

**CENTRAL UNIVERSITY OF KERALA  
DEPARTMENT OF COMPUTER SCIENCE  
M.Sc. COMPUTER SCIENCE**

ELECTIVE COURSE					
COURSE CODE	COURSE TITLE	CONTACT HRS/WEEK			CREDITS
		LEC	LAB	TUT	
CSC5006	Web Mining and Social Networking	2	2	1	4

Lec = Lecture, Tut = Tutorial, Lab = Practical

This is a participatory and problem solving **skill development course**.

Course Objective:

The objective of the course is to provide theoretical and practical aspects of techniques for data mining applied on Internet related data and social networking.

By completing this course, students will obtain the following course/learning outcomes:

1. Knowledge gained:
  - (i) Modeling of web content mining, web structure mining and web usage mining.
  - (ii) development of architecture and its related algorithms commonly used in web mining applications
2. Skill gained:
  - (iii) Skills in sentiment analysis, targeted marketing, linguistic forensics, topic/trend-detection-tracking and multi-document summarization
  - (iv) Skills to analyze the patterns involved in social media data
3. Competency gained:
  - (v) Solve practical web mining problems using tools and techniques

Prerequisites: Basic knowledge of data mining

Grading:

Lab implementation	– 10%
Assignment/Quiz/presentation	– 10%
Class Test	– 10%
Lab test	– 10%
Final Exam	– 60%

### CSC5006 – Web Mining and Social Networking

#### Module 1

Introduction: Data Mining and Web Mining, web Community and Social network Analysis. Theoretical Backgrounds: Web Data Model, Textual linkage and usage expressions, Similarity functions, Eigenvector, SVD, tensor expression and decomposition, Basic concepts of social networks.

#### Module 2

Web Mining: Web content mining: Vector space model, web search, feature enrichment of short texts, latent semantic indexing, automatic topic extraction from web documents, opinion search and opinion span. Web Linkage Mining: Web search and hyperlink, co-citation and bibliographic coupling, Page rank and HITS algorithm, web community discovery, web graph measurement and modelling, using link information for web page classification.

#### Module 3

Web usage mining: Modelling web usage interface using clustering, WUM using probabilistic latent semantic analysis, finding user access pattern, co-clustering analysis of weblogs using bipartite spectral projection approach, web usage mining applications.

#### Module 4

Extracting and analyzing web social networks: Extracting evaluation of web community from a series of web achieve, temporal analysis on semantic graph using three way tensor decomposition, analysis of communities and their evaluations in dynamic networking, Socio-Sense: A system for analyzing the societal behavior from web archive.

#### References:

1. Guandong Xu Yanchun Zhang Lin Li, *Web Mining and Social Networking*, Springer, 2011.
2. Aggarwal, Charu C, *Social network data analytics*, Springer, 2011.
3. Lee Giles, Mark Smith, *Advances in Social Network Mining and Analysis*, Springer 2008.
4. Bing Liu, *Web Data Mining*, Springer, 2011.