Homework 2 Basic Machine Learning
For the deadline see Canvas
Version: Mon 10th Sept, 2018 at 10:56.

Introduction

1. This is a group assignment, so sign up in groups of two students.
2. Each group has to submit a **pdf** with their answers and explanation. **Please put your names and group number at the top of the hand in**.
3. For questions about this homework assignment use the Discussion Board on Canvas.
4. Of course you may use a calculator or a programming environment such as Matlab or Python. **But your report should not contain any code. Explain your computations and results in English!**
5. It is allowed to incorporate handwritten notes or derivations or drawings in your submission as long as these are readable!
6. Explain your answers!

Exercise 1: Perceptron Learning rule (10 points)

Consider a 2-dimensional learning problem and for simplicity assume that we only have two data points: $(2, 3)$ with target -1 and $(-2, 0)$ with target 1. The weights of the perceptron are given by $\mathbf{w} = (w_0, w_1, w_2) = (1, 2, 1)$.

Part a

Draw the decision line corresponding to this perceptron, indicate which points will be classified as positive and draw the two data points

Part b

How will the weights of the perceptron be adapted in one applies the perceptron learning rule (batch version) for one iteration with learning rate $\eta = 0.5$.

Part c

Explain in your own words the difference between batch (normal) gradient descent versus stochastic gradient descent (Book section 3.1.3 on sequential learning and page 194).

Part d

Repeat part b if one applies stochastic learning for one iteration (one complete sweep through the dataset), the order of data points is first $(-2, 0)$ and afterwards $(2, 3)$.

Exercise 2: Logistic classification/discrimination (book section 4.3.2) (10 points)

Part a

Write in pseudo code (stepwise textual description or textual recipe) the algorithm for the logistic learning algorithm, batch version. You may skip the data loading part etcetera. Just

focus on the heart of the algorithm which concerns the initialization of the weight vector $\mathbf{w}$ and the updates of $\mathbf{w}$ based on the dataset until convergence. Be explicit about your convergence criteria.

Part b

Consider a 2-dimensional classification problem and assume that the weights for the logistic classifier are given by $\mathbf{w} = (w_0, w_1, w_2) = (1, 2, 2)$. Consider the data point $x = (-1, 1)$ and assume that this data point is misclassified. What is the formula for updating the weights and what will be the new weights if one applies stochastic gradient descent one time with learning parameter equal to 0.6.

Exercise 3: Error functions and gradient descent (10 points)

Part a

Assume that we use the error function

$$E(\mathbf{w}) = \sum_{n=1}^{N} (t_n - y_n)^4$$

instead of the error function (4.90) of the book. What will be the formula for the gradient $\nabla E(\mathbf{w})$ of E with respect to $\mathbf{w}$?

Part b

Repeat the same question as in part a but now for the error function

$$E(\mathbf{w}) = \sum_{n=1}^{N} |t_n - y_n|$$

where $|x|$ is the absolute value of x.

Exercise 4: Multiple choice questions (2 bonus points)

Design two multiple choice (MC) questions concerning Chapter 4. Clearly indicate what knowledge or skill you want to test with the MC questions.
An example of a multiple choice question which tests the knowledge of Bayes Law:

Which of the following formulas represents Bayes Law? Mark them all!

(a) $p(x|C) = P(C|x)P(C)/p(x)$

(b) $P(C|x) = p(x|C)P(C)/p(x)$

(c) $p(x) = P(C|x)p(x|C)/P(C)$

(d) $P(C|x)p(x) = P(x|C)P(C)$