Introduction to Machine Learning

cphbusiness

tdi@cphbusiness.dk

Intended Learning Outcomes

☐ To get acquainted with Machine Learning as BI engine
- to get familiar with the terminology
- to understand the machine learning tasks, processes and implementations
- to understand supervised learning

☐ To make difference between exploratory and predictive analytics

TDI@CPHBUS

Machine Learning

- AI technology dealing with predicting the future based on the past, for instance,
    - predict the weather at specific location based on historical data
    - predict how much a person would like a movie that she hasn't seen, based on her ratings of movies that she has seen in the past
    - predict the future sells and prices of a product

- Making informed guesses about some unobserved property of an object, based on observed properties of that object

- Called machine learning, as the computer program 'learns' from the available observed data and later implements the learning outcomes for prediction even in a new, unobserved situation

Machine Learning Foundation

- **The idea** is that the unobserved feature, the *Output,* is an unknown function *F()* of the observed features, the *Input*

$$Output = F(Input)$$

- **The objective** is to reveal this function and later use it for predictions

$$Input = Features$$
$$Output = Label$$

- ML **algorithms** are used to extract patterns from data for the purpose of granting computers the powers to predict and draw inferences

Model

□ Realistic and relatively precise representation of an object or an event

□ Types

- business model
- physical model
- mathematical model

□ In machine learning

- mathematical models
- the function telling how the output variable depends on the input variable/s

$$\begin{cases} u(x) = g_1(x) + \int_0^1 (\frac{1}{2}t^2)u^2(t)dt + \int_0^1 xv(t)dt, \\ v(x) = g_2(x) + \int_0^1 u(t)dt + \int_0^1 \frac{1}{4}v^3(t)dt. \end{cases}$$

Machine Learning

Supervised:

knowing the labels before the learning phase

Unsupervised:

not knowing any labels before the training
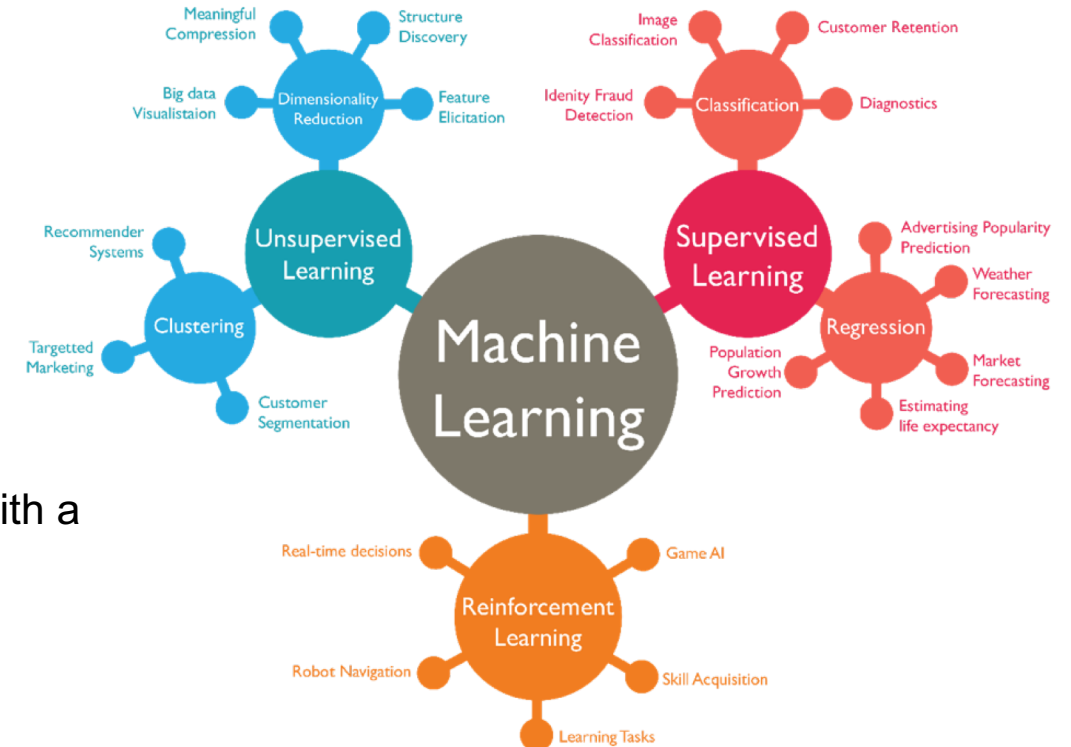
Classification

categorisation, association with a predefined class
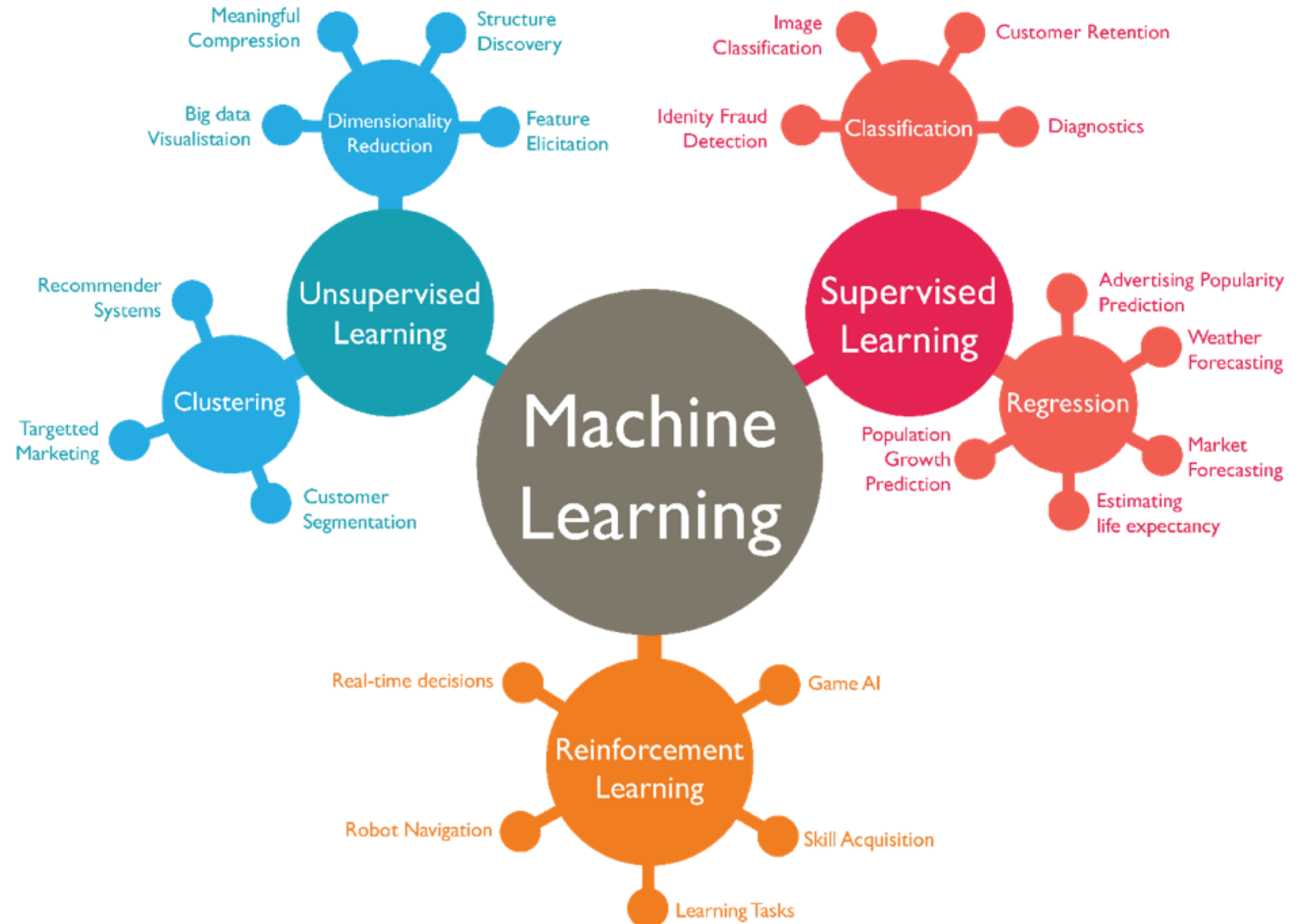- predicts qualitative values

Regression

estimation of a new, unknown real value
- predicts quantitative values
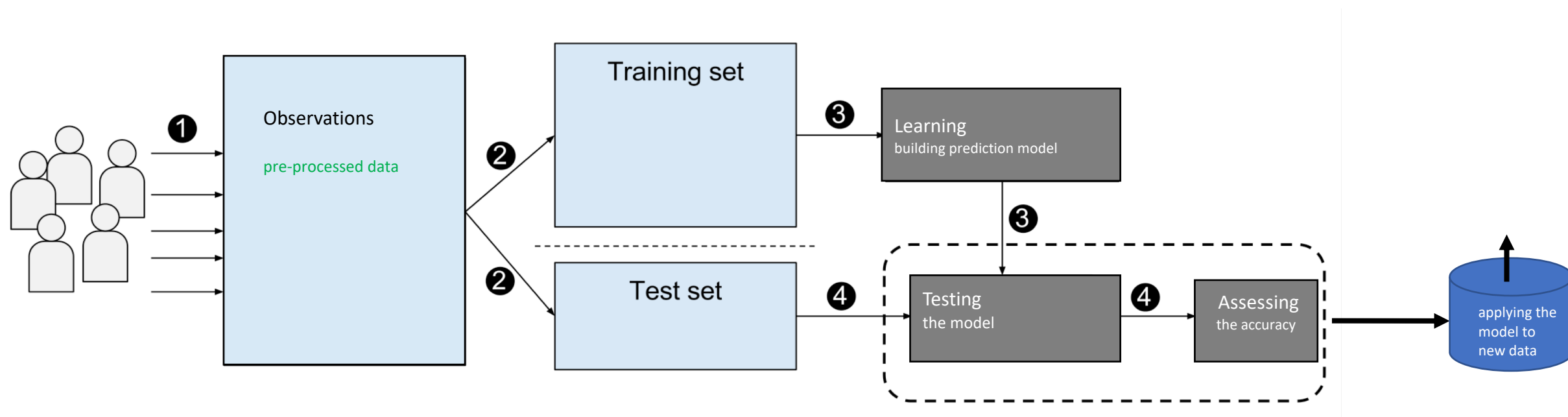
Clustering

categorisation, association with newly defined groups, clusters

Training Machine Learning Model

☐   Get data

☐   Visualize it

☐   Prepare it

☐   Split it into train and test subset

☐   Train the train subset

☐   Test prediction on the test subset

☐   Plot the regression line

☐   Validate with unknown data