

CSCI 5702/7702 - Data Mining and Analytics

Fall 2015 - Assignment 5

Due: 12/04/2015 8:59pm

1. Table 1 consists of training data from an employee database. The data have been generalized. For example, “31 ... 35” for *age* represents the age range of 31 to 35. For a given row entry, *count* represents the number of data tuples having the values for *department*, *status*, *age*, and *salary* given in that row. Let *status* be the class-label attribute.
 - A. Design a **multilayer feed-forward neural network** for the given data. Label the node in the input and output layers.
 - B. Using the multilayer feed-forward neural network obtained in (a), show the weight values after one iteration of the backpropagation algorithm, given the training instance “(sales, senior, 31 ... 35, 46K ... 50K)”. Indicate your initial weight values and biases and the learning rate used.

<i>Department</i>	<i>Status</i>	<i>Age</i>	<i>Salary</i>	<i>Count</i>
<i>Sales</i>	<i>Senior</i>	<i>31 ... 35</i>	<i>46K ... 50K</i>	<i>30</i>
<i>Sales</i>	<i>Junior</i>	<i>26 ... 30</i>	<i>26K ... 30K</i>	<i>40</i>
<i>Sales</i>	<i>Junior</i>	<i>31 ... 35</i>	<i>31K ... 35K</i>	<i>40</i>
<i>Systems</i>	<i>Junior</i>	<i>21 ... 25</i>	<i>46K ... 50K</i>	<i>20</i>
<i>Systems</i>	<i>Senior</i>	<i>31 ... 35</i>	<i>66K ... 70K</i>	<i>5</i>
<i>Systems</i>	<i>Junior</i>	<i>26 ... 30</i>	<i>46K ... 50K</i>	<i>3</i>
<i>Systems</i>	<i>Senior</i>	<i>41 ... 45</i>	<i>66K ... 70K</i>	<i>3</i>
<i>Marketing</i>	<i>Senior</i>	<i>36 ... 40</i>	<i>46K ... 50K</i>	<i>10</i>
<i>Marketing</i>	<i>Junior</i>	<i>31 ... 35</i>	<i>41K ... 45K</i>	<i>4</i>
<i>Secretary</i>	<i>Senior</i>	<i>46 ... 50</i>	<i>36K ... 40K</i>	<i>4</i>
<i>Secretary</i>	<i>Junior</i>	<i>26 ... 30</i>	<i>26K ... 30K</i>	<i>6</i>

Table 1 – the employee dataset

2. Consider the dataset in Figure 1, which has points from two classes c_1 (triangles) and c_2 (circles). Answer the questions below. Unless mentioned otherwise, assume we are dealing with the perfectly separable case.

- A. Find the equations for the two hyperplanes h_1 and h_2 .
- B. Show all the support vectors for h_1 and h_2 .
- C. Which of the two hyperplanes shown (h_1 and h_2) is better at separating the two classes based on the margin computation.

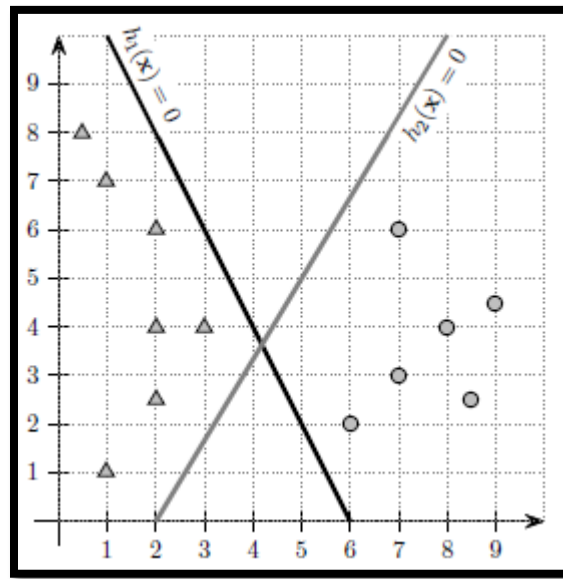


Figure 1

Submission Guideline

Please submit your assignment through **Canvas** before the deadline; late submissions are not accepted! You are allowed to submit your assignment multiple times, but only the last submission (**before the deadline**) will be recorded and graded.

Your submission should be a single **pdf** file named **<Your CU Denver Portal ID>-A5.pdf**. For example, if your CU Denver Portal ID is "john", the file name would be: **john-A5.pdf**.

Notes:

- ❖ Please download your assignment after submission and make sure it is not corrupt. We won't be responsible for the corrupted submissions and will not be able to take a resubmission after the deadline.

You are highly encouraged to ask your question on **Piazza** under the "a5" folder. Please **DO NOT** include your solutions in the comments you share on Piazza. Feel free to help other students with general questions.