

---

# Machine Learning Formulas<sup>\*</sup>

García Prado, Sergio  
sergio@garciparedes.me

Taboada Rodero, Ismael José  
ismaeljose.taboada@alumnos.uva.es

May 28, 2017

## Abstract

[TODO ]

## 1 Introduction

[TODO ]

## 2 Supervised Learning

[TODO ]

**Overfitting** [TODO ]

**Error Measures** [TODO ]

- **Error Rate:** [TODO ]
- **Resubstitution error:** [TODO ]
- **Generalization Error:** [TODO ]

**Experimental Strategies** [TODO ]

- **Holdout:** [TODO ]
- **Repeated Holdout:** [TODO ]
- **Cross Validation:** [TODO ]
- **Repeated Cross Validation:** [TODO ]

### 2.1 Decision Trees

Decision trees are classifiers that works splitting the instances between the branches of a tree according to its attributes and assigning the feature's attribute in the leafs of the tree. An scheme of this structure is been included in the figure 1. Decision trees can represent logic functions. It tend to ovefit the data so the advanced heuristics try to avoid this problem.

---

<sup>\*</sup>This document is being maintained in <https://github.com/garciparedes/machine-learning-formulas>



### 2.1.3 C4.5

C4.5 is an extension of ID3 tree. It includes some improvements like branch pruning, working with numerical attributes and unknown values. The pruning works in post-pruning mode, i.e. the tree is generated and after is pruned. The heuristic use to determine if branch is pruned is the improve of pesimistic error rate. There are two pruning operations:

- **Subtree Replacement:** Consist of remove the selected subtree and assign the most frecuent feature to this leaf.
- **Subtree raising:** Consis on replace a tree by one of it's childrens and redistribute the instances. The operation is expensive.

Unknown values are trated spliting the selection between branches and assigning probabilities to each of them. The feature where the instance is classified is the one who maximized the probability. During the train period attributes with unknown values are penalized.

## 2.2 Rule Based Systems

[TODO ]

### 2.2.1 1R

[TODO ]

### 2.2.2 PRISM

[TODO ]

### 2.2.3 IREP

[TODO ]

### 2.2.4 RIPPER

[TODO ]

### 2.2.5 PART

[TODO ]

## 2.3 Instance Based Learning

[TODO ]

### 2.3.1 K - Nearest Neighbors

[TODO ]

### 2.3.2 IB3

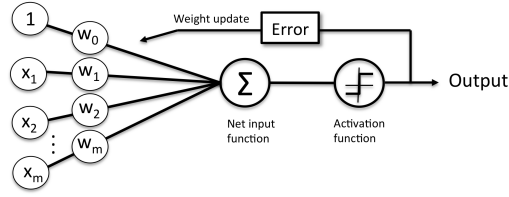
[TODO ]

## 2.4 Bayes Learning

[TODO ]

### 2.4.1 Naive Bayes

[TODO ]

**Figure 2:** *General concept of perceptron***2.4.2 K2**

[TODO ]

**2.4.3 TAN**

[TODO ]

**2.5 Linear Classifiers**

[TODO ]

**2.5.1 Linear Regression**

[TODO ]

**2.5.2 Logistic Regression**

[TODO ]

**2.5.3 Support Vector Machines**

[TODO ]

**2.6 Neural Networks****2.6.1 Perceptron model**

Perceptron's structures

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}, x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \quad (6)$$

Ouput equation

$$z = w_1x_1 + \cdots + w_mx_m = \mathbf{w}^T \mathbf{x} \quad (7)$$

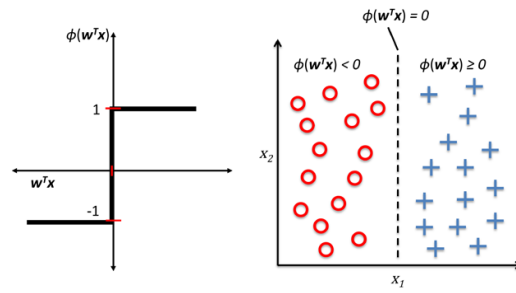
Activation function

$$\phi(z) = \begin{cases} 1 & \text{if } z \geq \theta \\ -1 & \text{otherwise} \end{cases} \quad (8)$$

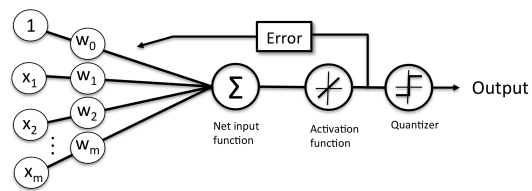
Update of weight vector

$$w_j := w_j + \Delta w_j \quad (9)$$

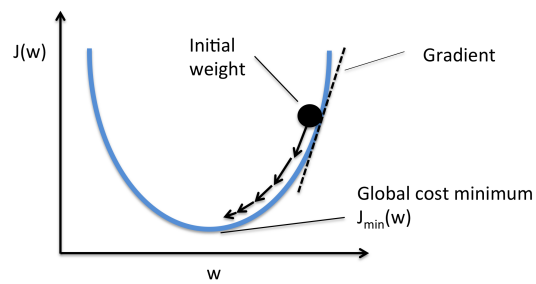
$$\Delta w_j = \eta(y^{(i)} - \hat{y}^{(i)})x_j^{(i)} \quad (10)$$



**Figure 3:** *Activation function*



**Figure 4:** *General concept of adaline perceptron*



**Figure 5:** *Gradient descent*

### 2.6.2 Adaptive linear neurons (ADALINE)

Cost function

$$J(w) = \frac{1}{2} \sum_i (y^{(i)} - \phi(z^{(i)}))^2 \quad (11)$$

Update of weight vector

$$w_j := w_j + \Delta w_j \quad (12)$$

$$\Delta w_j = -\eta \nabla J(w) = \eta \sum_i (y^{(i)} - \phi(z^{(i)})) x_j^{(i)} \quad (13)$$

Features standarization

$$x'_j = \frac{x_j - \mu_j}{\sigma_j} \quad (14)$$

### 2.6.3 Single Layer Neural Networks

Simple Perceptron [TODO ]

ADALINE [TODO ]

### 2.6.4 Multi Layer Perceptron

[TODO ]

### 2.6.5 Radial Basis Functions

[TODO ]

### 2.6.6 Convolutional Neural Networks

[TODO ]

### 2.6.7 Recurrent Neural Networks

[TODO ]

## 3 Unsupervised Learning

[TODO ]

## References

- [CCAG17] Teodoro Calonge Cano and Carlos Javier Alonso González. Técnicas de Aprendizaje Automático, 2016/17.
- [ML13] Big ML. A New Way to Visualize Decision Trees. <https://blog.bigml.com/2013/04/19/a-new-way-to-visualize-decision-trees/>, 2013.
- [Ols16] Dr. Randal S. Olson. Python Machine Learning, 2016.