Mid-term examination
BSCS 6A
AI Lab – Noor Saqib

CSCL4101- AI Lab

I. Using the KNN algorithm on the given dataset('cancer_prediction_data.csv) perform the following steps: (12 Marks)

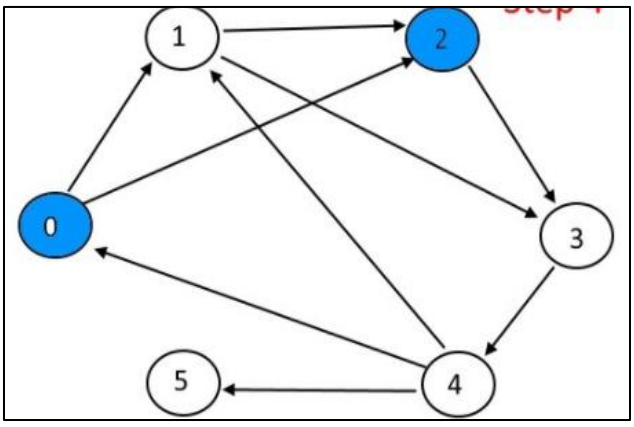
a. Load the dataset using PANDAS. (1 mark)
b. Write a code to view the first five rows of the imported dataset. (1 mark)
c. Break the dataset into attributes i.e (4 marks)
   i. 'Clump_thickness'
   ii. 'Uniformity_Cell_Size'
   iii. 'Uniformity_Cell_Shape'
   iv. 'Marginal_Adhesion'
   v. 'Single_Epithelial_Cell_Size'
   vi. 'Bland_Chromatin'
   vii. 'Normal_Nucleoli'
   viii. 'Mitoses',

   and label i.e., 'Class'.
d. Write down the code to only print the attributes. (1 mark)
e. Write down the code to print the label i.e., 'Class' only. (1 mark)
f. Break the dataset into training and testing data while keeping the testing data as 30%. (1 mark)
g. Keep K=3 when applying the KNeighborsClassifier. (1 mark)
h. Predict the values of the label of the testing data. (1 mark)
i. Print the predicted values. (1 mark)

II. Calculate the Euclidean distance between the following points using the present libraries and print the distance matrix: (2 marks)

point1 = [-43, 172];
point2 = [-44, 171]

III. Write the Python code for Breadth First Search for the following tree: (5 marks)



IV. File Handling:

Write the code for the following: (6 marks)

1. Open/create a file named 'yourname-registrationnumber.csv' (1 mark)
2. Insert the following values in the same file: (3 marks)

C_ID	CustomerName	ContactName	Address	City	Postal Code
89	White Clover Markets	Karl Jablonski	305 - 14th Ave. S. Suite 3B	Seattle	98128
90	Wilman Kala	Matti Karttunen	Keskuskatu 45	Helsinki	21240
91	Wolski	Zbyszek	ul. Filtrowa 68	Walla	01-012

3. Write the code to display the values present in the same file. (2 marks)