> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Question

An Introduction to Machine Learning

Fabio A. González Ph.D.

Depto. de Ing. de Sistemas e Industrial Universidad Nacional de Colombia, Bogotá

August 26, 2010

Content

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ● ●

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

1 Patterns and Generalization

Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems

Supervised Non-supervised Active On-line

3 Learning Techniques

4 Main Questions

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns Overfitting/ Overlearning

Learning Problems

Learning Techniques

Main Questions

1 Patterns and Generalization

Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line



4 Main Questions

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

1 Patterns and Generalization Generalizing from patterns

Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line



4 Main Questions

> Fabio A. González

What is a pattern?

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ○ □ ○ のへぐ

Ph.D. Patterns and

Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions • Data regularities

What is a pattern?

▲ロト ▲周 ト ▲ ヨ ト ▲ ヨ ト つ Q ()

Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Techniques

Main Questions

- Data regularities
- Data relationships

What is a pattern?

▲ロト ▲周ト ▲ヨト ▲ヨト 三三 - の久()

Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions • Data regularities

- Data relationships
- Redundancy

What is a pattern?

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

González Ph.D.

Fabio A.

Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

- Data regularities
- Data relationships
- Redundancy
- Generative model

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

Learning a Boolean function

x_1	x_2	f_1	f_2	 f_{16}
0	0	0	0	 1
0	1	0	0	 1
1	0	0	0	 1
1	1	0	1	 1

• How many Boolean functions of n variables are?

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

Learning a Boolean function

x_1	x_2	f_1	f_2	 f_{16}
0	0	0	0	 1
0	1	0	0	 1
1	0	0	0	 1
1	1	0	1	 1

- How many Boolean functions of n variables are?
- How many candidate functions are removed by a sample?

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

Learning a Boolean function

x_1	x_2	f_1	f_2	 f_{16}
0	0	0	0	 1
0	1	0	0	 1
1	0	0	0	 1
1	1	0	1	 1

- How many Boolean functions of n variables are?
- How many candidate functions are removed by a sample?

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

• Is it possible to generalize?

Inductive bias

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

Learning Fabio A. González Ph.D.

An Introduction to Machine

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions • In general, the learning problem is *ill-posed* (more than one possible solution for the same particular problem, solutions are sensitive to small changes on the problem)

Inductive bias

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

Patterns and Generalization

An Introduction to Machine

> Learning Fabio A. González Ph.D.

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Techniques

Main Questions

- In general, the learning problem is *ill-posed* (more than one possible solution for the same particular problem, solutions are sensitive to small changes on the problem)
- It is necessary to make additional assumptions about the kind of pattern that we want to learn

Inductive bias

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ○○○

Patterns and Generalization

An Introduction to Machine

> Learning Fabio A. González Ph.D.

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Techniques

Main Questions

- In general, the learning problem is *ill-posed* (more than one possible solution for the same particular problem, solutions are sensitive to small changes on the problem)
- It is necessary to make additional assumptions about the kind of pattern that we want to learn
- **Hypothesis space**: set of valid patterns that can be learnt by the algorithm

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

1 Patterns and Generalization

Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line



4 Main Questions

> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing fro patterns

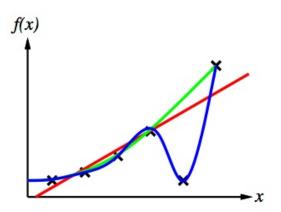
Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

What is a good pattern?



> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

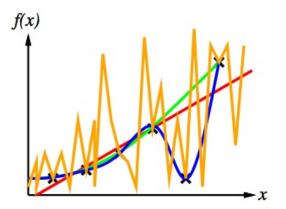
Overfitting/ Overlearning

Learning Problems

Learning Technique

Main Questions

What is a good pattern?



> Fabio A. González Ph.D.

Patterns and Generalization

Generalizing from patterns

Overfitting/ Overlearning

Learning Problems

Learning Techniques

Main Questions

Occam's Razor

from Wikipedia:

Occam's razor (also spelled Ockham's razor) is a principle attributed to the 14th-century English logician and Franciscan friar William of Ockham. The principle states that the explanation of any phenomenon should make as few assumptions as possible, eliminating, or "shaving off", those that make no difference in the observable predictions of the explanatory hypothesis or theory. The principle is often expressed in Latin as the lex parsimoniae (law of succinctness or parsimony).

"All things being equal, the simplest solution tends to be the best one."

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised Non-supervised Active On-line

Learning Techniques

Main Questions Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems

Supervised Non-supervised Active On-line



4 Main Questions

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Active

On-line

Learning Technique

Main Questions • Supervised learning

• Non-supervised learning

- Semi-supervised learning
- Active learning
- On-line learning

Types

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Active On-line

Learning Techniques

Main Questions Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised

Non-supervised Active

On-line



4 Main Questions

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

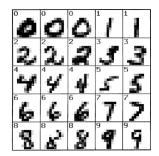
Supervised

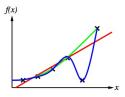
Non-supervise Active On-line

Learning Techniques

Main Questions Fundamental problem: to find a function that relates a set of inputs with a set of outputs

Supervised learning





イロト 不得 トイヨト イヨト 二日

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervise Active On-line

Learning Techniques

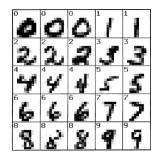
Main Questions

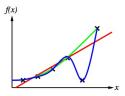
Fundamental

problem: to find a function that relates a set of inputs with a set of outputs

• Typical problems:

Supervised learning





・ロト・西ト・ヨト ヨー シック

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervise Active On-line

Learning Techniques

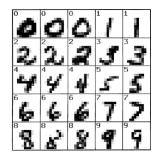
Main Questions

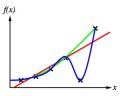
• Fundamental

problem: to find a function that relates a set of inputs with a set of outputs

- Typical problems:
 - Classification

Supervised learning





イロト 不得 トイヨト イヨト 二日

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervise Active On-line

Learning Techniques

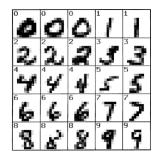
Main Questions

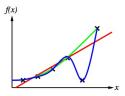
• Fundamental

problem: to find a function that relates a set of inputs with a set of outputs

- Typical problems:
 - Classification
 - Regression

Supervised learning





◆□ → ◆◎ → ◆○ → ◆○ → ○□

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Commissed

Non-supervised Active On-line

Learning Techniques

Main Questions Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems

Supervised

Non-supervised

Active

On-line



4 Main Questions

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervised

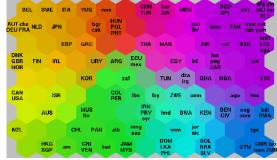
Active

Learning Technique

Main Question

Non-supervised learning

• There are not labels for the training samples



◆□▶ ◆□▶ ◆□▶ ◆□▶ ◆□ ● ● ●

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervised

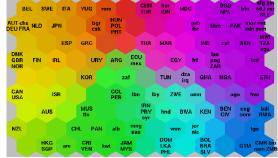
On-line

Learning Technique

Main Questions

Non-supervised learning

- There are not labels for the training samples
- Fundamental problem: to find the subjacent structure of a training data set



◆ロ ▶ ◆母 ▶ ◆臣 ▶ ◆臣 ▶ ○臣 ○ のへで

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervised

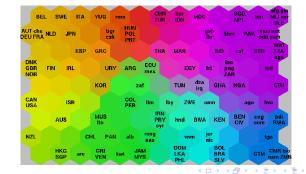
Active On-line

Learning Techniques

Main Questions

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, data compression



> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

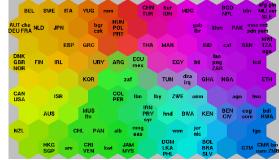
Non-supervised Active On-line

Learning Techniques

Main Questions

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, data compression
- Some samples may have labels, in that case it is called semi-supervised learning



◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervise

On-line

Learning Techniques

Main Questions Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems

Supervised Non-supervised

Active

On-line



4 Main Questions

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervised

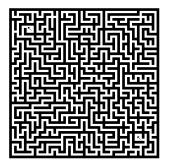
Active On-line

Learning Techniques

Main Questions

Active/reinforcing learning

• Generally, it happens in the context of an agent acting in an environment



> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

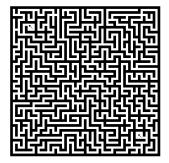
Active

Learning Techniques

Main Questions

Active/reinforcing 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



> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

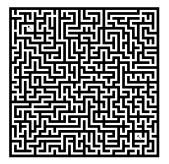
Active

Learning Techniques

Main Questions

Active/reinforcing 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)



> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

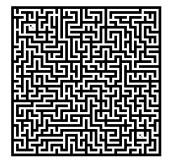
Active

Learning Techniques

Main Questions

Active/reinforcing 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)



▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problem:

Supervised Non-supervised Active

Learning Techniques

Main Questions Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems

Supervised Non-supervised Active

On-line

3 Learning Techniques

4 Main Questions

Fabio A.

On-line learning

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ○ □ ○ のへぐ

González Ph.D.

Generalization

Learning Problems

Supervised

Non-supervise

On-line

Learning Technique

Main Questions

• Only one pass through the data

On-line learning

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervise

On-line

Learning Technique

Main Questions

- Only one pass through the data
 - big data volume

On-line learning

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ● ●

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervise

On-line

Learning Technique

Main Questions

- Only one pass through the data
 - big data volume
 - real time

> Fabio A. González Ph.D.

On-line learning

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

Patterns and Generalization

Learning Problems

Supervised

Non-supervised

On-line

Learning Techniques

Main Questions • Only one pass through the data

• big data volume

real time

• It may be supervised or unsupervised

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Supervised

Non-supervised

On-line

Learning Techniques

Main Questions

On-line learning

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

- 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

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line

3 Learning Techniques

4 Main Questions

How to State the Learning Problem? How to Solve the Learning Problem? How to Measure the Quality of a Solution?

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

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

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

Outline

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ のQ@

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Technique

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure th Quality of a Solution? Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line

3 Learning Techniques

4 Main Questions

How to State the Learning Problem? How to Solve the Learning Problem? How to Measure the Quality of a Solution?

Outline

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Technique

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure th Quality of a Solution? Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

Learning Problems Supervised Non-supervised Active On-line

3 Learning Techniques

4 Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem? How to Measure the Quality of a Solution?

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

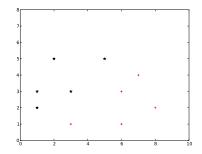
Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Two Class Classification Problem



• The idea is to buid a linear classifier function, $f: \mathbb{R}^2 \to \mathbb{R}$, such that:

$$f(x,y) = \begin{cases} < 0 & \text{if } (x,y) \in C_0 \\ > 0 & \text{if } (x,y) \in C_1 \end{cases}$$

Loss Function

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution? • Training set: $S = \{((x_1, y_1), l_1), \dots, ((x_n, y_n), l_n)\}$

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ○ □ ○ のへぐ

Loss Function

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ◆○◆

Learning Fabio A. González Ph.D.

An Introduction to Machine

Patterns and Generalization

Learning Problems

Learning Technique

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

- Training set: $S = \{((x_1, y_1), l_1), \dots, ((x_n, y_n), l_n)\}$
- Loss function:

$$L(f,S) = \frac{1}{2} \sum_{(x_i,y_i) \in S} (f(x_i,y_i) - l_n)^2$$

Loss Function

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Fabio A. González Ph.D.

An Introduction to Machine

Learning

Patterns and Generalization

Learning Problems

Learning Techniques

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

- Training set: $S = \{((x_1, y_1), l_1), \dots, ((x_n, y_n), l_n)\}$
- Loss function:

$$L(f,S) = \frac{1}{2} \sum_{(x_i,y_i) \in S} (f(x_i,y_i) - l_n)^2$$

• Are there other alternative loss functions?

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

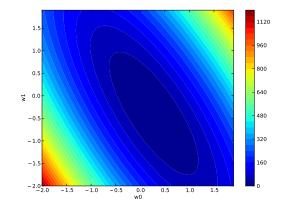
Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Square Error Loss



◆□▶ ◆□▶ ◆□▶ ◆□▶ ▲□ ◆ ◆ ◆

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

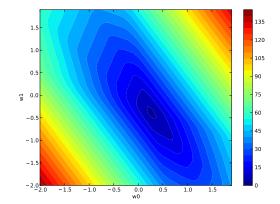
Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

L_1 Error Loss



> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Learning as Optimization

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへぐ

• General optimization problem:

 $\min_{f\in H} L(f,S)$

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Learning as Optimization

• General optimization problem:

 $\min_{f\in H} L(f,S)$

• Two Class 2D Classification:

$$H = \{f : f(x, y) = w_2 x + w_1 y + w_0, \forall w_0, w_1, w_2 \in \mathbb{R}\}$$

$$\min_{f \in H} L(f, S) = \min_{W \in \mathbb{R}^3} \frac{1}{2} \sum_{(x_i, y_i) \in S} (w_2 x_i + w_1 y_i + w_0 - l_i)^2$$

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ◆○◆

Outline

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure th Quality of a Solution? Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line

3 Learning Techniques

4 Main Questions

How to State the Learning Problem? How to Solve the Learning Problem? How to Measure the Quality of a Solution?

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Technique

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure th Quality of a Solution?

Iterative optimization of the loss function:

initialize $W^0 = w_0, w_1, w_2$ $k \leftarrow 0$

repeat

$$k \leftarrow k + 1$$

 $W^k \leftarrow W^{k-1} - \eta(k) \nabla L(f_{W^{k-1}}, S)$
until $|\eta(k) \nabla L(f_{W^{k-1}}, S)| < \Theta$

Gradient Descent

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ● ●

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

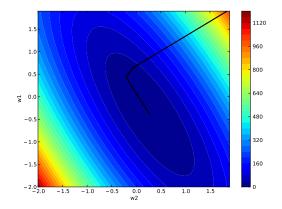
Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Gradient Descent Iteration Example (1)



▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● のへぐ

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

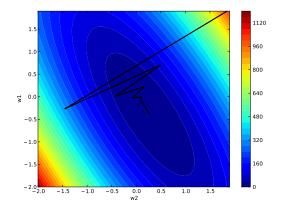
Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Gradient Descent Iteration Example (2)



▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

Outline

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

An Introduction to Machine Learning

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Technique

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution? Patterns and Generalization Generalizing from patterns Overfitting/ Overlearning

2 Learning Problems Supervised Non-supervised Active On-line

Learning Techniques

4 Main Questions

How to State the Learning Problem? How to Solve the Learning Problem? How to Measure the Quality of a Solution?

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Training Error vs Generalization Error

• The loss function measures the error in the training set

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ - つへ⊙

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Training Error vs Generalization Error

• The loss function measures the error in the training set

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

• Is this a good measure of the quality of the solution?

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Technique

Main Question

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

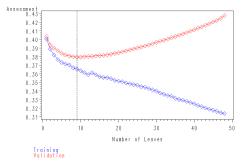
Training Error vs Generalization Error

• The loss function measures the error in the training set

◆□ → ◆◎ → ◆○ → ◆○ → ○□

• Is this a good measure of the quality of the solution?

Average Square Error (Gini index)



Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

• Generalization error:

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

Generalization Error

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

• Generalization error:

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

• How to control the generalization error during training?

Generalization Error

◆□▶ ◆□▶ ◆□▶ ◆□▶ □ - つへ⊙

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

• Generalization error:

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

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

Generalization Error

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ◆○◆

Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

• Generalization error:

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

Generalization Error

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ● ●

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

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution?

Regularization

• Vapnik, 1995:

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

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

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

> Fabio A. González Ph.D.

Patterns and Generalization

Learning Problems

Learning Techniques

Main Questions

How to State the Learning Problem?

How to Solve the Learning Problem?

How to Measure the Quality of a Solution? Alpaydin, E. 2004 Introduction to Machine Learning (Adaptive Computation and Machine Learning). The MIT Press. (Cap 1,2)

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○ ● ●