

Outline

1 Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

2 Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

3 Learning Problems

Supervised

Non-supervised

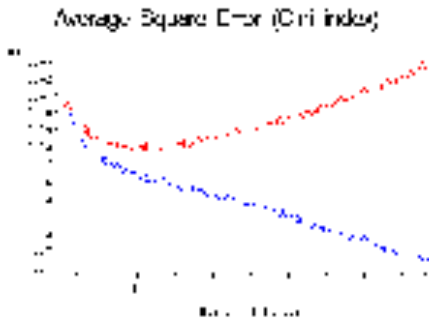
Active

On-line

4 Learning Techniques

Training error vs generalization error

- The loss function measures the error in the training set
- Is this a good measure of the quality of the solution?



•

Over-fitting and under-fitting

Introduction

Patterns and Generalization

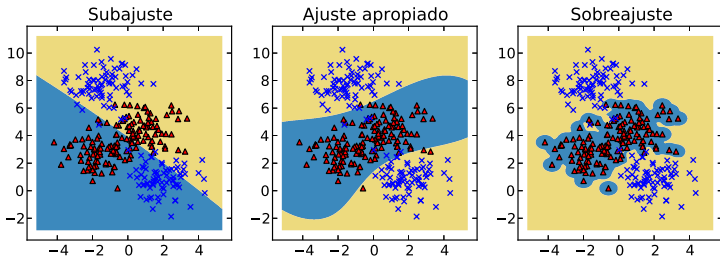
Generalizing from patterns

Overfitting/
Overlearning

How to Measure the Quality of a Solution?

Learning Problems

Learning Techniques



Generalization error

- Generalization error:

$$E[(L(f_w, S))]$$

- How to control the generalization error during training?
 - Cross validation
 - Regularization

- Vapnik, 1995:

$$R(\alpha) = \int \frac{1}{2} |y - f(\mathbf{x}, \alpha)| dP(\mathbf{x}, y)$$

$$R_{emp}(\alpha) = \frac{1}{2l} \sum_{i=1}^l |y_i - f(\mathbf{x}_i, \alpha)|$$

$$R(\alpha) \leq R_{emp}(\alpha) + \sqrt{\left(\frac{h(\log(2l/h) + 1) - \log(\eta/4)}{l} \right)}$$

Outline

1 Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

2 Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

3 Learning Problems

Supervised

Non-supervised

Active

On-line

4 Learning Techniques

Types

- Supervised learning
- Non-supervised learning
- Semi-supervised learning
- Active/reinforcement learning
- On-line learning

Outline

1 Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

2 Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

3 Learning Problems

Supervised

Non-supervised

Active

On-line

4 Learning Techniques

Supervised learning

- **Fundamental problem:** to find a function that relates a set of inputs with a set of outputs
- Typical problems:
 - Classification
 - Regression



Outline

① Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

② Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

③ Learning Problems

Supervised

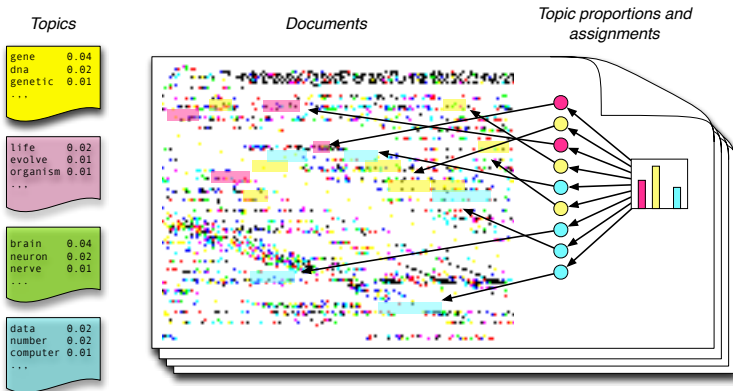
Non-supervised

Active

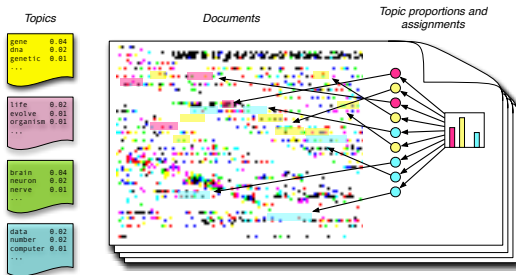
On-line

④ Learning Techniques

Non-supervised learning



Non-supervised learning



- There are not labels for the training samples
- **Fundamental problem:** to find the subjacent structure of a training data set
- Typical problems: clustering, probability density estimation, dimensionality reduction, latent topic analysis, data compression
- Some samples may have labels, in that case it is called semi-supervised learning

Outline

1 Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

2 Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

3 Learning Problems

Supervised

Non-supervised

Active

On-line

4 Learning Techniques

Active/reinforcement learning

- Generally, it happens in the context of an agent acting in an environment
- The agent is not told whether it has make the right decision or not
- The agent is punished or rewarded (not necessarily in an immediate way)
- **Fundamental problem:** to define a policy that allows to maximize the positive stimulus (reward)



<https://www.youtube.com/watch?v=iqXKQf2BOSE>

Outline

1 Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

2 Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

3 Learning Problems

Supervised

Non-supervised

Active

On-line

4 Learning Techniques

On-line learning

- Only one pass through the data
 - big data volume
 - real time
- It may be supervised or unsupervised
- **Fundamental problem:** to extract the maximum information from data with minimum number of passes

Outline

1 Introduction

Example

How to State the Learning Problem?

How to Solve the Learning Problem?

2 Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

How to Measure the Quality of a Solution?

3 Learning Problems

Supervised

Non-supervised

Active

On-line

4 Learning Techniques

Representative techniques

- Computational
 - Decision trees
 - Nearest-neighbor classification
 - Graph-based clustering
 - Association rules
- Statistical
 - Multivariate regression
 - Linear discriminant analysis
 - Bayesian decision theory
 - Bayesian networks
 - K-means
- Computational-Statistical
 - SVM
 - AdaBoost
- Bio-inspired
 - Neural networks
 - Genetic algorithms
 - Artificial immune systems



Alpaydin, E. 2010 Introduction to Machine Learning (Adaptive Computation and Machine Learning). The MIT Press. (Chap 1,2)

Introduction

Patterns and
Generalization

Learning
Problems

**Learning
Techniques**