Machine Learning & Data Mining

Prova intermedia 10 maggio 2016

1) Concept learning

1.a) Indicate which are the main complaints about the algorithm FIND-S.

1.b) Let G be the current set of maximally general hypotheses (G) in the execution of the Candidate Elimination algorithm: describe how G is modified when a negative example d is processed by the algorithm.

2) Alberi di decisione

2.a) Specify the definition of misclassification error, and describe the method to decide the best split using this definition.

2.b) Suppose that:

if we split using the binary attribute A we obtain Node1: 4 examples of class C1 and 5 examples of class C2 Node2: 4 examples of class C1 and 6 examples of class C2

If we split using the binary attribute B we have Node1': 3 examples of class C1 and 7 examples of class C2 Node2': 5 examples of class C1 and 4 examples of class C2

Demostrate using the misclassification error if it is better to split A or B.

2.c) Using the notions of Trainining Error (Error_{train}) and of Generalization Error over the full distribution D of the problem instances (Error_D), specify formally when a hypothesis h *overfits* the training data.

2.d) Why the validation set and the test set are different sets? (Answer by describing the purpose of these two sets).

3) Valutazione di algoritmi

Describe an algorithm to compare two learning algorithms L_{A} and L_{B} using the k-fold cross validation method.

4) Reti Neurali

4.1) Explain what is the definition of "momentum" that can be used in the back propagation algorithm.4.2) Specify the Gradient Descent Rule to update the weights of a linear unit.

5) Learning Bayesiano

5.1) Suppose that we have a new instance x with possible classifications "yes" and "no", and three possible hypotheses such that: P(h1|D) = 0.5, P(h2|D) = 0.2, P(h3|D) = 0.3 P("no"|h1) = 0, P("yes"|h1) = 1 P("no"|h2) = 0, P("yes"|h2) = 1P("no"|h3) = 1, P("yes"|h3) = 0

Which is the value v* computed by the Optimal Bayes Classifier for this example? Why?

5.2) Which is the assumption made by the Naive Bayes Classifier that makes it much more applicable than the Optimal Bayer Classifier?