

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2024

M.Tech-I Semester (CSE/IT/ECE/CE/BT/BI)

COURSE CODE (CREDITS): L-11M1WCI431

MAX. MARKS: 35

COURSE NAME: Advanced Web Mining

COURSE INSTRUCTORS: Dr. Nishant Sharma

MAX. TIME: 2 Hours

Note: (a) All questions are compulsory.

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems.

Q.No	Question	CO	Marks
Q1	<p>(a) Explain different compression schemes. Consider an unambiguous lossless compression scheme for 2-bit binary numbers. It is claimed that this scheme can reduce size of any 2-bit numbers by atleast 1 bit. Prove that the above mentioned scheme never uses less space than uncompressed encoding.</p> <p>(b) For the above scheme, find a sequence of numbers that take more space when it is compressed as compared to uncompressed.</p> <p>(c) What information do we need to know about the data that is to be compressed? Why is this information required?</p>	[CO-3], [CO-4]	[4+3+2]
Q2	<p>Explain following terms in detail:</p> <p>(a) Query-based stemming</p> <p>(b) Noisy channel model</p> <p>(c) Relevance feedback</p> <p>(d) Query based local search</p>	[CO-4]	[2.5*4]
Q3.	Describe similarities and differences between a social media tag and anchor text? Discuss how a nearest neighbor approach can be applied for a collaborative filtering algorithm? Discuss meaning and significance of mean squared error, Euclidean distance and correlation similarity.	[CO-5]	[2+4+3]
Q4.	<p>Discuss below-mentioned optimized techniques for query processing:</p> <p>(a) List skipping</p> <p>(b) Conjunctive processing</p> <p>(c) MaxScore</p> <p>(d) Early termination</p> <p>(e) Threshold methods</p>	[CO-5]	[1.5*4 + 1]