COURSE DESCRIPTION FORM

INSTITUTION National University of Computer and Emerging Sciences

PROGRAM (S) TO BE EVALUATED Computer Science

A. Course Description

(Fill out the following table for each course in your computer science curriculum. A filled-out form should not be more than 2-3 pages.)

Course Code	CS4051
Course Title	Information Retrieval
Credit Hours	3
Prerequisites by Course(s) and Topics	Data Structures, Programming
Assessment Instruments with Weights (homework, quizzes, midterms, final, programming assignments, lab work, etc.)	Assessment with the weight.

Assessment Type	**Weight**
Assignments	20
Quizzes	08
Mid-Terms	20 (10 each)
Project	12
Final	40

Course Coordinator	Dr. Muhammad Rafi
URL (if any)	
Current Catalog Description	Introduction to standard concepts in information retrieval, indexing and various retrieval methods, information retrieval evaluation, compression, metasearch, machine learning for IR, collaborative filtering clustering, spam filtering, news filtering, topic detection and tracking
Textbook (or **Laboratory Manual** for Laboratory Courses)	An Introduction to Information Retrieval By Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütz, Cambridge University Press Cambridge, England
Reference Material	

Course Goals

A. Course Learning Outcomes (CLOs)

1. Understand the basic concepts and techniques in Information Retrieval.
2. Understand how statistical models of text can be used to solve problems in IR, with a focus on how the vector-space model and language models are implemented and applied to document retrieval problems.
3. Understand how statistical models of text can be used for other IR applications, for example clustering, classification and extraction of information.
4. Appreciate the importance of data structures, such as an index, to allow efficient access to the information in large bodies of text, identify the storage and index requirements for different IR Models
5. Have experience of building a document retrieval system, through the practical sessions, including the implementation of a relevance feedback mechanism.
6. Understand the issues involved in providing an IR service on a web scale, including distributed index construction and user modelling for recommendation engines.

B. Program Learning Outcomes

For each attribute below, indicate whether this attribute is covered in this course or not. Leave the cell blank if the enablement is little or non-existent.

1. Academic Education:	To prepare graduates as computing professionals	✔
2. Knowledge for Solving Computing Problems:	Apply knowledge of computing fundamentals, knowledge of a computing specialization, and mathematics, science, and domain knowledge appropriate for the computing specialization to the abstraction and conceptualization of computing models from defined problems and requirements.	✔
3. Problem Analysis:	Identify, formulate, research literature, and solve complex computing problems reaching substantiated conclusions using fundamental principles of mathematics, computing sciences, and relevant domain disciplines.	✔
4. Design/ Development of Solutions:	Design and evaluate solutions for complex computing problems, and design and evaluate systems, components, or processes that meet specified needs with appropriate consideration for public health and safety, cultural, societal, and environmental considerations.	✔
5. Modern Tool Usage:	Create, select, adapt and apply appropriate techniques, resources, and modern computing tools to complex computing activities, with an understanding of the limitations.	✔

6. Individual and Team Work:	Function effectively as an individual and as a member or leader in diverse teams and in multi-disciplinary settings.	✔
7. Communication:	Communicate effectively with the computing community and with society at large about complex computing activities by being able to comprehend and write effective reports, design documentation, make effective presentations, and give and understand clear instructions.	
8. Computing Professionalism and Society:	Understand and assess societal, health, safety, legal, and cultural issues within local and global contexts, and the consequential responsibilities relevant to professional computing practice.	
9. Ethics:	Understand and commit to professional ethics, responsibilities, and norms of professional computing practice.	✔
10. Life-long Learning:	Recognize the need, and have the ability, to engage in independent learning for continual development as a computing professional.	✔

C. Relation between CLOs and PLOs
(CLO: Course Learning Outcome, PLOs: Program Learning Outcomes)

		PLOs									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CLOs	1	✔	✔		✔		✔				
	2	✔	✔	✔		✔	✔				
	3	✔	✔	✔		✔	✔				
	4	✔	✔		✔		✔				
	5	✔	✔		✔	✔	✔			✔	✔
	6	✔	✔	✔			✔				

Topics Covered in the Course, with Number of Lectures on Each Topic (assume 15-week instruction and one-hour lectures)

1. Topics to be covered:

List of Topics	No. of Weeks	Contact Hours	CLO
Introduction to IR course, IR Problem and its Components, Basic IR model – Boolean Information Retrieval, Extended Boolean Model, Example of Commercial Systems – WestLaw	**1**	**3**	**1,2**
Document Processing, Term vocabulary and positing lists, Stemming & Lemmatization, posting list processing via skip lists, Phrase Query, positional indexing, bi-word indexing, Combining different indexing techniques.	**1**	**3**	**1, 2,4**
Dictionary and Tolerant Retrieval, Search Structures for Dictionary, Wildcard queries, permuterm index, k-gram index, Spelling Correction, Edit Distance, Phonetic Correction	**1**	**3**	**2,4**
Index construction, single pass, in-memory, distributed indexing, dynamic indexing. & Heaps law, Zipf's law, dictionary indexing, fixed length and variable length coding.	**1**	**3**	**2,4**
Vector space Model	**1**	**3**	**2,5**
Midterm I	**1**	**1**	
Evaluation in IR	**1**	**3**	**3,5**
Relevance Feedback	**1**	**3**	**5**
Basic Web Search, Crawler and Indexing	**1**	**3**	**6**
Link Analysis	**1**	**3**	**4,6**
Midterm II	**1**	**1**	
Text Classification in Vector Space Model	**1**	**3**	**2**

	Text Clustering	**1**	**3**	**3**
	NEWS Processing / Filtering / Recommendation	**1**	**3**	**2,6**
	Question/Answer Processing	**1**	**3**	**6**
	Advanced Topic: Neural Information Retrieval	**1**	**3**	
	Total	**16**	**48**	

Laboratory Projects/Experiments Done in the Course				
Programming Assignments Done in the Course	Yes, In C++, Java, Python etc.			
Class Time Spent on (in credit hours)	**Theory**	**Problem Analysis**	**Solution Design**	**Social and Ethical Issues**
	2	0.5	0.5	0
Oral and Written Communications	Every student is required to submit at least __1___ written reports of typically _5____ pages and to make __1___ oral presentations of typically __10___ minute's duration. Include only material that is graded for grammar, spelling, style, and so forth, as well as for technical content, completeness, and accuracy.			

Instructor Name: Dr. Muhammad Rafi

Instructor Signature:

Dated: Spring 2023