Machine Learning (Assignment # 1)
Deadline: Thursday (11:59 pm), September 1, 2022

Note: *Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. It is an individual assignment.*

Instructions

- Submit your source code to GitHub and represent the work through word document properly (provide your screenshots as well. The screenshot should have both the code and the output). Briefly describe the execution of your code. Comment your code.
- Make Video and provide its link in word document (2 – 3 min video showing the demo of the assignment, with brief voice over on the code and logic explanation. Make sure the video is public and accessible, otherwise assignment would be graded partially)

Question 1

The following is a list of 10 students ages:

ages = [19, 22, 19, 24, 20, 25, 26, 24, 25, 24]

- Sort the list and find the min and max age
- Add the min age and the max age again to the list
- Find the median age (one middle item or two middle items divided by two)
- Find the average age (sum of all items divided by their number)
- Find the range of the ages (max minus min)

Question 2

- Create an empty dictionary called dog
- Add name, color, breed, legs, age to the dog dictionary
- Create a student dictionary and add first_name, last_name, gender, age, marital status, skills, country, city and address as keys for the dictionary
- Get the length of the student dictionary
- Get the value of skills and check the data type, it should be a list
- Modify the skills values by adding one or two skills
- Get the dictionary keys as a list
- Get the dictionary values as a list

Question 3

- Create a tuple containing names of your sisters and your brothers (imaginary siblings are fine)
- Join brothers and sisters tuples and assign it to siblings
- How many siblings do you have?

- Modify the siblings tuple and add the name of your father and mother and assign it to family_members

Question 4

it_companies = {'Facebook', 'Google', 'Microsoft', 'Apple', 'IBM', 'Oracle', 'Amazon'}

A = {19, 22, 24, 20, 25, 26}

B = {19, 22, 20, 25, 26, 24, 28, 27}

age = [22, 19, 24, 25, 26, 24, 25, 24]

- Find the length of the set it_companies
- Add 'Twitter' to it_companies
- Insert multiple IT companies at once to the set it_companies
- Remove one of the companies from the set it_companies
- What is the difference between remove and discard
- Join A and B
- Find A intersection B
- Is A subset of B
- Are A and B disjoint sets
- Join A with B and B with A
- What is the symmetric difference between A and B
- Delete the sets completely
- Convert the ages to a set and compare the length of the list and the set.

Question 5

The radius of a circle is 30 meters.

- Calculate the area of a circle and assign the value to a variable name of _area_of_circle_
- Calculate the circumference of a circle and assign the value to a variable name of _circum_of_circle_
- Take radius as user input and calculate the area.

Question 6

“I am a teacher and I love to inspire and teach people”

- How many unique words have been used in the sentence? Use the split methods and set to get the unique words.

Question 7

Use a tab escape sequence to get the following lines.

Name Age Country City

Asabeneh 250 Finland Helsinki

Question 8

Use the string formatting method to display the following:

radius = 10

area = 3.14 * radius ** 2

"The area of a circle with radius 10 is 314 meters square."

Question 9

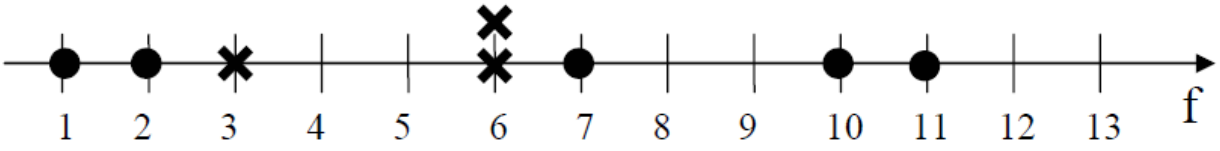
Write a program, which reads weights (lbs.) of N students into a list and convert these weights to kilograms in a separate list using Loop. N: No of students (Read input from user)

Ex: L1: [150, 155, 145, 148]

Output: [68.03, 70.3, 65.77, 67.13]

Question 10

The diagram below shows a dataset with 2 classes and 8 data points, each with only one feature value, labeled f. Note that there are two data points with the same feature value of 6. These are shown as two x's one above the other.



1. Divide this data equally into two parts. Use first part as training and second part as testing. Using KNN classifier, for K=3, what would be the predicted outputs for the test samples? Show how you arrived at your answer.
2. Compute the confusion matrix for this and calculate accuracy, sensitivity and specificity values.