JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST-1 EXAMINATION- February, 2023

B.Tech. (βI) VI Semester

COURSE CODE: 18B1WBI632(3)   MAX. MARKS: 15

COURSE NAME: Data Warehousing and Mining for Bioinformatics

COURSE INSTRUCTORS: Dr. Ekta Gandotra   MAX. TIME: 1 Hour

Note: All questions are compulsory. Marks are indicated against each question in square brackets.

Q1.   Describe the steps involved in data mining when viewed as a process of knowledge   [4]   CO1
       discovery.

Q2.   a. Briefly outline how to compute the dissimilarity between objects described by   [2]   CO3
       nominal attributes.

       b. For the following vectors, x and y, calculate the indicated similarity or distance   [3]   CO3
       measures.

       i. $x = (0, 1, 1, 1)$, $y = (1, 0, 1, 0)$ Cosine, Correlation, Jaccard

       ii. $x=(20, 1, 34, 9)$, $y= (17, 0, 30, 8)$ Euclidean, Manhattan, Supremum

Q3.   Consider the following data for an attribute "**Age**":   [3]   CO3

       13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40,
       45, 46, 52, 70.

       a.   Give five number summary for the given data.

       b.   Use smoothing by bin boundary to smooth the above data, using a bin depth of 4.

Q4.   Consider the following table representing the rate of economic growth (X) and the rate   [3]   CO1
       of return (Y). Using the covariance formula, determine whether economic growth and
       returns have a positive or inverse relationship.

Economic Growth % (X)	Returns % (Y)
2.1	8
2.5	12
4.0	14
3.6	10