Information Retrieval (CS60092) – Class Test

Date: 24 March, 2026

Name: _____ Roll no.: _____

Each question carries 1 mark. Mark only 1 option. If you mark more than one, it will not be evaluated.

1. Why can both PageRank and TextRank be modeled using a Markov process over a graph? A. They rely on term frequency B. They use supervised learning C. They compute stationary distributions D. They require labeled data	**2. A page has high PageRank but is irrelevant to a query. What explains this?** A. Overfitting B. Query-independence C. Graph sparsity D. Token mismatch
3. Why does LSI struggle with polysemy? A. It is nonlinear B. It ignores dimensionality C. It is linear D. It uses embeddings	**4. When is Learning-to-Rank most useful?** A. No data B. Combining multiple signals C. Small corpora D. Static ranking
5. High BLEU but low ROUGE indicates: A. High recall B. Low precision C. Low coverage D. Perfect summary	**6. Why does web-scale data improve NLP models?** A. Removes bias B. Stabilizes probabilities C. Reduces size D. Eliminates noise
7. Why add CRF on BiLSTM in NER? A. Faster training B. Global consistency C. Better embeddings D. Smaller model	**8. Why are span-based models preferred in DyGIE++?** A. Faster B. Joint modeling C. Less data D. Simpler
9. Why does standard RAG fail multi-hop reasoning? A. No embeddings B. Random retrieval C. No relational reasoning D. Slow inference	**10. Why is BLEU weak for summarization?** A. Too slow B. Precision-only C. Requires labels D. Needs graphs
11. Why does CRF decoding use Viterbi? A. Faster embeddings B. Optimal sequence C. Less memory D. Parallelization	**12. What is a key limitation of OpenIE?** A. Needs schema B. Noisy relations C. Slow training D. Requires labels
13. Why do neural rankers outperform classical ones? A. More features B. Learn representations C. Less data D. Simpler models	**14. Why is web data biased?** A. Too small B. Uniform distribution C. Overrepresentation D. Clean data