Pattern Analysis and Machine Intelligence

M. Matteucci, L. Malagò, D. Eynard
26/06/2012

Answer the following questions identifying the key aspects and try not to exceed the 1.5 page limit per question.

- Use only the 3 sheets provided by the teacher
- **Write your answers on different sheets according to the question**
- Write your name and Student ID on each sheet you turn in
- English is the official language, however Italian is allowed
- Either pen and pencil are allowed no other technical mean to support yourself is allowed

In case you have special needs (e.g., being graded within a given time) please write it on top of your assignment and ***tell it to the teacher!***.

Question 1: Logistic Regression (Answer on sheet 1)

With reference to the Logistic Regression method/model for classification, describe:

1. Its assumptions and analytical form
2. Its (dis)advantages wrt plain Linear Regression (on the indicator matrix), Linear Discriminant Analysis, and Optimal Separating Hyperplanes
3. How do we train this classifier from data?
4. Does it work for multi-class problems?
   If yes how? If no, why?

Question 2: Local Regression Methods (Answer on sheet 1)

With reference to Local Linear Regression and Local Polinomial Regression:

1. Describe the idea of Local Linear Regression & Local Polinomial Regression.
2. What do we mean by "Equivalent Kernel" in Local Regression?
3. What do we mean by width of the Kernel in kernel smoothing methods?
4. How can we select the width of the kernel used for Local Regression?

Question 3: Clustering (Answer on sheet 2)

Suggest one or more algorithms for each of the questions and motivate your answers. If you are in trouble finding an answer, you can also suggest algorithms you would suggest NOT to use for the different taks, motivating why they are not suitable for them.

1 What clustering algorithm would you suggest to use if you need
  1.1 to detect not only groups, but also outliers in the data?
  1.2 to deal with clusters of different densities?
  1.3 to deal with data items which cannot be plotted in a Euclidean space?

2. What is the difference between similarity (or dissimilarity) and closeness (or distance) between two data points? Can you easily pass from one to the other?

Question 4: Regression (Answer on sheet 3)

Introduce and discuss briefly ridge regression:

1. What are the main differences compared to ordinary least squares? Motivate why such differences have been introduced.
2. Describe the two equivalent formulations for the minimization problem solved in ridge regression: constrained formulation and the Lagrangian formulation, and in the case of constrained formulation, present a geometric interpretation (do not forget to label the axis)
3. Derive mathematically the closed formula for the ridge regression estimators (do not forget to state the dimensions of the matrix X and vector y)
4. Discuss differences and similarities between ridge regression and the lasso.