OBJECTIVE

To use an open source search engine framework and explore its capabilities, represent documents in different ways and discuss its effect on similarity calculations and on search, modify Page Rank and HITS algorithms or Personalization, Semantic or any other aspect, design and implement an innovative feature in a search engine and explain the search components affected by the innovation, design a smart information management system with Information Retrieval components.

UNIT – I INTRODUCTION


UNIT – II INFORMATION RETRIEVAL


UNIT – III WEB SEARCH ENGINE – INTRODUCTION AND CRAWLING


UNIT – IV WEB SEARCH – LINK ANALYSIS AND SPECIALIZED SEARCH


UNIT – V DOCUMENT TEXT MINING

Information filtering; organization and relevance feedback – Text Mining - Text classification and clustering - Categorization algorithms: naive Bayes; decision trees; and nearest neighbor - Clustering algorithms: agglomerative clustering; k-means; expectation maximization (EM)

TOTAL: 45
TEXT BOOKS:


REFERENCES: