Seat No.: _____ **Enrolment No._____**

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–VI (NEW) - EXAMINATION – SUMMER 2018

Subject Code:2161603 Date:03/05/2018

Subject Name:Data Compression and data Retrival

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.**

Q.1	**(a)**	Define: self-information, entropy, lossy compression	**03**
	(b)	State the models for lossless compression and explain any one in	**04**
	(c)	Explain update procedure of adaptive huffman codding with suitable example.	**07**
Q.2	**(a)**	Given an alphabet A = {a_1, a_2, a_3, a_4}, find the first order entropy in the following case: P(a_1) = 0.505, P(a_2) = ¼, P(a_3) = 1/8 and P(a_4) =0.12.	**03**
	(b)	Determine the minimum variance Huffman code with the given probabilities. P(a_1) = 0.2, P(a_2) = 0.4, P(a_3) = 0.2, P(a_4) = 0.1 and P(a_5) = 0.1	**04**
	(c)	The probability model is given by P(a_1) = 0.2, P(a_2) = 0.3 and P(a_3) = 0.5. Find the real valued tag for the sequence a_1a_1 a_3 a_2 a_3 a_1. (Assume cumulative probability function: F(0) = 0)	**07**
		OR	
	(c)	Explain Tunstall codes with suitable example.	**07**
Q.3	**(a)**	Determine whether the following codes are uniquely decodable or not. 1. {0, 01, 11, 111} 2. {0, 10, 110, 111}	**03**
	(b)	Explain Linde-Buzo-Gray algorithm in detail.	**04**
	(c)	Explain stemming and lemmatization with suitable example.	**07**
		OR	
Q.3	**(a)**	Generate the golomb code for m = 9 and n = 8 & 9.	**03**
	(b)	Explain incident matrix and inverted index with suitable example.	**04**
	(c)	Write pseudocode for integer arithmetic encoding and decoding algorithm.	**07**
Q.4	**(a)**	Explain nonuniform quantization.	**03**
	(b)	Explain skip pointer with suitable example.	**04**
	(c)	Explain LZ77 with suitable example.	**07**
		OR	
Q.4	**(a)**	Explain pdf optimized quantization.	**03**
	(b)	Explain phrase queries with suitable example.	**04**
	(c)	Explain LZ78 with suitable example.	**07**
Q.5	**(a)**	Explain structured vector quantizers.	**03**
	(b)	Explain challenges in XML information retrieval.	**04**
	(c)	Explain usage of discrete cosine transform (DCT) in JPEG.	**07**
		OR	
Q.5	**(a)**	Explain pyramid vector quantization.	**03**
	(b)	Explain tokenization in detail.	**04**
	(c)	Explain audio compression technique with suitable diagram.	**07**

****