Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–VII (NEW) EXAMINATION – WINTER 2023

Subject Code:3170718 Date:08-12-2023

Subject Name: Information Retrieval

Time: 10:30 AM TO 01:00 PM Total Marks:70

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q.1	**(a)**	Explain the nature of unstructured and semi-structured text.	**03**
	(b)	What is inverted index and Boolean queries in IR?	**04**
	(c)	What is Index compression? Explain lexicon compression and postings lists compression.	**07**
Q.2	**(a)**	Describe Zipf's Law.	**03**
	(b)	Differentiate between dynamic indexing and positional indexes.	**04**
	(c)	Discuss the Boolean retrieval in detail with diagram.	**07**
		OR	
	(c)	What is text encoding? Explain tokenization and stemming with example.	**07**
Q.3	**(a)**	Explain n-gram indexes.	**03**
	(b)	Discuss the Postings size estimation.	**04**
	(c)	Describe how to perform document length normalization.	**07**
		OR	
Q.3	**(a)**	Explain Vector space scoring.	**03**
	(b)	Describe any two retrieval models in details.	**04**
	(c)	Describe relevance feedback and query expansion with example.	**07**
Q.4	**(a)**	What is the need of Filtering against spamming?	**03**
	(b)	Write Short notes on Support vector machine classifiers.	**04**
	(c)	What is k-means clustering? Explain with suitable example.	**07**
		OR	
Q.4	**(a)**	Define the terms: TFIDF, Okapi, and F-measure.	**03**
	(b)	Write Short notes on: kappa measure and interjudge agreement.	**04**
	(c)	Describe Vector space classification using hyper planes.	**07**
Q.5	**(a)**	What is web crawling?	**03**
	(b)	Describe Cross language information retrieval.	**04**
	(c)	Describe Hierarchical agglomerative clustering with suitable example.	**07**
		OR	
Q.5	**(a)**	Define the terms Page Rank, HITS and Boosting.	**03**
	(b)	Which is the fastest search engine for retrieval of information? Justify.	**04**
	(c)	Explain XML retrieval? Also write down the major challenges in XML retrieval?	**07**

****