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.

5	In question paper weightage of each module will be proportional to number of respective lecture hours as mentioned in the syllabus.

Useful Digital Links	
1	Data sets for Machine Learning algorithms: https://www.kaggle.com/datasets
2	Machine Learning repository- https://archive.ics.uci.edu/ml/index.php
3	Machine Learning from Coursera
4	https://towardsdatascience.com/machine-learning/home
5	https://onlinecourses.nptel.ac.in/noc21_cs85/preview

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

References:	
1	Bill Franks , —Taming The Big Data Tidal Wave: Finding Opportunities In HugeData StreamsWithAdvancedAnalytics|,Wiley
2	Chuck Lam, —Hadoop inAction|, Dreamtech Press
3	Jared Dean, —Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners|,Wiley India Private Limited, 2014.
4	Jiawei Han and Micheline Kamber, —Data Mining: Concepts and Techniques|, Morgan Kaufmann Publishers, 3rd ed, 2010.
5	Lior Rokach and Oded Maimon, —Data Mining and Knowledge Discovery Handbook|, Springer, 2nd edition,2010.
6	Ronen Feldman and James Sanger, —The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data|, Cambridge University Press, 2006.
7	Vojislav Kecman, —Learning and Soft Computing|, MITPress, 2010.

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 additional40% syllabus is completed. Duration of each test shall be one hour.	
End Semester Theory Examination:	
1	Question paper will consist of 6 questions, each carrying 20 marks.
2	The students need to solve a total of 4 questions.
3	Question No.1 will be compulsory and based on the entire syllabus.
4	Remaining question (Q.2 to Q.6) will be selected from all the modules.

Useful Links	
1	https://nptel.ac.in/courses/106104189
2	https://www.coursera.org/specializations/big-data#courses
3	https://www.digimat.in/nptel/courses/video/106106169/L01.html
4	https://www.coursera.org/learn/nosql-databases#syllabus
5	https://www.coursera.org/learn/basic-recommender-systems#syllabus

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:

	from theory to practice, IBM Press, 2013.
3	Alexander Clark, Chris Fox, Shalom Lappin — The Handbook of Computational Linguistics and Natural Language Processing, John Wiley and Sons, 2012.
4	Nitin Indurkhya and Fred J. Damerau, —Handbook of Natural Language Processing, Second Edition, Chapman and Hall/CRC Press, 2010.
5	Niel J le Roux and SugnetLubbe, A step by step tutorial: An introduction into R application and programming.
6	Steven Bird, Ewan Klein and Edward Loper, Natural language processing with Python: analyzing text with the natural language toolkit, O'Reilly Media, 2009.

Digital References :	
1	http://www.cse.iitb.ac.in/~cs626-449
2	http://cse24-iiith.virtual-labs.ac.in/#
3.	https://nptel.ac.in/courses/106105158

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.	
End Semester Theory Examination:	
1	Question paper will consist of 6 questions, each carrying 20 marks.
2	The students need to solve a total of 4 questions.
3	Question No.1 will be compulsory and based on the entire syllabus.
4	Remaining question (Q.2 to Q.6) will be selected from all the modules.

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.

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 question need to be solved.
5	In question paper weightage of each module will be proportional to the number of respective lecture hours as mention in the syllabus.

Digital Useful Links	
1	Blockchain By Example, Bellaj Badr, Richard Horrocks, Xun (Brian) Wu, November 2018, Implement decentralized blockchain applications to build scalable Dapps.
2	Blockchain for Business, https://www.ibm.com/downloads/cas/3EGWKGX7.
3	https://www.hyperledger.org/use/fabric
4	NPTEL: https://onlinecourses.nptel.ac.in/noc19_cs63/preview

Course Code	Course Name	Credits
ILO 7013	Management Information System	03

Objectives:

1. The course is blend of Management and Technical field.
2. Discuss the roles played by information technology in today's business and define various technology architectures on which information systems are built
3. Define and analyze typical functional information systems and identify how they meet the needs of the firm to deliver efficiency and competitive advantage
4. Identify the basic steps in systems development

Outcomes: Learner will be able to…

1. Explain how information systems Transform Business
2. Identify the impact information systems have on an organization
3. Describe IT infrastructure and its components and its current trends
4. Understand the principal tools and technologies for accessing information from databases to improve business performance and decision making
5. Identify the types of systems used for enterprise-wide knowledge management and how they provide value for businesses

Sr. No.	Detailed Contents	Hrs
01	Introduction To Information Systems (IS): Computer Based Information Systems, Impact of IT on organizations, Importance of IS to Society. Organizational Strategy, Competitive Advantages and IS	4
02	Data and Knowledge Management: Database Approach, Big Data, Data warehouse and Data Marts, Knowledge Management Business intelligence (BI): Managers and Decision Making, BI for Data analysis and Presenting Results	7
03	Ethical issues and Privacy: Information Security. Threat to IS, and Security Controls	7
04	Social Computing (SC): Web 2.0 and 3.0, SC in business-shopping, Marketing, Operational and Analytic CRM, E-business and E-commerce – B2B B2C. Mobile commerce.	7
05	Computer Networks Wired and Wireless technology, Pervasive computing, Cloud computing model.	6
06	Information System within Organization: Transaction Processing Systems, Functional Area Information System, ERP and ERP support of Business Process. Acquiring Information Systems and Applications: Various System development life cycle models.	8

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. Kelly Rainer, Brad Prince, Management Information Systems, Wiley
2. K.C. Laudon and J.P. Laudon, Management Information Systems: Managing the Digital Firm, 10th Ed., Prentice Hall, 2007.
3. D. Boddy, A. Boonstra, Managing Information Systems: Strategy and Organization, Prentice Hall, 2008