

## Detailed Syllabus

<b>Subject Code</b>	19M12CS111	<b>Semester odd</b>	<b>Semester Sixth Session 2019- 2020</b> <b>Month from July to December</b>
<b>Subject Name</b>	Web Intelligence		
<b>Credits</b>	3	<b>Contact Hours</b>	3
<b>Faculty (Names)</b>	<b>Coordinator(s)</b>	1. Dr. Anuja Arora	
	<b>Teacher(s) (Alphabetically)</b>	1. Anuja Arora    2. Neetu sardana	

### Course Outcomes:

At the completion of the course, students will be able to,

- Outline the various web technologies, methods, and models used to design an intelligent web. [understanding Level 2]
- Make use of web caching strategies at varied level: user, web server, and gateway server. [Apply Level 3]
- Analyze and Model the users' browsing behavior on web. [Apply and Analyse-3,4]
- Evaluate various Web content mining algorithms, Web language models and learning to rank models to handle complex Web. [Evaluate Level 5]
- Design and develop the computational intelligent web algorithms to handle complex real problems [Create Level 6].

Module No.	Subtitle of the Module	Topics in the module	No. of Lectures for the module
1.	Web Content Feature Engineering	Frequency Filter, POS Tag, Unigram, Ngram, Collocation, Levenstein Distance, KL-Divergence, T-Test	2
2	Keyterms Extraction Methods	TF-IDF, SGRank, SGRank-IDF, Single Rank,	3
3	Web Content Searching Techniques	Word-Word occurrence matrix; Probabilistic models: Bayes model, BM25 Ranking model; Word2Vec, CBoW, Skip Gram Model Link Based Search Algorithm , Power Iteration Method for ranking nodes on web, Handling Spider Traps and Dead ends, Topic Sensitive Page Ranking.	8
4	Ranking Algorithm	Point wise ranking, Pair wise Ranking, Listwise ranking,  Metrics for Learning to rank : CG, DCG, NDCG, P@K, MAP, AP	3
5	Web caching Algorithm	LRV, FIFO, LRU, Random, OPT	3

6	Matrix Factorization Techniques	Matrix decomposition, Eigenvalue decomposition, Non-Negative matrix factorization, Singular value decomposition, objective functions , UV decomposition, CUR decomposition	3
7	Tensor Factorization	Multidimensional Matrix factorization, Matricization, Tucker decomposition, High Order SVD, clustHOSVD, other methods	3
9	Spamming on Web	Term Spamming, Link Spamming, Combating Link Spam, Google Trust Rank, Spam Mass.	3
10	Advertising on the Web	Issues in On-Line Advertising, Off-line Vs On-Line Algorithms, Greedy Algorithms The Matching Problem and Adwords Problem, The Balance Algorithm	4
11	Web Usage Mining	Introducing Web Logs ,Data Cleaning and Preprocessing :Page view Identification , User Identification, Sessionization, Path Completion ,Data Integration , Cluster Analysis and Visitor Segmentation, Analysis and predictions of Sequential and Navigational Patterns Markov Chains, Modelling web browsing using All Kth Markov chains, Hidden Markov Model, Turney algorithm.	8
Total number of Lectures			40

#### Evaluation Criteria

Components	Maximum Marks
T1	20
T2	20
End Semester Examination	35
TA	25 (Attendance/ Class Assignments/Quiz/ Internal assessment & Mini-Project)
<b>Total</b>	<b>100</b>

**Project Based Learning:** Students will develop small size project in order to build an intelligent web concept in a group of 2-3. Basically, small size projects are given to students in form of assignments to provide solution out of topics discussed in the course. Understanding usage of appropriate methodology, then implementation of those selected methodology to handle real scenario intelligent web problem and evaluation of applied methodology using various performance measures is the prime concept to enhance students' knowledge towards intelligent web.

Recommended Reading material: Author(s), Title, Edition, Publisher, Year of Publication etc. ( Text books, Reference Books, Journals, Reports, Websites etc. in the IEEE format)

1.	Liu, Bing. Web data mining. Springer-Verlag Berlin Heidelberg, 2007.
----	--

2.	Soumen Chakrabarti,. Mining the Web: Discovering knowledge from hypertext data. Morgan Kaufmann, 2003.
3.	Scime, Anthony, ed. Web mining: applications and techniques. IGI Global, 2005.
4.	Hitzler, Pascal, Markus Krotzsch, and Sebastian Rudolph. Foundations of semantic web technologies. CRC Press, 2011.
5.	King, Andrew B. Website optimization. " O'Reilly Media, Inc.", 2008.
6.	Segaran, Toby. Programming collective intelligence: building smart web 2.0 applications. "O'Reilly Media, Inc.", 2007.
7.	Aggarwal Charu.C, Social Network Data Analytics, Springer Science+Business Media, LLC 2011
8.	Easley, David, Jon Kleinberg. Networks, Crowds, and Markets: Reasoning about a Highly Connected World. New York, NY: Cambridge University Press, 2010.
9.	Jackson Matthew O,. Social and Economic Networks. Princeton, NJ: Princeton University Press, 2008
10	Borgatti Stephon. P., Everett Martin G and Johnson Jeffery C , Analyzing Social Networks, Sage Publications, 2013