

CSE7212c – Python CUTe

Section I: Python

(Total Marks: 25)

1. Given a string for example: “This is the first python test”, get the list of vowels and consonants separately. (Marks: 2)
2. Given an array, get the kth largest element in the array. Write a function to get this. The inputs to the function are array and the kth value.

Hint: Use list.sort()

Example:

Input arr=[2,4,12,8,6,9,10], k=2

Output: 10

(Marks: 2)

3. There is an array in which the elements are repeated twice and one element is repeated thrice. Using NumPy, find that repeated element.

Example:

Input arr = [2,4,6,2,6,4,8,9,7,7,8,9,8]

Output = 8

(Marks: 3)

4. i) Create Numpy arrays for the following points: A(1,2,3) and B(2,4,5).
ii) Multiply the two arrays (element-wise multiplication)
iii) Add the two arrays and divide them by 2. Call it as C (element-wise addition)
iv) Find the distance between the C and the other two points:

Distance between two points $P(x_1, y_1)$ and $Q(x_2, y_2)$ is given by:

$$d(P, Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad \{\text{Distance formula}\}$$

(HINT: np.sum(np.square(A-B)))

(Marks: 6)

5. Using the list comprehensions, get the list of squares of numbers in the range of 10 and find their sum. (Marks: 2)

Section II: Pandas

Read the ‘bank.csv’ file as DataFrame with Pandas and answer the following questions: (Marks: 10)

- a. What is the most common occupation in the ‘job’ column?
- b. What is the mean ‘balance’ for people with marital status as married and also education as tertiary?
- c. What is the mean ‘age’ for different ‘education’ categories?
- d. Drop the following columns from the dataframe: job, day, month, pdays, previous and poutcome.
- e. Recode categorical variables with 2 levels as 0 & 1.