Program Structure for Fourth Year Computer Engineering

UNIVERSITY OF MUMBAI (With Effect from 2022-2023)

Semester VII

Course Code	Course Name	Teaching Scheme (Contact Hours)		Credits Assigned		
		Theory	Pract. Tut.	Theory	Pract.	Total
CSC701	Machine Learning	3	--	3	--	3
CSC702	Big Data Analytics	3	--	3		3
CSDC 701X	Department Level Optional Course-3	3	--	3	--	3
CSDC 702X	Department Level Optional Course-4	3	--	3	--	3
ILO 701X	Institute Level Optional Course-1	3	--	3	--	3
CSL701	Machine Learning Lab	--	2	--	1	1
CSL702	Big Data Analytics Lab	--	2	--	1	1
CSDL 701X	Department Level Optional Course-3 Lab	--	2	--	1	1
CSDL 702X	Department Level Optional Course-4 Lab	--	2	--	1	1
CSP701	Major Project 1	--	6#	--	3	3
Total		**15**	**14**	**15**	**7**	**22**

Course Code	Course Name	Examination Scheme							
		Theory					Term Work	Pract. & oral	Total
		Internal Assessment			End Sem Exam	Exam. Duration (in Hrs)			
		Test 1	Test 2	Avg					
CSC701	Machine Learning	20	20	20	80	3	--	--	100
CSC702	Big Data Analysis	20	20	20	80	3	--	--	100
CSDC 701X	Department Level Optional Course-3	20	20	20	80	3	--	--	100
CSDC 702X	Department Level Optional Course-4	20	20	20	80	3	--	--	100
ILO 701X	Institute Level Optional Course-1	20	20	20	80	3	--	--	100
CSL701	Machine Learning Lab	--	--	--	--	--	25	25	50
CSL702	Big Data Analytics Lab	--	--	--	--	--	25	25	50
CSDL 701X	Department Level Optional Course-3 Lab						25	-	25
CSDL 702X	Department Level Optional Course-4 Lab	--	--	--	--	--	25	-	25
CSP701	Major Project 1	--	--	--	--	--	50	25	75
Total		--	--	**100**	**400**	--	**150**	**75**	**725**

Course Code:	Course Title	Credit
CSC701	Machine Learning	3

Prerequisite: Engineering Mathematics, Data Structures, Algorithms	
Course Objectives:	
1	To introduce the basic concepts and techniques of Machine Learning.
2	To acquire in depth understanding of various supervised and unsupervised algorithms
3	To be able to apply various ensemble techniques for combining ML models.
4	To demonstrate dimensionality reduction techniques.
Course Outcomes:	
1	To acquire fundamental knowledge of developing machine learning models.
2	To select, apply and evaluate an appropriate machine learning model for the given
3	To demonstrate ensemble techniques to combine predictions from different models.
4	To demonstrate the dimensionality reduction techniques.

Module		Content	Hrs
1		Introduction to Machine Learning	04
	1.1	Machine Learning, Types of Machine Learning, Issues in Machine Learning, Application of Machine Learning, Steps in developing a Machine Learning Application.	
	1.2	Training Error, Generalization error, Overfitting, Underfitting, Bias-Variance trade-off.	
2		Learning with Regression and Trees	09
	2.1	Learning with Regression: Linear Regression, Multivariate Linear Regression, Logistic Regression.	
	2.2	Learning with Trees: Decision Trees, Constructing Decision Trees using Gini Index (Regression), Classification and Regression Trees (CART)	
	2.3	Performance Metrics: Confusion Matrix, [Kappa Statistics], Sensitivity, Specificity, Precision, Recall, F-measure, ROC curve	
3		Ensemble Learning	06
	3.1	Understanding Ensembles, K-fold cross validation, Boosting, Stumping, XGBoost	
	3.2	Bagging, Subagging, Random Forest, Comparison with Boosting, Different ways to combine classifiers	
4		Learning with Classification	08
	4.1	Support Vector Machine Constrained Optimization, Optimal decision boundary, Margins and support vectors, SVM as constrained optimization problem, Quadratic Programming, SVM for linear and nonlinear classification, Basics of	

		Kernel trick.	
	4.2	Support Vector Regression, Multiclass Classification	
5		Learning with Clustering	07
	5.1	Introduction to clustering with overview of distance metrics and major clustering approaches.	
	5.2	Graph Based Clustering: Clustering with minimal spanning tree Model based Clustering: Expectation Maximization Algorithm, Density Based Clustering: DBSCAN	
6		Dimensionality Reduction	05
	6.1	Dimensionality Reduction Techniques, Principal Component Analysis, Linear Discriminant Analysis, Singular Valued Decomposition.	
		Total	39

Textbooks:	
1	Peter Harrington, "Machine Learning n Action", DreamTech Press
2	Ethem Alpaydın, "Introduction to Machine Learning", MIT Press
3	Tom M. Mitchell, "Machine Learning" McGraw Hill
4	Stephen Marsland, "Machine Learning An Algorithmic Perspective", CRC Press
References:	
1	Han Kamber, —Data Mining Concepts and Techniques‖, Morgan Kaufmann Publishers
2	Margaret. H. Dunham, —Data Mining Introductory and Advanced Topics, Pearson Education
3	Kevin P. Murphy , Machine Learning — A Probabilistic Perspective‖
4	Samir Roy and Chakraborty, —Introduction to soft computing‖, Pearson Edition.
5	Richard Duda, Peter Hart, David G. Stork, "Pattern Classification", Second Edition, Wiley Publications.
Assessment:	
Internal Assessment:	
Assessment consists of two class tests of 20 marks each. The first class test is to be conducted when approximately 40% syllabus is completed and the second class test when an additional 40% syllabus is completed. Duration of each test shall be one hour.	
End Semester Theory Examination:	
1	Question paper will comprise a total of six questions.
2	All question carries equal marks
3	Questions will be mixed in nature (for example supposed Q.2 has part (a) from module 3 then part (b) will be from any module other than module 3)
4	Only Four questions need to be solved.

Course Code	Course Name	Credit
CSC702	Big Data Analysis	03

Prerequisite: Database, Data mining.	
Course Objectives: The course aims:	
1	To provide an overview of the big data platforms, its use cases and Hadoop ecosystem.
2	To introduce programming skills to build simple solutions using big data technologies such as MapReduce, Scripting for No SQL and R
3	To learn the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.
4	To enable students to have skills that will help them to solve complex real-world problems for decision support.
Course Outcomes:	
1	Understand the building blocks of Big Data Analytics.
2	Apply fundamental enabling techniques like Hadoop and MapReduce in solving real world problems.
3	Understand different NoSQL systems and how it handles big data.
4	Apply advanced techniques for emerging applications like stream analytics.
5	Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications, etc.
6	Apply statistical computing techniques and graphics for analyzing big data.

Module		Detailed Content	Hours
1		**Introduction to Big Data and Hadoop**	2
	1.1	Introduction to Big Data - Big Data characteristics and Types of Big Data	
	1.2	Traditional vs. Big Data business approach	
	1.3	Case Study of Big Data Solutions	
	1.4	Concept of Hadoop, Core Hadoop Components; Hadoop Ecosystem	
2		**Hadoop HDFS and MapReduce**	8
	2.1	Distributed File Systems: Physical Organization of Compute Nodes, Large-Scale File-System Organization.	
	2.2	MapReduce: The Map Tasks, Grouping by Key, The Reduce Tasks, Combiners, Details of MapReduce Execution, Coping With Node Failures.	
	2.3	Algorithms Using MapReduce: Matrix-Vector Multiplication by MapReduce, Relational-Algebra Operations, Computing Selections by MapReduce, Computing Projections by MapReduce, Union ,Intersection, and Difference by MapReduce	

	2.4	Hadoop Limitations	
3		**NoSQL**	10
	3.1	Introduction to NoSQL, NoSQL Business Drivers	
	3.2	NoSQL Data Architecture Patterns: Key-value stores, Graph stores, Column family (Bigtable)stores, Document stores, Variations of NoSQL architectural patterns, NoSQL Case Study	
	3.3	NoSQL solution for big data, Understanding the types of big data problems; Analyzing big data with a shared-nothing architecture; Choosing distribution models: master-slave versus peer-to-peer; NoSQL systems to handle big data problems.	
4		**Mining Data Streams**	11
	4.1	The Stream Data Model: A Data-Stream-Management System, Examples of Stream Sources, Stream Queries, Issues in Stream Processing.	
	4.2	Sampling Data techniques in a Stream	
	4.3	Filtering Streams: Bloom Filter with Analysis.	
	4.4	Counting Distinct Elements in a Stream,Count-Distinct Problem, Flajolet-Martin Algorithm, Combining Estimates, Space Requirements	
	4.5	Counting Ones in a Window: The Cost of Exact Counts,The Datar-Gionis-Indyk-Motwani Algorithm, Query Answering in theDGIM Algorithm, Decaying Windows.	
5		**Real-Time Big Data Models**	4
	5.1	A Model for Recommendation Systems, Content-Based Recommendations, Collaborative Filtering	
	5.2	Case Study: Product Recommendation	
	5.3	Social Networks as Graphs, Clustering of Social-Network Graphs, Direct Discovery of Communities in a social graph	
6		**Data Analytics with R**	4
	6.1	Exploring Basic features of R, Exploring RGUI, Exploring RStudio, Handling Basic Expressions in R, Variables in R, Working with Vectors, Storing and Calculating Values in R, Creating and using Objects, Interacting with users, Handling data in R workspace, Executing Scripts, Creating Plots, Accessing help and documentation in R	
	6.2	Reading datasets and Exporting data from R, Manipulating and Processing Data in R, Using functions instead of script, built-in functions in R	
	6.3	Data Visualization: Types, Applications	

Textbooks:	
1	Cre Anand Rajaraman and Jeff Ullman —Mining of Massive Datasets|, Cambridge UniversityPress
2	Alex Holmes —Hadoop in Practice|, Manning Press, Dreamtech Press.
3	Dan Mcary and Ann Kelly —Making Sense of NoSQL| – A guide for managers and the rest of us, Manning Press.
4	DT Editorial Services, "Big Data Black Book", Dreamtech Press
5	EMC Education Services,"Data Science and Big Data Analytics",Wiley

Course Code	Course Name	Credit
CSDC7013	Natural Language Processing	03

Pre-requisite: Theory of Computer Science, System Programming &Compiler Construction	
Course Objectives: The course aims	
1	To define natural language processing and to learn various stages of natural language processing.
2	To describe basic concepts and algorithmic description of the main language levels: Morphology, Syntax, Semantics, and Pragmatics &Discourse analysis.
3	To design and implement various language models and POS tagging techniques.
4	To design and learn NLP applications such as Information Extraction, Question answering.
5	To design and implement applications based on natural language processing.
Course Outcomes: Students will be able	
1	To describe the field of natural language processing.
2	To design language model for word level analysis for text processing.
3	To design various POS tagging techniques and parsers.
4	To design, implement and test algorithms for semantic and pragmatic analysis.
5	To formulate the discourse segmentation and anaphora resolution.
6	To apply NLP techniques to design real world NLP applications.

Module		Detailed Content	Hours
1	1.1	**Introduction to NLP**	3
		Origin & History of NLP; Language, Knowledge and Grammar in language processing; Stages in NLP;Ambiguities and its types in English and Indian Regional Llanguages; Challenges of NLP;Applications of NLP	
	1.2	**Self-Learning topics:** Variety types of tools for regional languages pre-processing and other functionalities	
2	2.1	**Word Level Analysis**	9
		Basic Terms: Tokenization, Stemming, Lemmatization; Survey of English Morphology, Inflectional Morphology, Derivational Morphology; Regular expression with types; Morphological Models: Dictionary lookup, finite state morphology; Morphological parsing with FST (Finite State Transducer);Lexicon free FST Porter Stemmer algorithm; Grams and its variation: Bigram, Trigram; Simple (Unsmoothed) N-grams; N-gram Sensitivity to the Training Corpus; Unknown Words: Open versus closed vocabulary tasks; Evaluating N-grams: Perplexity;	

		Smoothing: Laplace Smoothing, Good-Turing Discounting;	
	2.2	**Self-Learning topics:** Noisy channel models, various edit distance, Advance Issues in Language Modelling	
3	3.1	**Syntax analysis**	10
		Part-Of-Speech tagging(POS); Tag set for English (Upenn Treebank); Difficulties /Challenges in POS tagging; Rule-based, Stochastic and Transformation-based tagging; Generative Model: Hidden Markov Model (HMM Viterbi) for POS tagging; Issues in HMM POS tagging; Discriminative Model: Maximum Entropy model, Conditional random Field (CRF);Parsers: Top down and Bottom up; Modelling constituency; Bottom Up Parser: CYK, PCFG (Probabilistic Context Free Grammar), Shift Reduce Parser; Top Down Parser: Early Parser, Predictive Parser	
	3.2	**Self-Learning topics:** Evaluating parsers, Parsers based language modelling, Regional languages POS tree banks	
4	4.1	**Semantic Analysis**	7
		Introduction, meaning representation; Lexical Semantics; Corpus study; Study of Various language dictionaries like WorldNet, Babelnet; Relations among lexemes & their senses –Homonymy, Polysemy, Synonymy, Hyponymy; Semantic Ambiguity; Word Sense Disambiguation (WSD); Knowledge based approach(Lesk's Algorithm), Supervised (Naïve Bayes, Decision List),Introduction to Semi-supervised method (Yarowsky) Unsupervised (Hyperlex)	
	4.2	**Self-Learning topics:** Dictionaries for regional languages, Distributional Semantics, Topic Models	
5	5.1	**Pragmatic & Discourse Processing**	5
		Discourse: Reference Resolution, Reference Phenomena, Syntactic & Semantic constraint on coherence; Anaphora Resolution using Hobbs and Cantering Algorithm	
	5.2	**Self-Learning topics:** Discourse segmentation, Conference resolution	
6	6.1	**Applications of NLP**	5
		Case studies on (preferable in regional language):Machine translation; Text Summarization; Sentiment analysis; Information retrieval; Question Answering system	
	6.2	**Self-Learning topics:** Applications based on Deep Neural Network with NLP such as LSTM network, Recurrent Neural network etc.	

Textbooks:	
1	Daniel Jurafsky, James H. and Martin, Speech and Language Processing, Second Edition, Prentice Hall, 2008.
2	Christopher D.Manning and HinrichSchutze, Foundations of Statistical Natural Language Processing, MIT Press, 1999.
References:	
1	Siddiqui and Tiwary U.S., Natural Language Processing and Information Retrieval, Oxford University Press, 2008.
2	Daniel M Bikel and ImedZitouni — Multilingual natural language processing applications

Course Code:	Course Title	Credit
CSDC7022	Blockchain	3

Prerequisite: Cryptography and System Security	
Course Objectives:	
1	Understand blockchain platforms and its terminologies.
2	Understand the use of cryptography required for blockchain.
3	Understand smart contracts, wallets, and consensus protocols.
4	Design and develop blockchain applications
Course Outcomes:	
1	Explain blockchain concepts.
2	Apply cryptographic hash required for blockchain.
3	Apply the concepts of smart contracts for an application.
4	Design a public blockchain using Ethereum.
5	Design a private blockchain using Hyperledger.
6	Use different types of tools for blockchain applications.

Module		Content	Hrs
1		**Introduction to Blockchain**	6
	1.1	What is a blockchain, Origin of blockchain (cryptographically secure hash functions), Foundation of blockchain: Merkle trees	
	1.2	Components of blockchain, Block in blockchain, Types: Public, Private, and Consortium, Consensus Protocol, Limitations and Challenges of blockchain	
2		**Cryptocurrency**	6
	2.1	Cryptocurrency: Bitcoin, Altcoin, and Tokens (Utility and Security), Cryptocurrency wallets: Hot and cold wallets, Cryptocurrency usage, Transactions in Blockchain, UTXO and double spending problem	
	2.2	Bitcoin blockchain: Consensus in Bitcoin, Proof-of-Work (PoW), Proof-of-Burn (PoB), Proof-of-Stake (PoS), and Proof-of-Elapsed Time (PoET), Life of a miner, Mining difficulty, Mining pool and its methods	
3		**Programming for Blockchain**	8
	3.1	Introduction to Smart Contracts, Types of Smart Contracts, Structure of a Smart Contract, Smart Contract Approaches, Limitations of Smart Contracts	
	3.2	Introduction to Programming: Solidity Programming – Basics, functions, Visibility and Activity Qualifiers, Address and Address Payable, Bytes and Enums, Arrays-Fixed and Dynamic Arrays, Special Arrays-Bytes and strings, Struct, Mapping, Inheritance, Error handling	
	3.3	Case Study – Voting Contract App, Preparing for smart contract development	

4		**Public Blockchain**	**8**
		Introduction to Public Blockchain, Ethereum and its Components, Mining in Ethereum, Ethereum Virtual Machine (EVM), Transaction, Accounts, Architecture and Workflow, Comparison between Bitcoin and Ethereum	
		Types of test-networks used in Ethereum, Transferring Ethers using Metamask, Mist Wallet, Ethereum frameworks, Case study of Ganache for Ethereum blockchain. Exploring etherscan.io and ether block structure	
5		**Private Blockchain**	**8**
	5.1	Introduction, Key characteristics, Need of Private Blockchain, Smart Contract in a Private Environment, State Machine Replication, Consensus Algorithms for Private Blockchain - PAXOS and RAFT, Byzantine Faults: Byzantine Fault Tolerant (BFT) and Practical BFT	
	5.2	Introduction to Hyperledger, Tools and Frameworks, Hyperledger Fabric, Comparison between Hyperledger Fabric & Other Technologies	
	5.3	Hyperledger Fabric Architecture, Components of Hyperledger Fabric: MSP, Chain Codes, Transaction Flow, Working of Hyperledger Fabric, Creating Hyperledger Network, Case Study of Supply Chain Management using Hyperledger	
6		**Tools and Applications of Blockchain**	**3**
		Corda, Ripple, Quorum and other Emerging Blockchain Platforms, Blockchain in DeFi: Case Study on any of the Blockchain Platforms.	

Textbooks:	
1	Blockchain Technology, Chandramouli Subramanian, Asha A. George, Abhillash K. A and Meena Karthikeyen, Universities Press.
2	Mastering Ethereum, Building Smart Contract and Dapps, Andreas M. Antonopoulos Dr. Gavin Wood, O'reilly.
3	Imran Bashir, Mastering Blockchain: A deep dive into distributed ledgers, consensus protocols, smart contracts, DApps, cryptocurrencies, Ethereum, and more, 3rd Edition, Packt Publishing
References:	
1	Blockchain for Beginners, Yathish R and Tejaswini N, SPD
2	Blockchain Basics, A non Technical Introduction in 25 Steps, Daniel Drescher, Apress.
3	Blockchain with Hyperledger Fabric,Luc Desrosiers, Nitin Gaur, Salman A. Baset, Venkatraman Ramakrishna, Packt Publishing

Assessment:

Internal Assessment:

Assessment consists of two class tests of 20 marks each. The first class test is to be conducted when approx. 40% syllabus is completed and second class test when additional 40% syllabus is completed. Duration of each test shall be one hour.

Course Code	Course Name	Credits
ILO 7016	Cyber Security and Laws	03

Objectives:

1. To understand and identify different types cybercrime and cyber law
2. To recognized Indian IT Act 2008 and its latest amendments
3. To learn various types of security standards compliances

Outcomes: Learner will be able to…

1. Understand the concept of cybercrime and its effect on outside world
2. Interpret and apply IT law in various legal issues
3. Distinguish different aspects of cyber law
4. Apply Information Security Standards compliance during software design and development

Sr. No.	Detailed Contents	Hrs
01	**Introduction to Cybercrime:** Cybercrime definition and origins of the world, Cybercrime and information security, Classifications of cybercrime, Cybercrime and the Indian ITA 2000, A global Perspective on cybercrimes.	4
02	**Cyber offenses & Cybercrime:** How criminal plan the attacks, Social Engg, Cyber stalking, Cyber café and Cybercrimes, Botnets, Attack vector, Cloud computing, Proliferation of Mobile and Wireless Devices, Trends in Mobility, Credit Card Frauds in Mobile and Wireless Computing Era, Security Challenges Posed by Mobile Devices, Registry Settings for Mobile Devices, Authentication Service Security, Attacks on Mobile/Cell Phones, Mobile Devices: Security Implications for Organizations, Organizational Measures for Handling Mobile, Devices-Related Security Issues, Organizational Security Policies and Measures in Mobile Computing Era, Laptops	9
03	**Tools and Methods Used in Cyberline** Phishing, Password Cracking, Key loggers and Spywares, Virus and Worms, Steganography, DoS and DDoS Attacks, SQL Injection, Buffer Over Flow, Attacks on Wireless Networks, Phishing, Identity Theft (ID Theft)	6
04	**The Concept of Cyberspace** E-Commerce , The Contract Aspects in Cyber Law ,The Security Aspect of Cyber Law ,The Intellectual Property Aspect in Cyber Law , The Evidence Aspect in Cyber Law , The Criminal Aspect in Cyber Law, Global Trends in Cyber Law , Legal Framework for Electronic Data Interchange Law Relating to Electronic Banking , The Need for an Indian Cyber Law	8
05	**Indian IT Act.** Cyber Crime and Criminal Justice: Penalties, Adjudication and Appeals Under the IT Act, 2000, IT Act. 2008 and its Amendments	6
06	**Information Security Standard compliances** SOX, GLBA, HIPAA, ISO, FISMA, NERC, PCI.	6

Assessment:

Internal Assessment for 20 marks:

Consisting **Two Compulsory Class Tests**

First test based on approximately 40% of contents and second test based on remaining contents (approximately 40% but excluding contents covered in Test I)

End Semester Examination:

Weightage of each module in end semester examination will be proportional to number of respective lecture hours mentioned in the curriculum.

1. Question paper will comprise of total **six questions**, **each carrying 20 marks**
2. **Question 1** will be **compulsory** and should **cover maximum contents of the curriculum**
3. **Remaining questions will be mixed in nature** (for example if Q.2 has part (a) from module 3 then part (b) will be from any module other than module 3)
4. Only **Four questions need to be solved**.

REFERENCES:

1. Nina Godbole, Sunit Belapure, *Cyber Security*, Wiley India, New Delhi
2. The Indian Cyber Law by Suresh T. Vishwanathan; Bharat Law House New Delhi
3. The Information technology Act, 2000; Bare Act- Professional Book Publishers, New Delhi.
4. Cyber Law & Cyber Crimes By Advocate Prashant Mali; Snow White Publications, Mumbai
5. Nina Godbole, *Information Systems Security,* Wiley India, New Delhi
6. Kennetch J. Knapp, *Cyber Security &Global Information Assurance* Information Science Publishing.
7. William Stallings, *Cryptography and Network Security,* Pearson Publication
8. Websites for more information is available on : The Information Technology ACT, 2008- TIFR : https://www.tifrh.res.in
9. Website for more information , A Compliance Primer for IT professional : https://www.sans.org/reading-room/whitepapers/compliance/compliance-primer-professionals-33538

Lab Code	Lab Name	Credit
CSL70011	Machine Learning Lab	1

Prerequisite: Data Structures, Analysis of Algorithms	
Lab Objectives:	
1	To introduce the basic concepts and techniques of Machine Learning.
2	To acquire in depth understanding of various supervised and unsupervised algorithms
3	To be able to apply various ensemble techniques for combining ML models.
4	To demonstrate dimensionality reduction techniques.
Lab Outcomes: At the end of the course, the students will be able to	
1	To implement an appropriate machine learning model for the given application.
2	To implement ensemble techniques to combine predictions from different models.
3	To implement the dimensionality reduction techniques.

Suggested List of Experiments	
Sr. No.	**Title of Experiment**
1	To implement Linear Regression.
2	To implement Logistic Regression.
3	To implement Ensemble learning (bagging/boosting)
4	To implement multivariate Linear Regression.
5	To implement SVM
6	To implement PCA/SVD/LDA
7	To implement Graph Based Clustering
8	To implement DB Scan
9	To implement CART
10	To implement LDA

Term Work:	
1	Term work should consist of 6 experiments.
2	Journal must include one mini project/case study on any machine learning application.
3	The final certification and acceptance of term work ensures the satisfactory performance of laboratory work and minimum passing marks in term work.
4	Total 25 Marks (Experiments & Assignments: 15-marks, Attendance: 05-marks, mini project: 05-marks)
Oral & Practical exam.	
Based on the entire syllabus **CSC7011** Machine Learning **and CSL7011**: Machine Learning Lab	

Lab Code	Lab Name	Credit
CSL7012	Big Data Analytics Lab	1

Prerequisite: C Programming Language.	
Lab Objectives: Students will be able to	
1	Solve Big Data problems using Map Reduce Technique and apply to various algorithms.
2	Identify various types of NoSQL databases and execute NOSQL commands
3	Understand implementation of various analytic techniques using Hive/PIG/R/Tableau, etc.
4	Apply streaming analytics to real time applications.
Lab Outcomes:	
1	To interpret business models and scientific computing paradigms, and apply software tools for big data analytics.
2	To implement algorithms that uses Map Reduce to apply on structured and unstructured data
3	To perform hands-on NoSql databases such as Cassandra, HadoopHbase, MongoDB, etc.
4	To implement various data streams algorithms.
5	To develop and analyze the social network graphs with data visualization techniques.

Suggested List of Experiments **(Select a case study and perform the experiments 1 to 8.).**	
Star (*) marked experiments are compulsory.	
Sr. No.	**Name of the Experiment**
1*	Hadoop HDFS Practical: -HDFS Basics, Hadoop Ecosystem Tools Overview. -Installing Hadoop. -Copying File to Hadoop. -Copy from Hadoop File system and deleting file. -Moving and displaying files in HDFS. -Programming exercises on Hadoop
2	Use of Sqoop tool to transfer data between Hadoop and relational database servers. a. Sqoop - Installation. b. To execute basic commands of Hadoop eco system componentSqoop.
3*	To install and configure MongoDB/ Cassandra/ HBase/ Hypertable to execute NoSQL commands
4	Experiment on Hadoop Map-Reduce: -Write a program to implement a word count program using MapReduce.
5	Experiment on Hadoop Map-Reduce: -Implementing simple algorithms in Map-Reduce: Matrix multiplication, Aggregates, Joins, Sorting, Searching, etc
6	Create HIVE Database and Descriptive analytics-basic statistics.
7*	Data Stream Algorithms (any one): - Implementing DGIM algorithm using any Programming Language - Implement Bloom Filter using any programming language Implement Flajolet Martin algorithm using any programming language
8	Social Network Analysis using R (for example: Community Detection Algorithm)
9	Data Visualization using Hive/PIG/R/Tableau/.
10	Exploratory Data Analysis using Spark/ Pyspark.

11*	Mini Project: One real life large data application to be implemented (Use standard Datasets available on the web). - Streaming data analysis – use flume for data capture, HIVE/PYSpark for analysis of twitter data, chat data, weblog analysis etc. - Recommendation System (for example: Health Care System, Stock Market Prediction, Movie Recommendation, etc.) SpatioTemporal DataAnalytics

Useful Links:	
1	https://www.coursera.org/learn/hadoop#syllabus
2	https://www.coursera.org/learn/introduction-mongodb#syllabus
3	https://www.coursera.org/learn/data-visualization-tableau?specialization=data-visualization#syllabus
4	https://www.coursera.org/learn/introduction-to-big-data-with-spark-hadoop#syllabus

Term Work:	
1	Term work should consist of 8 experiments.
2	The final certification and acceptance of term work ensures satisfactory performance of laboratory work and minimum passing marks in term work.
3	The final certification and acceptance of term work ensures satisfactory performance of laboratory work and minimum passing marks in term work. Total 25 Marks (Experiments: 15-marks, Attendance Theory & Practical: 05-marks, Assignment: 05-marks)
Oral & Practical exam	
	Based on the entire syllabus of and CSC702 : Big Data Analytics and CSL702 Big Data Analytics Lab

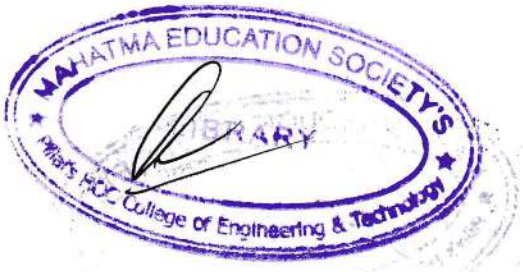
Lab Code	Lab Name	Credit
CSDL7013	**Natural Language processing Lab**	**1**

Prerequisite: Java/Python	
Lab Objectives: The course aims	
1	To understand the key concepts of NLP.
2	To learn various phases of NLP.
3	To design and implement various language models and POS tagging techniques.
4	To understand various NLP Algorithms
5	To learn NLP applications such as Information Extraction, Sentiment Analysis, Question answering, Machine translation etc.
6	To design and implement applications based on natural language processing
Lab Outcomes: Learners will be able	
1	Apply various text processing techniques.
2	Design language model for word level analysis.
3	Model linguistic phenomena with formal grammar.
4	Design, implement and analyze NLP algorithms.
5	To apply NLP techniques to design real world NLP applications such as machine translation, sentiment analysis, text summarization, information extraction, Question Answering system etc.
6	Implement proper experimental methodology for training and evaluating empirical NLP systems.

Suggested List of Experiments **(Select a case study and perform the experiments 1 to 8.)**.	
Star (*) marked experiments are compulsory.	
Sr. No.	**Name of the Experiment**
1	Study various applications of NLP and Formulate the Problem Statement for Mini Project based on chosen real world NLP applications: [Machine Translation, Text Categorization, Text summarization, chat Bot, Plagarism, Spelling & Grammar checkers, Sentiment / opinion analysis, Question answering, Personal Assistant, Tutoring Systems, etc.]
2	Apply various text preprocessing techniques for any given text : Tokenization and Filtration & Script Validation.
3	Apply various other text preprocessing techniques for any given text : Stop Word Removal, Lemmatization / Stemming.
4	Perform morphological analysis and word generation for any given text.
5	Implement N-Gram model for the given text input.
6	Study the different POS taggers and Perform POS tagging on the given text.
7	Perform Chunking for the given text input.
8	Implement Named Entity Recognizer for the given text input.
9	Implement Text Similarity Recognizer for the chosen text documents.

10	Exploratory data analysis of a given text (Word Cloud)
11	Mini Project Report: For any one chosen real world NLP application.
13	Implementation and Presentation of Mini Project

Term Work:	
1	Study various applications of NLP and Formulate the Problem Statement for Mini Project based on chosen real world NLP applications: [Machine Translation, Text Categorization, Text summarization, chat Bot, Plagarism, Spelling & Grammar checkers, Sentiment / opinion analysis, Question answering, Personal Assistant, Tutoring Systems, etc.]
2	Apply various text preprocessing techniques for any given text: Tokenization and Filtration & Script Validation.
3	Apply various other text preprocessing techniques for any given text: Stop Word Removal, Lemmatization / Stemming.



Lab Code	Lab Name	Credit
CSDL7022	Blockchain Lab	1

Prerequisite: Cryptography and Network Security	
Lab Objectives:	
1	To explore Blockchain concepts.
2	To implement public and private Blockchain.
3	To create applications using Blockchain.
Lab Outcomes: At the end of the course, the students will be able to	
1	Creating Cryptographic hash using merkle tree.
2	Design Smart Contract using Solidity.
3	Implementing ethereum blockchain using Geth.
4	Demonstrate the concept of blockchain in real world application.

Suggested List of Experiments	
Sr. No.	**Title of Experiment**
1	Cryptography in Blockchain, Merkle root tree hash
2	Creating Smart Contract using Solidity and Remix IDE.
3	Creating Transactions using Solidity and Remix IDE
4	Embedding wallet and transaction using Solidity
5	Blockchain platform ethereum using Geth.
6	Blockchain platform Ganache.
7	Case Study on Hyperledger
8	Case Study on Other Blockchain platforms.
9	Creating a blockchain Application

Term Work:	
1	Term work should consist of 8 experiments and one mini project.
2	Journal must include at least 2 assignments on content of theory and practical of "Blockchain Lab"
3	The final certification and acceptance of term work ensures satisfactory performance of laboratory work and minimum passing marks in term work.
4	Total 25 Marks (Experiments: 15-marks, Attendance Theory & Practical: 05-marks, Assignments: 05-marks)