DATS 6202 Term 2018-Fall

Machine Learning I

Quiz 3 October 2, 2018

Quiz 3: Solutions

Full Name: _						
GWID:						

• DATS 6202, Instructor: Yuxiao Huang

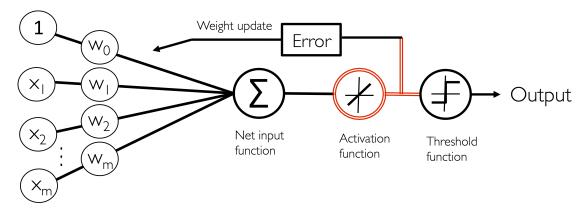
Material Covered

- Perceptron
- Adaline

Note

- The quiz has 100 points.
- The quiz period is 20 minutes.
- The quiz is closed book and closed notes.
- The quiz is closed electronics (e.g., no laptops, netbooks, OLPCs, tablets, iPads, calculators, cellular phones, iPhones, Nexi, iPods, Zunes, Kindles, Nooks).
- There is only one correct answer for each Multiple Choice Question.
- For each Calculation question (if there is any), you must show the essential steps. **No mark** will be given if only the result is provided.

1 Description (50 points)



Adaptive Linear Neuron (Adaline)

Picture courtesy of the book *Raschka S. and Mirjalili V.* (2017). *Python Machine Learning. 2nd Edition.*

Consider the Adaline model in the figure above.

1. What does the Net input function do? Write down the equation that summarizes the function.

It calculates the weighted sum of the input. Specifically, it outputs z where

$$z = w_0 + w_1 \cdot \mathbf{x}_1 + w_2 \cdot \mathbf{x}_2 + \ldots + w_m \cdot \mathbf{x}_m.$$

2. What does the Activation function do?

It is an identity function that takes as input the output of the net input function (i.e., z), and outputs the exact same input. In simple words, it does nothing.

3. What is the Error? In other words, it is the difference between which two items?

It is the difference between the real class label and the output of the Activation function. That is

$$y-z$$

4. Change one thing in the structure of Adaline to get the structure of logistic regression?

Change the Activation function into the sigmoid function.

5. Change one thing in the structure of Adaline to get the structure of linear regression?

Remove the Threshold function.

2 Calculation (50 points)

Assume the Threshold function in Adaline produces outcomes that are either 0 or 1. Specifically,

$$\hat{y} = \begin{cases} 1 & \text{if } z \ge 0\\ 0 & \text{otherwise} \end{cases}.$$

Here.

- \hat{y} is the predicted class label
- z is the output of the Activation function

Assume there are two features in the data, x_1 and x_2 , and two class labels, 0 and 1.

1. The value of the parameters obtained by gradient descent are as follows:

$$w_0 = -3$$
 and $w_1 = 2$ and $w_2 = 2$.

Calculate the value of z and \hat{y} given the value of x_1 and x_2 :

$$egin{array}{llll} x_1 = 0 & \mbox{and} & x_2 = 0, & z = & \mbox{and} & \hat{y} = \\ x_1 = 0 & \mbox{and} & x_2 = 1, & z = & \mbox{and} & \hat{y} = \\ x_1 = 1 & \mbox{and} & x_2 = 0, & z = & \mbox{and} & \hat{y} = \\ x_1 = 1 & \mbox{and} & x_2 = 1, & z = & \mbox{and} & \hat{y} = \\ \end{array}$$

$$egin{array}{llll} x_1 = 0 & \mbox{and} & x_2 = 0, & z = -3 & \mbox{and} & \hat{y} = 0 \\ x_1 = 0 & \mbox{and} & x_2 = 1, & z = -1 & \mbox{and} & \hat{y} = 0 \\ x_1 = 1 & \mbox{and} & x_2 = 0, & z = -1 & \mbox{and} & \hat{y} = 0 \\ x_1 = 1 & \mbox{and} & x_2 = 1, & z = 1 & \mbox{and} & \hat{y} = 1 \end{array}$$

2. The value of the parameters obtained by gradient descent are as follows:

$$w_0 = -1$$
 and $w_1 = 2$ and $w_2 = 2$.

Calculate the value of z and \hat{y} given the value of x_1 and x_2 :

$$egin{array}{llll} x_1 = 0 & \mbox{and} & x_2 = 0, & z = & \mbox{and} & \hat{y} = \\ x_1 = 0 & \mbox{and} & x_2 = 1, & z = & \mbox{and} & \hat{y} = \\ x_1 = 1 & \mbox{and} & x_2 = 0, & z = & \mbox{and} & \hat{y} = \\ x_1 = 1 & \mbox{and} & x_2 = 1, & z = & \mbox{and} & \hat{y} = \\ \end{array}$$

THIS PAGE INTENTIONALLY LEFT BLANK

(You may use it as scratch paper, but do submit it as part of your completed exam.)