Dr. Jagpreet

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
TEST-1 EXAMINATION (FEB 2020)
B-Tech (8th SEM)

Course Code: 11B1WCI832 Max. Marks: 15

Course Name: INFORMATION RETERIVAL AND DATA MINING Max. Time: 1 HRS

Course Credit: 3

Note: All questions are compulsory

Q. No. 1 Proximity is typically defined between a pair of objects. **[1 * 3] [CO-1]**

(a) Define two ways in which you might define the proximity among a group of objects.

(b) How might you define the distance between two sets of points in Euclidean space?

(c) How might you define the proximity between two sets of data objects? (Make no assumption about the data objects, except that a proximity measure is defined between any pair of objects.)

Q. No. 2 For the following vectors, x and y, calculate the indicated similarity or distance measures. **[2 * 3] [CO-1]**

(a) x: (0,- 1,0, 1) , y: (1,0,- 1,0) Cosine, Euclidean

(b) x : (1,1 ,0,1 ,0,1) , y : (1,1 ,1 ,0,0,1) Correlation and Jaccard.

(c) x : (2, -7,0,2,0, -3) , y : (-1, 1,- 1,0,0, -1) Cosine, Correlation.

Q. No. 3 Describe how you would create visualizations to display information that describes the following types of systems. **[2 * 3] [CO-2]**

(a) Computer networks. Be sure to include both the static aspects of the network, such as connectivity, and the dynamic aspects, such as traffic.

(b) The distribution of specific plant and animal species around the world for a specific moment in time.

(c) The change in occupation of workers in a particular country over the last thirty years. Assume that you have yearly information about each person that also includes gender and level of education.

Be sure to address the following issues:

- Representation. How will you map objects, attributes, and relationships to visual elements?
- Arrangement. Are there any special considerations that need to be taken into account with respect to how visual elements are displayed? Specific examples might be the choice of viewpoint, the use of transparency, or the separation of certain groups of objects.
- Selection. How will you handle a large number of attributes and data objects?