

An Overview of Machine Learning

A horizontal bar spanning the width of the slide, divided into five colored segments: dark blue, light blue, green, orange, and dark blue.A faint, stylized silhouette of a mountain range in the background at the bottom of the slide.

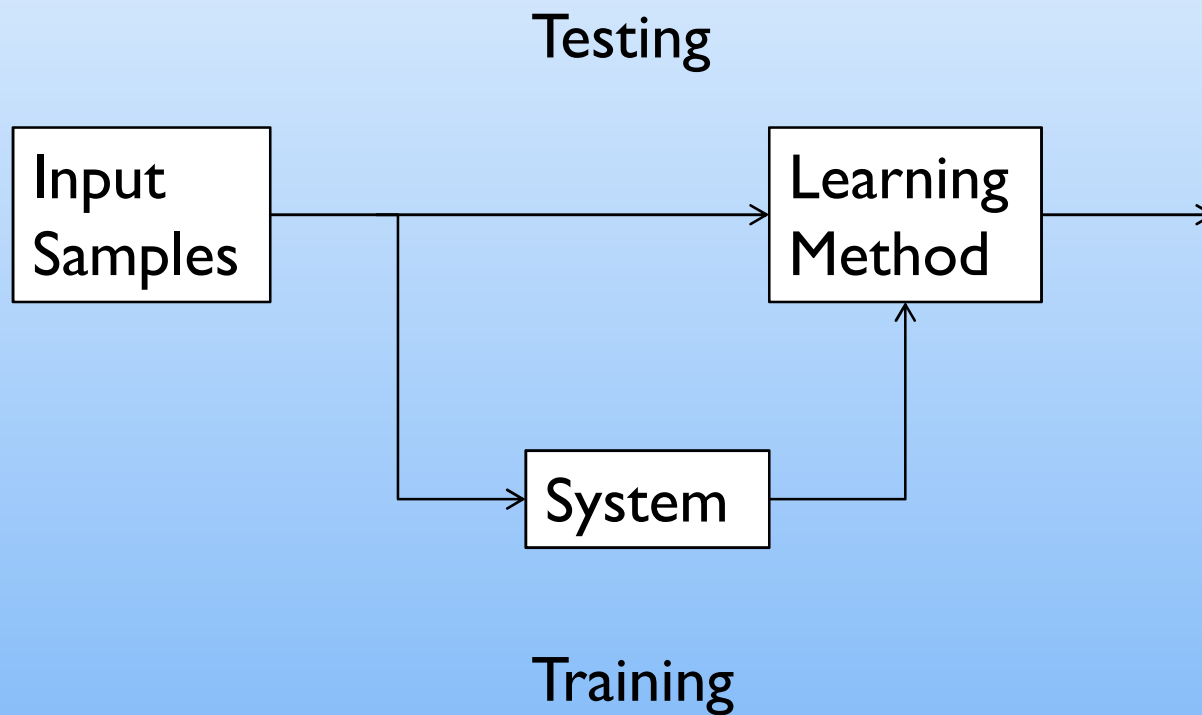
Outline & Content

- ▣ What is machine learning?
- ▣ Learning system model
- ▣ Training and testing
- ▣ Performance
- ▣ Algorithms
- ▣ Machine learning structure
- ▣ What are we seeking?
- ▣ Learning techniques
- ▣ Applications
- ▣ Conclusion

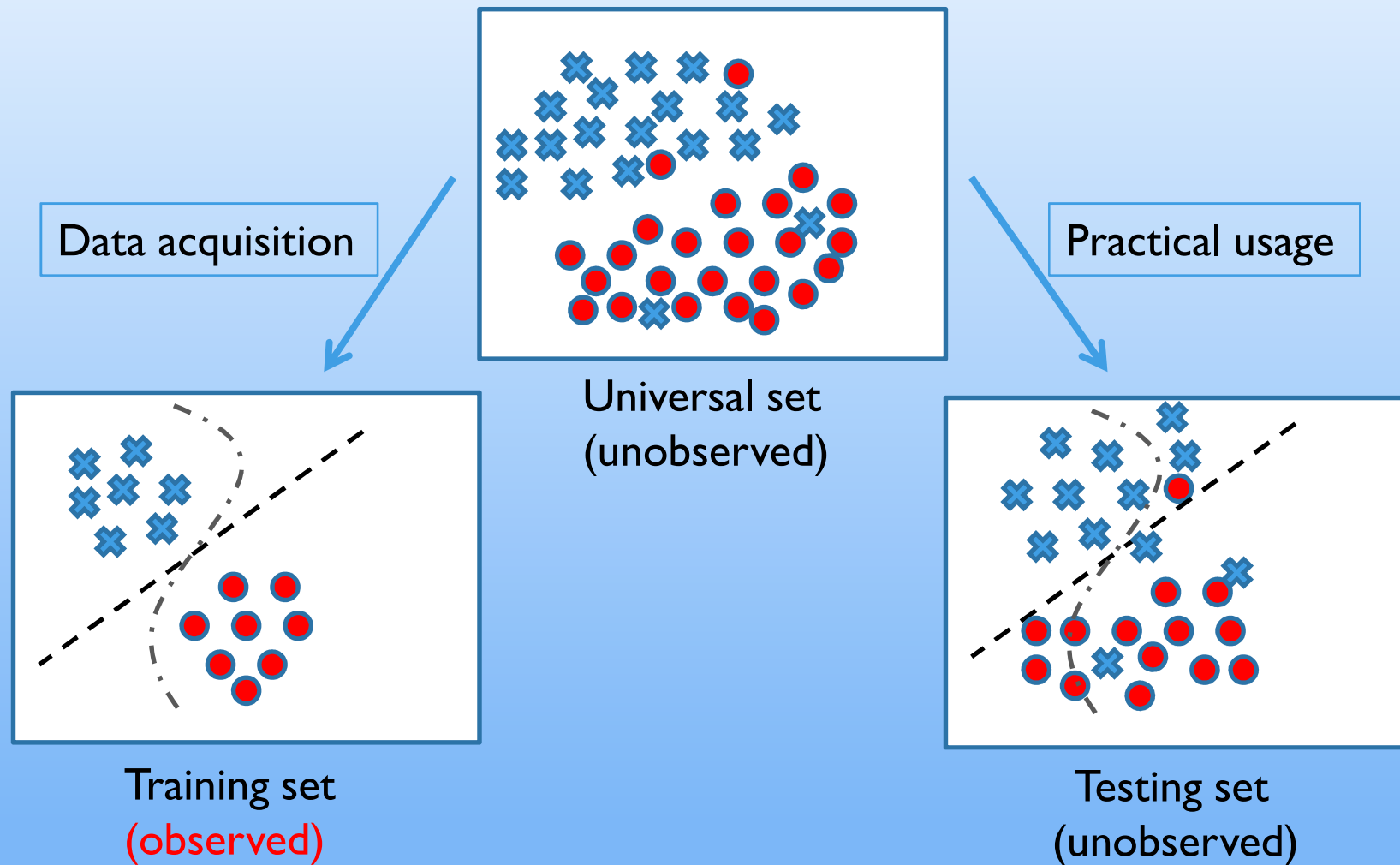
What is machine learning?

- 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.

Learning system model

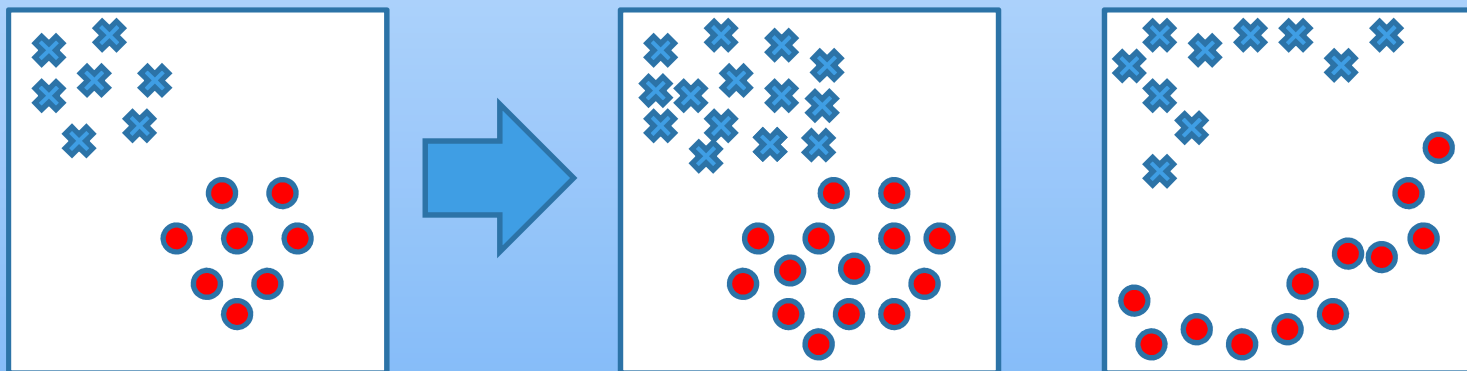


Training and testing



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

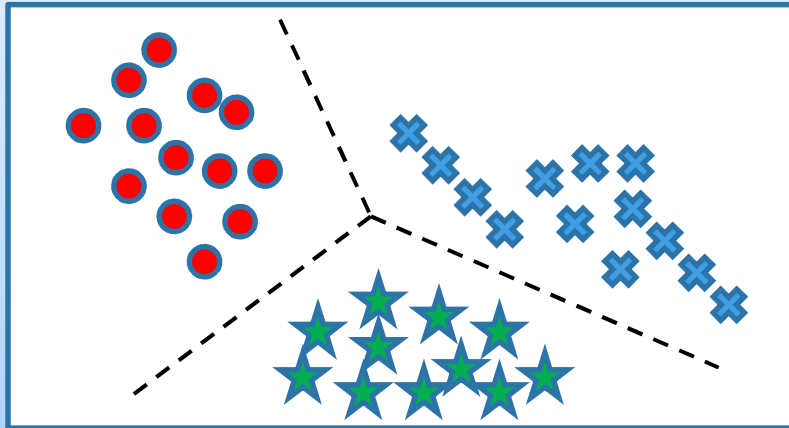
Algorithms

- 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.

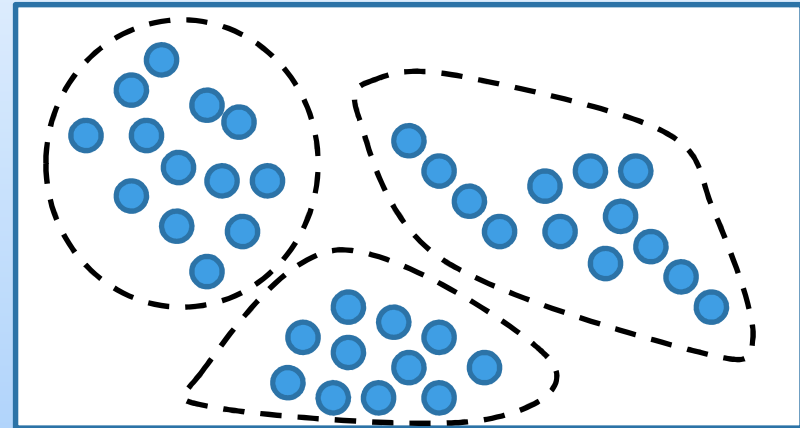
Algorithms

- **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)

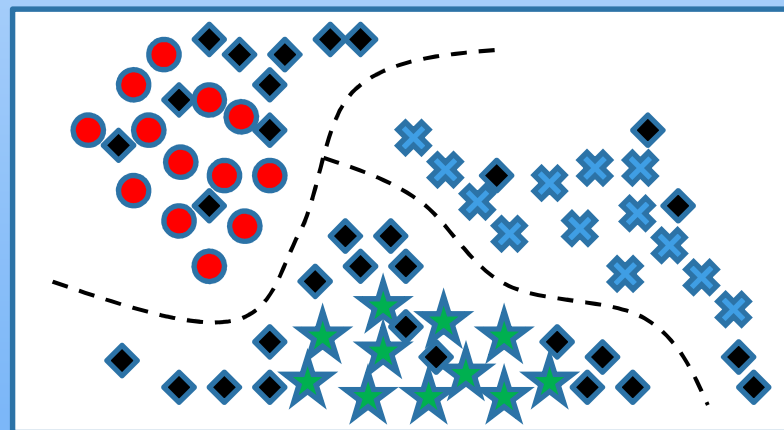
Algorithms



Supervised learning



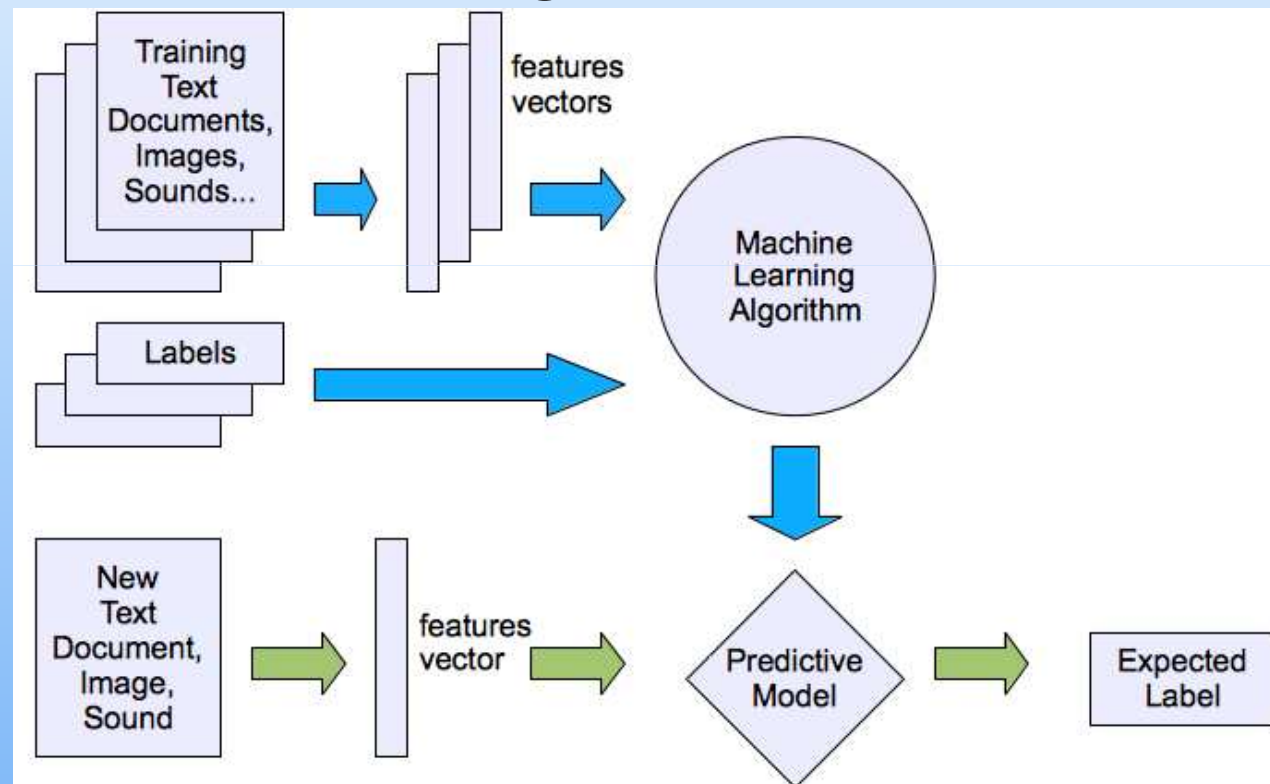
Unsupervised learning



Semi-supervised learning

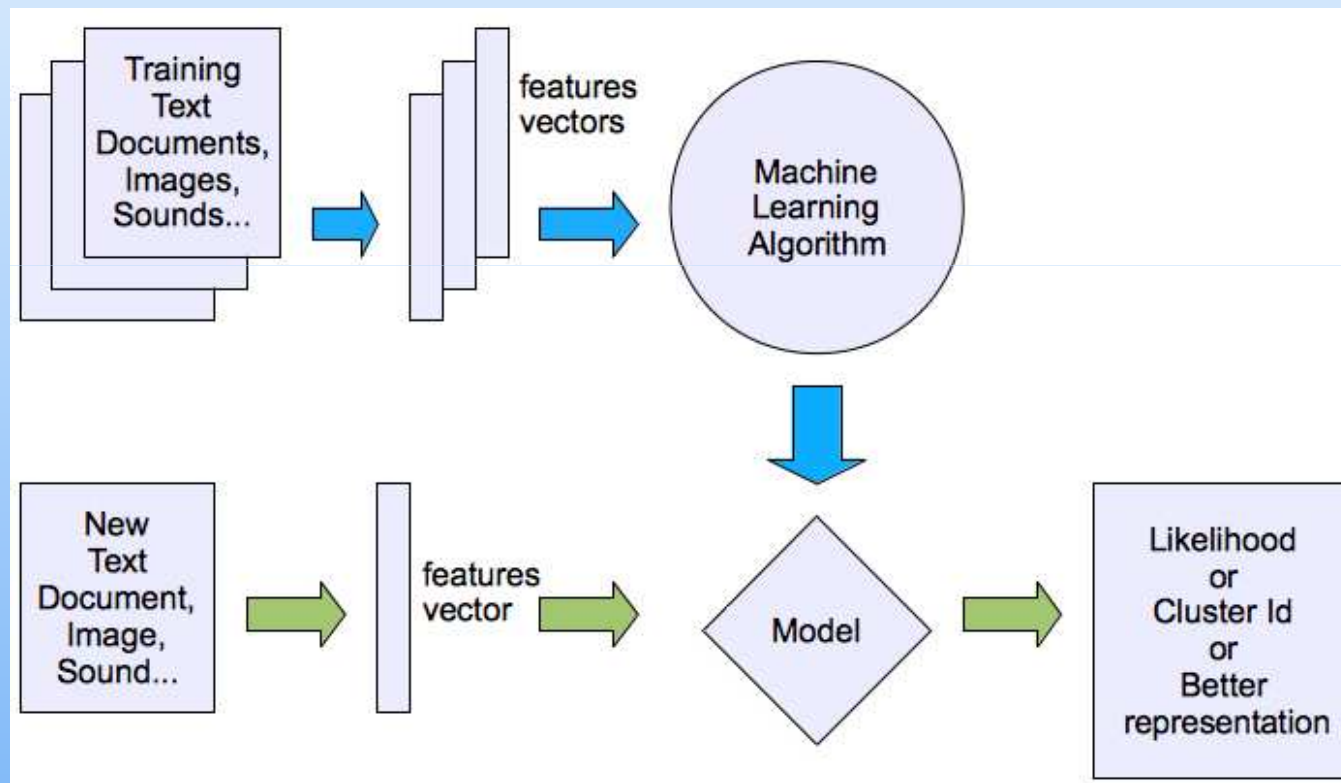
Machine learning structure

▣ Supervised learning



Machine learning structure

■ Unsupervised learning



What are we seeking?

- Supervised: Low E-out or maximize probabilistic terms

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

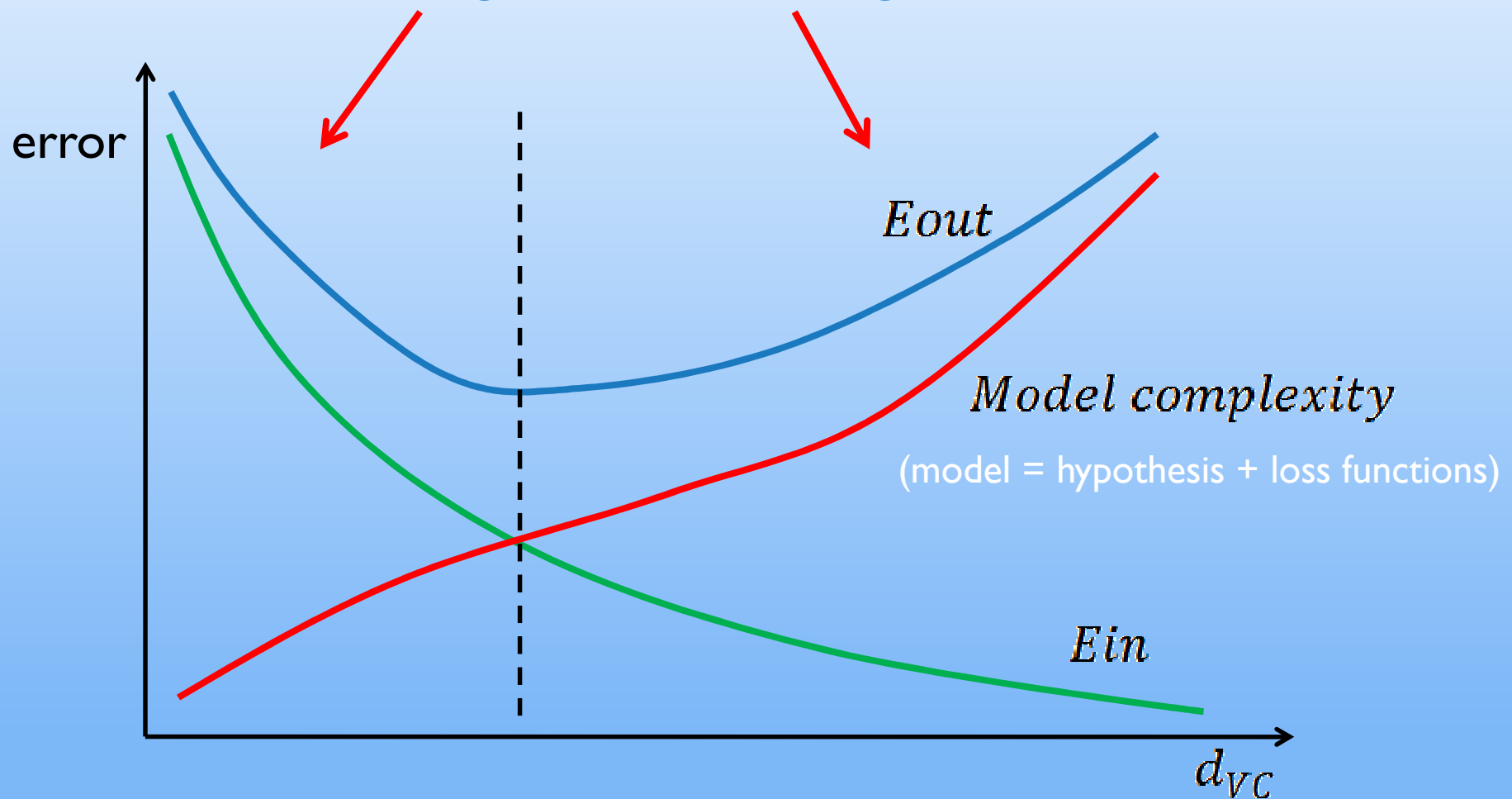
E-in: for training set
E-out: for testing set

$$E_{out}(g) \leq E_{in}(g) \pm O\left(\sqrt{\frac{d_{VC}}{N} \ln N}\right)$$

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

What are we seeking?

Under-fitting VS. Over-fitting (fixed N)

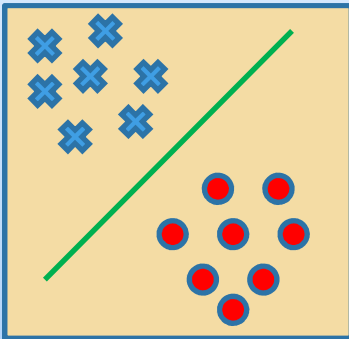


Learning techniques

- 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

- Linear classifier



$$g(x_n) = \text{sign}(w^T x_n)$$

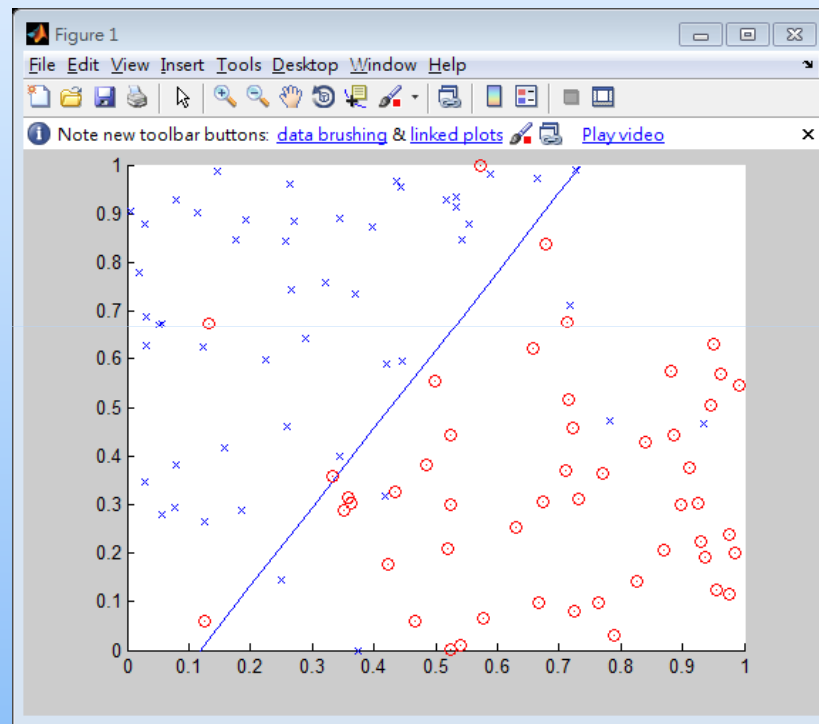
, where w is an d -dim vector (learned)

▣ Techniques:

- ▣ Perceptron
- ▣ Logistic regression
- ▣ Support vector machine (SVM)
- ▣ Ada-line
- ▣ Multi-layer perceptron (MLP)

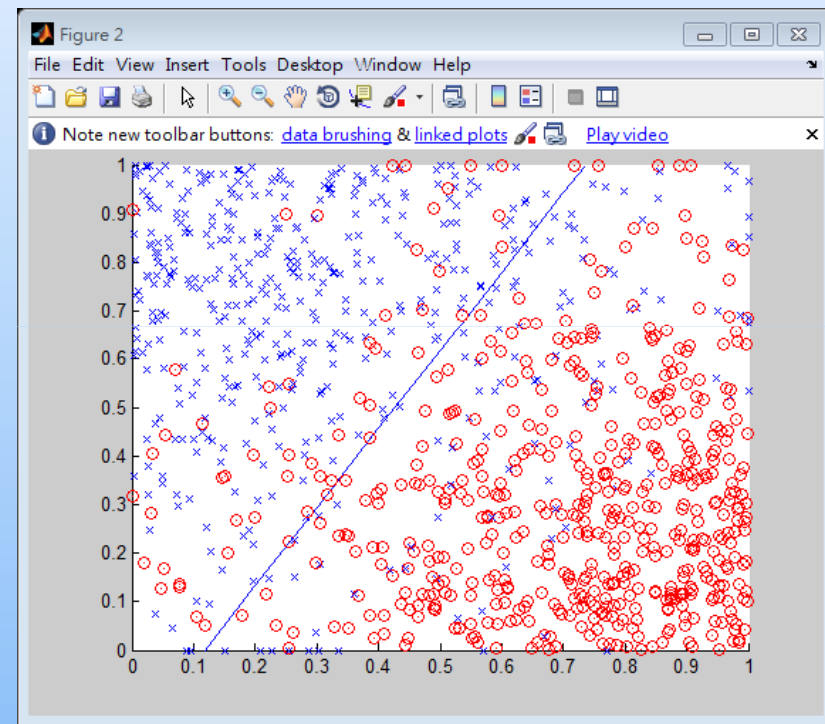
Learning techniques

Using perceptron learning algorithm(PLA)



Training

Error rate: 0.10

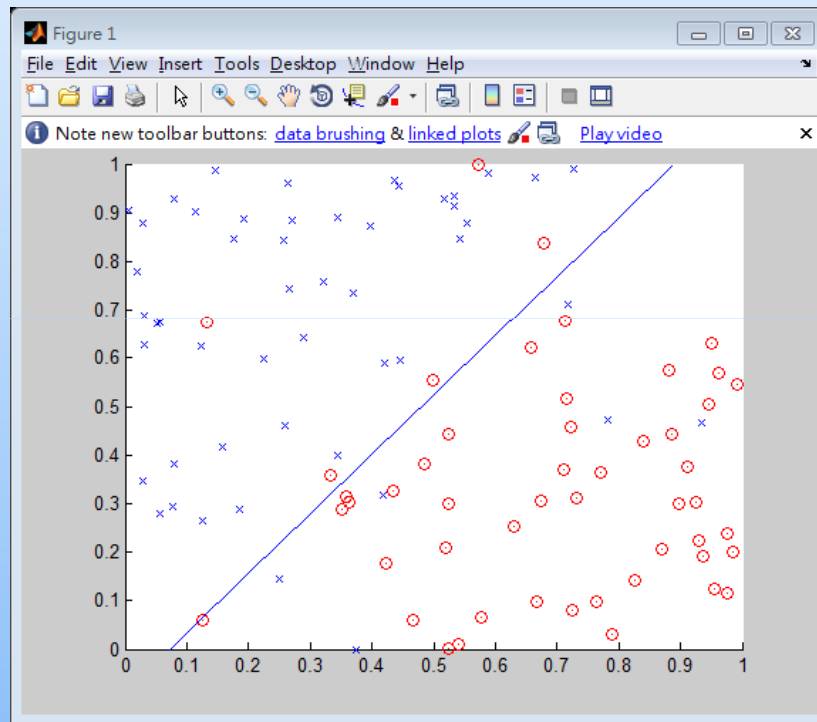


Testing

Error rate: 0.156

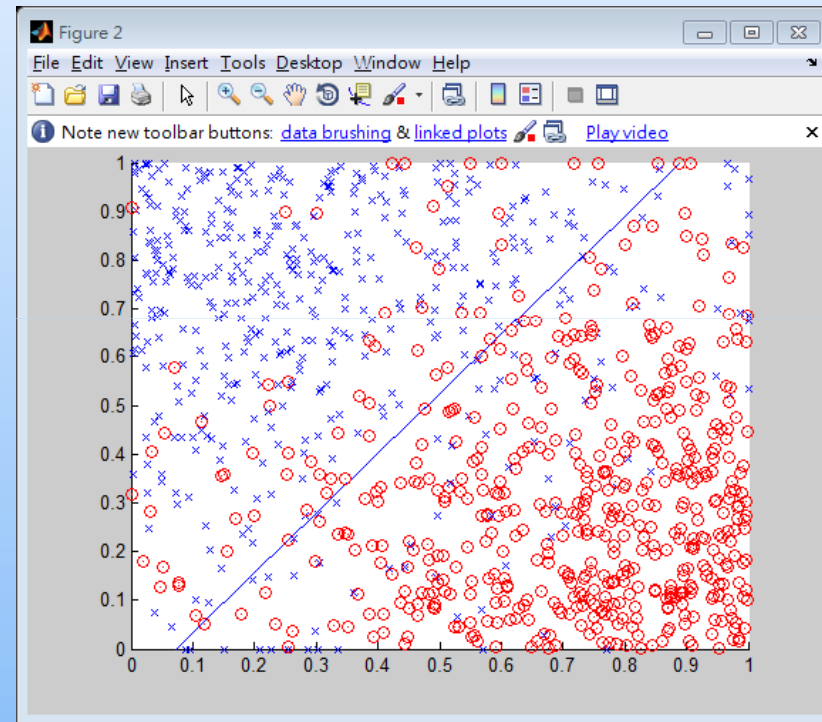
Learning techniques

Using logistic regression



Training

Error rate: 0.11

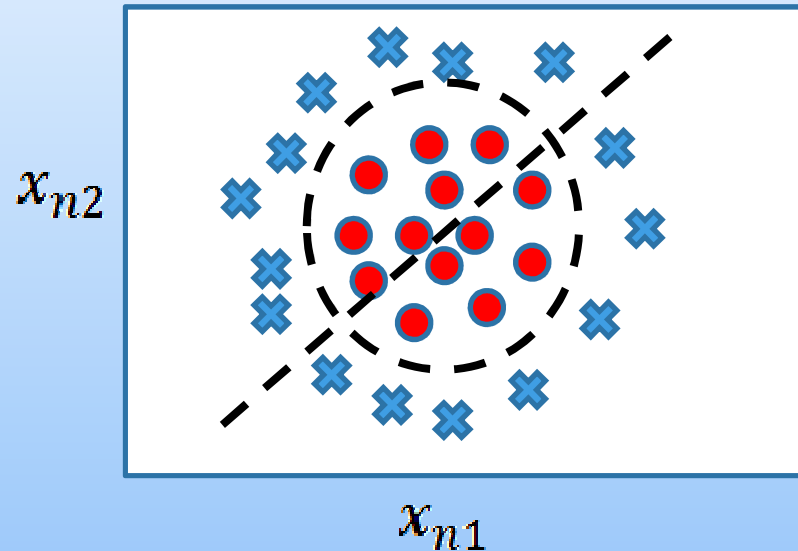


Testing

Error rate: 0.145

Learning techniques

- Non-linear case



$$x_n = [x_{n1}, x_{n2}]$$



$$x_n = [x_{n1}, x_{n2}, x_{n1} * x_{n2}, x_{n1}^2, x_{n2}^2]$$

$$g(x_n) = \text{sign}(w^T x_n)$$

- ▣ Support vector machine (SVM):
 - ▣ Linear to nonlinear: **Feature transform** and **kernel function**

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

Applications

- ▣ Face detection
- ▣ Object detection and recognition
- ▣ Image segmentation
- ▣ Multimedia event detection
- ▣ Economical and commercial usage

Conclusion

We have a simple overview of some techniques and algorithms in machine learning. Furthermore, there are more and more techniques apply machine learning as a solution. In the future, machine learning will play an important role in our daily life.

Reference

[1] W. L. Chao, J. J. Ding, “Integrated Machine Learning Algorithms for Human Age Estimation”, NTU, 2011.