

INDIAN STATISTICAL INSTITUTE  
THIRD SEMESTER EXAMINATION, 2018/20 SESSION  
**PAPER - 13: INFORMATION STORAGE, RETRIEVAL AND DBMS**  
(MASTER IN LIBRARY INFORMATION SCIENCE)

20<sup>th</sup> November 2019 (10:00-13:00) (3 Hours)

This Question paper consists of 2 pages. **Attempt Questions and/or Sub-Questions to score maximum of 100 marks.** Please print all your answers in the Answer Booklet provided. Scientific Calculator is allowed.

**QUESTION 1.** Write briefly about

- |   |           |
|---|-----------|
| (i) Advanced Data Storage Devices                               | [4 marks] |
| (ii) K-Means Clustering   | [5 marks] |
| (iii) Information Retrieval in Library Information Science      | [4 marks] |
| (iv) Inverse Document Frequency (IDF)                           | [4 marks] |
| (v) Delimiter Space and it's geometric significance             | [4 marks] |
| (vi) Information dimension computation via multifractal spectra | [4 marks] |

**QUESTION 2.**

A. How to binarize grayscale image data? What is the role of image histogram in the image thresholding? [4 marks]

B. Figure 1 is a four-bit greyscale image,  $f(x,y)$ , of size 7 x 7 pixels, depicting greyscale values of a picture in a document. These greyscale values range between 1 and 16 (4-bit image).

1	2	7	6	5	7	13
13	11	9	15	10	16	6
12	11	9	14	9	12	12
4	6	14	7	11	7	9
1	2	7	6	14	7	6
6	9	5	3	7	2	6
8	7	9	10	14	11	13

Figure 1

Convert the data shown in Figure 1 into binary form by choosing a median value as the threshold value [6 marks]

C. Compute  $\sum_{x,y} f(x,y)$ , and plot a histogram for the above 7 X 7 multi-valued image

[6 marks]

**QUESTION 3.** Compute the Moran's Index for the following unstructured data. Explain the interpretation of the index from the perspective of Information Science. [12 marks]

	+15	+9
+5	-5	-4
	+8	+12

Figure 2

**QUESTION 4.** First pages of articles appearing in technical periodicals possess unique geometry of delimiter spaces. Rectangular Granulometries is a technique to quantify the geometric complexity of such a delimiter space. Let delimiter-space ( $M$ ) and text ( $M^c$ ) represent respectively foreground (in white shade) and background (in black shade). Explain how the geometry of  $M$  could be quantitatively characterized via rectangular granulometries. Give full set of equations involved in document summarization. **[12 marks]**

**QUESTION 5.** How to automatically summarize the documents via probabilistic approach based K-Mixture Model. Explain a scenario with the support of equations. **[10 marks]**

**QUESTION 6.** Create data of your own choice in two groups, and two variables in each group. Demonstrate the computation of Mahalanobis Distance that measures the separation between the two groups. Write your conclusions about the relevance of Mahalanobis Distance in the context of information retrieval. **[12 marks]**

**QUESTION 7.** Explain under what situation one employs the following morphological interpolations? **[9 marks]**

- a. Skeletonization by Influence Zones (SKIZ), and Weighted SKIZ
- b. Binary Morphological Median
- c. Grayscale Morphological Median Function

Explain the above three morphology-based interpolations with the support of illustrations and equations. **[12 marks]**

**QUESTION 8.** Let  $f^1$ ,  $f^2$ , and  $f^3$  be three spatial fields with the similar size configurations. Write a simple morphology-based algorithm to compute the ranks for pairing the three spatial fields. Give an example of your choice to explain illustratively with support of equations. **[12 marks]**

**END OF PAPER**