CS 584-04: Machine Learning

Fall 2018 Assignment 5

# INSTRUCTIONS

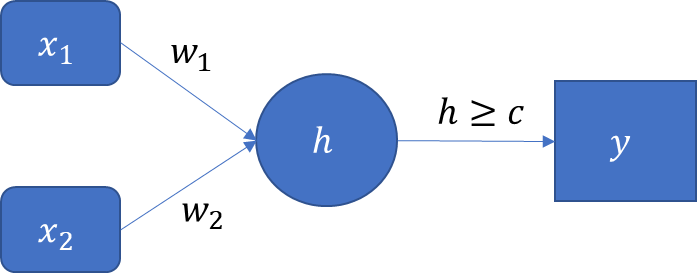
1. Students should complete this assignment independently.
2. Students must submit their answers to Blackboard before 11:59 PM on November 18, 2018.

# Question 1 (40 points)

Logical operators (i.e., NOT, AND, OR, XOR, etc.) are the building blocks of any computational device. Logical functions return only two possible values, TRUE or FALSE, based on the truth or false values of their input values. For example, the operator AND returns TRUE only when all the input values are TRUE. If at least one of the input values is FALSE, then it returns FALSE. If we denote TRUE by 1 and FALSE by 0, then the logical AND function can be represented by the following table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 0 | 1 | 1 |
|  | 0 | 1 | 0 | 1 |
| AND | 0 | 0 | 0 | 1 |

This function can be implemented by a perceptron with two binary inputs:



The activation function for the Output layer has this form: if . Otherwise, .

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.



X1



Y



X2

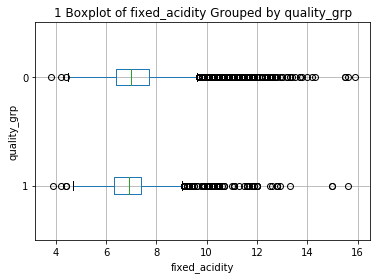


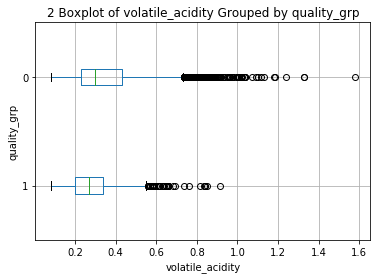
w1 = 1, w2 = 1 w3 = -1, w4 = -1, w5 =1 w6 =1

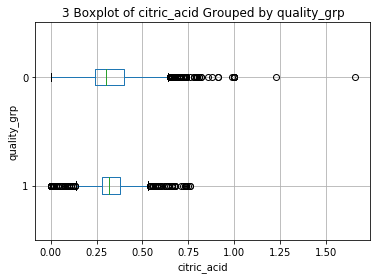
# Question 2 (60 points)

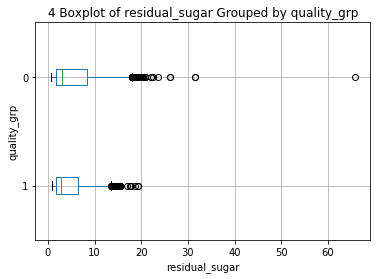
The Center for Machine Learning and Intelligent Systems at the University of California, Irvine manages the Machine Learning Repository (<https://archive.ics.uci.edu/ml/index.php>). We are going to analyze the Wine Quality dataset which is made available in the WineQuality.csv. The target variable is quality\_grp which is binary. Its values are 0 or 1. The input attributes are: (1) fixed\_acidity, (2) volatile\_acidity, (3) citric\_acid, (4) residual\_sugar, (5) chlorides, (6) free\_sulfur\_dioxide, (7) total\_sulfur\_dioxide, (8) density, (9) pH, (10) sulphates, and (11) alcohol.

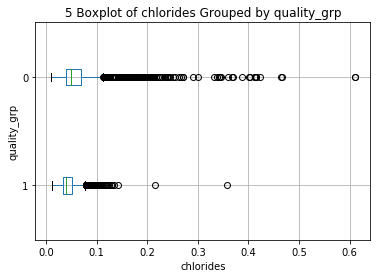
1. (10 points). Generate a horizontal box-plot for each input attribute, grouped by the target variable quality\_grp.

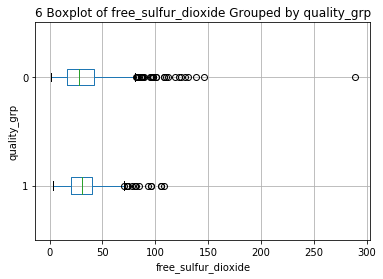


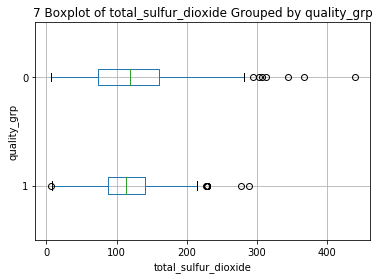


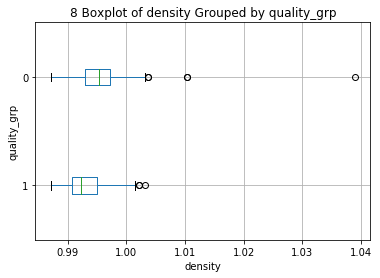


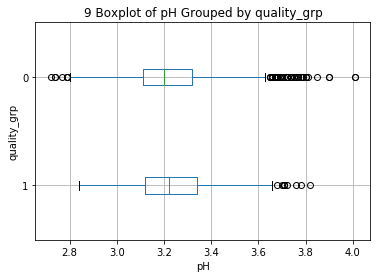


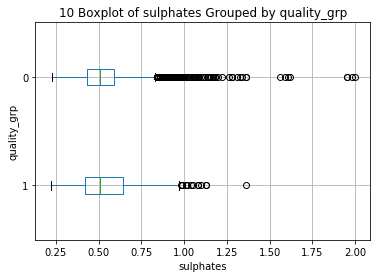


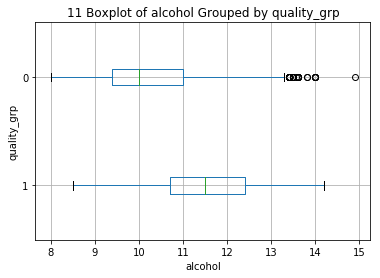




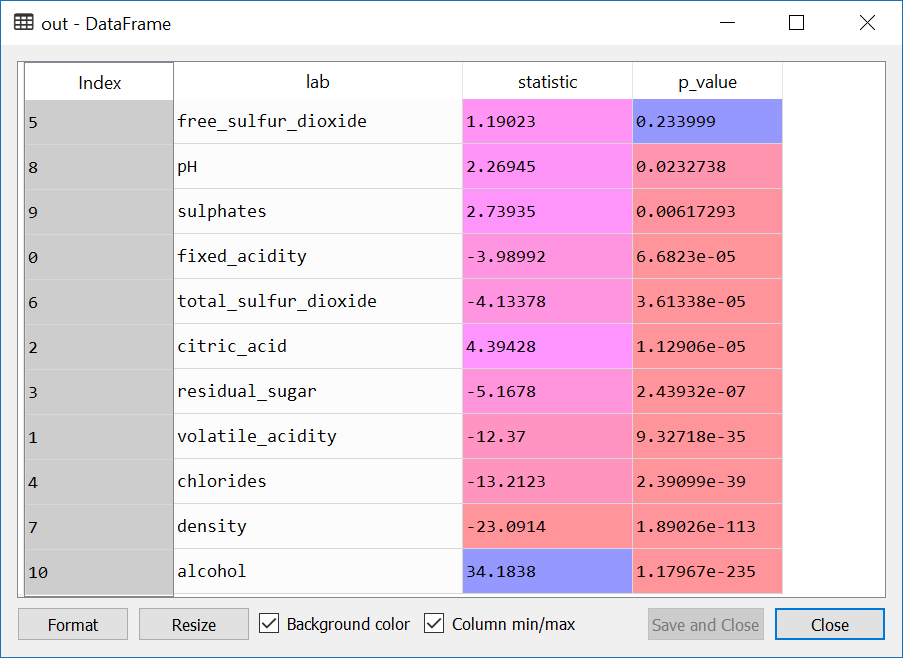








1. (10 points). The scipy.stats module has the ttest\_ind function for comparing two independent samples using the Student’s *t* test. Use this function to calculate the two-sided *p*-value of the Student’s *t* test. The group variable is the target variable quality\_grp. List the names of the input 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?

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

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

Mean Accuracy = 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).

Intercept:

[-2.11236983]

Weight Coefficients:

[[-0.83531774 -0.65083327 -0.9685864 0.26397236]]

x1, x2, x3, x4 values are from the columns of [volatile\_acidity chlorides density 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']

[ 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']

[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']

[ 0.4 0.065 0.99699 11.3 ]

quality\_group [0]