MANGALORE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE
MSc Computer Science

CSS 504: INFORMATION RETRIEVAL SYSTEMS		
Hours/Week: 4 **Credits : 4**		**I.A. Marks: 30** **Exam. Marks: 70**

Course Outcomes:

CO1: Understanding the basics of Information retrieval like what is a corpus, what is precision and recall of an IR system

CO2: Understanding the data structures like Inverted Indices used in Information retrieval systems

CO3: Understanding the basics of web search

CO4: Understanding the different techniques for compression of an index including the dictionary and its posting list

CO5: Understanding the different components of an Information retrieval system

CO6: Developing the ability of develop a complete IR system from scratch.

	UNIT-I	**12 Hrs.**

Boolean retrieval. The term vocabulary and postings lists. Dictionaries and tolerant retrieval. Index construction. Index compression.

	UNIT-II	12 Hrs.

XML retrieval. Probabilistic information retrieval. Language models for information retrieval. Text classification. Vector space classification.
Support vector machines and machine learning on documents. Flat clustering. Hierarchical clustering. Matrix decompositions and latent semantic indexing.

	UNIT-III	12 Hrs.

XML retrieval. Probabilistic information retrieval. Language models for information retrieval. Text classification. Vector space classification.
Support vector machines and machine learning on documents. Flat clustering. Hierarchical clustering. Matrix decompositions and latent semantic indexing.

	UNIT-IV	12 Hrs.

Web search basics. Web crawling and indexes. Link analysis.

REFERENCE BOOKS:
1. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütz, Introduction to Information Retrieval , Cambridge University Press, 2008.
2. Gerald J. Kowalski, Mark T. Maybury, Information Storage and Retrieval Systems: Theory and Implementation, Springer publication, 2008.
3. Ricardo Baeza-Yates, Modern Information Retrieval, Pearson Education, 2009.
4. David A Grossman and Ophir Frieder, Information Retrieval: Algorithms and Heuristics, 2nd Edition, Springer, 2004.
5. William B Frakes, Ricardo Baeza Yates, Information Retrieval Data Structures and Algorithms, Pearson Education, 1992.