PRAGYA VERMA – (101276204)

SEIS 763: 01 Machine Learning

Assignment #1

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#1 - Read in the CSV file “ML\_HW\_Data\_FisherIris.csv” into a matrix named as “Iris”. Please do NOT output the whole matrix in our answer.  
import pandas as pd  
import numpy as np

Iris = pd.read\_csv("C:/Users/pragy/Downloads/ML\_HW\_Data\_FisherIris.csv",header=None, delimiter=",")

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#2 - Display total number of rows and total number of columns of the matrix “Iris”.  
  
row,columns = Iris.shape  
  
print("Rows:", row)  
print("Columns:", columns)

Ans:

Rows: 150

Columns: 5

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#3. Display all the row numbers (i.e. record numbers) that have the 5th column less than 0.

Row\_Numbers = Iris[Iris[4] <0].index  
print("Row Numbers:", Row\_Numbers)

Ans:

Row Numbers: Index([10, 23, 58, 89, 108, 136], dtype='int64')  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#4. Remove the rows with the 5th column less than 0 from the “Iris” matrix. Please do NOT output the whole resulting matrix in our answer.

Iris = Iris.drop(Iris[Iris[4]<0].index)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#5. Display total number of rows and total number of columns of the “Iris” matrix again.  
  
row,columns = Iris.shape  
  
print("Row:", row)  
print("Columns:", columns)

Ans:

Row: 144

Columns: 5

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#6. Copy the first 4 columns in the new “Iris” matrix into a new matrix “X”. Please do NOT output the whole resulting matrix in our answer

X = Iris.iloc[:, :4]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#7. Copy the 5th columns in the new “Iris” matrix into a new variable (or matrix) “Y”. Please do NOT output the whole resulting matrix in our answer.  
  
Y = Iris.iloc[:, 4:]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#8. Display the maximum value and the minimum value of EACH column in “X”.  
  
Xminmax = X.agg([min, max])  
print(Xminmax)

Ans:

0 1 2 3

min 1 10 20 43

max 25 69 44 79

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#9. Display total number of elements (i.e. items) in the third column of the matrix “X” that are greater than 36.  
  
totalitems = len(X[(X[2]>36)])  
print("Total no of Elements in the third column > 36 = ", totalitems)

Total no of Elements in the third column > 36 = 15