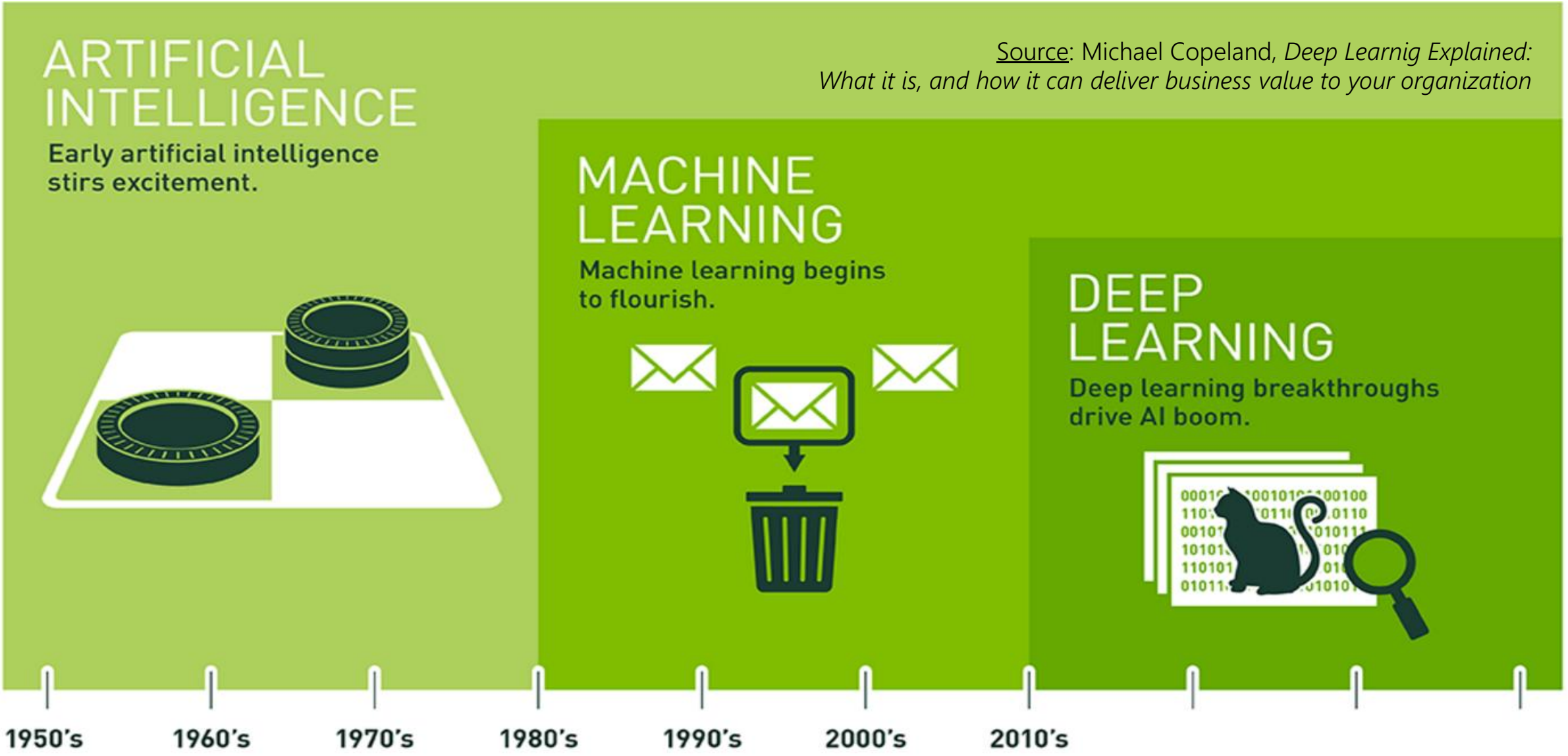
- Intro to Machine Learning -

Matteo Matteucci, PhD (matteo.matteucci@polimi.it)
Artificial Intelligence and Robotics Laboratory
Politecnico di Milano

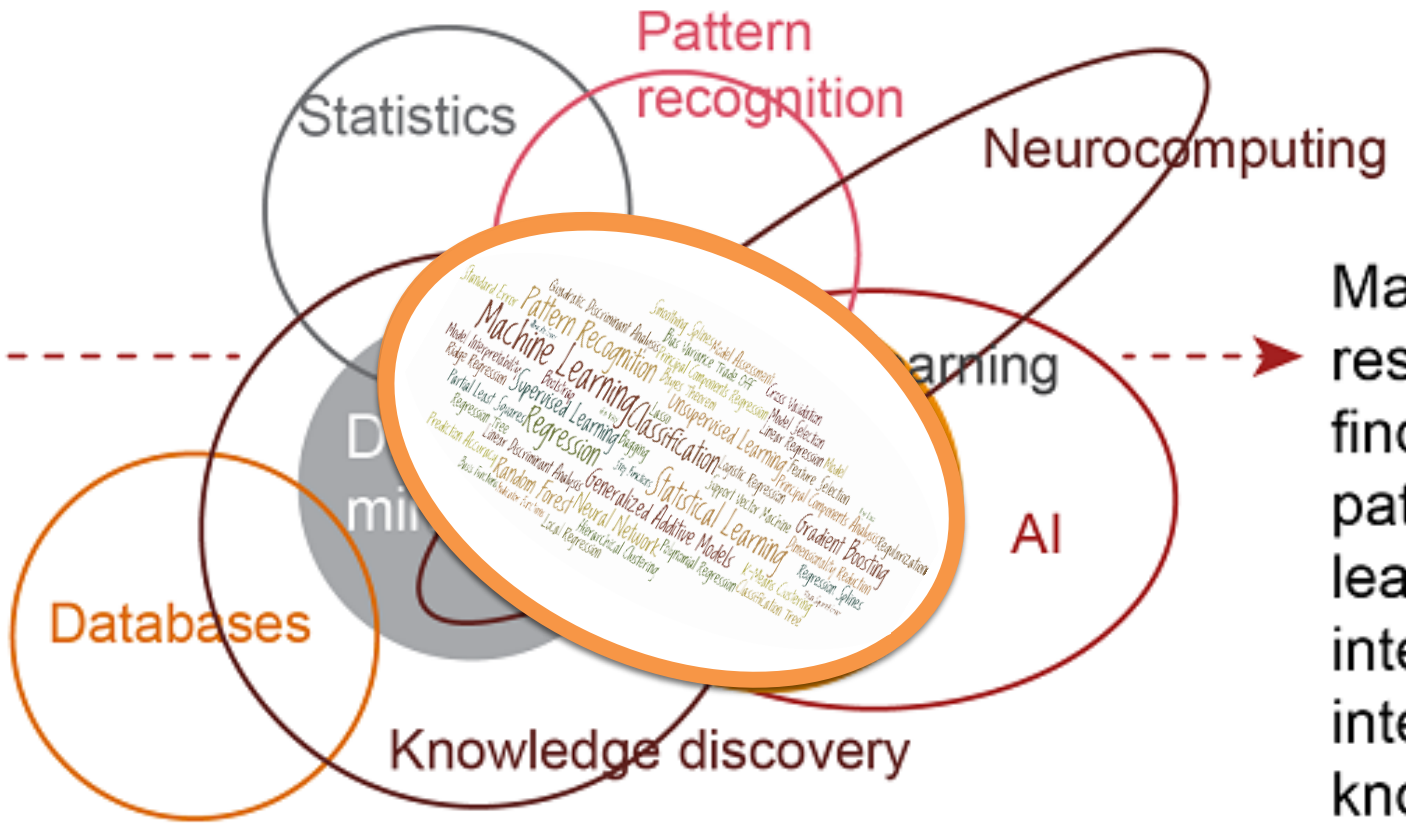
AIRLAB
ARTIFICIAL INTELLIGENCE AND ROBOTICS LAB

«Deep Learning is not AI, nor Machine Learning»

Source: Michael Copeland, *Deep Learnig Explained: What it is, and how it can deliver business value to your organization*



Machine Learning



Machine learning is a category of research and algorithms focused on finding patterns in data and using those patterns to make predictions. Machine learning falls within the artificial intelligence (AI) umbrella, which in turn intersects with the broader field of knowledge discovery and data mining.

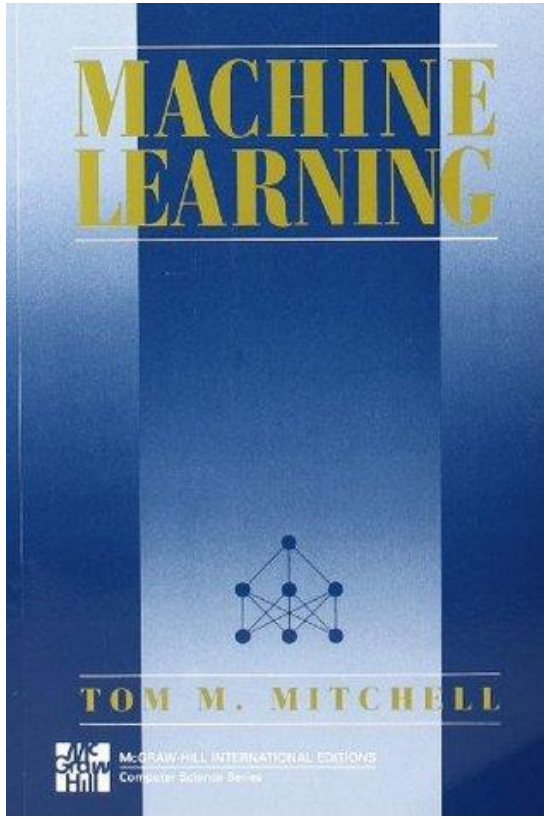
Source: SAS, 2014 and PwC, 2016 *and myself, 2017*

Machine Learning

Machine Learning (Tom Mitchell – 1997)

T = Regression/Classification/…
E = Data
P = Errors/Loss

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



Machine Learning Paradigms

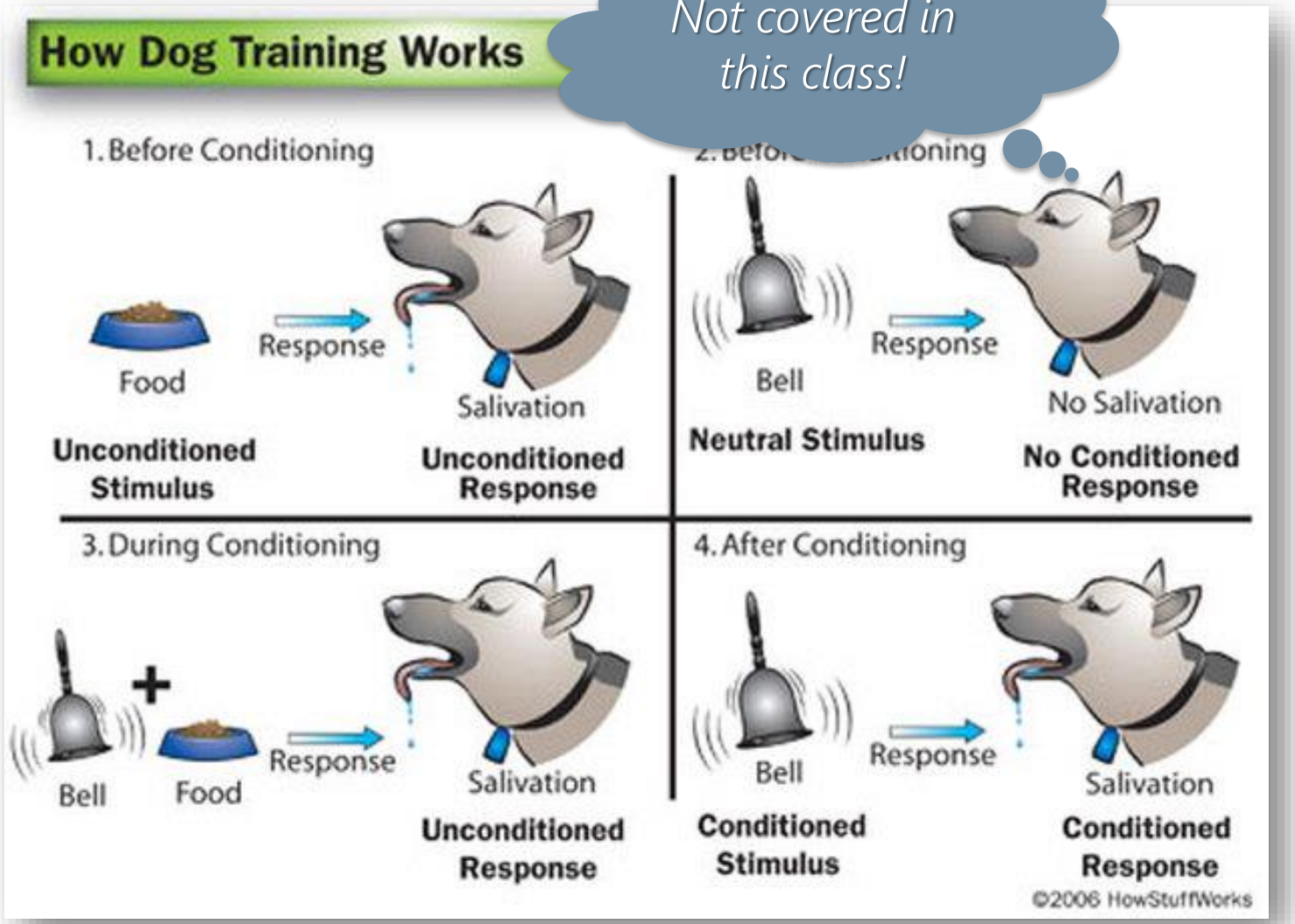
Imagine you have a certain experience E, i.e., data, and let's name it

$$D = x_1, x_2, x_3, \dots, x_N$$

- ***Supervised learning***: given the desired outputs $t_1, t_2, t_3, \dots, t_N$ learn to produce the correct output given a new set of input
- ***Unsupervised learning***: exploit regularities in D to build a representation to be used for reasoning or prediction
- ***Reinforcement learning***: producing actions $a_1, a_2, a_3, \dots, a_N$ which affect the environment, and receiving rewards $r_1, r_2, r_3, \dots, r_N$ learn to act in order to maximize rewards in the long term

Reinforcement Learning is Wellknown

Not covered in this class!



Machine Learning Paradigms

Imagine you have a certain experience E, i.e., data, and let's name it

$$D = x_1, x_2, x_3, \dots, x_N$$

- ***Supervised learning***: given the desired outputs $t_1, t_2, t_3, \dots, t_N$ learn to produce the correct output given a new set of input
- ***Unsupervised learning***: exploit regularities in D to build a representation to be used for reasoning or prediction
- ***Reinforcement learning***: producing actions $a_1, a_2, a_3, \dots, a_N$ which affect the environment, and receiving rewards $r_1, r_2, r_3, \dots, r_N$ learn to act in order to maximize rewards in the long term

Supervised learning: Classification

Cars

Motorcy

Learning is about modeling …

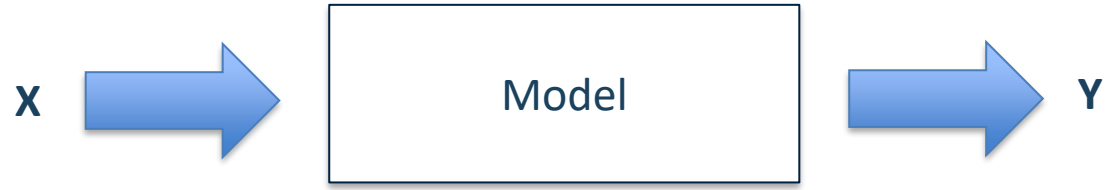
Hand-crafted Features

Learned Classifier

Motorcycle

Terminology in Classification

- Input
- Features
- Observations
- Independent Variables



- Output
- Class
- Dependent Variable

- Classifier
- Inductive Hypothesis
- Learning Machine
- …

Supervised learning: Regression

12000 $

15000 $

6000 $

2000 $

8000 $

22000 $

4000 $

28000 $

6000 $

35000 $

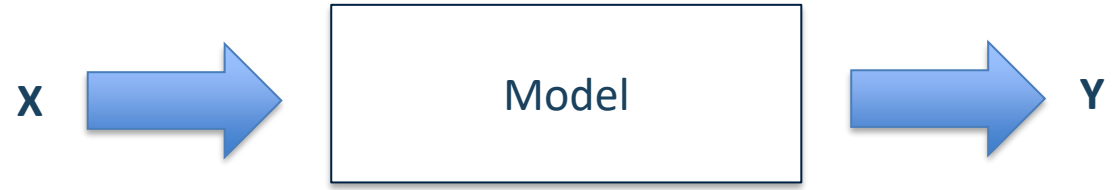
Hand-crafted Features

Learned Regressor

3800 $

Terminology in Regression

- Input
- Predictor
- Observations
- Independent Variable



- Output
- Prediction
- Response
- Dependent Variable

- Model
- Function
- Inductive Hypothesis
- Learning Machine
- …

Machine Learning Paradigms

Imagine you have a certain experience E, i.e., data, and let's name it

$$D = x_1, x_2, x_3, \dots, x_N$$

- ***Supervised learning***: given the desired outputs $t_1, t_2, t_3, \dots, t_N$ learn to produce the correct output given a new set of input
- ***Unsupervised learning***: exploit regularities in D to build a representation to be used for reasoning or prediction
- ***Reinforcement learning***: producing actions $a_1, a_2, a_3, \dots, a_N$ which affect the environment, and receiving rewards $r_1, r_2, r_3, \dots, r_N$ learn to act in order to maximize rewards in the long term

Unsupervised learning: Clustering




















Unsupervised learning: Clustering



















Unsupervised learning: Clustering




















Unsupervised learning: Clustering




















Unsupervised learning: Clustering




















Unsupervised learning: Clustering



















Unsupervised learning: Clustering



















Notation in Brief

In this case the training dataset is given by a set of input records

$$D =< x_1 >< x_2 >< x_3 >< \cdots >< x_N >$$

The task is to produce a partitioning of the data which highlights some knowledge about its organization.

Sometimes this knwoledge is named «patterns» …

Association Rules

Is there any pattern of interest in our trasactions?

Bread Peanuts Milk Fruit Jam	Bread Jam Soda Chips Milk Fruit	Steak Jam Soda Chips Bread	
Jam Soda Chips Milk Bread	Fruit Soda Chips Milk	Fruit Soda Peanuts Milk	Fruit Peanuts Cheese Yogurt

Association Rules

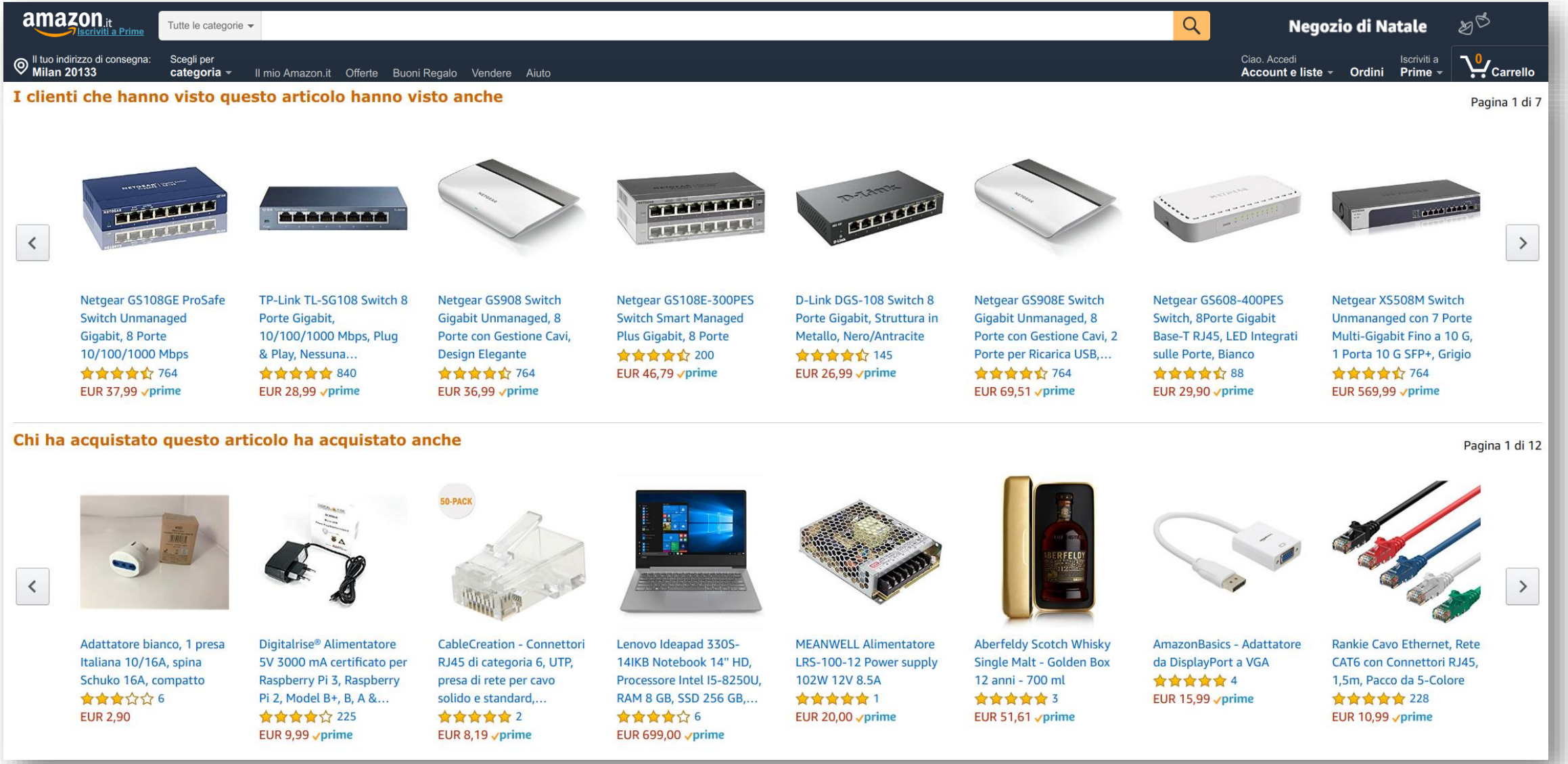
A large supermarket chain, Wal-Mart, did an analysis customers' buying habits.

They found a statistically significant correlation between purchases of beer and purchases of nappies (diapers in US).

It was theorized that the reason for this was that fathers were stopping off at Wal-Mart to buy nappies for their babies, and since they could no longer go down to the pub as often, would buy beer as well.

As a result of this finding, the supermarket chain is alleged to have the nappies next to the beer, resulting in increased sales of both.

This is for Real !!!



What are Machine Learning Major Tasks?

Most common tasks regard:

- Classification: predicting an item class / category / outcome
- Estimation/Regression: predicting a continuous value
- Clustering: finding clusters / groups in data
- Associations: detect frequent occurring events…
- Visualization: to facilitate human discovery
- Summarization: describing a group of data in a meaningful way
- Deviation Detection: finding changes in normal data patterns
- Link Analysis: finding relationship (e.g., social media, page-rank)

But many appears as time goes by …

- Outlier analysis, rare event analysis
- Trend and evolution analysis, sequential pattern mining
- Text Mining, Graph Mining, Data Streams
- Sentiment analysis, Reputation analysis, Opinion mining
- …

POLITECNICO
MILANO 1863

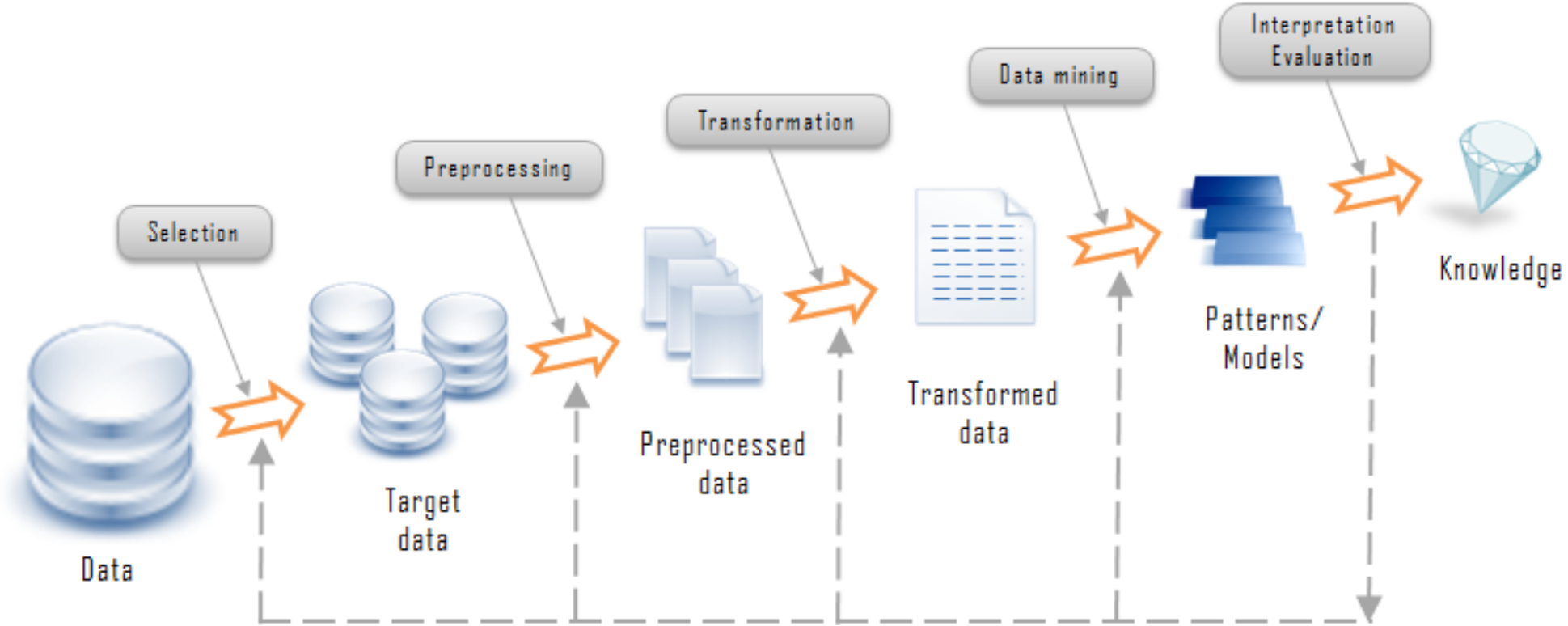
Machine Learning

- The Machine Learning Process -

Matteo Matteucci, PhD (matteo.matteucci@polimi.it)
Artificial Intelligence and Robotics Laboratory
Politecnico di Milano

AIRLAB
ARTIFICIAL INTELLIGENCE AND ROBOTICS LAB

Behind the courtain (1/4)



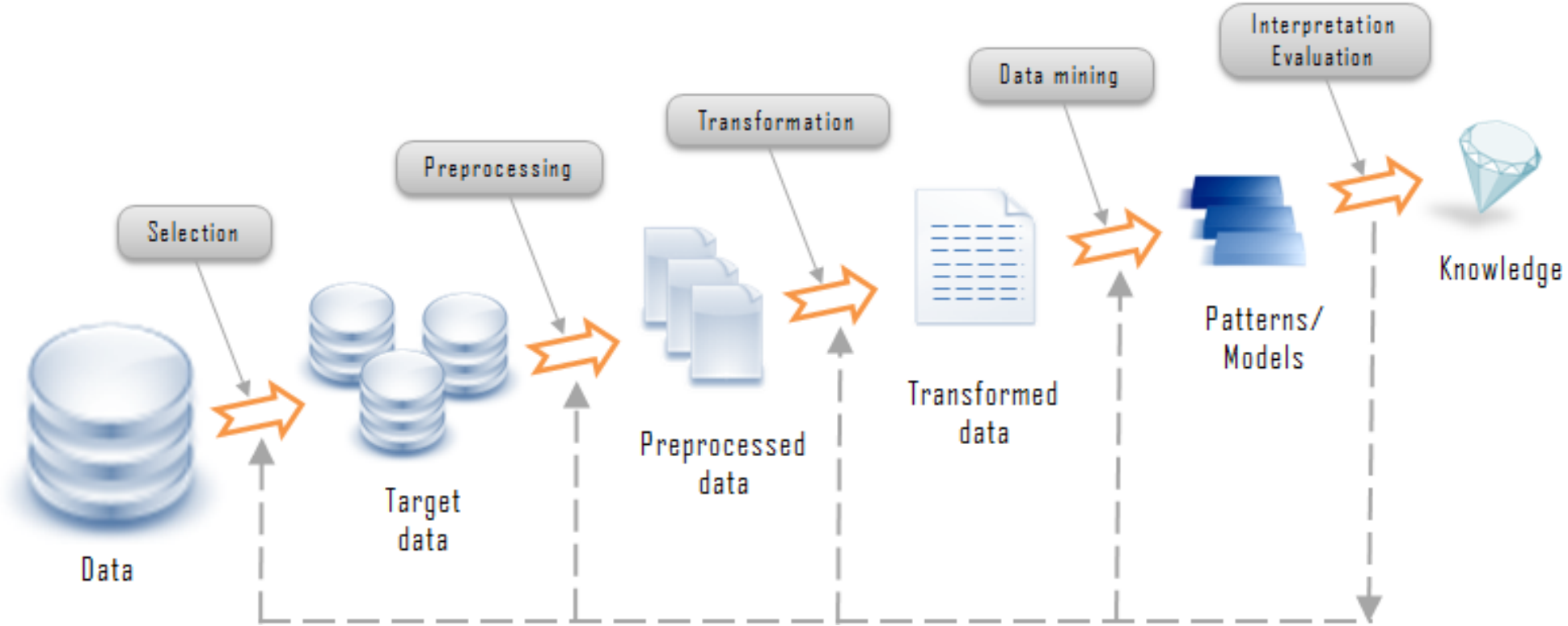
Selection

- What are data we actually need / have to answer the posed question?

Cleaning / Preprocessing

- Are there any errors or inconsistencies in the data we need to eliminate?

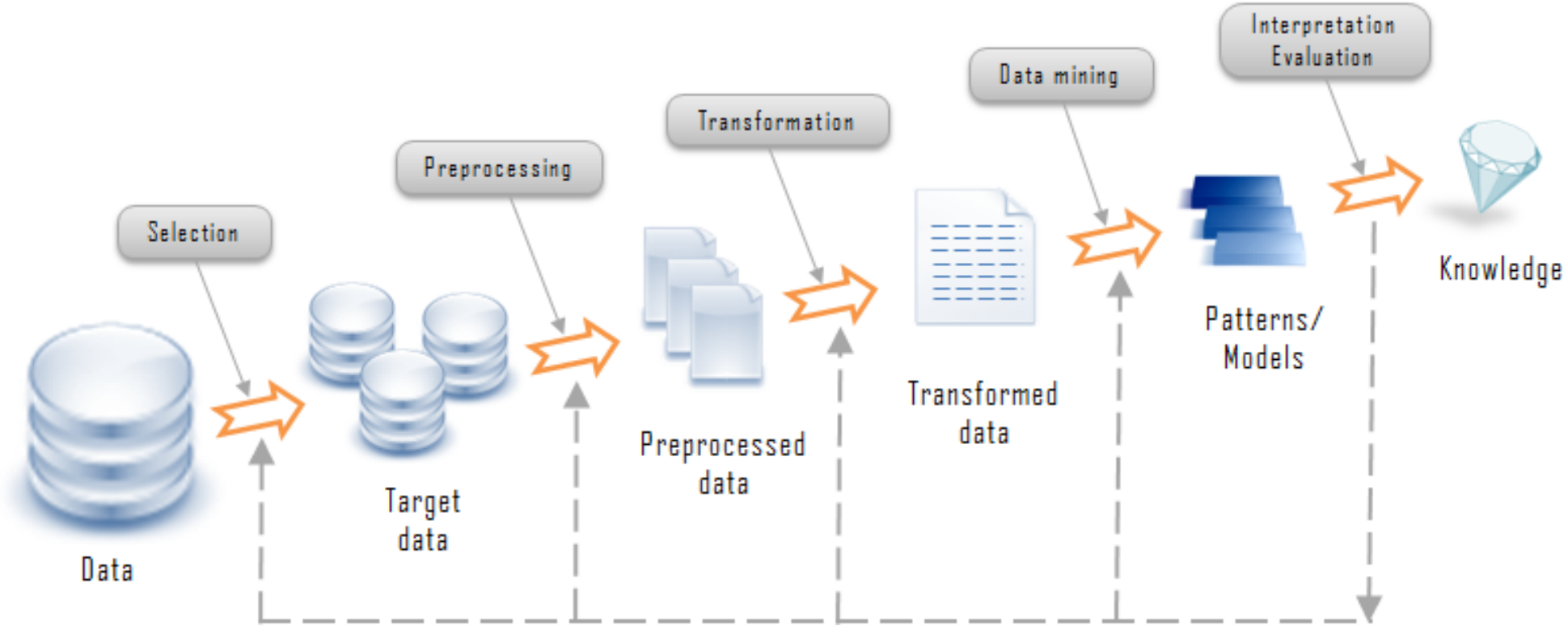
Behind the courtain (2/4)



Transformation

- Some variables might be eliminated because equivalent to others
- Some variables might be elaborated to create new variables (e.g. birthday to age, daily measures into weekly/monthly measures, log?)

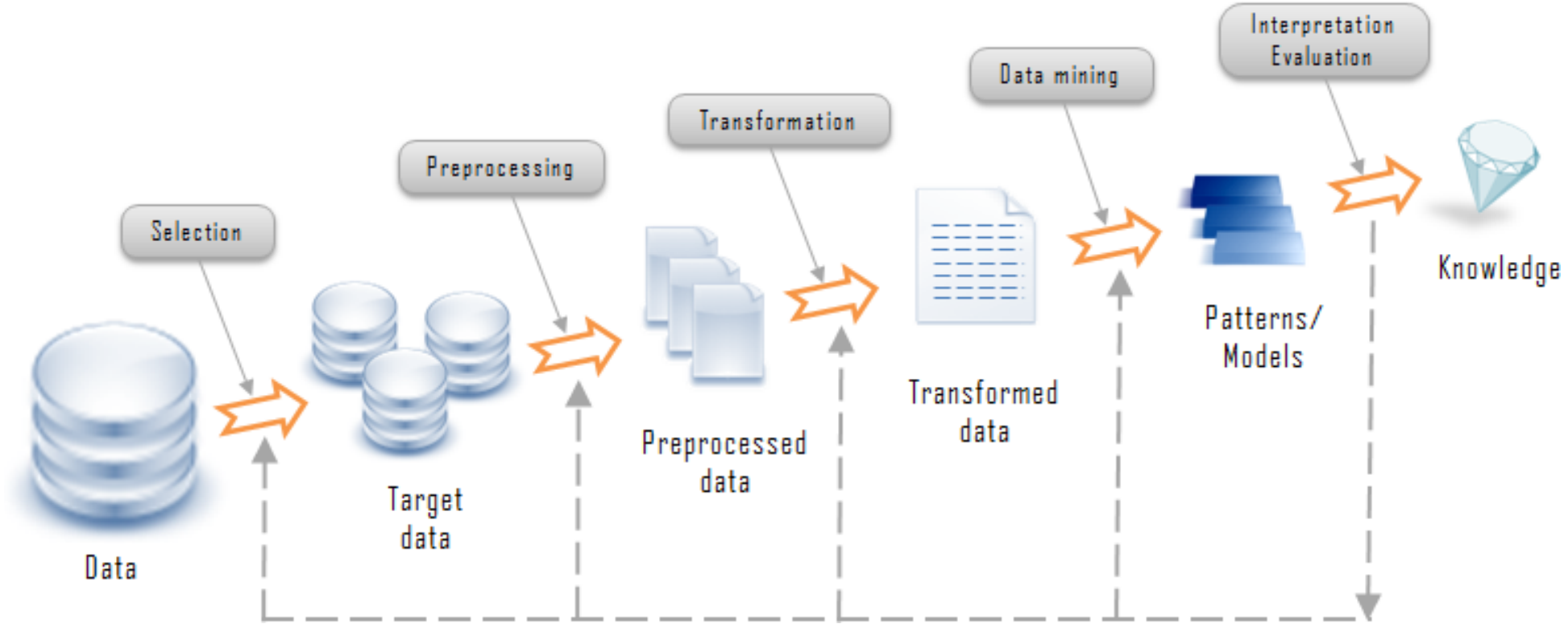
Behind the courtain (3/4)



Mining / Learning

- Select the mining /learning approach: classification, regression, association, etc.
- Choose and apply the mining algorithm(s)

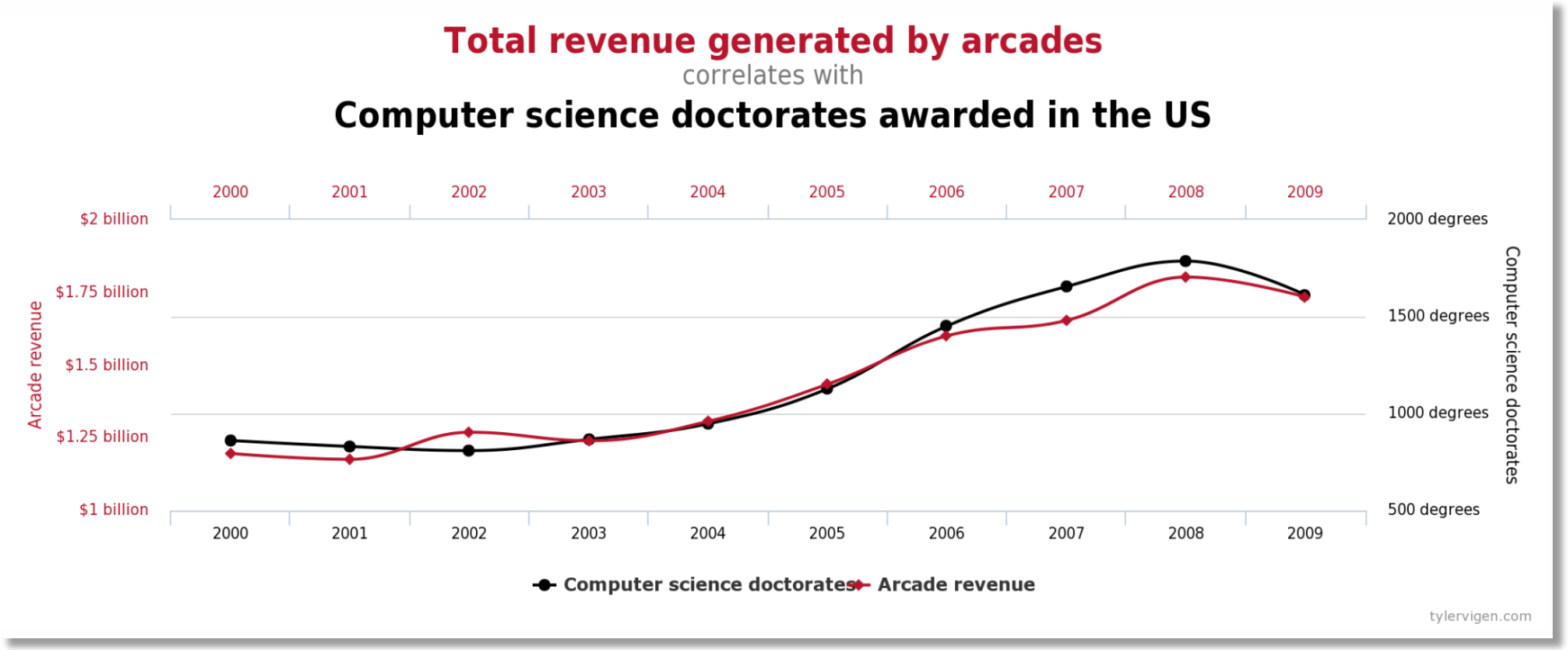
Behind the courtain (4/4)



Validation / Interpretation

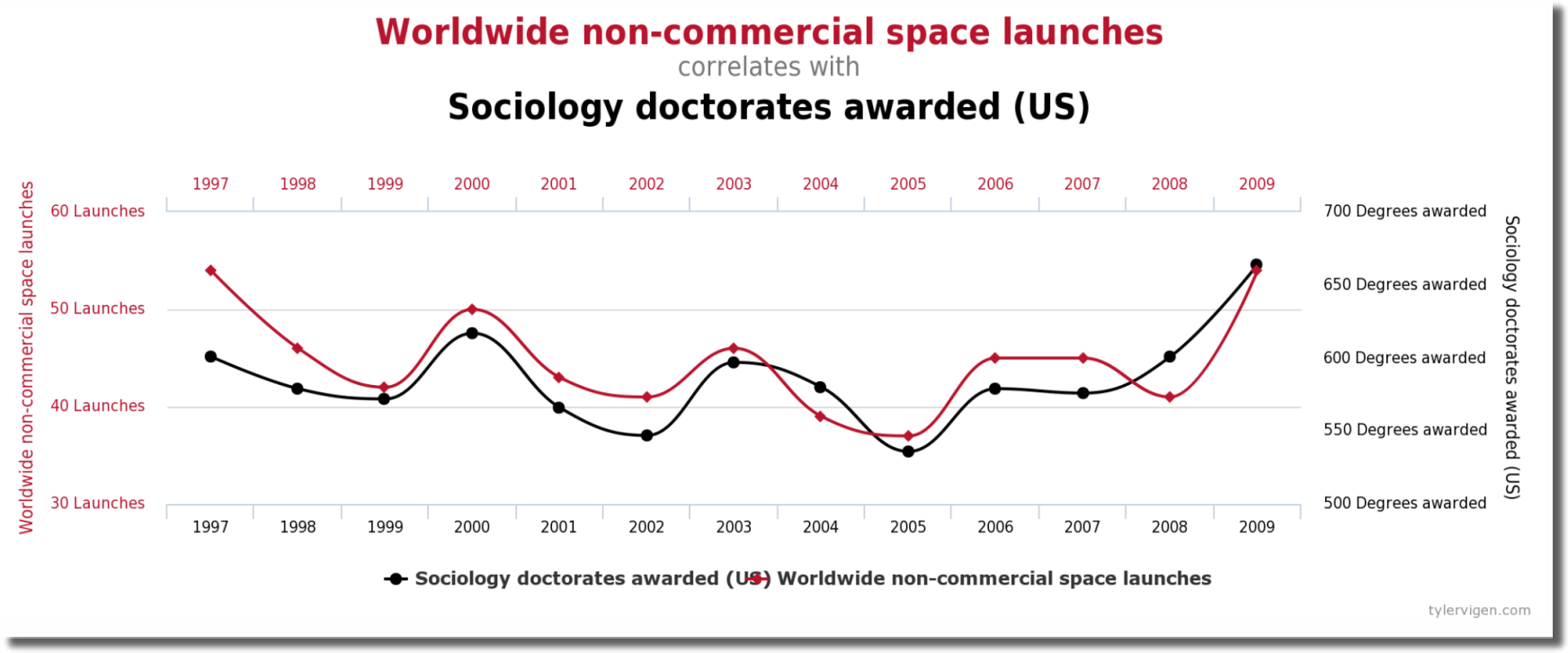
- How good are the results? Can we improve on those?
- Are the patterns we discovered sound? According to what criteria?
- Are the criteria sound? Can we explain the result?

Watch Out Correlations!



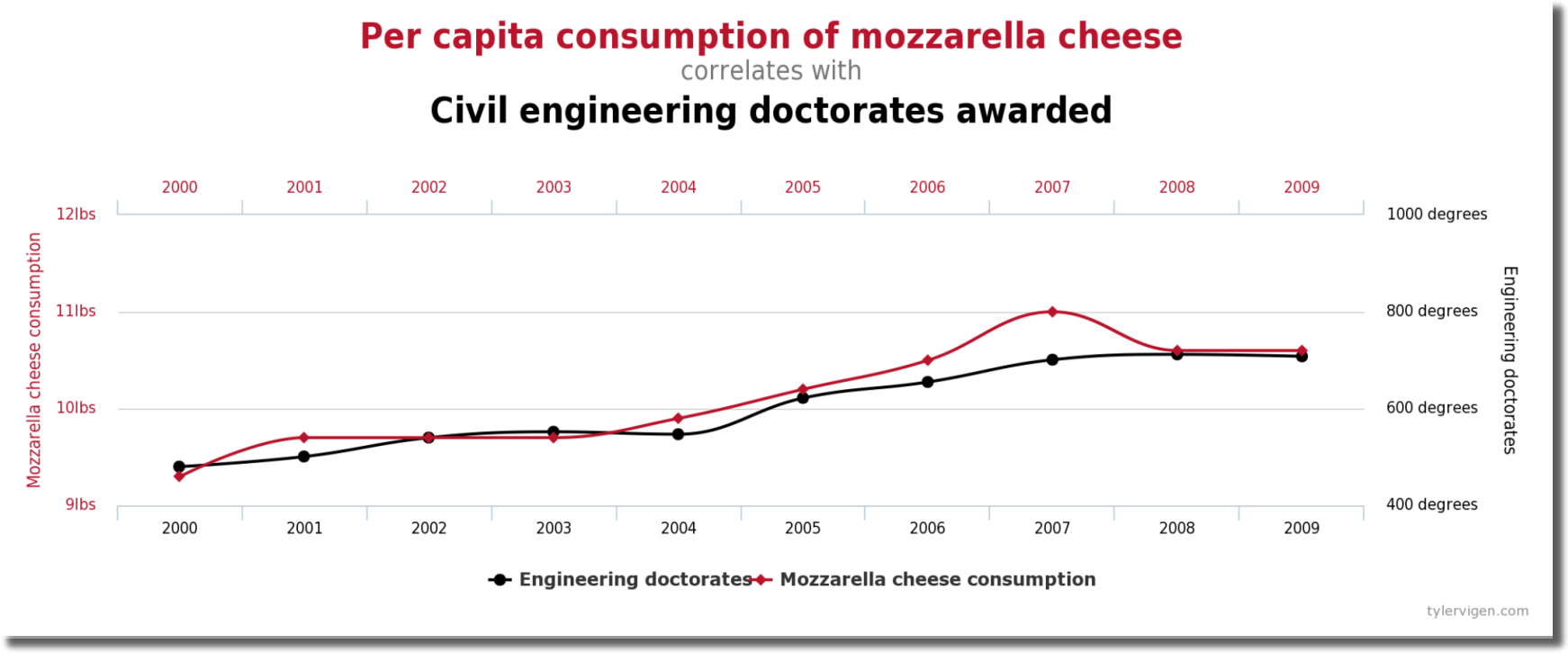
http://www.tylervigen.com

Watch Out Correlations!



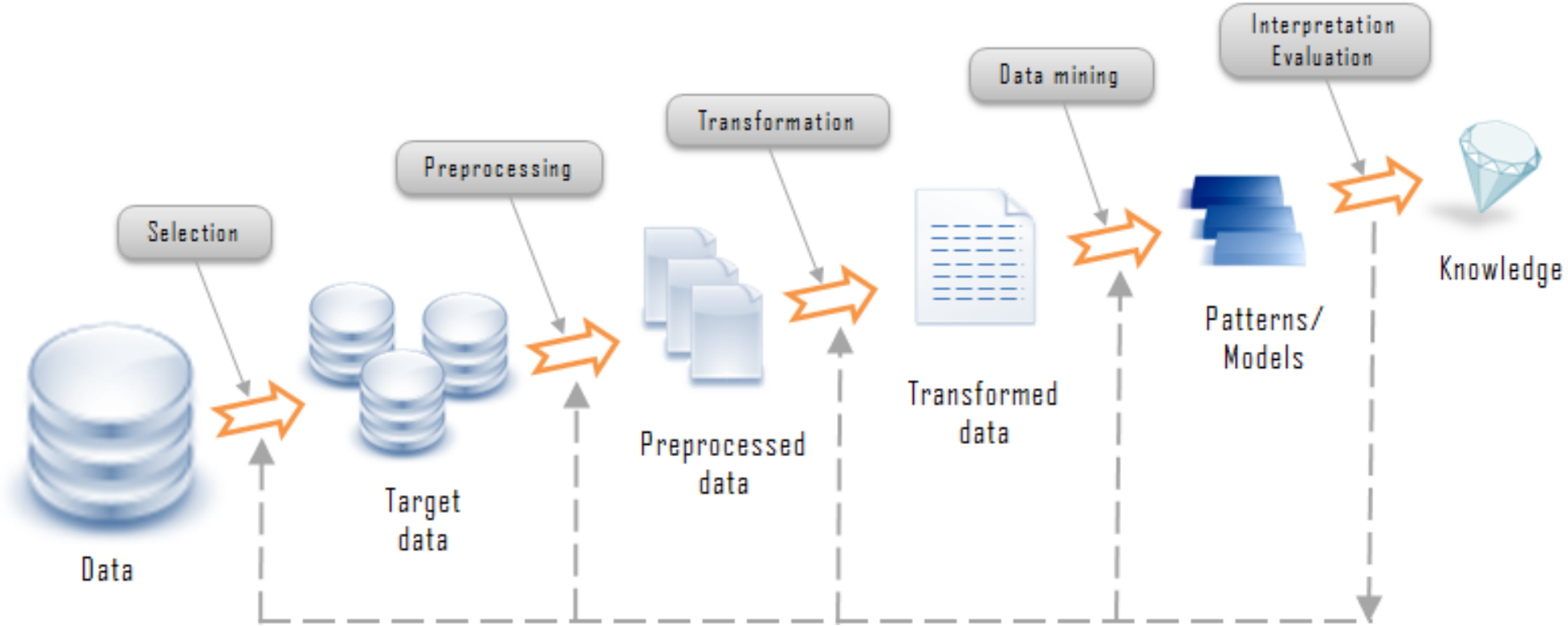
http://www.tylervigen.com

Watch Out Correlations!



http://www.tylervigen.com

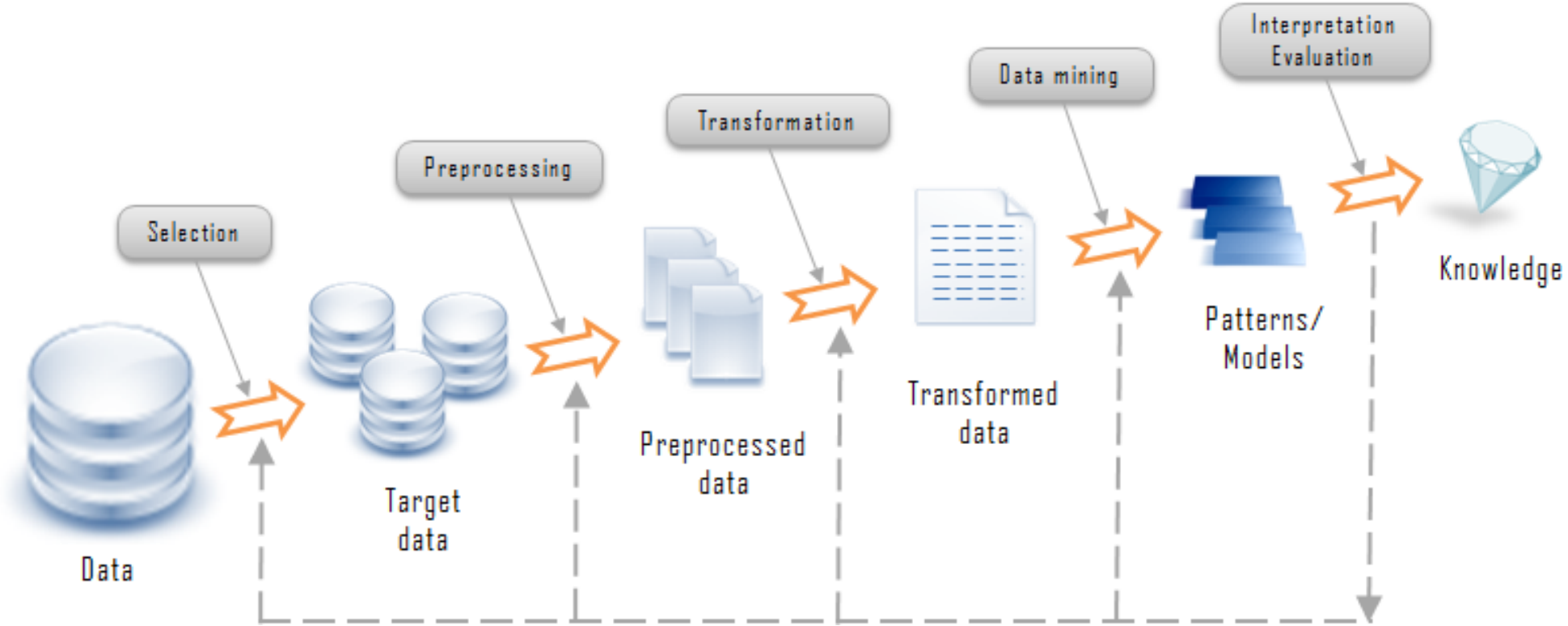
The human in the loop (1/3)



Learn about the application domain to extract relevant prior knowledge and define the goals for the mining / learning

Prepare data for the mining (data selection + cleaning + reduction + transformation)

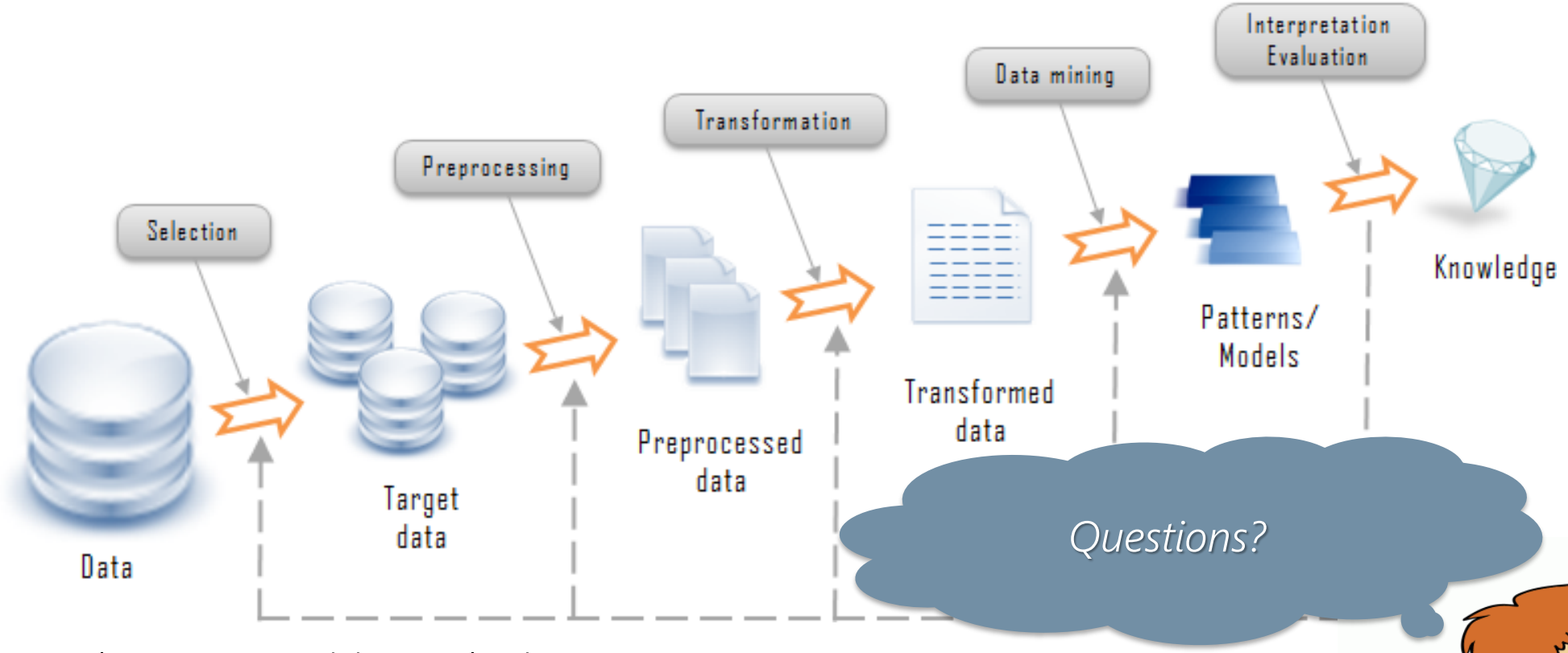
The human in the loop (2/3)



Mining / Learning

- Select the mining approach: classification, regression, association, clustering, etc. (this is related to the potential use of the result)
- Choose the mining algorithm(s) and perform mining / learning

The human in the loop (3/3)



Pattern evaluation and knowledge presentation

- Visualization, transformation, removing redundant patterns, etc.

Depending on the outcome

- Use of discovered knowledge or repeat the process from any of the previous steps