Dr. Jagpreet Sidhu.

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
TEST-1 EXAMINATION (FEB 2019)
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 Classify the following attributes as binary, discrete, or continuous. Also classify them as qualitative (nominal or ordinal) or quantitative (interval or ratio). Some cases may have more than one interpretation, so briefly indicate your reasoning if you think there may be some ambiguity. **[.5 *10]** **[CO-1]**

Example: Age in years. Answer: Discrete, quantitative, ratio

1. Brightness as measured by a light meter.
2. Brightness as measured by people's judgments.
3. Angles as measured in degrees between 0 and 360.
4. Bronze, Silver, and Gold medals as awarded at the Olympics.
5. Number of patients in a hospital.
6. ISBN numbers for books. (Look up the format on the Web.)
7. Ability to pass light in terms of the following values: opaque, translucent, transparent.
8. Distance from the center of campus.
9. Density of a substance in grams per cubic centimetre.
10. Coat check number. (When you attend an event, you can often give your coat to someone who, in turn, gives you a number that you can use to claim your coat when you leave.)

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

(a) x : (1, 1, 1, 1), y : (2,2,2,2) 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 (a) How might you address the problem that a histogram depends on the number and location of the bins, Illustrate the concept with suitable data and figures? **[2+2]** **[CO-2]**

(b) Discuss the differences between dimensionality reductions based on aggregation and dimensionality reduction based on techniques such as PCA?