INTRODUCTION TO MACHINE LEARNING (ML)

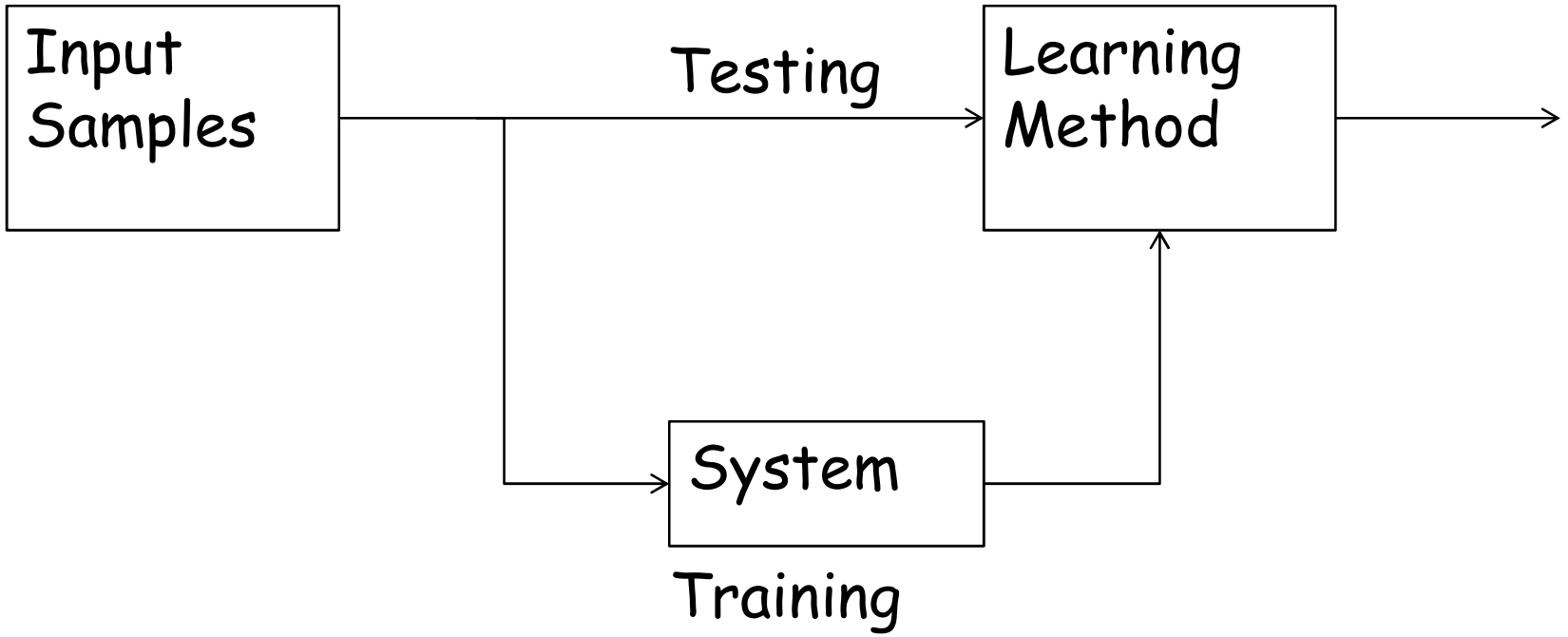
Outline & Content

- What is machine learning?
- Learning system model
- Training and testing
- Performance
- Algorithms
- Machine learning structure
- Learning techniques
- Applications

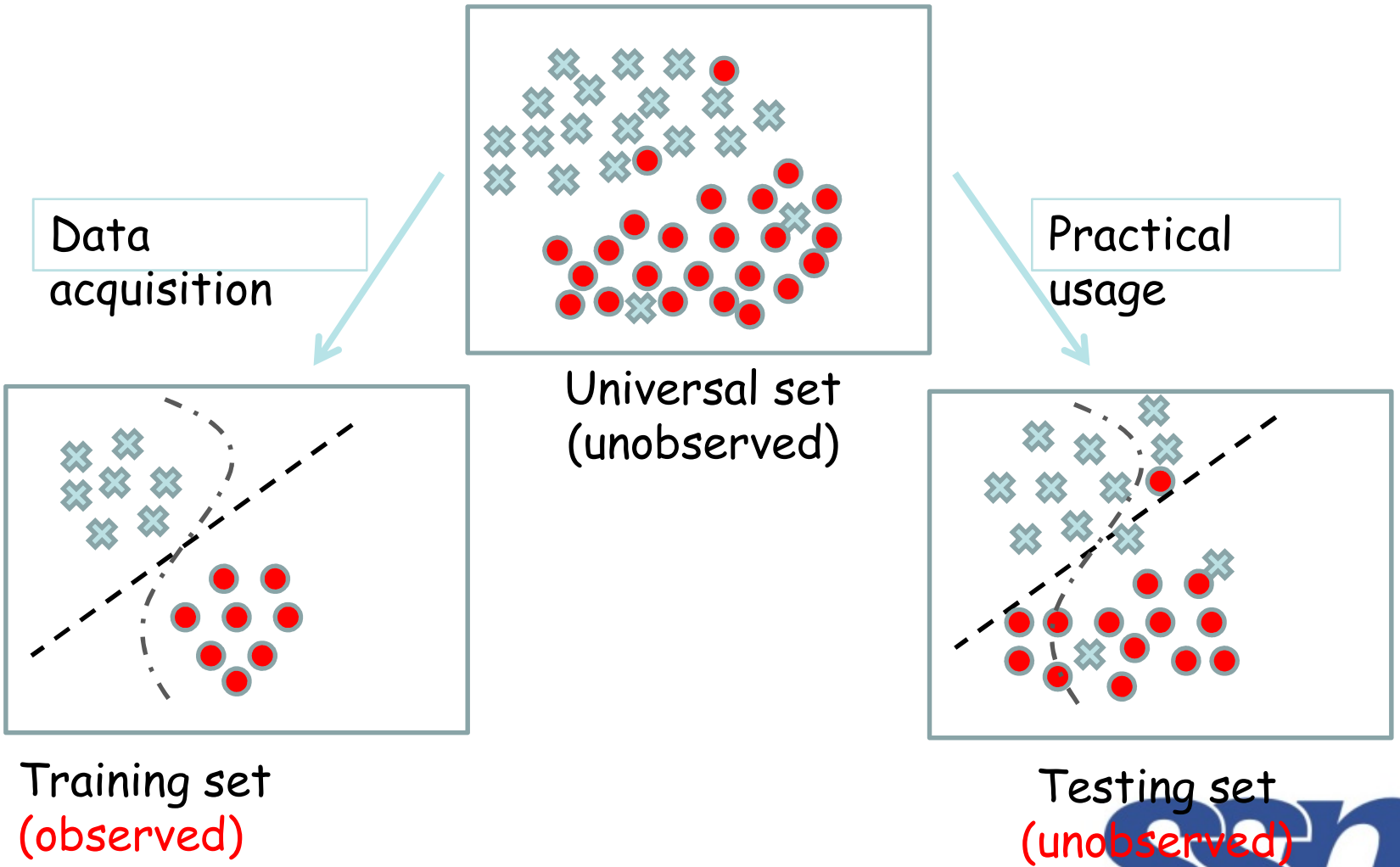
What is machine learning?

- A branch of **artificial intelligence**, concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data.
- As intelligence requires knowledge, it is necessary for the computers to acquire knowledge.
- Machine learning is concerned with using the right features to build the right models that achieve the right task

Learning system model



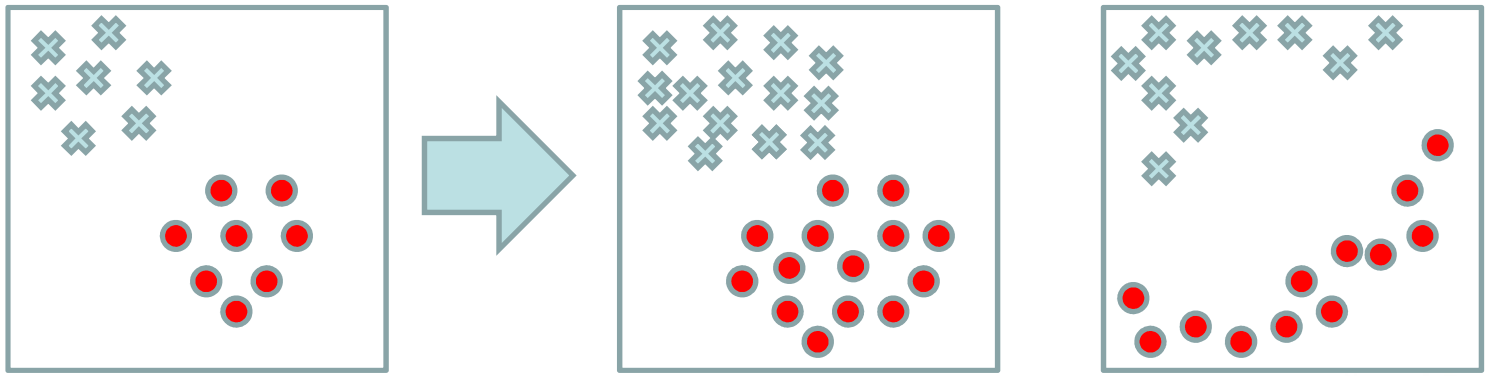
Training and testing



Training and testing

- Training is the process of making the system able to learn.

- No free lunch rule:
  - Training set and testing set come from the same distribution
  - Need to make some assumptions or bias



Performance

- There are several factors affecting the performance:
  - **Types of training** provided
  - The form and extent of any initial **background knowledge**
  - The **type of feedback** provided
  - The **learning algorithms** used

- Two important factors:
  - Modeling
  - Optimization

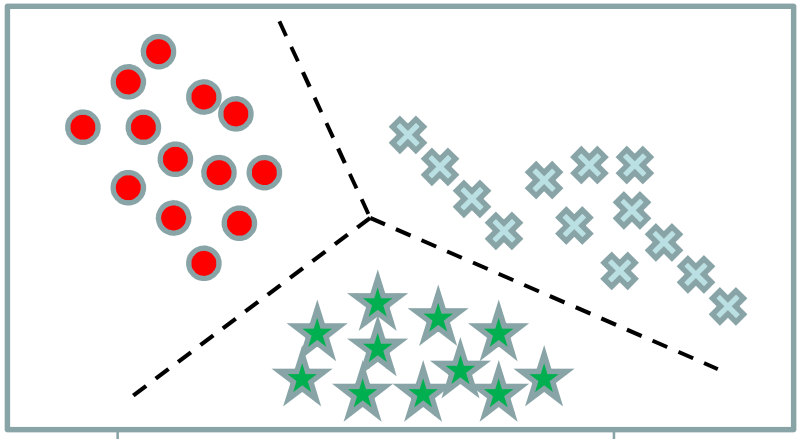
Algorithms

- The success of machine learning system also depends on the algorithms.
- The algorithms control the search to find and build the knowledge structures.
- The learning algorithms should extract useful information from training examples.

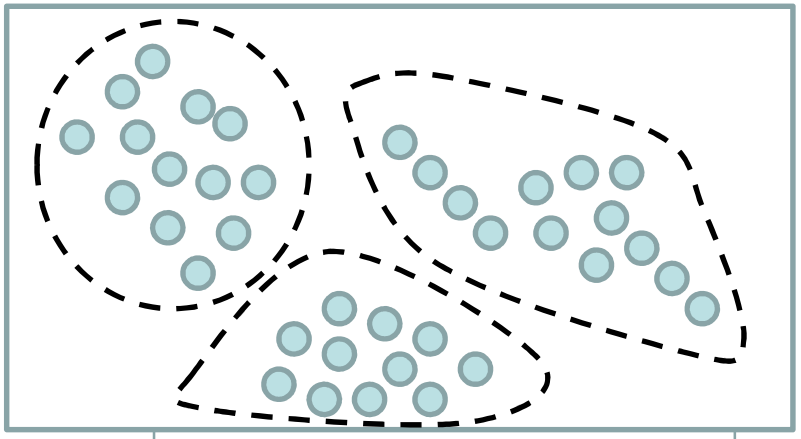
Algorithms

- **Supervised learning** ($\{x_n \in R^d, y_n \in R\}_{n=1}^N$)
  - Prediction
  - Classification (discrete labels), Regression (real values)
- **Unsupervised learning** ($\{x_n \in R^d\}_{n=1}^N$)
  - Clustering
  - Probability distribution estimation
  - Finding association (in features)
  - Dimension reduction
- **Semi-supervised learning**
- **Reinforcement learning**
  - Decision making (robot, chess machine)

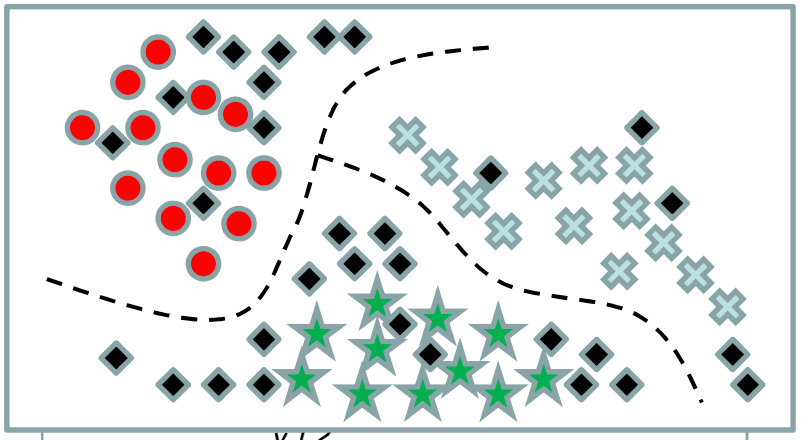
Algorithms



Supervised learning

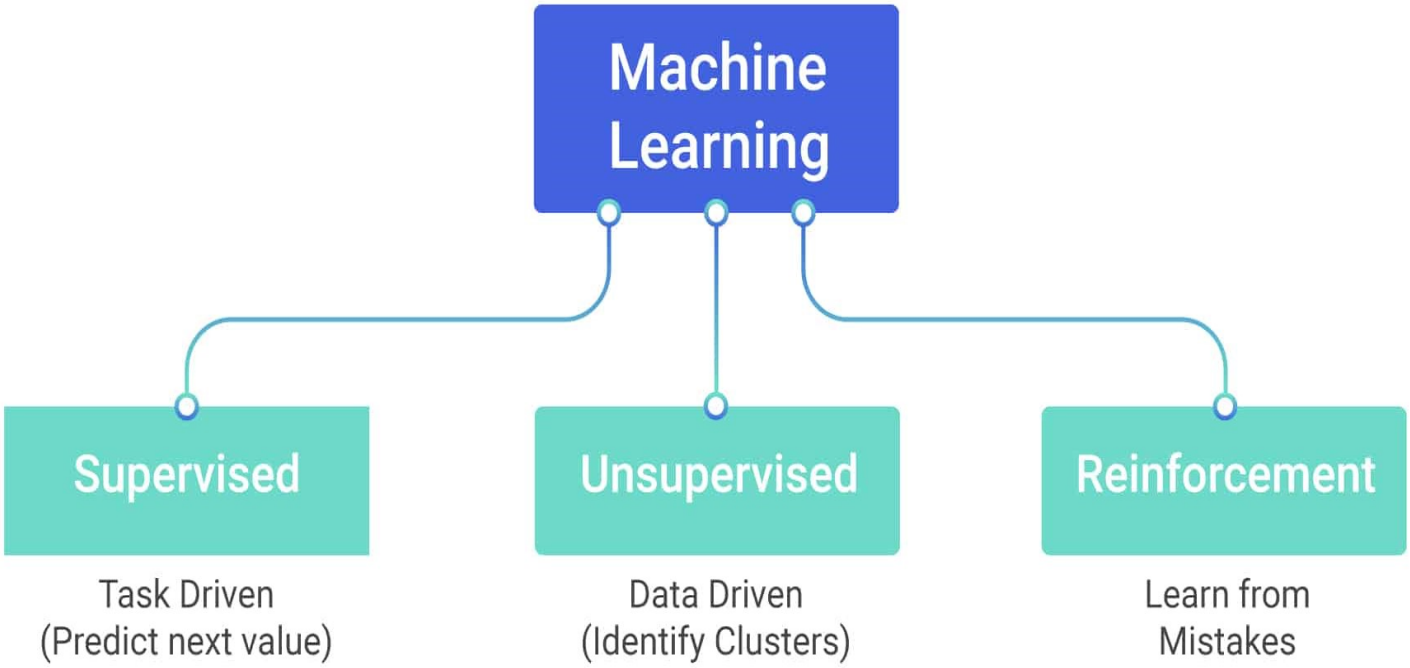


Unsupervised learning



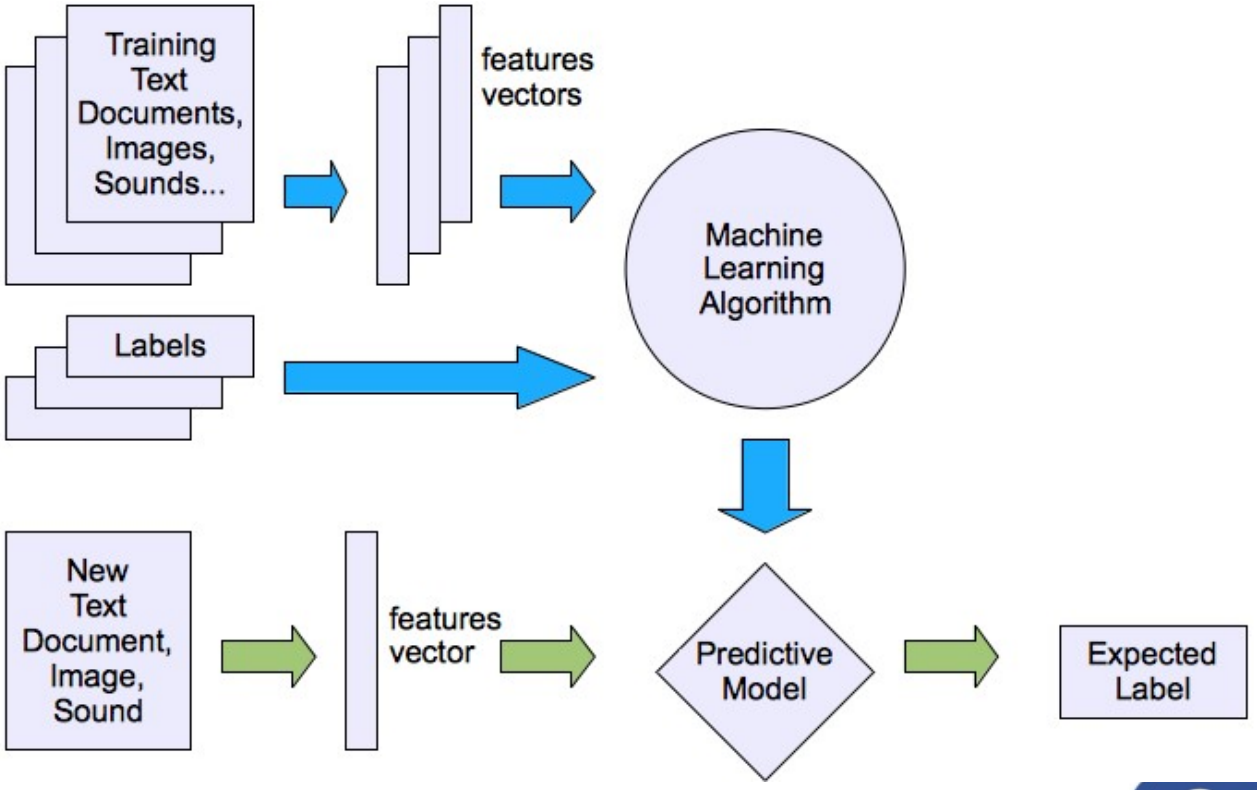
Semi-supervised learning

ML Types


Machine Learning
Supervised
Unsupervised
Reinforcement
Task Driven
(Predict next value)
Data Driven
(Identify Clusters)
Learn from
Mistakes

Machine learning structure

- Supervised learning

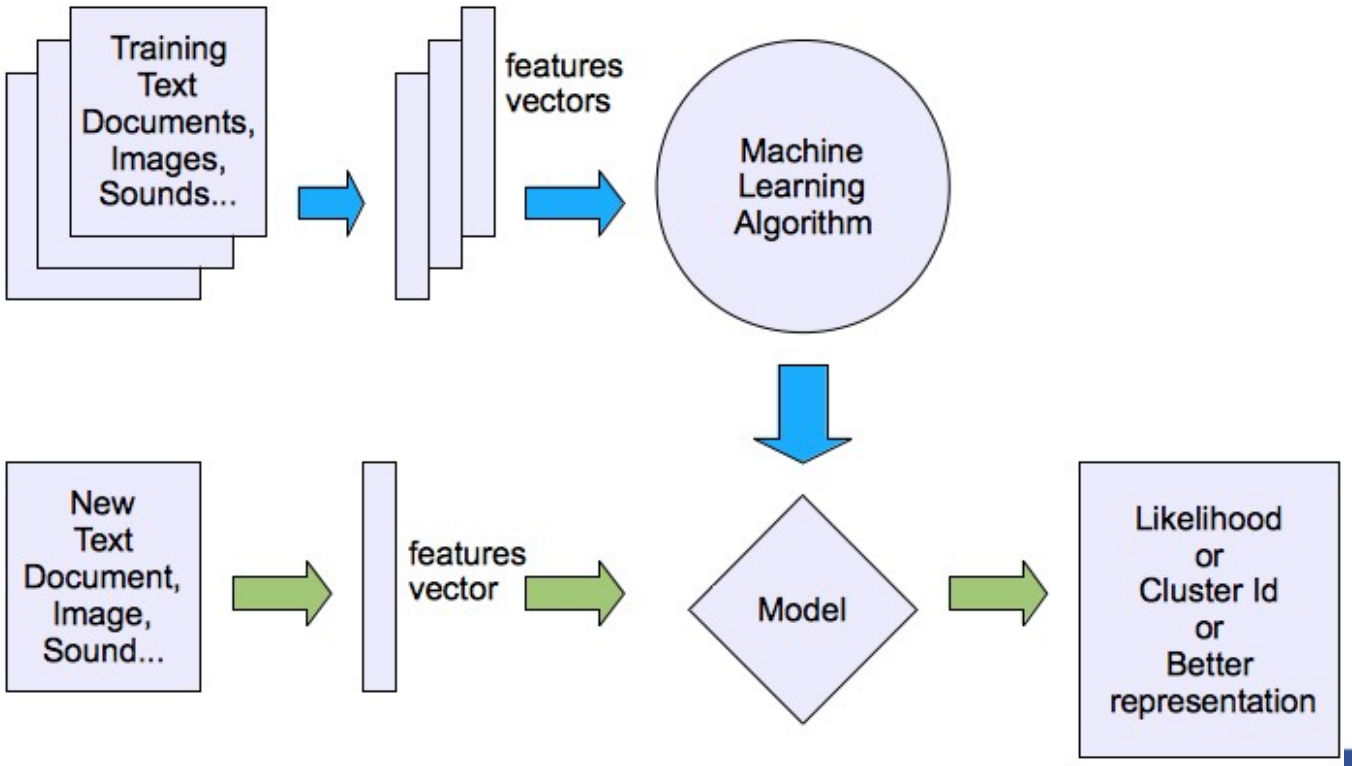


Supervised learning

- Training data includes desired outputs
- Supervised learning requires labelled training data.
- To evaluate the model, you also need labelled test data that is distinct from training data.
- Example: To train a spam filter, you need a training set of e-mails labelled spam and ham – classification.
- Learning from training examples labelled with true function values – regression.

Machine learning structure

- Unsupervised learning



Unsupervised learning

- Training data does not include desired outputs – class labels of data are unknown.
- Unsupervised learning works with unlabelled data and so there is no test data as such.
- Given a set of data, the task is to establish the existence of classes or clusters in the data.
- Example: The partition of data into clusters (instances similarity), learning associations (things that tend to occur together) and identifying hidden variables

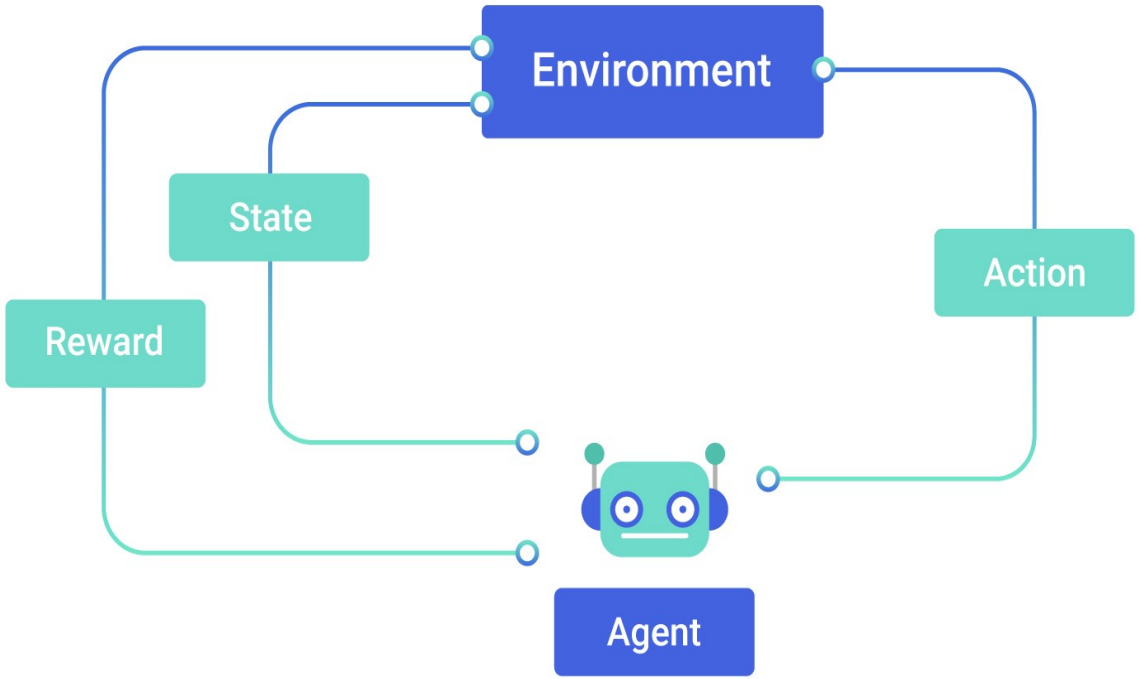
Semi-supervised learning

- Training data includes a few desired outputs.
- Data is cheap, but labelled data is expensive.
- Use small labelled training set to build an initial model, which is then refined using the unlabelled data

Reinforcement learning

- In reinforcement learning framework, provide only a reward function which indicates to the learning algorithm that when it is doing correct and when it is doing poorly.
- Then it is the job of learning algorithm to figure out how to choose actions over time so as to obtain large rewards.
- Example: autonomous helicopter flight, robot legged locomotion, chess play, etc.,

Reinforcement learning



What are we seeking?

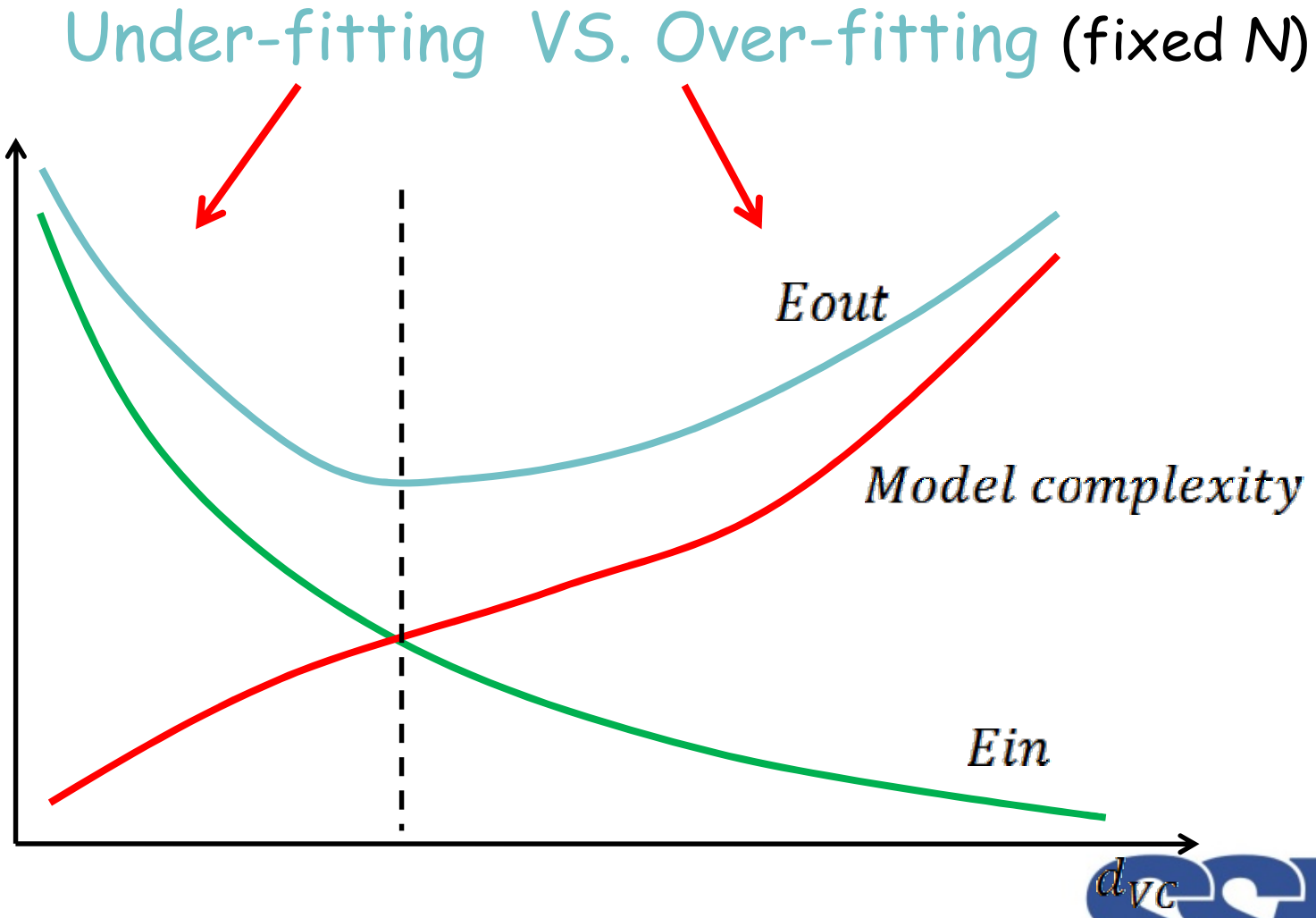
- Supervised: Low E-out or maximize probabilistic terms

$$error = \frac{1}{N}\sum_{n=1}^{N}[y_n \neq g(x_n)]$$

E-in: for training set
E-out: for testing set
E-out(g)<=E-in(g)

- Unsupervised: Minimum quantization error, Minimum distance, MAP, MLE(maximum likelihood estimation)

What are we seeking?



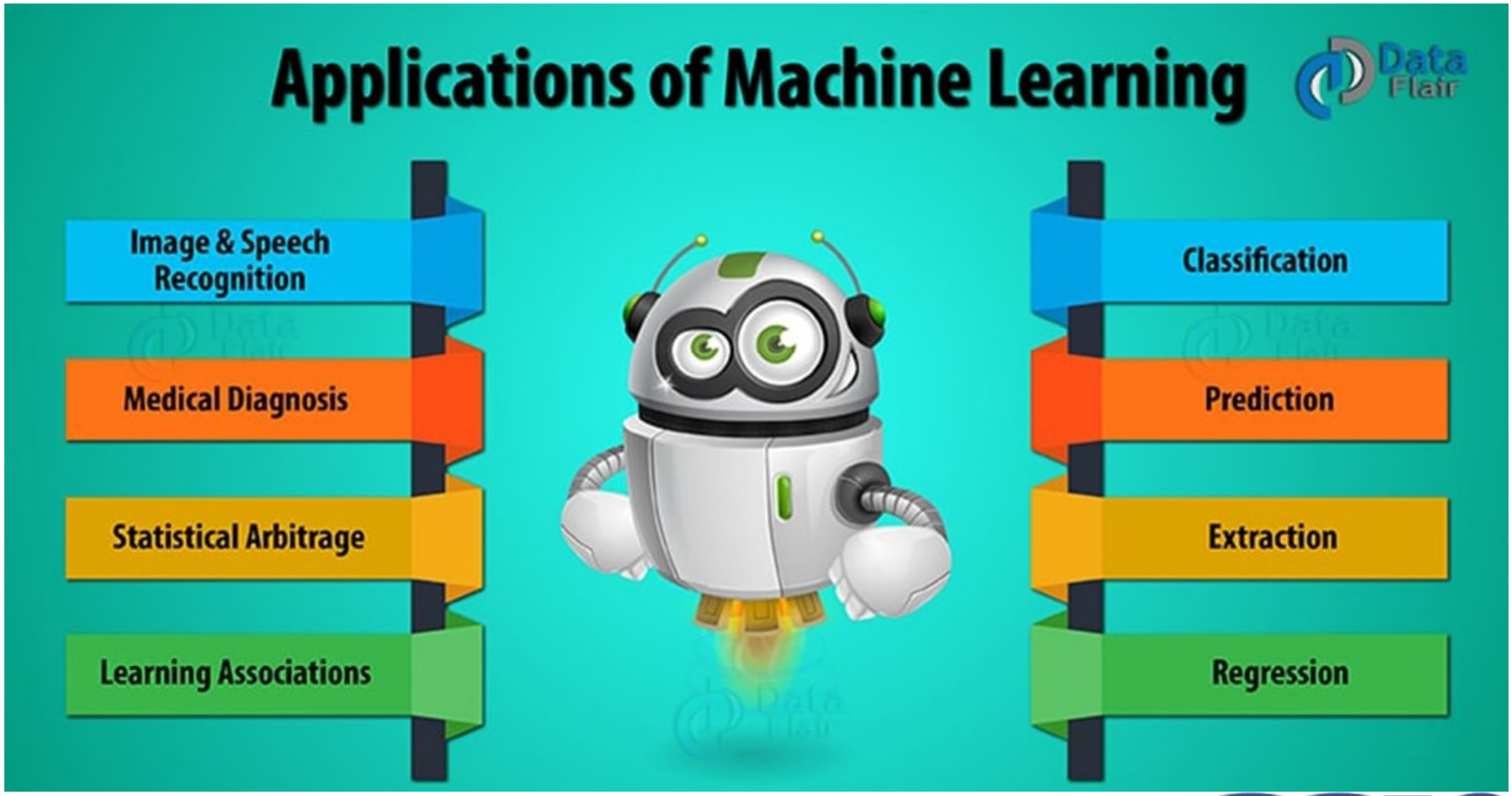
Learning techniques

- Supervised learning categories and techniques
  - **Linear classifier** (numerical functions)
  - **Parametric** (Probabilistic functions)
    - Naïve Bayes, Gaussian discriminant analysis (GDA), Hidden Markov models (HMM), Probabilistic graphical models
  - **Non-parametric** (Instance-based functions)
    - *K*-nearest neighbors, Kernel regression, Kernel density estimation, Local regression
  - **Non-metric** (Symbolic functions)
    - Classification and regression tree (CART), decision tree
  - **Aggregation**
    - Bagging (bootstrap + aggregation), Adaboost, Random forest

Learning techniques

- Unsupervised learning categories and techniques
  - **Clustering**
    - K-means clustering
    - Spectral clustering
  - **Density Estimation**
    - Gaussian mixture model (GMM)
    - Graphical models
  - **Dimensionality reduction**
    - Principal component analysis (PCA)
    - Factor analysis

Machine Learning Applications



Machine Learning Applications

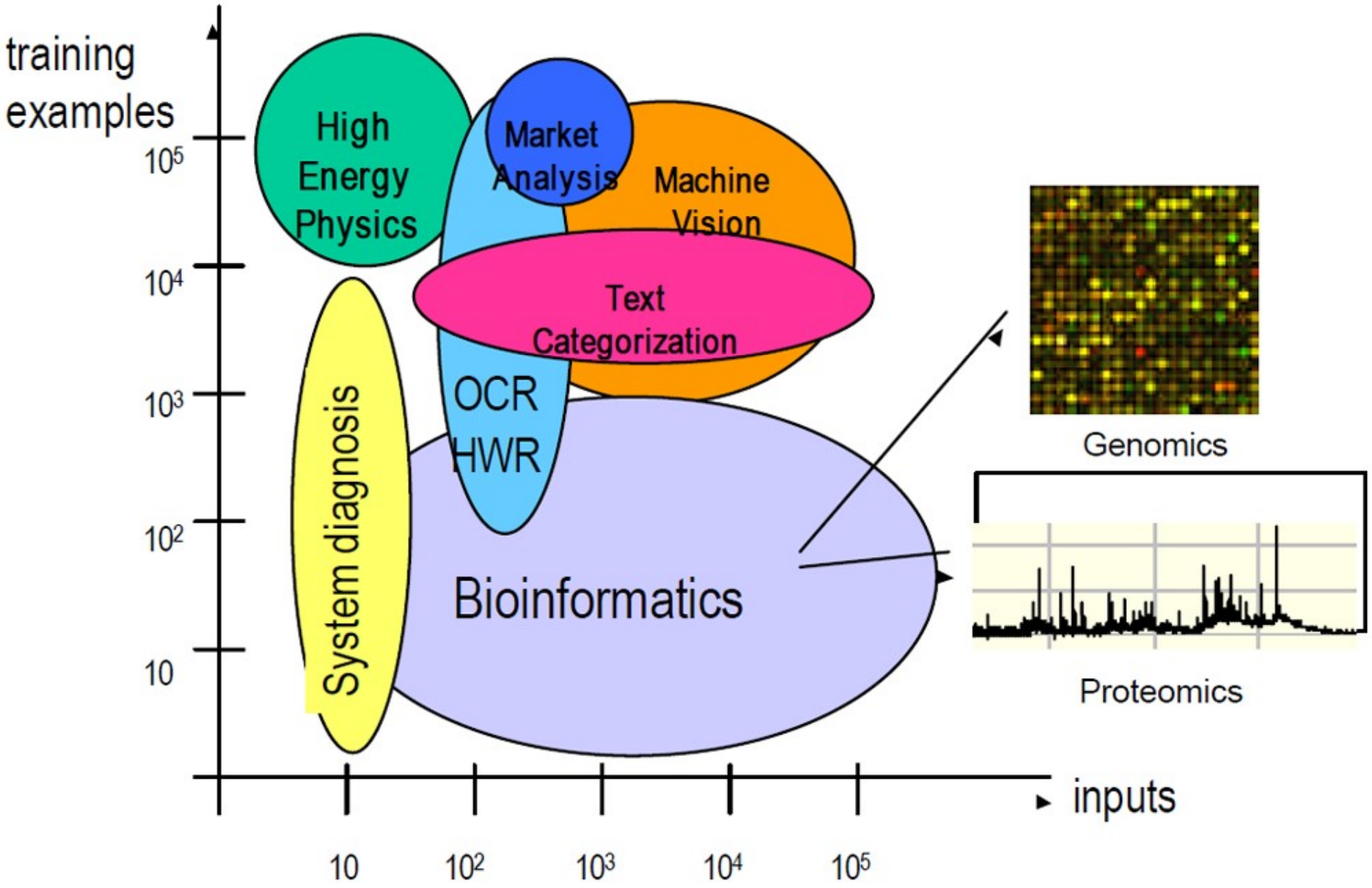


Image Categorization - Training

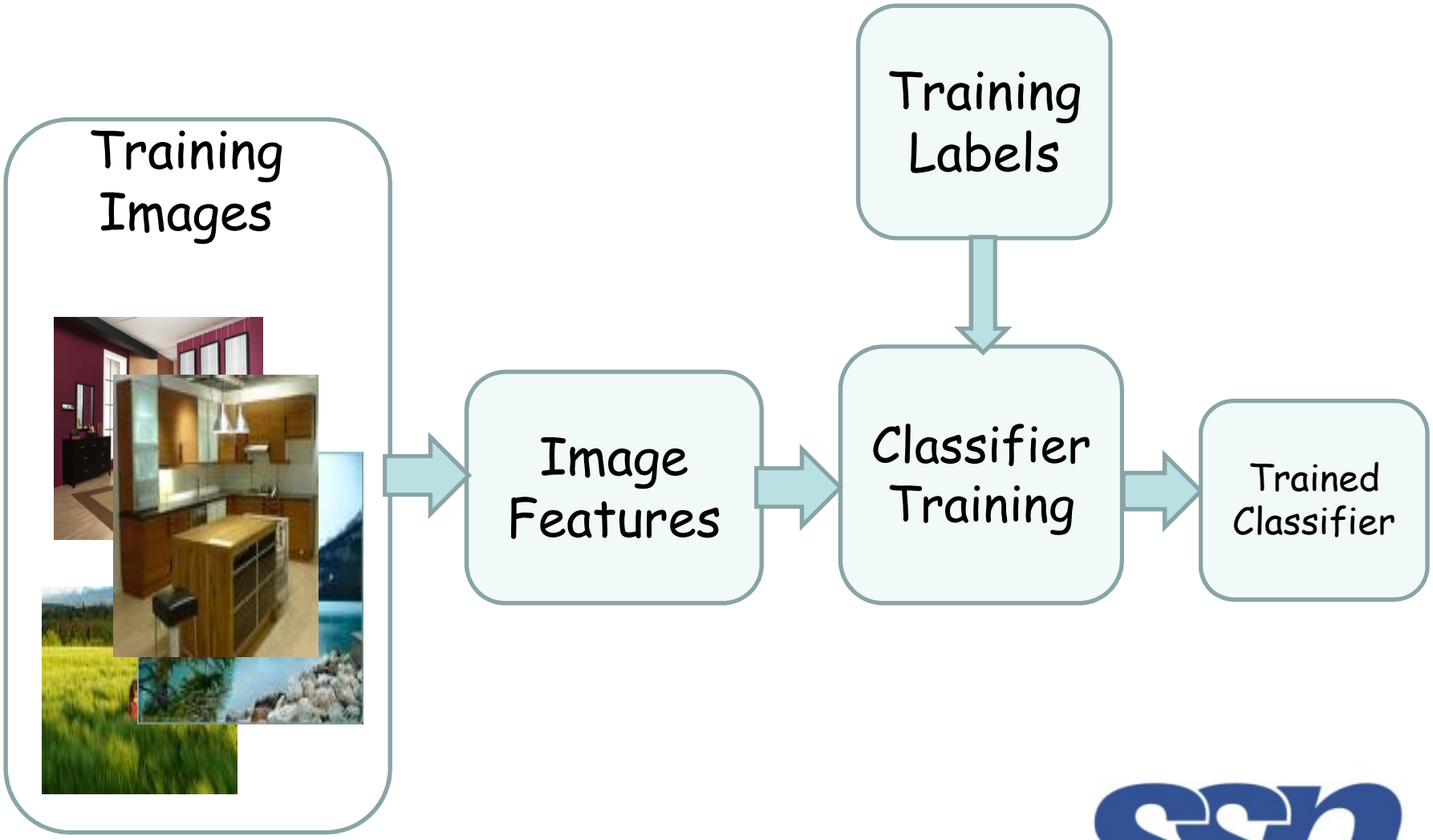


Image Categorization - Testing

Test Image

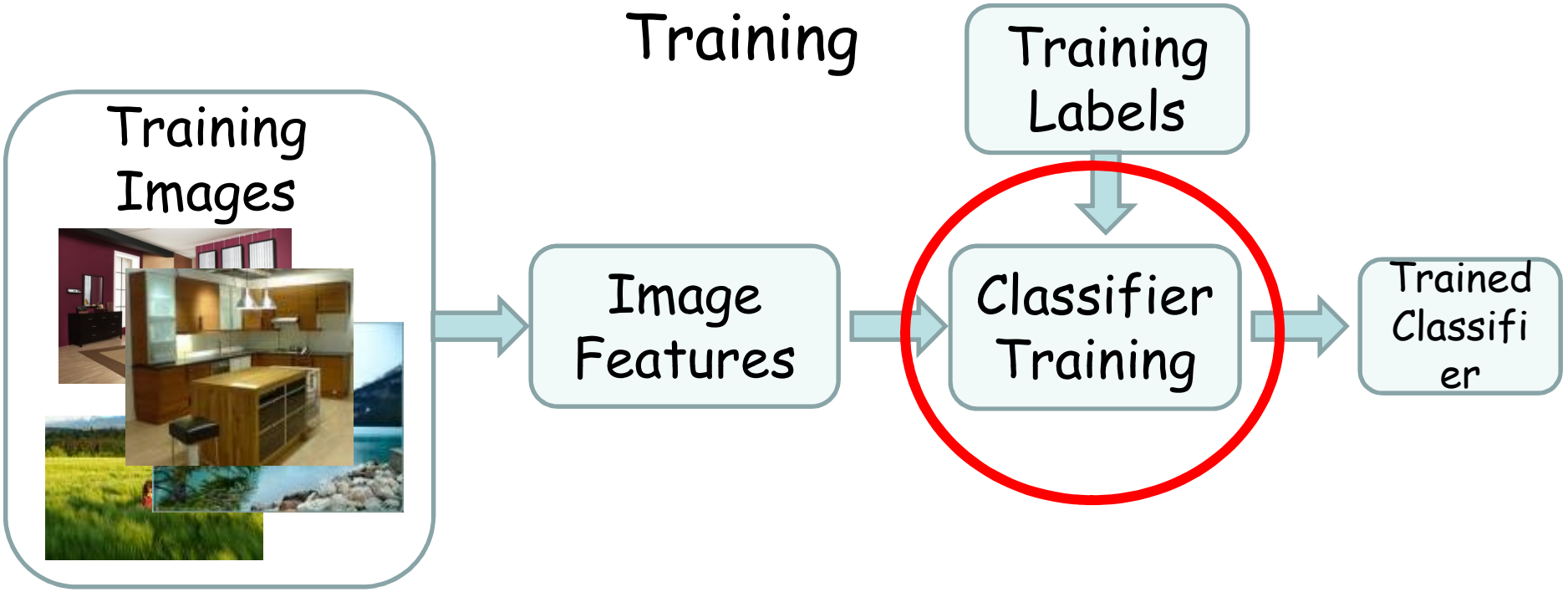
Image Features

Trained Classifier

Prediction

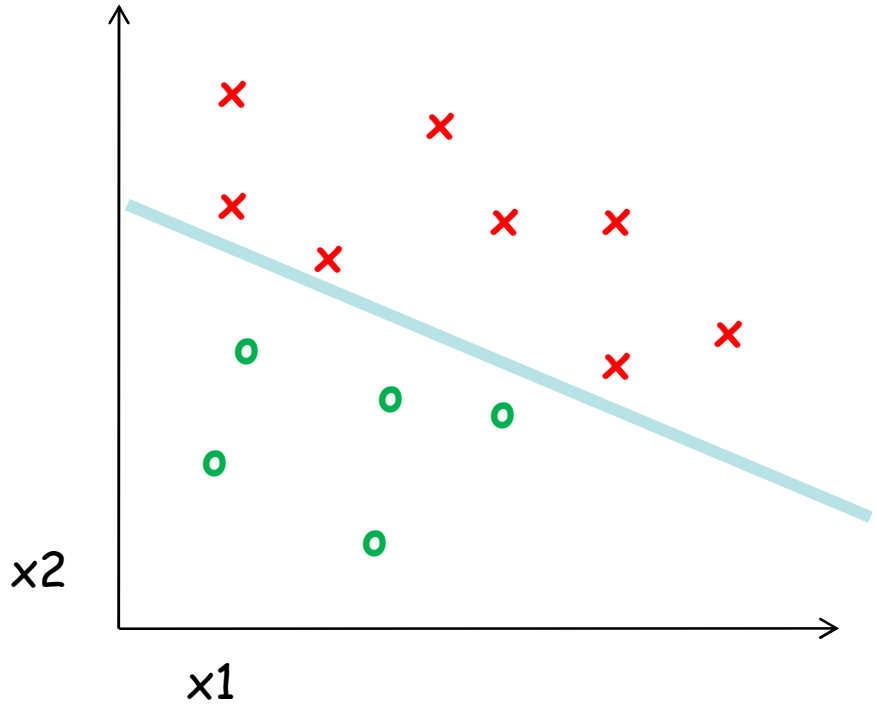
Outdoor

Classifiers



Learning a classifier

Given some set of features with corresponding labels, learn a function to predict the labels from the features



One way to think about it…

- Training labels dictate that two examples are the same or different, in some sense
- Features and distance measures define visual similarity
- Classifiers try to learn weights or parameters for features and distance measures so that visual similarity predicts label similarity

Machine Learning Problems

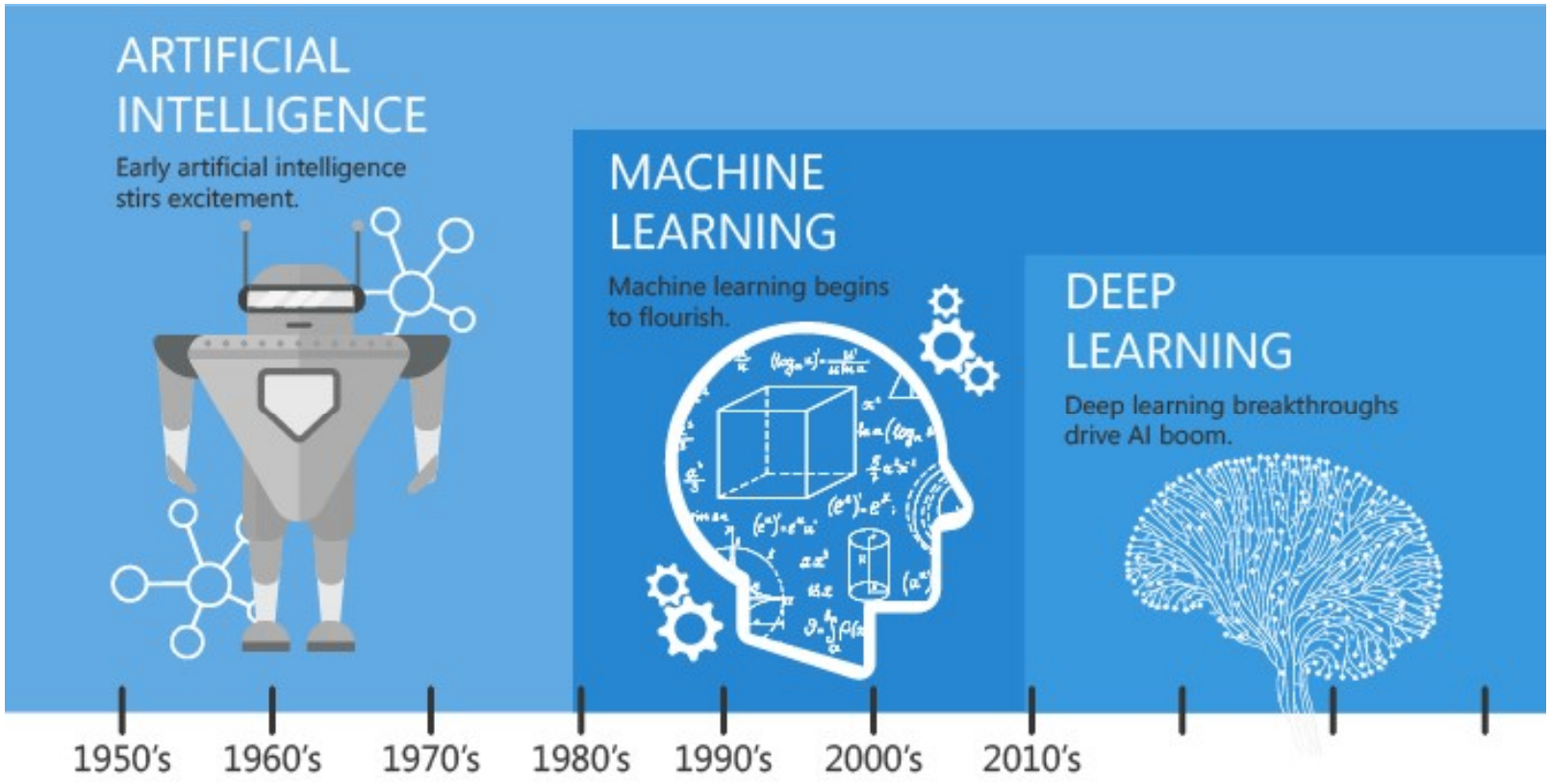
	Supervised Learning	*Unsupervised Learning*
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

Many classifiers to choose from

- SVM
- Neural networks
- Naïve Bayes
- Bayesian network
- Linear / Logistic regression
- Decision tree
- K-nearest neighbor
- Ensemble approach
- Boosting
- Bagging
- Random Forest

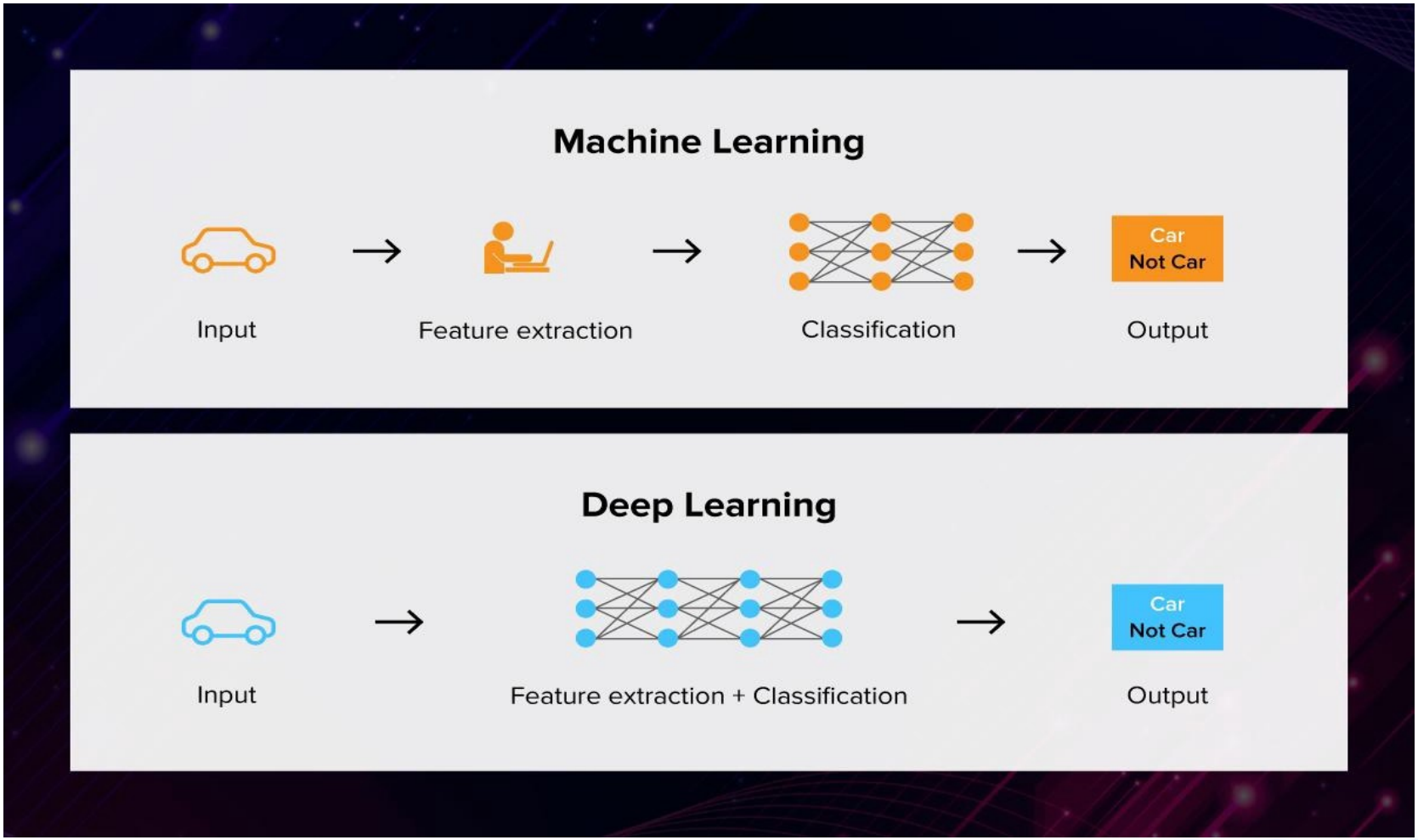
Which is the best one?

AI-ML-DL

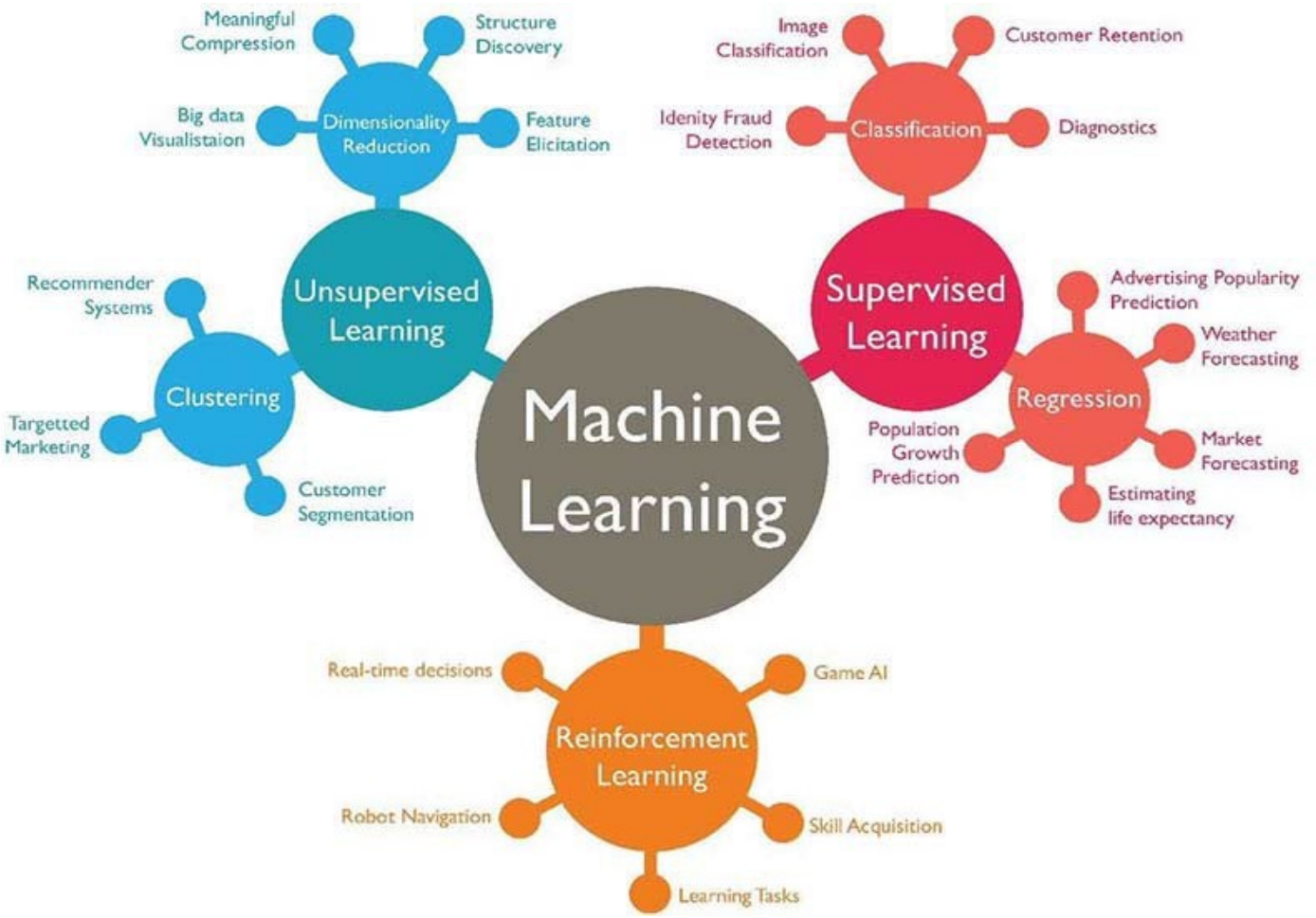


Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

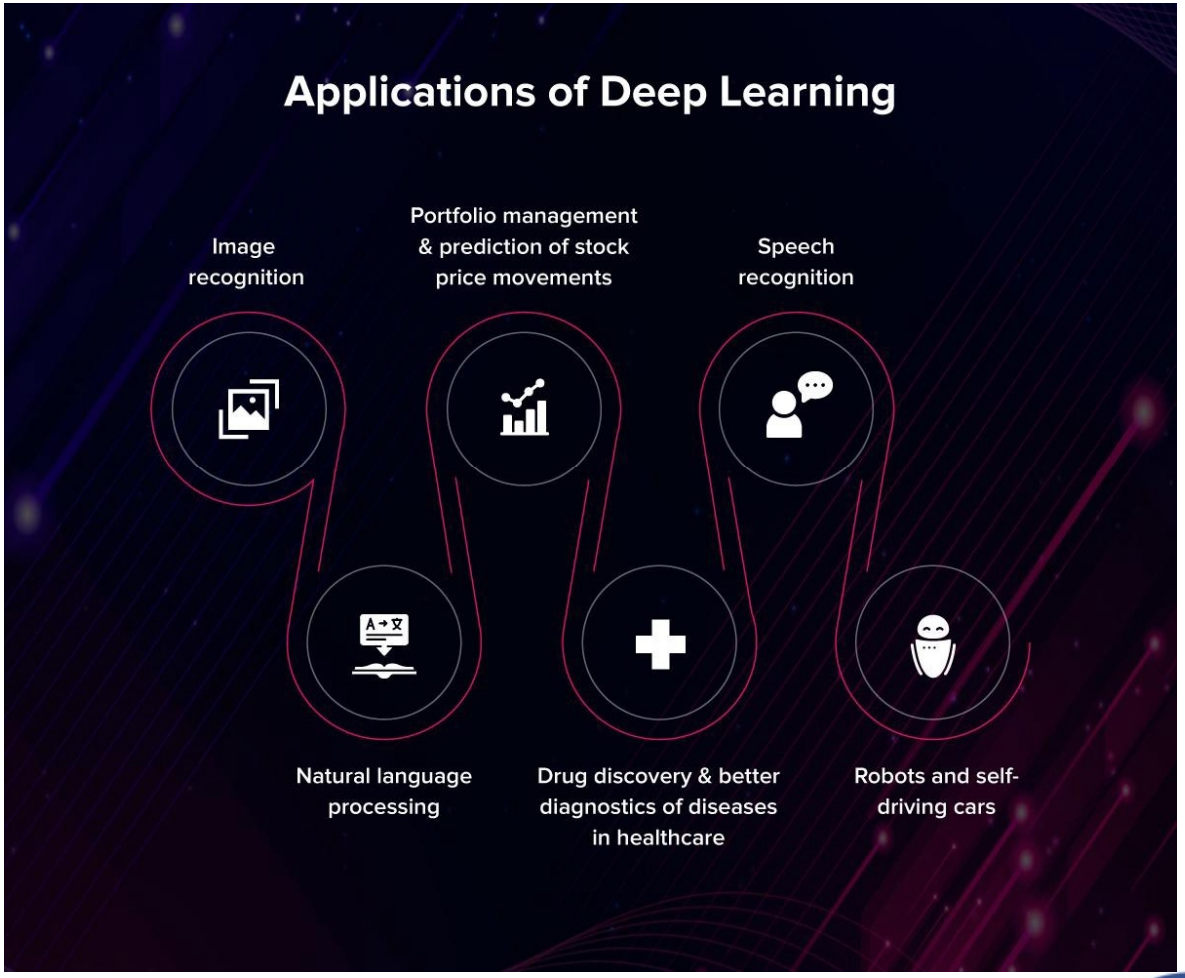
AI-ML-DL-Classification



ML Applications



DL Applications



Check your understanding

- Dataset without class labels are suitable for ……. learning.
- Learning through reward function is called …….
- Machine learning is a sub domain of…….
- Learning without feature extraction is called…….
- Examples for probabilistic based learning algorithms are ……and.....

Summary

- AI-ML-DL
- ML types
- Applications of ML
- Algorithm names of ML
- ML problem types