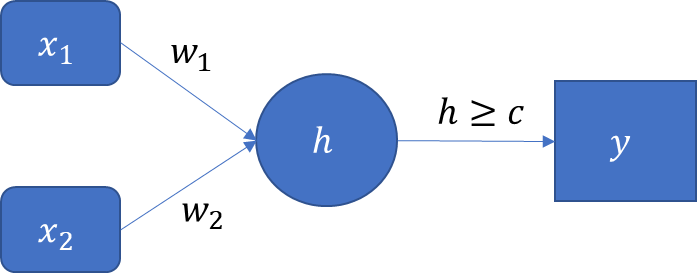
CS 584-04: Machine Learning

Fall 2018 Assignment 5

# Question 1 (40 points)



1. (10 points). If we restrict the values of the parameters , , and to positive integers, then specify the lowest possible values for these parameters such that the perceptron can implement the logical AND function.

W1 = 1

W2 = 1

C = 2

1. (10 points). If we restrict the values of the parameters , , and to positive integers, then specify the lowest possible values for these parameters such that the perceptron can implement the logical OR function which can be represented by the following table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 0 | 1 | 1 |
|  | 0 | 1 | 0 | 1 |
| OR | 0 | 1 | 1 | 1 |

W1 = 1

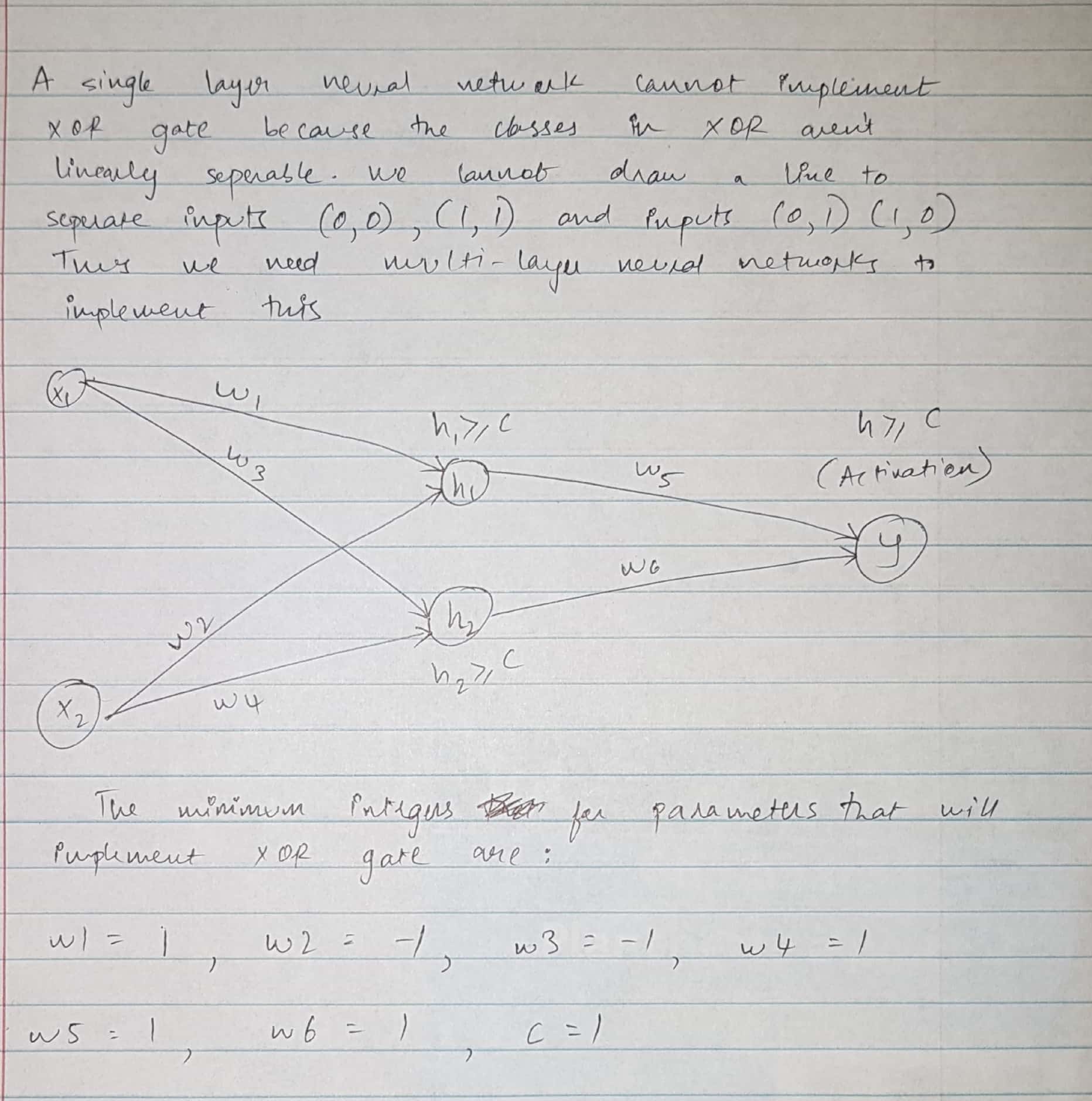
W2 = 1

C = 1

1. (20 points). The logical XOR function (i.e., the Exclusive OR) returns TRUE only when one argument is TRUE and another is FALSE. Otherwise, it returns FALSE. This can be represented by the following table:

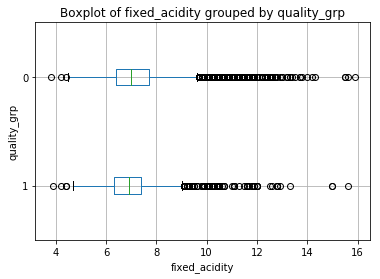
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 0 | 1 | 1 |
|  | 0 | 1 | 0 | 1 |
| XOR | 0 | 1 | 1 | 0 |

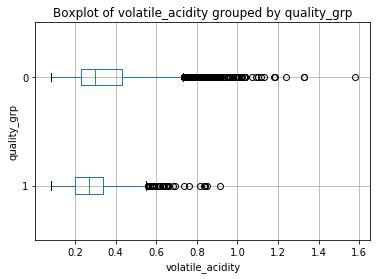
Consider a neural network which has two neurons in a single hidden layer. Specify the four synaptic weights and a threshold value such that the neural network can implement the XOR function. The parameters are still integers, but we allow negative integers.

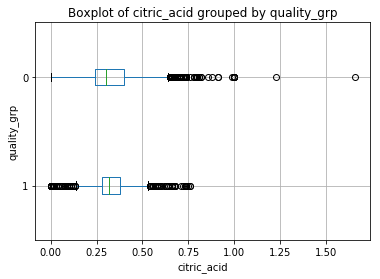


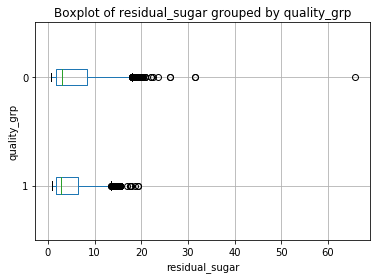
# Question 2 (60 points)

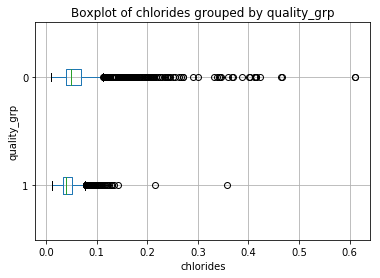
1. (10 points). Generate a horizontal box-plot for each input attribute, grouped by the target variable quality\_grp.

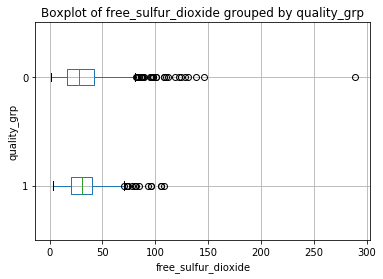


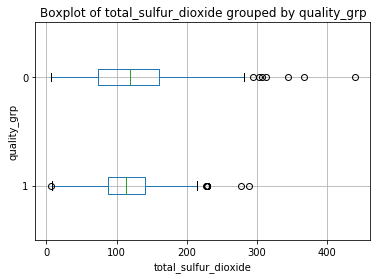


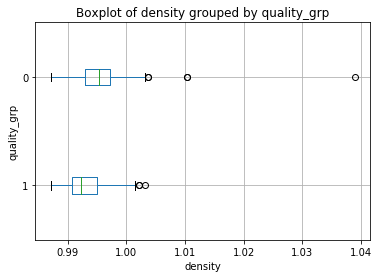


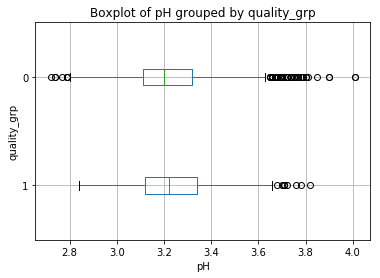


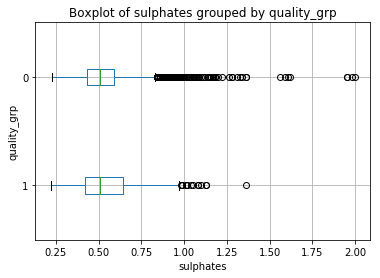


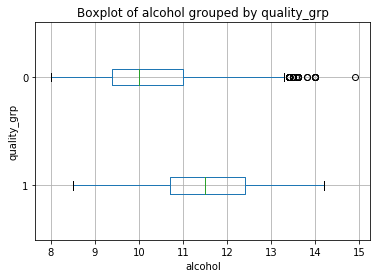




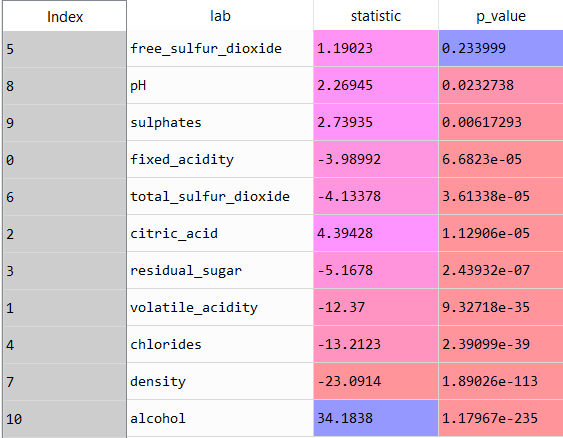








1. (10 points). The scipy.stats module has the ttest\_ind function for comparing two independent samples using the Student’s *t* test. Use this function to calculate the two-sided *p*-value of the Student’s *t* test. The group variable is the target variable quality\_grp. List the names of the input attribute, their *t* statistics, and their two-sided *p*-values. The rows are in descending order of the two-sided *p*-values.



1. (10 points). Perform the Support Vector Machine analysis using the svm.LinearSVC function. The random state value is specified to 20181111. The maximum number of iterations is specified to 10000. When the algorithm does not converge, we may need to remove some variables. You will first use all eleven input attributes. If the algorithm does not converge, you will remove the variable which has the highest *p*-value. If that does not help the algorithm converge, then the variable which has next highest *p*-value is removed, and so on. What input attributes are retained such that the algorithm can converge for the first time?

Input attributes that are retained such that the algorithm can converge for the first time:

['volatile\_acidity', 'chlorides', 'density', 'alcohol']

1. (5 points). What is the Mean Accuracy of your model in (c)?

Mean Accuracy of the model in (c) = 0.8123749422810528

1. (5 points). What is the hyperplane? You need to present the hyperplane in this format . Include only the attributes that you use in (c).

In geometry, a hyperplane is a subspace whose dimension is one less than that of its ambient space. If a space is 3-dimensional then its hyperplanes are the 2-dimensional planes, while if the space is 2-dimensional, its hyperplanes are the 1-dimensional lines. This notion can be used in any general space in which the concept of the dimension of a subspace is defined. In machine learning, hyperplanes are a key tool to create support vector machines.

Let be a vector of scalars and at least one of them is not zero.

Let be a vector in the the *p*-dimensional space of real numbers.

The set for a constant is a hyperplane. In other words, a hyperplane is a subspace of .

* 1. is the inner product of the vectors and .

If *p* = 1, then this point is a hyperplane.

If *p* = 2, then this line is a hyperplane when both and .

* 1. If or (but not both), then this hyperplane reduces to a point.

If *p* = 3, then a plane is a hyperplane when , , and .

* 1. If one of the scalars is zero, then this hyperplane reduces to a line.
  2. If two of the scalars are zero, then this hyperplane reduces to a point.

In general, a hyperplane lies in a subspace which is at least one dimension less than that of .

Intercept = [-2.11236983]

Weight Coefficients = [[-0.83531774 -0.65083327 -0.9685864 0.26397236]]

X1 = values of volatile\_acidity

X2 = values of chlorides

X3 = density

X4 = alcohol

Hyperplane = (-2.11236983) + (-0.83531774)( volatile\_acidity) + (-0.65083327)( chlorides) + (-0.9685864)( density) + (0.26397236)(alcohol)

1. (10 points). When the attributes are at their overall means, what will be the predicted category for quality\_group? List the attributes’ overall means with your answer.

['volatile\_acidity', 'chlorides', 'density', 'alcohol']

Overall means:

|  |  |  |  |
| --- | --- | --- | --- |
| volatile\_acidity | chlorides | density | alcohol |
| 0.339666 | 0.05603386 | 0.99469663 | 10.49180083 |

Quality\_group = 0

1. (5 points). When the attributes are at their overall 25th percentiles, what will be the predicted category for quality\_group? List the attributes’ overall 25th percentiles with your answer.

['volatile\_acidity', 'chlorides', 'density', 'alcohol']

Overall means:

|  |  |  |  |
| --- | --- | --- | --- |
| volatile\_acidity | chlorides | density | alcohol |
| 0.23 | 0.038 | 0.99234 | 9.5 |

Quality\_group = 0

1. (5 points). When the attributes are at their overall 75th percentiles, what will be the predicted category for quality\_group? List the attributes’ overall 75th percentiles with your answer.

['volatile\_acidity', 'chlorides', 'density', 'alcohol']

Overall means:

|  |  |  |  |
| --- | --- | --- | --- |
| volatile\_acidity | chlorides | density | alcohol |
| 0.4 | 0.065 | 0.99699 | 11.3 |

Quality\_group = 0