Indian Institute of Technology Kharagpur
Class Test II 2025-26

Date of Examination: _____ Feb, 2026 _____   Duration: 45 Minutes

Subject No.: CS60092   Subject: Information Retrieval

Department/Center/School: Computer Science   Credits: 3   Full marks: 30

Name: _____

Roll Number: _____

Question:	1	2	3	Total
Points:	10	10	10	30
Score:				

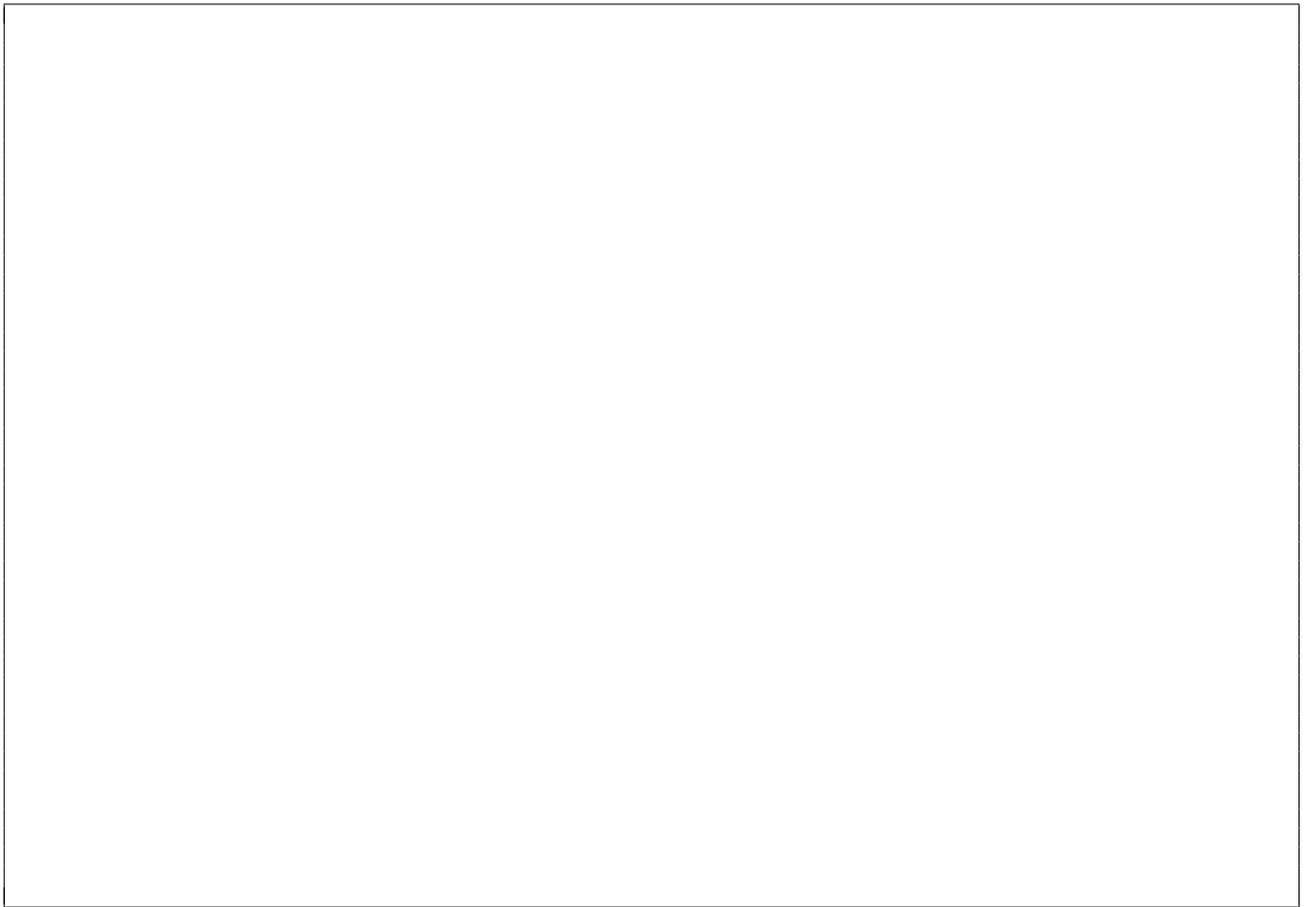**Write the answers in the space provided after the questions only.**

1. (a) (10 points) Consider the following sorted dictionary with 20 bytes allocated per term. Use blocking with block size 5 followed by front coding to compression it. What is the compression ratio?

   `automobile, automatic, automation, autonomy, autumn`

2. (10 points) Write the formula for document length normalization in the BM25 model. Calculate the normalized term frequency for a term with (unnormalized) term frequency as x in document D3. The document lengths are:

   `D1:  120, D2:  20, D3:  30, D4:  40, D5:  20`

   Consider $k_1 = 1.2$ and $b = 0.75$

3. (10 points) Which parameter from the BIM can be evaluated through relevance feedback. Write the formula for it, clearly mention all inputs.

Rough space