Detailed Syllabus

Subject Code	14M1NCI334	**Semester odd**	**Semester Sixth Session 2018- 2019** **Month from July to December**
Subject Name	Web Algorithms		
Credits	**3**	**Contact Hours**	**3**

Faculty (Names)	**Coordinator(s)**	1. Anuja Arora
	Teacher(s) (Alphabetically)	1. Anuja Arora 2. Neetu Sardana

SNO	Description	Cognitive Level (Bloom Taxonomy)
C121.1	Outline web caching strategies at varied level-user, web server, and gateway server	Understand Level (Level 2)
C121.2	Implement and evaluate various recommendation algorithm, link prediction algorithms, and News Feed Algorithms to provide solutions for various social media and E-Commerce Applications.	Evaluate Level (Level 5)
C121.3	Apply Statistical techniques like Newman Girvan, K-Lin, CPM, Max-Min cut, etc to discover web based communities.	Apply Level (Level 3)
C121.4	Apply Link Analysis using Page Rank, HITS, and link structure of web using Power, Zipf, or Pareto Law.	Analysis Level (Level 4)
C121.5	Design information Diffusion on web using information cascade model, linear threshold model, epidemic models etc.	Create Level (Level 6)

Module No.	**Subtitle of the Module**	**Topics in the module**	**No. of Lectures for the module**
1.	Searching , crawling and indexing Algorithms	Link Based Search Algorithm, Web Crawling, Indexing, Searchiing, Zone Indexing, Term-Frequency, Link Analysis Algorithm, Page rank vector , alpha and power method, user clicks, Naive bayes classifier, hybrid approaches, Doc rank. Website performance optimization	5
2.	Google Patent Algorithm Facebook Algorithms	Page Rank Algorithm, Random Surfer Algorithm, Pigon rank Algorithm, Hilltop Algorithm, Topic Sensitive Page Rank algorithm, Google Panda, Google penguin and google Hummingbird Algorithm, Spamdexing, Author rank News feed Algorithm(NFO), Edge Rank Algorithm.	5

3	Web caching Algorithm	LRV, FIFO, LRU, Random, OPT	3
4	Web Ontology	Semantic Modeling, Resource description Language(RDF), Web Ontology Language(OWL), ontology query, ontology rules	3
5	Recommendation Algorithms	Collaborative Filtering, Item-to-Item recommendation, Memory Based Recommendation,	4
6	Graph Theory and Web	Directed and Undirected graphs, Connectivity, Component, Path, diameter, Geodesic path, Giant Component, SCC, WCC, bipartite graphs, Clique, Subgraphs. Network Measurements, Graph Structure of Web	4
7	Evolution of Social Network	Random network: Erdos-Renyi and Barabasi-Albert, Small World Phenomenon, Power Law, Ziff Law, Pareto Principle,	4
8	Community detection	Community Discovery Algorithms: Newmann Girvan Algorithm, K-Lin Algorithm, Ford Fulkerman Algorithm, CPM.	5
9	Game Theory	Nash equilibrium, Pareto optimality, Social Optimality.	2
10	Centrality	Degree centrality, Betweenness centrality, Closeness Centrality, Eigen vector centrality. Prestige, Proximity Prestige, Rank Prestige, Co-citation, Bibliographic coupling.	2
11	Information Diffusion	Stability Analysis, Dynamic Analysis, Decision Based Models of Cascades, Probabilistic Models of Information Flow, Epidemic Models.	4
		Total number of Lectures	**41**

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.	Chakrabarti, Soumen. 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.	Charu.C. Aggarwal, 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

Detailed Syllabus
Lecture-wise Breakup

Subject Code	**17M1NCI131**	**Semester Odd (specify Odd/Even)**	**Semester – 2nd Month from Jul to Dec**
Subject Name	Flexible Computer Networks		
Credits	**3**	**Contact Hours**	**3**

Faculty (Names)	**Coordinator(s)**	1. Sangeeta Mittal
	Teacher(s) (Alphabetically)	1. Sangeeta Mittal

Course Outcomes

CO#	Course Outcome	Cognitive Level (Bloom's Taxonomy)
1.	Explain the current network-traffic characteristics and modern networking scenarios	Understanding (level - 2)
2.	Assess limitations of classical networking techniques in supporting recent applications	Analyzing (level-4)
3.	Explain Software Defined Network architecture, need and concepts	Understanding (level - 2)
4.	Experiment with Openflow based southbound API in Mininet emulator	Applying(level-3)
5.	Evaluate SDN using Pox and OpenDaylight SDN Controllers	Evaluating(level-5)
6.	Build traffic engineering modules for load balancing, quality of service and multicast data transport in SDN	Creating(level-6)

Module No.	Subtitle of the Module	Topics in the module	No. of Lectures for the module
1.	Modern Networking Elements	Fast Ethernet , Gigabit WiFi, 4G/5G Cellular , Cloud Computing , IoT	3
2.	Basics of Modern Network Traffic	Types of Network Traffic, Real time characteristics, Big Data, Cloud Computing and Mobile Traffic , QoS and QoE – Difficulties in achieving them	4
3.	Drivers and Components of Flexible Networking	Evolving Requirements SDN and NFV	2
4.	Introduction to Software Defined Network (SDN)	Architecture , Characteristics, Standards, Open Development Initiatives	3
5.	SDN Data Plane and Open Flow	Data Plane Functions, OpenFlow logical network Device – Flow Tables, Group Tables, Openflow Protocol	6
6.	SDN Control Plane	Control Plane Architecture , OpenDaylight Project – Architecture and APIs	6
7.	SDN Application Plane	Application Plane Architecture, Data center networking and Information center networking over SDN	6
8.	Network Function	Virtualization Approach, NFV use	4

	Virtualization (NFV) – Concepts	cases, NFV and SDN	
9.	NFV Infrastructure	Virtualized Network Functions, Virtual LAN, Virtual Tenant Network	6
		Total number of Lectures	**40**

Detailed Syllabus

Lecture-wise Breakup

Subject Code	17M11CS112	**Semester Odd (specify Odd/Even)**	**Semester** Even **Session** 2018 - 19 **Month from** July to December
Subject Name	Machine Learning and Data Mining		
Credits	3	**Contact Hours**	3

Faculty (Names)	**Coordinator(s)**	Bharat Gupta
	Teacher(s)	Bharat Gupta

COURSE OUTCOMES		**COGNITIVE LEVELS**
C112.1	Differentiate between Classification, Clustering and Association Rules techniques.	C2
C112.2	Apply and Compare different classification techniques, e.g., k-Nearest Neighbours, Naïve Bayes, ID3 Decision Trees, Support Vector Machine, Ensemble methods , etc.	C3
C112.3	Apply and compare different clustering techniques, e.g., k-means, k-mediods, etc.	C3
C112.4	Apply Apriori algorithm to generate the frequently used rules in a market basket analysis.	C3
C112.5	Apply different dimensionality reduction techniques e.g. PCA, SVD, Factor Analysis, Linear Discriminant Analysis, etc., in big data scenarios.	C3
C112.6	Use Artificial Neural Network techniques, i.e., Back propagation, Feed forward Network, Kohonen Self-Organising Feature Maps, Learning Vector Quantization, etc, for solving classification and clustering problems.	C3

Module No.	Subtitle of the Module	Topics in the module	No. of Lectures for the module
1	Introduction	Introduction to Machine Learning, Data Mining and Knowledge Discovery in Data Bases, Data Types	2
2	Classification	Introduction to classification, k-Nearest Neighbours, Naïve Bayes, Decision Trees	6
3	Regression	Linear Regression with One Variable, Linear Regression with Multiple Variables, Logistic Regression	4
4.	Clustering	Introduction, Different type of Clustering Methods, Partitioning Clustering Methods, Hierarchical Clustering Methods, k-means, k-medoids	6
5.	Association Rules	Frequent itemsets, Apriori algorithm, Association rules	4

6.	Dimensionality Reduction	Introduction, Subset Selection, PCA, SVD, Factor Analysis, Multidimensional Scaling, Linear Discriminant Analysis	8
7.	Artificial Neural Methods	Cost Function, Back propagation, Feed forward Network, Network training, Error Propagation, Application of Neural Networks	8
8.	Ensemble Methods	Ensemble methods of classification-Bagging, Boosting, and Random Forest	4
Total number of Lectures			42

Evaluation Criteria

Components	**Maximum Marks**
T1	**20**
T2	**20**
End Semester Examination	**35**
TA	**25 (Attendance (10), Quiz performance (15))**
Total	**100**

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

1.	Jiawei Han, Micheline Kamber, Data Mining, Morgan Kaufmann Publishers,Elsevier,2005
2.	Kimball R. and Ross M ,The Data Warehouse Toolkit", Wiley
3.	Pujari, Arun K,Data mining and statistical analysis using SQL, Universities press
4.	Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining
5.	Soumen Chakrabarti, Mining the Web: Discovering knowledge from hypertext data", Morgan Kaufmann, Elsevier
6.	Alex, Berson,Stephen J.Smith, Data Warehousing, data mining and OLAP , McGraw-Hill,2004
7.	Inmon W.H.,Building the Data Warehouse ,4th Edition, Wiley
8.	Anahory S. and Murray D, Data Warehousing in the Real World, Addison-Wesley
9.	Margaret H. Dunham, Data Mining: Introductory and Advanced Topics, Prentice Hall,2003
10.	Mattison R. ,Web Warehousing and Knowledge Management", Tata McGraw-Hill.
11.	David Hand, Heikki Mannila and Padhraic Smyth ,Principles of Data Mining,PHI

12.	Transactions on Database Systems (ACM)
13	IEEE Transactions on Knowledge & Data Engineering
14	The VLDB Journal The International Journal on Very Large Data Bases

Detailed Syllabus
Lecture-wise Breakup

Course Code	17M11CS121	**Semester** ODD **(specify Odd/Even)**	**Semester** Ist **Session** 2018 -2019 **Month from** July 2018- December 2018
Course Name	E-Commerce and Social Web		
Credits	3-0-0	**Contact Hours**	3

Faculty (Names)	**Coordinator(s)**	Dr. Sandeep Kumar Singh
	Teacher(s) (Alphabetically)	…

COURSE OUTCOMES		**COGNITIVE LEVELS**
C120.1	Compare and categorize different commercial models of E-commerce.	Understand Level (Level 2)
C120.2	Design and develop marketing strategies based on interactions and insights from Social web to enhance revenue promote brand and reach out to customers.	Create Level (Level 6)
C120.3	Make Use of Open source API s from various social networking sites.	Apply Level (Level 3)
C120.4	Outline suggestions and recommendations for Social Shopping.	Understand Level (Level 2)
C120.5	Measure the effect of different Social media marketing strategies using Social Media metrics.	Apply Level (Level 3)

Module No.	**Title of the Module**	**Topics in the Module**	**No. of Lectures for the module**
1.	Introduction and overview of e-Commerce	Definition and models of e-Commerce and examples. Selection of an E-commerce type and business model. Business models based on (1) Transaction Parties (2) Transaction Types. Case Studies of Indian context.	3
2.	Introduction to Social Web	Social Media : An Overview, Social Media Analytics: An Overview, SOCIAL MEDIA TEXT ANALYTICS, Twitter as Marketing Tool	5
3.	Social Web Landscape	Social Web overview, data-types, format, Text cleaning, tagging and storage, Social media techniques, tools and platforms, data visualization of data, research, applications and challenges in social Web.	3
4.	Introduction to Social e- Commerce	Introduction to Social Commerce, Supporting Theories and Concepts for Social Commerce, Tools and Platforms for Social Commerce	3
5.	Social Web Analysis	Analyzing Social web, Nodes, Edges and Network measures, Centrality, Power and Bottlenecks, Concept of Cliques, Clusters and Components, Viral marketing, Graph data in real world, Business use of Social web, Privacy in Social web, Influencer Outreach	5
6.	Social Shopping and Social Marketing	Social Media Marketing, Social Shopping: Concepts, Benefits, and Models, Customer Engagement and Metrics, Basic Social Marketing Strategies- Physical goods, Digital goods, Services, Affiliate Marketing, Guerrilla Marketing	5
7.	Programming using	Introduction to OAuth protocol, Programming and Crawling Social media using Twitter 4j Facebook API, LinkedIn API ,	6

	API and RSS feeds	Google +, Reddit, API, Instagram API	
8	Twitter and Face book Data Analytics for Viral Marketing	Topic-based Clusters in Egocentric Networks on Facebook, Changes in Tie Strength Through Site Use on Facebook, Patterns of Responses to Resource Requests on Facebook, Exploring requests for help on Facebook, Analysis of User-Generated Content on Facebook, Predicting Clicks on Ads,Predicting the quality of new contributors to the Facebook crowdsourcing system	8
9.	Social Search Engine Optimization	Optimizing for Web Search, Using Photo-Sharing Sites for SEO, Optimizing for Social Search Engines	6
10.	Creating Suggestions and Recommendations	Perform web-market segmentation, making recommendations: collaborative filtering and content based filtering approaches, creating suggestions and building recommendation engines, Understanding recommendation engines based on users, items, and content, Finding recommendations about friends, articles, and news stories, Creating recommendations for sites similar to Netflix	6
		Total number of Lectures	**45**

Evaluation Criteria

Components	**Maximum Marks**
T1	20
T2	20
End Semester Examination	35
TA	25 (Assignments and Attendance)
Total	**100**

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.	Michael P Papazoglou and Pieter M.A. Ribbers, " e-Business- Organizational and technical foundation" , John Wiley and Sons, 2006.
2.	Efraim Turban , David King, Dennis Viehland, Jae Lee, "Electronic Commerce A Managerial Perspective 2006", 4ed, Pearson Education International edition, 2006.
3.	Stephen Chen, "Strategic management of e-business", second edition, John Wiley and Sons, 2005.
4.	RS Prasad, "Cyber crime: An Introduction", ICFAI Books, ICFAI University, 2004.
5.	RS Prasad, "Cyber crime: Combat Strategies", ICFAI Books, ICFAI University, 2004.
6.	RS Prasad, "CRM Present and Future", ICFAI Books, ICFAI University, 2005.
7.	Elaine Lawrence et al, "Internet commerce – Digital models for Business", John Wiley and Sons, 2003.
8.	Abhijit Choudhury and Jean-Pierre Kuilboer, "E-business and E-Commerece Infrastructure – Technologies supporting E-Business Initiative", McGraw Hill, 2002.
9.	Henry Chan et al, E-Commerece – fundamentals and applications", John Wiley and Sons, 2001.
10.	Programming Collective Intelligence: Building Smart Web 2.0 Applications by Toby Segaran
11.	Algorithms of the Intelligent Web Haralambos Marmanis, Dmitry Babenko
12.	Recommender Systems: An Introduction Dietmar Jannach (Author), Markus Zanker (Author), Alexander Felfernig (Author), Gerhard Friedrich
13.	Recommender Systems Handbook Francesco Ricci (Editor), Lior Rokach
14.	Recommendation Systems in Software Engineering Martin P. Robillard (Editor), Walid Maalej (Editor), Robert J

	Walker (Editor), Thomas Zimmermann
15.	Web Analytics 2.0 Avinash Kaushik
16.	Analyzing Social Web Jeneffir Golbeg
17.	Predictive Analytics Eric Segel

Detailed Syllabus

Lab-wise Breakup

Course Code	17M15CS112	**Semester: ODD**	**Semester: I Session** 2018 -2019 Month from: July-Dec
Course Name	Machine Learning and Data Mining Lab		
Credits	1	**Contact Hours**	**2**

Faculty (Names)	**Coordinator(s)**	Bharat Gupta
	Teacher(s) (Alphabetically)	Bharat Gupta

COURSE OUTCOMES		**COGNITIVE LEVELS**
C173.1	Understanding basic syntax in Python	Understanding Level (C2)
C173.2	Understanding Control Flow and looping in Python	Understanding Level (C2)
C173.3	Apply and Compare different classification techniques, Logistic Regression e.g., k-Nearest Neighbours, Support Vector Machine, etc.	Apply Level (C3)
C173.4	Apply clustering techniques k-Means on a dataset	Apply Level (C3)
C173.5	Apply dimensionality reduction technique e.g. PCA on a dataset.	Apply Level (C3)
C173.6	Analyse the real world problem to identify the appropriate data science techniques for classification, clustering and Association rules	Analyse Level (C4)

Module No.	**Title of the Module**	**List of Experiments**	**CO**
1.	Python basic syntax	Practicing basic python commands	CO1
2.	Control Flow and looping in Python	1. Write a python program that displays the sum of all digits for a user entered number. 2. Write a python function leap_year that prints all the leap years between ranges. The user will enter lower and upper year boundary inside the function. 3. Write a program that outputs all possible strings formed by using the characters a, c, t, o, and g. a particular character can appear only once and all the characters should be used in the formation of string. 4. Write a python script that takes input from file representing a paragraph, and writes to a file named	CO1

		out.txt with all the stop words (a, an, the) removed. 5. Write a recursive function in python to print a Fibonacci series. The Fibonacci sequence is the series of numbers: 0,1,1,2,3,5,8,13,21,34,…etc 6. Write a program for sorting the integer data by using quick sort.	
3.	K-NN	Implement the KNN (K Nearest Neighbours) algorithm in python. Your program should have different functions as follows: 1. HandleData: Open the dataset from CSV and split into test/train (datasets). A ratio of 67/33 for train/test is a standard ratio used for splitting data. 2. Similarity: Calculate the distance between two data instances. The Euclidean distance is used for calculating the difference. It is defined as the square root of the sum of the squared differences between the two arrays of numbers. Only first 4 attributes are used for calculating the distance. 3. Neighbours: Locate k most similar data instances. 4. Response: Generate a response from a set of data instances. It is a function for getting the majority voted response from a number of neighbors. It devises a predicted response based on those neighbors. 5. Accuracy: Summarize the accuracy of predictions. An easy way to evaluate the accuracy of the model is to calculate a ratio of the total correct predictions out of all predictions made, called the classification accuracy. 6. Main: Take split = 0.67, k=3.	CO3
4.	Weka Toolkit	1. Apply the KNN algorithm in Weka tool on the iris dataset. Compare the results of your implemented algorithm with algorithm of Weka tool. 2. Implement the linear Regression. The data will be taken as input from the file. Select the appropriate dataset from the website https://archive.ics.uci.edu/ml/index.php". Justify the reason why the dataset has been selected. b) Apply the Linear regression in Weka tool on the same dataset. Compare the results of your implemented algorithm with algorithm of Weka tool.	CO3
5.	Clustering	Remove the label column of the Parkinson_dataset.csv dataset and implement the following: a) Perform K-Means clustering and Hierarchical clustering. b) Use Manhattan distance c) Use Average merging Strategy in Hierarchical	CO4

		clustering. d) Use three different K values in K-Mean clustering. e) Validate using RMSE and compare both the techniques.	
6	Logistic regression and SVM	Divide the Parkinson_dataset.csv dataset in training and testing dataset randomly and implement the following: a. Classify the disease using Logistic regression and SVM b. Find out the accuracy of classification Model. c. Perform 5-fold cross- validation. d. Compare the result of both techniques using matplotlib.	CO3
7	scikit-learn toolkit	Implementation of the following algorithms in scikit-learn a. Principal components analysis (PCA) b. Decomposing signals in components (matrix factorization problems) c. K-means	CO5
8	Mini Project	1. Specify the broad topic of your mini project based on the Machine Learning and Data mining. 2. Study minimum 6 quality research papers based on the selected topic. 3. Do the SWOT analysis of selected research papers/reports. 4. Identify the research problem. 5. Propose your novelty/improvement in terms of algorithm/new feature. 6. Design the architecture for the proposed problem. 7. Design the test bed. 8. Design a set of experiments to be carried out for the proposed problem. 9. Perform the experimental analysis (in Python language only). 10. Prepare your report. 11. Write a short research paper based on your contribution.	CO6

Evaluation Criteria

Components	Maximum Marks
Lab Test1	20
Lab Test2	20
Mini Project, Regularity, performance	60
Total	100

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.	Jiawei Han, Micheline Kamber, Data Mining, Morgan Kaufmann Publishers,Elsevier,2005
2.	Kimball R. and Ross M ,The Data Warehouse Toolkit", Wiley
3.	Pujari, Arun K,Data mining and statistical analysis using SQL, Universities press
4.	Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining
5.	Soumen Chakrabarti, Mining the Web: Discovering knowledge from hypertext data", Morgan Kaufmann, Elsevier
6.	Margaret H. Dunham, Data Mining: Introductory and Advanced Topics, Prentice Hall,2003
7.	Mattison R. ,Web Warehousing and Knowledge Management", Tata McGraw-Hill.
8.	David Hand, Heikki Mannila and Padhraic Smyth ,Principles of Data Mining,PHI
9.	Transactions on Database Systems (ACM)
10.	IEEE Transactions on Knowledge & Data Engineering
11.	The VLDB Journal The International Journal on Very Large Data Bases

Detailed Syllabus
Lecture-wise Breakup

Course Code	17M21CS111	**Semester Odd**	**Semester I Session 2018_2019** Month from July 2018 to Dec 2018
Course Name	Cloud Based Big Data Systems I		
Credits	3	**Contact Hours**	3

Faculty (Names)	**Coordinator(s)**	Dr Parmeet Kaur
	Teacher(s) (Alphabetically)	1. Dr Parmeet Kaur

COURSE OUTCOMES		**COGNITIVE LEVELS**
CO111.1	Outline the concept and challenge of big data and how cloud technology is useful to store or analyze the big data	Outline Level 2
CO111.2	Compare techniques of big data distribution in clouds – Partitioning and Replication.	Compare Level 4
CO111.3	Outline Hadoop architecture and MapReduce framework.	Outline Level 2
CO111.4	Explain Cloud NoSQL- Cassandra architecture, transaction processing and repair mechanisms for big data storage.	Explain Level 2
CO111.5	Apply Cassandra CQL commands to define, query and manipulate a NoSQL database.	Apply Level 3
CO111.6	Design and develop a simple application and connect with a NoSQL database, NewSQL database or Hadoop distributed file system. [Level 6]	Design Level 6

Module No.	**Title of the Module**	**Topics in the Module**	**No. of Lectures for the module**
1.	Introduction to Database Systems and Cloud Computing	Overview of Database Management Systems ,Basic principles of cloud computing, Classifying cloud services, and Basic terms and principles of DaaS (Database as a Service)	4
2.	Data Distribution: Partitioning and Replication	Data partitioning and replication techniques,Consistent Hashing, and Implementing highly available, scalable, and network partition tolerant cloud databases	6
3.	Trade-offs in Cloud Databases	Differences between conventional (relational) databases and cloud databases ACID database properties, CAP Conjecture, and BASE properties of cloud databases, NewSQL	4
4.	SQL based Cloud Databases	SQL compliancy, Transaction Control, Elasticity & Scalability. Case Study: nuoDB/ DB as a service	6
5.	Cloud NoSQL Databases	*Cloud Data Models*: Key-Value data model, Document data model, Column Family data model. Graph Data Model	2
6.	Cassandra Architecture and Cassandra Data	Internode Communication ,Data Distribution and Replication ,Partitioning , Snitches ,Basic features of Cassandra CDBMS, Formal definition of Cassandra column	6

	Model	family data model, Cassandra CQL query language and CQL data model: Key space, Table definition, Column, and Data Types	
7.	Cassandra Consistency Levels	Configuring Data Consistency -Write Requests, Read Requests	3
8.	Cassandra Repair Mechanisms , Transaction Processing	Hinted Handoff Writes, Anti-entropy Node Repair, Transactions and Concurrency Control, Light Weight Transactions	5
9.	Cassandra CQL Queries	The Syntax of the SELECT Statement Simple SELECT expressions ,Filtering Data using WHERE Clause ,Using Indexes ,Filtering Collections , Querying Tables with Columns of the counter Type Keyspace Design Heuristics	6
		Total number of Lectures	**42**

Evaluation Criteria

Components	**Maximum Marks**
T1	20
T2	20
End Semester Examination	35
TA	25 (Class Assignments:10, Project: 10, Attendance:5)
Total	**100**

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.	Furht B., Villanustre F. (2016) Introduction to Big Data. In: Big Data Technologies and Applications. Springer, Cham
2.	Li, Kuan-Ching, Hai Jiang, Laurence T. Yang, and Alfredo Cuzzocrea, eds. *Big data: Algorithms, analytics, and applications*. CRC Press, 2015.
3.	Buyya, Rajkumar, Christian Vecchiola, and S. Thamarai Selvi. *Mastering cloud computing: foundations and applications programming*. Newnes, 2013.
4.	Zomaya, Albert Y., and Sherif Sakr, eds. *Handbook of big data technologies*. Berlin: Springer, 2017.
5.	Sullivan, Dan. *NoSQL for mere mortals*. Addison-Wesley Professional, 2015.
6.	Lam, Chuck. *Hadoop in action*. Manning Publications Co., 2010.

Detailed Syllabus
Lecture-wise Breakup

Course Code	17M22CS123	**Semester: (specify Odd/Even)**	**Semester** ODD **Session** 2018-2019 **Month from** June 18 **to** Dec 18
Course Name	Natural Language Processing and Understanding		
Credits	3	**Contact Hours**	3-0-0

Faculty (Names)	**Coordinator(s)**	Dr. Shikha Jain
	Teacher(s) (Alphabetically)	Dr. Shikha Jain, Mr. Vimal Kumar

COURSE OUTCOMES		**COGNITIVE LEVELS**
CO1	Preprocess the natural language text through tokenization, lemmatization, stop-word removal and morphological analysis.	Apply (Level 3)
CO2	Apply and analyze various language models for data representation.	Analyze (Level4)
CO3	Select and apply various part-of-speech (POS) tagging approaches.	Apply (Level 3)
CO4	Apply non-probabilistic and probabilistic parsing techniques for checking the syntax of the natural language text.	Apply (Level 3)
CO5	Apply and analyze various contextual analysis techniques for meaningful information extraction.	Analyze (Level4)
CO6	Design and evaluate various NLP applications such as topic modelling, text classifications, word prediction, question answering system and machine translation.	Evaluate (Level 5)

Module No.	**Subtitle of the Module**	**Topics in the module**	**No. of Lectures for the module**
1.	Introduction	Intro to NLP, Challenges & Requirements	01
2.	Preprocessing	Tokenization, Lemmatization, Stop word removal, Morphological analysis	01
3	Word embedding	Word embedding and language models: N-gram, Vector representations (one hot vector, Bag of word, co-occurrence matrix), Word Embedding (word2vec : CBOW, skipgram; Glove), Advanced word vector representations (softmax, single layer networks);	07
4	Part of Speech Tagging	Hidden Markov Models, Maximum Entropy Markov Models & Conditional Random Fields; Smoothing;	05

5	Parsing	Context Free Grammars, Non-probabilistic Parsing, Probabilistic Parsing	05
6	Applications - I	Graph-based Methods for NLP Applications;	03
7	Unsupervised Language Discovery	Statistical Models of Semantics and Unsupervised Language Discovery: resolving ambiguity; Language modeling and Naive Bayes	04
8	Supervised Language Discovery	Supervised Language Discovery: text classification	02
9	Topic Modeling	LSI, PLSI, LDA	03
10	Applications - II	Word prediction: LSTM	04
		The pragmatics of questions and answers: Partition semantics and decision-theoretic;	02
11	Machine Translation	IBM model 1, 2 and 3	03
12	Case Study	Case Study: Apple Siri/ Amazon Alexa/ IBM Watson	02
		Total number of Lectures	**42**

Evaluation Criteria

Components	**Maximum Marks**
T1	20
T2	20
End Semester Examination	35
TA	25
Total	**100**

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.	Handbook of Natural Language Processing & Machine Translation by Olive, Joseph, Christianson, Caitlin, McCary, John (Eds.), Springer
2.	Natural Language Understanding by James Allen, Benjamin Cummins Publisher
3.	Foundations of Statistical NLP by Hinrich Schtze, Christopher D. Manning
4.	Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (second edition) D. Jurafsky and J. Martin
5.	Natural Language Processing with Python by Steven Bird, Ewan Klein, and Edward Loper

Detailed Syllabus
Lab-wise Breakup

Course Code	17M25CS111	**Semester** Even **(specify Odd/Even)**	**Semester** M Tech (DA) 1st sem **Session** 2018 -2019 **Month from July-Dec 2019**
Course Name	Cloud Based Big Data Systems Lab-I		
Credits	1	**Contact Hours**	2

Faculty (Names)	**Coordinator(s)**	Dr Parmeet Kaur
	Teacher(s) (Alphabetically)	Dr Parmeet Kaur

COURSE OUTCOMES		**COGNITIVE LEVELS**
CO1	Outline cloud based big data systems and classify them on basis of their features and applicability	Classify Level 2
CO2	Apply Hadoop file system shell commands to perform various Hadoop distributed file system (HDFS) operations	Apply Level 3
CO3	Develop a real-world application using the MapReduce framework	Develop Level 6
CO4	Apply Cassandra CQL commands to define, query and analyze a NoSQL database.	Apply Level 3
CO5	Apply NuoDB operations to insert and query data.	Apply Level 3

Module No.	**Title of the Module**	**List of Experiments**	**CO**
1.	Introduction To Cloud Based Big Data Systems	1. Investigate the Cloud Based Big Data Systems. 2. Use open source Big Data visualization Tool to form data analysis 3. Explore the various NoSQL systems	CO1
2.	Overview of Hadoop Framework	4. Apply Hadoop DFS commands for file manipulation.	CO2
3.	Map Reduce Framework	5. Write Map Reduce programs to solve big data problems	CO3
4.	Introduction to Cassandra	6. Install Cassandra 7. Perform Cassandra CRUD operations	CO4
5.	Aggregation with Cassandra	8. Perform data analysis with Cassandra aggregation operators	CO4
6.	Overview of NewSQL	9. Install NuoDB, understand its architecture 10. Define a big data store in NuoDB	CO5
7.	Working with NuoDB	11. Insert data in NuoDB 12. Query data from NuoDB	CO5

Evaluation Criteria

Components	**Maximum Marks**
1. Lab Test1	20
2. Lab Test 2	20
3. Lab Assignments	25

4. Project	25
5. Attendance	10
Total	**100**

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.	Furht B., Villanustre F. (2016) Introduction to Big Data. In: Big Data Technologies and Applications. Springer, Cham
2.	Li, Kuan-Ching, Hai Jiang, Laurence T. Yang, and Alfredo Cuzzocrea, eds. *Big data: Algorithms, analytics, and applications*. CRC Press, 2015.
3.	Buyya, Rajkumar, Christian Vecchiola, and S. Thamarai Selvi. *Mastering cloud computing: foundations and applications programming*. Newnes, 2013.
4.	Sullivan, Dan. *NoSQL for mere mortals*. Addison-Wesley Professional, 2015.
5.	Lam, Chuck. *Hadoop in action*. Manning Publications Co., 2010.

Detailed Syllabus

Lab-wise Breakup

Course Code	17M25CS113	**Semester** Even	**Semester** MTech-Ist year **Session** 2018 -2019 Month from July 2018 to Dec 2018
Course Name	Data Science Programming Lab-I		
Credits	…	**Contact Hours**	0-0-2

Faculty (Names)	**Coordinator(s)**	Neetu Sardana
	Teacher(s) (Alphabetically)	Neetu Sardana

COURSE OUTCOMES		**COGNITIVE LEVELS**
C172.1	Understand the fundamentals of R programming language.	Understand Level (Level 2)
C172.2	Importing different data formats and Apply data pre-processing techniques to handle missing values and noisy data values.	Apply Level (Level 3)
C172.3	Apply Data Visualization techniques for graphical representation and analysis of real world data	Apply Level (Level 3)
C172.4	Apply varied predictive modeling techniques and interpret the results.	Analyze Level (Level 4)

Module No.	**Title of the Module**	**List of Experiments**	**CO**
1.	Basics of R	Objects,Functions, Number & Vector, Matrices & Array,Factors, Conditional Statements, Loop, Scripts, R package. List, List Operation, Recursive List, Data Frame, Creating Data Frame, Data Frame Operations, lapply() and sapply() functions.	CO1
2.	String Handling	Introduction to String handling, String functions, String Manipulation, Regular Expressions & Pattern Matching, and Introduction to "stringr" package.	CO1
3.	Import & Export	Introduction, Saving & Loading R data, Import and Export to different file formats: Excel File, Binary File, XML File, JSON File. Analyzing data & Reshaping the data.	CO2
4.	Data Preprocessing using R	Data Preprocessing, forms of Data Preprocessing, Data Cleaning Techniques, Data Redundancy- chi square test, correlation analysis, covariance coefficient, Data Transformation, Data Reduction- Principal Component Analysis, Regression, R packages for Data Preprocessing:	CO2

		caret, dplyr.	
5.	Data Visualization	Visual Representation of statistical analysis, High level plotting commands- create plots with axes, titles, labels and others on the graphics device and Low level plotting commands- add new features like extra labels, point or line. Plots, Histogram, Scatter Plots, Pie chart, Box Plot, QQ Plot, customized Plotting. Introduction to data visualization packages: Ggobi & ggplot.	CO3
6.	Classification and Clustering Algorithm	Classification Techniques: Introduction to Classification, Regression, Naïve Bayes, Decision Tree, KNN, Ensemble Methods. Clustering Techniques: Introduction to Clustering, K-means, Hierarchical Clustering, DB Scan.	CO4

Evaluation Criteria

Components	**Maximum Marks**
Lab Test 1	20
Lab Test 2	20
Quiz 1	10
Evaluation 1	15
Evaluation 2	15
Mini Project	20
Total	**100**

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.	Paul Teetor.R Cookbook - Proven Recipes for Data Analysis, Statistics, and Graphics. O'Reilly, 2011.
2.	Alain F. Zuur, Elena N. Ieno, and Erik Meesters. A Beginner's Guide to R. Use R. Springer, 2009. ISBN: 978-0-387-93836-3.
3.	John Maindonald and John Braun. Data Analysis and Graphics Using R. Cambridge University Press, Cambridge, 2nd edition, 2007. ISBN 978-0-521-86116-8.
4.	Advanced R, by Hadley Wickham, ISBN 9781466586963.
5.	Using R for Introductory Statistics, by John Verzani, Chapman & Hall/CRC, 2004, ISBN 1584884509

Detailed Syllabus

Lecture-wise Breakup

Course Code	18M11GE111	**Semester** Odd	**Semester** I	**Session** 2018 -2019	
			Month from	July 2018 to Dec 2018	
Course Name	Research Methodology & Intellectual Property Rights				
Credits	2		**Contact Hours**	2-0-0	
Faculty (Names)	**Coordinator(s)**	Prof. B.P. Chamola			
	Teacher(s) (Alphabetically)	Prof. B.P. Chamola			

COURSE OUTCOMES		**COGNITIVE LEVELS**
After pursuing the above mentioned course, the students will be able to:		
CO1	understand the basic concepts and types of research	Understanding Level (C2)
CO2	define a research problem, its formulation, methodologies and analyze research related information	Analyzing Level (C4)
CO3	follow research ethics, understand IPR, patents and their filing related to their innovative works.	Understanding Level (C2)
CO4	understand and analyze the statistical data and apply the relevant test of hypothesis in their research problems	Analyzing Level (C4)

Module No.	**Title of the Module**	**Topics in the Module**	**No. of Lectures for the module**
1.	Research	What is research? Types of research. What is not research? How to read a Journal paper?	3
2.	Report writing	How to write report? Use of Mendeley in report writing. How to write a research paper? Problem identification and solving.	4
3.	Ethics, IPR and Research methodologies	Research ethics, patents, intellectual property rights, plagiarism regulation 2018. Steps in research process and common methodologies to attempt solution to research paper.	8

4.	Basics of statistics and probability distributions	Basic statistical concepts. Handling of raw data, Some common probability distributions.	7
5.	Test of hypothesis and regression analysis	Hypothesis testing. Parametric and non-parametric data, Introduction to regression analysis.	8
		Total number of Lectures (Course delivery method: open ended discussion, guided self-study, lectures)	**30**

Evaluation Criteria

Components	**Maximum Marks**
Viva –1 before T2 1 Exam	15
Viva –2 after End Sem.	20
End Semester Examination	35
Assignments	30 (Quiz, Assignments)
Total	**100**

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.	**Stuart Melville and Wayne Goddard**, Research methodology: An Introduction for Science & Engineering Students, Kenwyn, South Africa : Juta& Co. Ltd., 1996.
2.	**Kothari, C.R.,** Research Methodology: Methods and Techniques, New Age International, New Delhi, 2009.
3.	**Kumar, Ranjit,** Research Methodology: A Step by Step Guide for Beginners, 2nd Edition, Sage Publications Ltd., 2005.
4.	**Ramappa, T.,** Intellectual Property Rights Under WTO, S. Chand, New Delhi, 2008.
5.	**Wayne Goddard and Stuart Melville,** Research Methodology: An Introduction, Kenwyn, South Africa : Juta& Co, 2001.

Detailed Syllabus
Lecture-wise Breakup

Course Code	**18M12CS117**	**Semester (Odd)**	**Semester** I **Session** 2018 -2019 **Month from July - December**
Course Name	Blockchain Technology and Applications		
Credits	03	**Contact Hours**	(L+T) (3+1)

Faculty (Names)	**Coordinator(s)**	Dr. P. Raghu Vamsi
	Teacher(s) (Alphabetically)	Dr. P. Raghu Vamsi

COURSE OUTCOMES		**COGNITIVE LEVELS**
C141.1	Understand the structure of a blockchain and why/when it is better than a simple distributed database	Understand Level (Level 2)
C141.2	Analyze the incentive structure in a blockchain based system and critically assess its functions, benefits and vulnerabilities	Evaluate Level (Level 5)
C141.3	Evaluate the setting where a blockchain based structure may be applied, its potential and its limitations	Apply Level (Level 3)
C141.4	Attain awareness of the new challenges that exist in monetizing businesses around blockchains and smart contracts	Analyze Level (Level 4)
C141.5	Describe and apply the differences between the most prominent blockchain structures and permissioned blockchain service providers, as well as rising alliances and networks	Apply Level (Level 3)

Module No.	**Title of the Module**	**Topics in the Module**	**No. of Lectures for the module**
1.	Blockchain Basics	What is Blockchain (BC), public ledgers, BC as public ledgers; BC history - Bitcoin and Cryptocurrency, BC 2.0, Smart contracts; BC architecture – Blocks in BC, transactions and distributed consensus; BC conceptualization - The Chain and the Longest Chain, Cryptocurrency to Blockchain 2.0, Permissioned Model of Blockchain.	4
2.	Cryptographic Primitives	Cryptographic Hash Function, Properties of a hash function, Hash pointer and Merkle tree, Digital Signature, Public Key Cryptography, A basic cryptocurrency	5
3.	Distributed Consensus	Distributed consensus in open environments, Consensus in a Bitcoin network; Bitcoin Consensus - Proof of Work (PoW) – basic introduction, Hashcash PoW, Beyond Consensus in Bitcoin - Bitcoin PoW, Attacks on PoW and the monopoly problem, Proof of Stake, Proof of Burn and Proof of Elapsed Time; Consensus in Bitcoin (The Miners) - The life of a Bitcoin Miner, Mining Difficulty, Mining Pool.	6
4.	Smart contracts – 1	Smart contracts, Solidity, REMIX IDE, Ethereum Blockchain, Ethereum Virtual Machine.	8
5.	Smart contracts – 2	Decentralized applications (Dapps), Truffle development, Design improvements, Application models and standards	7

6.	Use cases	Blockchain for Voting, Government Use-cases – Public distribution system, Blockchain for Tax Payments, Blockchain for Managing Land Registry Records	3
7.	Other Blockchain frameworks	IBM Hyperledge fabric	7-10
9.	Research aspects in Blockchain	Consensus protocols, Identity management, Strong and weak synchronization, avoiding forks, Mining improvements.	3
		Total number of Lectures	42-45

Evaluation Criteria

Components	**Maximum Marks**
T1	20
T2	20
End Semester Examination	35
TA	25
Total	**100**

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.	Drescher, Daniel. "Blockchain basics", Apress, 2017.
2.	Mougayar, William. "The business blockchain: promise, practice, and application of the next Internet technology", John Wiley & Sons, 2016.
3.	Dannen, Chris. "Introducing Ethereum and Solidity", Berkeley: Apress, 2017.
4.	Prusty, Narayan. "Building Blockchain Projects", Packt Publishing Ltd, 2017.
5.	Pilkington, Marc. "Blockchain technology: principles and applications" Research handbook on digital transformations, 2016.
6.	Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, Steven Goldfeder. Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction, Princeton University Press, 2016.
7.	Swan, Melanie, "Blockchain: Blueprint for a new economy", O'Reilly Media, Inc., 2015.
8.	Antonopoulos, Andreas M. "Mastering Bitcoin: unlocking digital cryptocurrencies", O'Reilly Media, Inc., 2014.