CSE 6363： Machine Learning, Spring 2020

Time: Tuesday 7-9:50pm Location: NH 202

Instructors: Dr. Chris Ding, chqding@uta.edu
Dr. Di Ming, di.ming@mavs.uta.edu
Office Hour: Tuesday, 10:30am-12:30pm, ERB 424. (or by appointment)

TA: Qicheng Wang, qicheng.wang@mavs.uta.edu
Office Hour: Tuesday & Thursday, 3:00pm-5:00pm, ERB 204.

Textbook:

Pattern Recognition and Machine Learning
Christopher Bishop

Course Schedule

Week 1 & 2.

Introductions
Three concrete examples:
1. Data Mining example: Market basket Data analysis
2. Pattern Recognition example: Handwritten letters recognition
3. Cancer prediction using DNA expressions recorded on microarrays

Fitting Curve to Data (textbook sec. 1.1)

Linear Regression

Probability

Week 3 & 4.

Naïve Bayes Classifier

Decision Tree, Mutual Information, Random Forest

Week 5 & 6 & 7.

Classification：
KNN.
Centroid Method.

Support Vector Machine (SVM):
Margin, Primal-Dual Problems, KKT Condition.
Hard SVM, Soft SVM, Kernel SVM.

Week 8.

Spring break (no class).

Week 9.

March 17th: **first computer quiz** (polynomial curve fitting & naïve Bayes classifier)

Week 10.

March 24th: **first written quiz** (including all the lectures before spring break)

Week 11 - 18.

TBD.

Homework 1-5

HW1: Textbook Exercise 1.1(p.6, p.58)

HW2: Show that when M=1, the results of HW1 is identical the results of linear regression.

HW3: Textbook Exercise 1.2.

HW4

A problem on a multiple-choice quiz is answered correctly with probability 0.9 if a student is prepared. An unprepared student guesses between 4 possible answers, so the probability of choosing the right answer is 1/4. Seventy-five percent of students prepare for the quiz. If Mr. X gives a correct answer to this problem, what is the chance that he did not prepare for the quiz?

HW5

At a plant, 20% of all the produced parts are subject to a special electronic inspection. It is known that any produced part which was inspected electronically has no defects with probability 0.95. For a part that was not inspected electronically this probability is only 0.7. A customer receives a part and finds defects in it. What is the probability that this part went through an electronic inspection?

HW1, HW2, HW3, HW4, HW5 are due on Feb 11th, 7:00pm.

Homework 6-7

HW6

Solve SVM for a data set with 3 data instances in 2 dimensions: (1,1,+), (0,-1,+), (-1,1,-). Here, the first 2 number are the 2-dimension coordinates, '+' in 3rd place is positive class, and '-' in 3rd place is negative class. Your task is to: (1) write down dual-problem using those 3 data instances. (2) compute alpha's. (3) compute w and b. (4) compute margin.

HW7

Solve SVM when data are non-separable, when minimizing the violations of the mis-classification, i.e., on those slack variables.

$$\min_{\mathbf{w}} \ \frac{1}{2}||\mathbf{w}||_2^2 + C\Big(\sum_{i=1}^{n}\xi_i\Big)^k$$

$$\text{s.t.} \ \ y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \quad i = 1, \cdots, n.$$

$$\xi_i \geq 0, \quad i = 1, \cdots, n.$$

Your task is to: derive the dual-problem of the above primal-problem when k=2.

HW6, HW7 are due on Mar 24th, 7:00pm.

Computer Project 1, due on Mar 17th, 7:00pm. ----------

You MUST write the codes YOURSELF.

For the implementation of "Polynomial Curve Fitting" and "Naïve Bayes Classifier", you are NOT allowed to use any Python Library.

Computer Project 1A: Write a computer program to generate the 10 data points as shown in Figure 1.2.

Hint 1A: for each data point (x_i, t_i), a random noise e_i is added to obtain $t_i = \sin(2\pi x_i) + e_i$. Details can be found in the textbook page #4.

Computer Project 1B: Write a computer program to solve the equations of Exercise 1.1, for the 10 data points you generated in part 1A. Plot the fitted curves and original data points as Figure 1.4, for M=0, 1, 3, 9.

Hint 1B: (1) transform original 1-dimensional feature to multi-dimensional features;
(2) use linear regression to solve equation (1.2).

Computer Project 1C: Write a computer program to solve the equations of Exercise 1.2, for the 10 data points you generated in part 1A. Here, M is fixed as 9. Show that as the lambda of Equation (1.4) increases, the overfitting of Figure 1.4 (the right-bottom figure) is reduced significantly, see Figure 1.7.

Hint 1C: (1) transform original 1-dimensional feature to multi-dimensional features;
(2) use linear regression with L_2-regularization to solve equation (1.4).

Computer Project 1D: Write a computer program to implement naïve Bayes classifier with Laplacian (add-1) smoothing, for the given "vertebrate.txt" dataset. Compute the multinomial distribution for each attribute of the data instances and prior probability. When a new data instance is presented, compute the class label using NAÏVE Bayes classification method.

Hint 1D: (1) training stage: compute the prior probability and likelihood.
(2) testing stage: compute the posterior probability using Bayes theorem.