| **Faculty of Computing and Informatics**  **Department of Information Systems**  **University of Gondar** | |
| --- | --- |
| Course Code | MSIS6122 |
| Course Title | **Modern Information Storage and Retrieval** |
| Credit Hour | 3 |
| ECTS Credits | 7 |
| Contact Hours (per week) | Lecture – 2 Lab -2 |
| Pre-requisite | Information Storage and Retrieval basics |
| Year/ Semester | 2/1 |
| Status of Course | Compulsory |
| Course Objectives | * To familiarize students with the basic theories and principles of information storage and retrieval * To introduce modern concepts of information retrieval systems. * To acquaint students with the various indexing, matching, organizing and evaluating strategies developed for information retrieval (IR) systems * To enable students understand current research issues and trends in IR |
| Course Contents | 1. **Introduction to ISR**  * IR and the retrieval process * Basic structure of an IR system  1. **Text/Document Operations and Automatic Indexing**  * Index term selection (Luhn’s selection and Zipf’s law in IR) * Document pre-processing (Lexical analysis, Stop word Elimination, stemming) * Term extraction (Term weighting and similarity measures)  1. **Indexing Structures**  * Inverted files * Tries, Suffix Trees and Suffix Arrays * Signature files  1. **IR Models**  * A Formal Characterization of IR Models * Boolean model and Vector space mode * Probabilistic model  1. **Retrieval Evaluation**  * Evaluation of IR systems * Relevance judgement * Performance measures (Recall, Precision, etc.)  1. **Query Languages**  * Keyword-based queries (Boolean queries, weighted queries, etc.), * Pattern matching * Structural queries  1. **Query Operations**  * Document clustering strategies * Relevance feedback * Query expansion  1. **Advanced topics**  * Parallel and distributed query processing * Machine learning in IR * Cross – language in IR  1. **Current Research Issues in IR**   Research in IR (Multimedia Retrieval, Web Retrieval, Question answering. etc.) |
| Literature | **References:**   * Ricardo A. Baeza-Yates, Berthier Ribeiro-Neto, Modern Information Retrieval, ACM Press, 1999. * Salton, G. and McGill, M. J. Introduction to Modern Information Retrieval, McGraw-Hill Co., 1983. * Robert R. Korfhage, Information Storage and Retrieval, John Wiley & Sons, 1997. * C. J. Van Rijsbergen. Information retrieval. London: Butterworths, 1979.(available at: http://www.dcs.gla.ac.uk/~iain/keith/) * Information Retrieval: Data Structures & Algorithms by W. B. Frakes and R. Baeza-Yates (Eds.) (Prentice-Hall) 1992, ISBN 0-13-463837-9. * Spärck Jones, K. & Willett, P. (eds.) . *Readings in information retrieval*. San Francisco: Morgan Kaufmann, 1997.   **IR Journals**   * Journal of the American Society for Information Science and Technology (JASIST) * Information Processing and Management (IP & M) * Information Retrieval (IR) * Journal of Documentation (JDoc) * Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval * Literary Linguistics Computing (LLC) |