SMAI-2020-Homework

September 2020

1 Objective Question [1 mark]

The Linear Discriminant Analysis (LDA) classifier is a supervised algorithm that computes the direction maximizing the ratio of between-class variance over within-class variance. (True/False).

2 Objective Question [2 marks]

You trained binary classifier model which gives very high accuracy on the training data, but much lower accuracy on validation data. The following may be true:

1. This is an instance of overfitting.
2. This is an instance of underfitting.
3. The training was not well regularized.
4. The training and testing examples are sampled from different distributions.

3 Subjective Question [2 marks]

What is the relation between Linear Discriminant Analysis and Bayes Rule?

4 Programming Question [3 marks]

In the tutorial you saw how to compute LDA for a two class problem. In this excercise we will work on a multi-class problem. We will be working with the famous Iris dataset that has been deposited on the UCI machine learning repository. Write the code that computes the multi-class LDA on the IRIS dataset and compute the overall accuracy of the classifier. Plot the decision boundary for the final classifier obtained. You are provided with 'LDA Excercise 1.ipynb' with the starter code.

5 Programming Question [2 marks]

In previous question, we computed the LDA for a multi-class problem, the IRIS dataset. In this problem, write the code that computes the PCA and LDA on the IRIS dataset. Compare the feature subspace that you obtain via PCA and LDA and summarize your observations in couple of sentences.