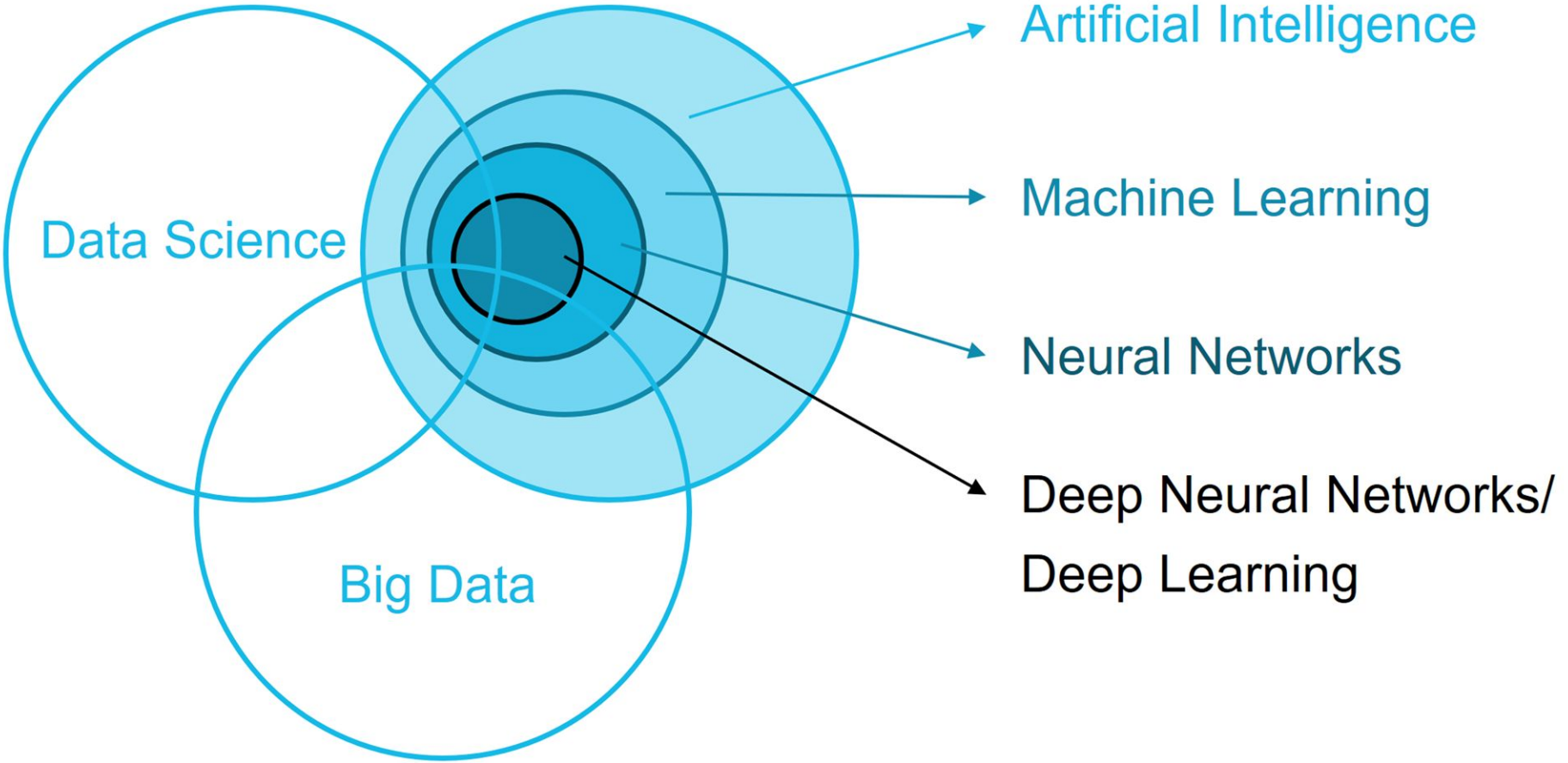
Introduction to ML

Machine Learning is the study

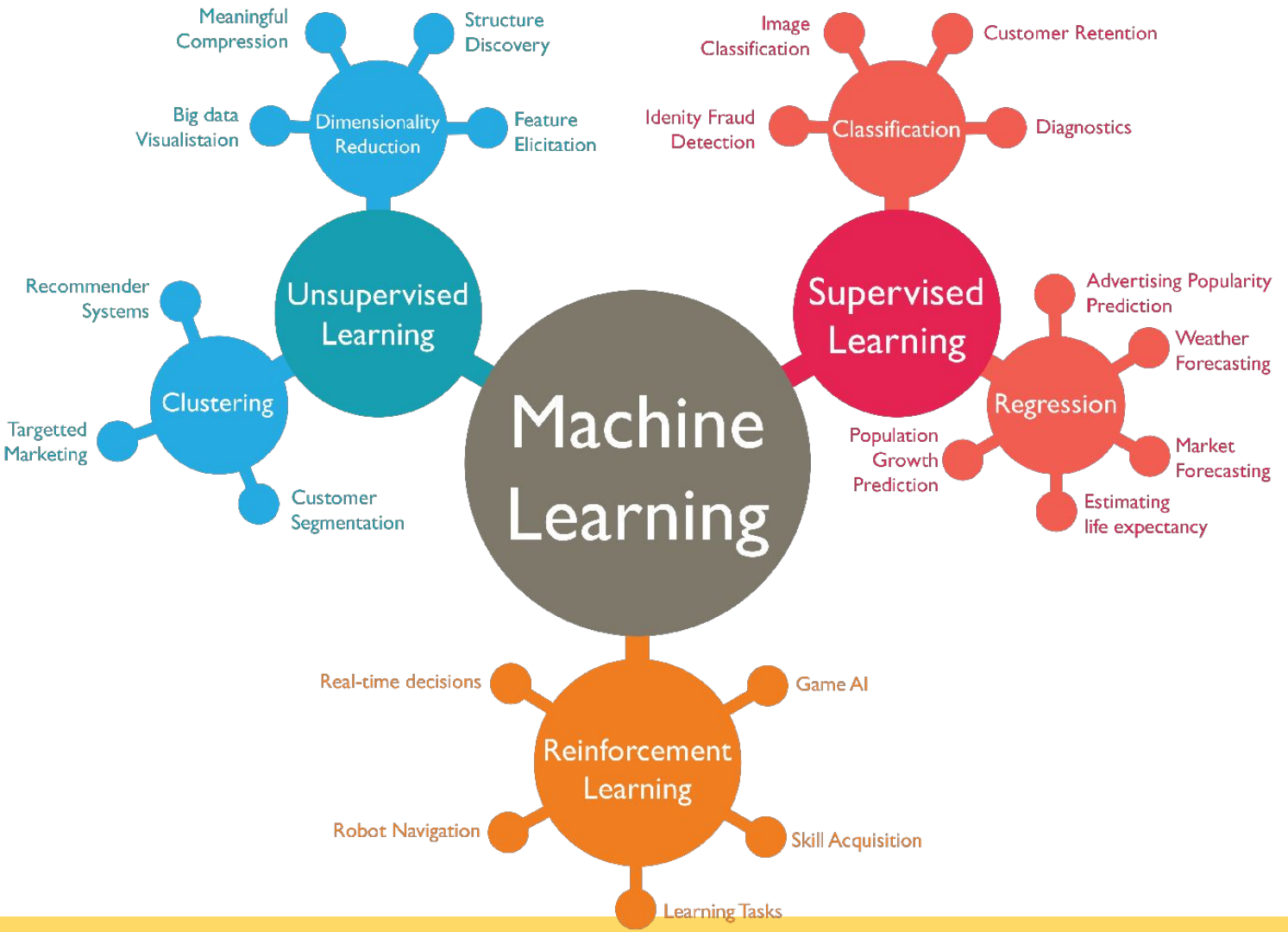
and production of algorithms

that can learn from

and make predictions on data



https://www.fokus.fraunhofer.de/en/fame/workingareas/ai


Machine Learning
Unsupervised Learning
Dimensionality Reduction
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Clustering
Recommender Systems
Targetted Marketing
Customer Segmentation
Supervised Learning
Classification
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Regression
Advertising Popularity Prediction
Weather Forecasting
Population Growth Prediction
Market Forecasting
Estimating life expectancy
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks

Supervised Learning

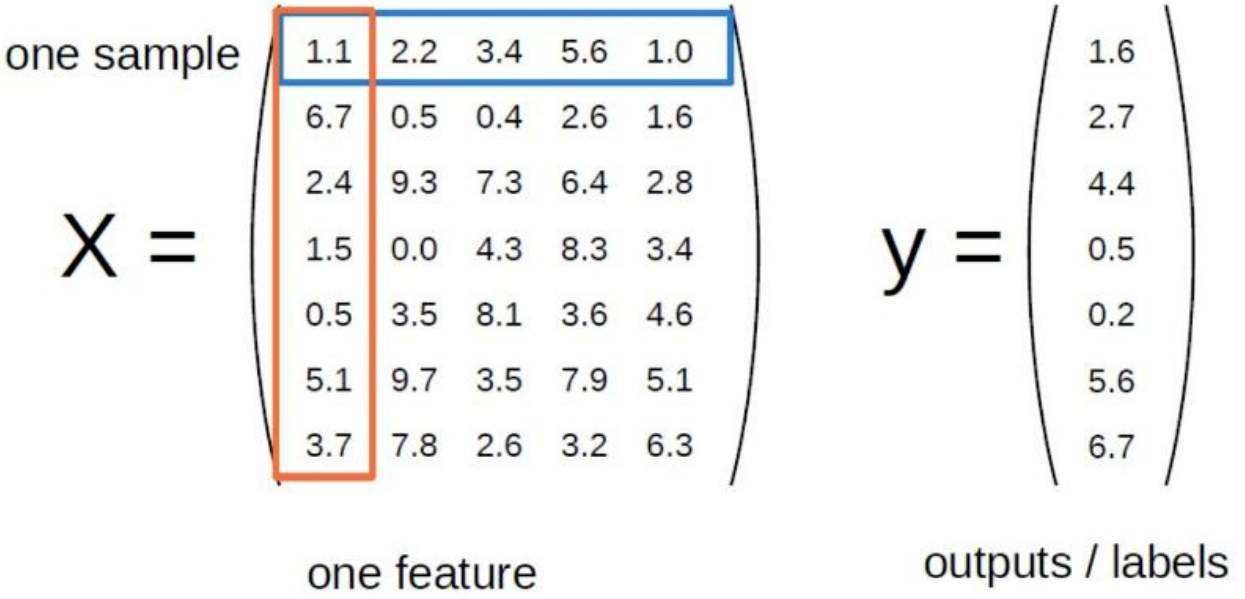
Formalism

- Hypothesis space: ***H***
- Vector spaces: ***X, y***
- Data: $\boldsymbol{X}_{train}$, $\boldsymbol{y}_{train}$
- Performance metric: L(f(X), y)
- **Assumption: *H*** is rich enough such that the true f* exists, f(X) ≃ y.
- **Goal:**

Find *f'* in ***H*** such that

$$L(f'(\mathbf{X}), \mathbf{y}) \simeq L(f^*(\mathbf{X}), \mathbf{y}) \simeq \min_{f \text{ in } H} L(f(\mathbf{X}), \mathbf{y})$$

Data Representation



Approximation

- First idea, minimize

$$L(f(X_{train}), y_{train}).$$

⇒ **Trivia:** what can go wrong?

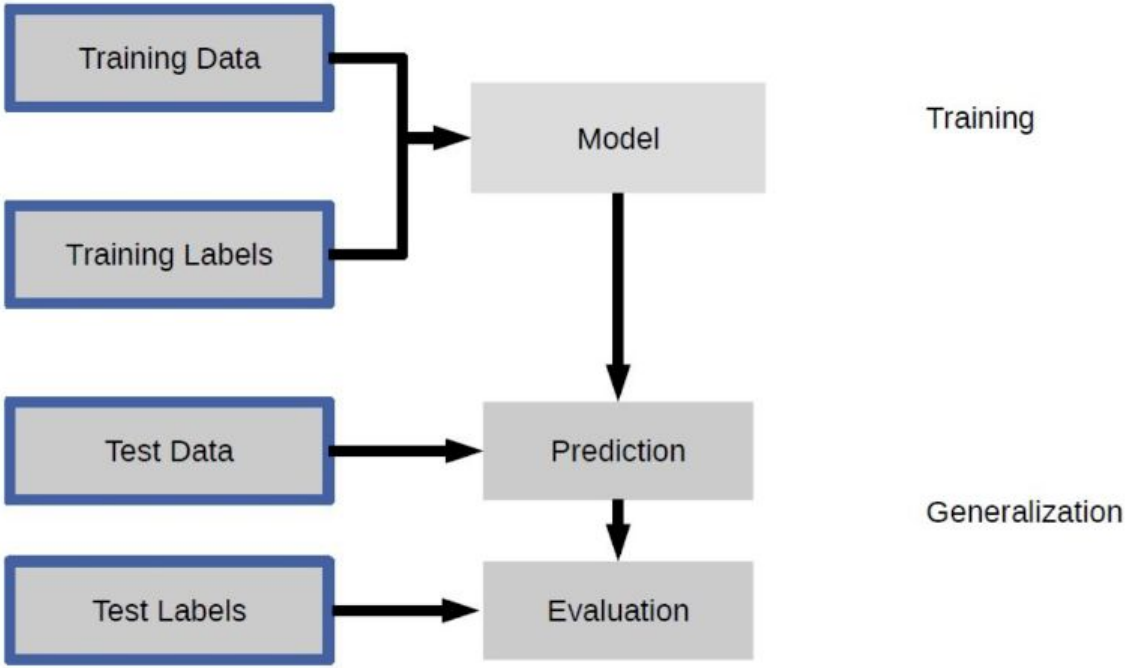
Generalization

- “Memorizing” the information we have can be optimistic but misleading.

- Our model should be able to work well on unseen circumstances.

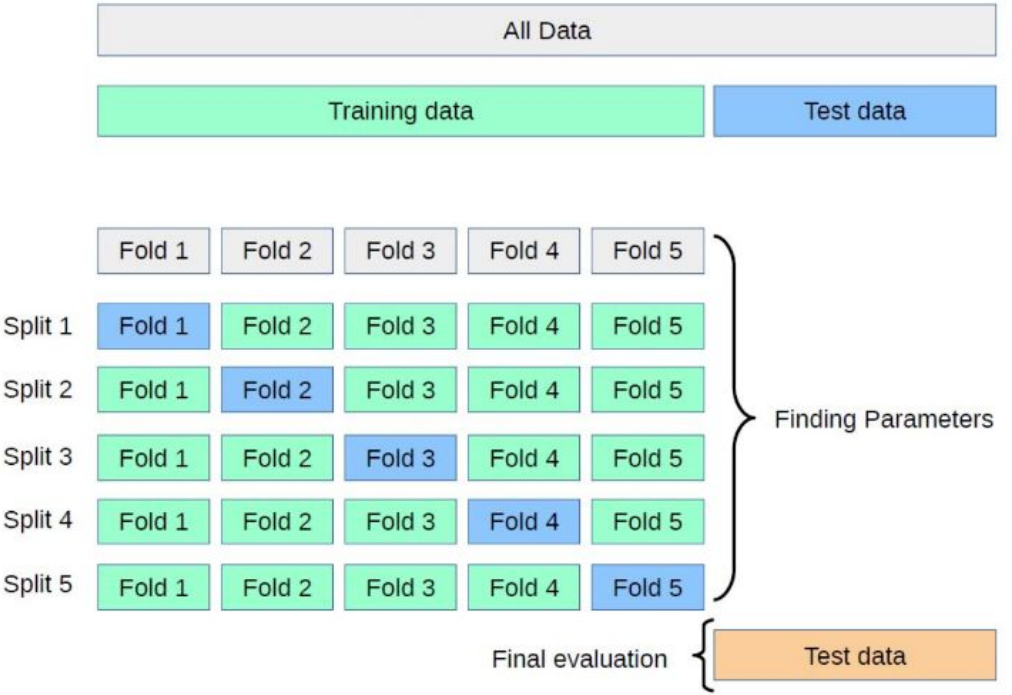
- Provided that our training data is a *representative sample* of the problem space, how can we do that?

Supervised Learning



Cross Validation

Holdout Validation



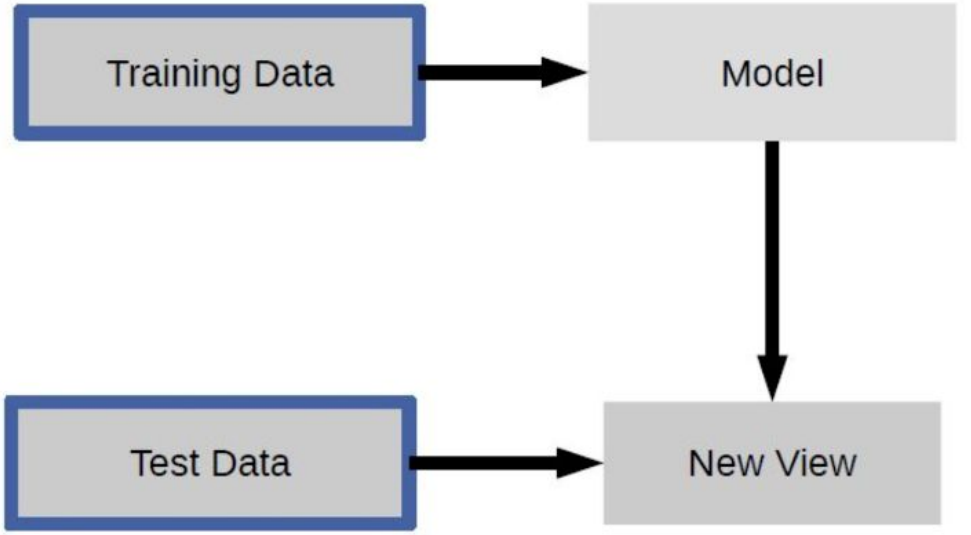
Supervised Learning Taxonomy

- Response type:
  - Classification
    - Binary classification
    - Multi-class
  - Regression
    - Single-output
    - Multi-objective
- Approximating family
  - Parametric
  - Non-parametric

Examples

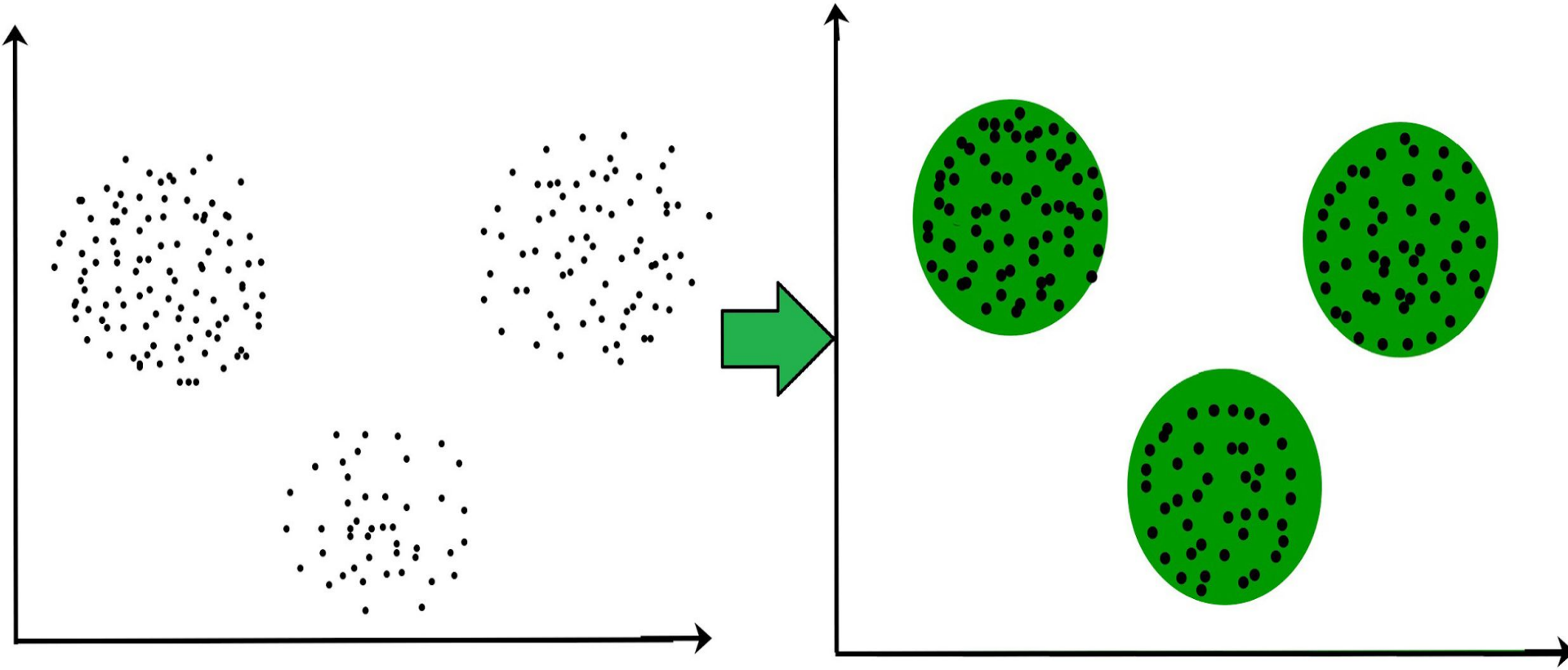
Unsupervised Learning

Unsupervised Learning



Examples

- Anomaly/change detection
- Customer Segmentation
- Clustering



Reinforcement Learning

Reinforcement Learning / Black-box Optimization

- Interaction of an agent in an unknown, difficult to describe environment.
- Sparse signal / Credit assignment problem
- **Requires access to real world (or at least to a simulator).**

Policy methods

- Parameterize the behaviour of the agent as a function of the state.
- Calculate obtained reward (or incurred loss).
- Generate a new parameter with an update rule that depends on the reward.


KBA
Google DeepMind
Challenge Match
8 - 15 March 2016
AlphaGo
ALPHAGO
00:00:54
AlphaGo
Lee Sedol

Recap

- Three types of machine learning
- Different input/output structure.
- Many applications!