Roll No. | | | | | | | | | | | | | Total No. of Pages : 01

Total No. of Questions : 08

M.Tech.(CSE Engg.) EI-I (2015 to 2017) (Sem.–3)

INFORMATION RETRIEVAL

Subject Code : MTCS-305

M.Code : 74162

Time : 3 Hrs. Max. Marks : 100

INSTRUCTION TO CANDIDATES :

1. **Attempt any FIVE questions out of EIGHT questions.**
2. **Each question carries TWENTY marks.**

1) Discuss index construction. What are the constraints that govern the design of indexing algorithms? (20)

2) Discuss the role of clustering in information retrieval. State its applications. Explain different criteria of clustering quality. (20)

3) What is content based ranking? Explain **tf-idf** weighting with a suitable example. Why is the **idf** of a term always finite? (20)

4) Consider the vector-space representation of documents and compare the cosine distance to the ordinary Euclidean distance. Show that for vectors of unit length the ranking induced by the two distances is the same. (20)

5) i) How do you select an information resource? What are different reference database types? What are different ways to search for information in a database? (10)

ii) How do you build your search? Discuss the ways to refine a search based on the initial results. And how the search results are evaluated? (10)

6) What are different challenges in XML retrieval? Discuss text central *vs.* data centric XML retrieval. State evaluation of XML retrieval. (20)

7) What is sentiment analysis? What are its applications? State different levels of sentiment analysis investigation. (20)

8) Explain in detail sentiment classification using any supervised learning technique like e.g., naïve Bayes classification, and support vector machines (SVM), etc. Write the formula where needed. (20)

NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student..