Homework CS 273A Intro ML

Due Tu Jan. 11 2011, noon in EEE dropbox (only pdfs)

1. Reading 1: Bishop, Sec 1.1, 2.5.2, 4.3.2, Appendix C.

2. Reading 2: Classnotes Chapters 1,2,3,4,5,7

3. Download the “Iris Dataset” into Matlab. Randomly divide the dataset into 70% training data and 30% test data.

   A) Implement the 1 Nearest Neighbor classification algorithm on the Iris Dataset and report the classification error (the number of misclassified test-cases divided by the total number test-cases).

   B) Derive the gradient descend update rules for logistic regression.

   C) Implement logistic regression on the Iris dataset and report classification error.

   D) Which of the two algorithms (1NN or LR) entertains a larger set of hypotheses?

4. Prove the relation: \( p(y \mid x_I, x_{II}) \propto p(x_{II} \mid y) p(y \mid x_I) \) under the conditional independence assumption for the naïve Bayes classifier.