# Machine Learning & Data Mining

Prova scritta 14 Giugno 2018

1) **Concept learning**

Describe the Find-S algorithm and indicate the main complaints of this algorithm.

2) **Decision trees**

Dynamic split construction: Describe a method to dynamically create a discrete attribute to test a continuous a attribute

3) **Overfitting and Evaluation of learning algorithms**

Describe what the *validation set* is and how it is used in the technique "Reduced-Error Pruning" applied to decision trees.

4) **Artificial neural networks**

Describe the difference between the "Batch Gradient Descent" algorithm and the "Incremental (Stochastic) Gradient Descent" algorithm.

5) **Bayesian learning**

Describe the Naive Bayes Algorithm.

6) **Clustering**

Total SSE is the sum of the SSE for each separate attribute.

(a) What does it mean if the SSE for one variable is low for all clusters?

(b) Low for just one cluster?

(c) High for all clusters?

(d) High for just one cluster?

(e) How could you use the per variable SSE information to improve your clustering?

(f) Discuss the advantages and disadvantages of treating clustering as an optimization problem.

7) **Regression**

We are given the following 11 observations on variables x and y:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **x** | 6 | 1 | 0 | 11 | 2 | 4 | 14 | 15 | 5 | 8 | 10 |
| **y** | 11.2 | 7.7 | 14 | 10.7 | 13.4 | 11.8 | 8.8 | 5.5 | 8.5 | 7.6 | 10 |

We want to see if a linear model gives a reasonable fit of the data, so we estimate the model with least squares. Compute the least squares estimates of $w_0$ and $w_1$.

Compute the proportion of the variance in y that is explained by the model (R2= 1- SSE/SST). Where SSE= $\Sigma e^2$ and SST=$\Sigma(y-\mu(y))^2$

8) **Ensemble methods**

Explain the training and testing phases of Boosting and compare it with Bagging.

9) **SVM**

Consider the following linearly separable two-dimensional input vectors.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **x** | 1 | 1 | 1 | 3 | 3 | 4 | 5 | 6 | 6 | 8 | 9 | 10 |
| **y** | 1 | 4 | 9 | 2 | 5 | -1 | 13 | 6 | 9 | 3 | 7 | -3 |
| **Class** | A | A | A | A | A | A | B | B | B | B | B | B |

Find the linear SVM that optimally separates the classes by maximizing the margin. Show the support vectors and compute the margin.