RASHTRASANT TUKADOJI MAHARAJ NAGPUR UNIVERSITY, NAGPUR
FOUR YEAR BACHELOR OF ENGINEERING (B. Tech.) DEGREE COURSE
SEMESTER: SEVENTH (C.B.C.S.)
BRANCH: COMPUTER SCIENCE & ENGINEERING(DATA SCIENCE)

Sr. No.	Subject	Teaching Scheme			Evaluation Schemes			Credits	Category
		L	T	P	CA	UE	TOTAL		
1	Data Analytics	3	-	-	30	70	100	3	PCC-CS
2	Data Mining and Visualization	3	-	-	30	70	100	3	PCC-CS
3	Elective-IV	3	-	-	30	70	100	3	PEC-CS
4	Elective-V	3	-	-	30	70	100	3	PEC-CS
5	Open Elective-II	3	-	-	30	70	100	3	OEC
6	Machine Learning and Data analytics Lab	-	-	2	25	25	50	1	PCC-CS
7	Data Mining and visualization Lab	-	-	2	25	25	50	1	PCC-CS
8	Project	-	-	6	25	25	50	2	PROJ-CS
9	Research Methodology(Audit Course)	2	-	-	50	-	-	Audit	HSMC
	Total	17	0	10	275	425	650	19	

Table title: B. Tech. Seventh Semester CSE(DS)

Elective-IV:
i)   Deep Learning   ii) Information Extraction and Retrieval
iii) Time Series Analysis and forecasting

Elective-V:
i) Big Data Analytics   ii) Data Modeling and Simulation
iii) Data Warehousing

Open Elective-II
i) Machine Learning

Sr. No.	Subject	Teaching Scheme			Evaluation Schemes			Credits	Category
		L	T	P	CA	UE	TOTAL		
1	Industry Project	-	-	16	75	75	150	8	PROJ-CS
2	MOOCS Course-1	-	-	-	-	-	-	3	OEC
3	MOOCS Course-2	-	-	-	-	-	-	3	OEC
	Total	0	0	16	75	75	150	14	

Table title: B. Tech. Eight Semester CSE(DS)

MOOCS Course-1:
i) Social Networks

ii) Reinforcement Learning iii) Introduction to Industry 4.0 and Industrial Internet of Things

MOOCS Course-2

i) Cyber Security and Privacy   ii) Block Chain and its Applications   iii) Ethical Hacking

Subject Name: Data Analytics

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 (Theory)	Hrs	3	-	3	30	70	100

Course Objectives:	
1	To provide strong foundation for data analytics and application area related to it.
2	To understand the underlying core concepts and emerging technologies in data analytics.

Course Outcomes	
At the end of this course students are able to:	
CO1	Explore the fundamental concepts of data analytics
CO2	Understand data analysis techniques for applications handling large data
CO3	Understand various machine learning algorithms used in data analytics process
CO4	Visualize and present the inference using various tools
CO5	Learn to think through the ethics surrounding privacy, data sharing and algorithmic decision-making

Unit 1	**INTRODUCTION** Data Analytics - Types – Phases - Quality and Quantity of data – Measurement - Exploratory data analysis - Business Intelligence.	7Hr
Unit 2	**BIG DATA** Big Data and Cloud technologies - Introduction to HADOOP: Big Data, Apache Hadoop, MapReduce - Data Serialization - Data Extraction - Stacking Data - Dealing with data.	7Hr
Unit 3	**DATA VISUALIZATION** Introduction to data visualization – Data visualization options – Filters – Dashboard development tools – Creating an interactive dashboard with dc.js - summary.	7Hr
Unit 4	**ANALYTICS AND MACHINE LEARNING** Machine learning – Modeling Process – Training model – Validating model – Predicting new observations –Supervised learning algorithms – Unsupervised learning algorithms.	7Hr
Unit 5	**ETHICS AND RECENT TRENDS** Data Science Ethics – Doing good data science – Owners of the data - Valuing different aspects of privacy - Getting informed consent - The Five Cs – Diversity – Inclusion – Future Trends.	8Hr

Text Books	
1	Dr Anil Maheshwari, Data Analytics Made Accessible, Publisher: Amazon.com Services LLC.
2	Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning: with Applications in R, Springer, 1st edition, 2013.
Reference Books	
1	Davy Cielen, Arno D. B. Meysman, Mohamed Ali, Introducing Data Science, Manning Publications Co., 1st edition, 2016
2	Eric Siegel, Predictive Analytics The Power to Predict Who Will Click, Buy, Lie, or Die, 2nd Ed., Wiley.

Subject Name: Data mining and Visualization

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 (Theory)	Hrs	3	-	3	30	70	100

Course Objectives:	
1	To preprocess and analyze data, to choose relevant models and algorithms for respective applications and to develop research interest towards advances in data mining
2	To understand the components involved in visualization design

Course Outcomes	
At the end of this course students are able to:	
CO1	Understand different types of data to be mined
CO2	Evaluate different models used for classification and Clustering
CO3	Categorize the scenario for applying different data mining techniques
CO4	understand the type of data impacts the type of visualization
CO5	Recognize the feasibility of Data Preparation and Univariate Graphs

Unit		Hr
Unit 1	Data Mining – Kinds of data to be mined – Kinds of patterns to be mined – Technologies – Targeted Applications - Major Issues in Data Mining – Data Objects and Attribute Types – Measuring Data similarity and dissimilarity - Data Cleaning –Data Integration - Data Reduction – Data Transformation – Data Discretization.	8Hr
Unit 2	**Clustering:** Introduction, Clustering, Cluster Analysis, Clustering Methods- K means, Hierarchical clustering, Agglomerative clustering, Divisive clustering, clustering and segmentation software, evaluating clusters.	7Hr
Unit 3	**Web Mining:** Introduction, Terminologies, Categories of Web Mining – Web Content Mining, Web Structure Mining, Web Usage Mining, Applications of Web Mining, and Agent based and Data base approaches, Web mining Software. Applications of Data mining.	7Hr
Unit 4	Data Preparation: Importing Data - Text files -Excel spreadsheets -Statistical packages - Databases - Cleaning Data: Selecting variables - Selecting observations - Creating/Recoding variables - Summarizing data - Using pipes -Reshaping data - Missing data - Introduction to ggplot2 -ggplot- geoms - grouping scales - facets - labels- themes - Placing the data and mapping options-Graphs as objects	7Hr
Unit 5	Univariate Graphs - Categorical: Bar Chart -Pie Chart - Tree Map- Quantitative - Histogram - Kernel Density plot - Dot Chart - Bivariate Graphs - Categorical vs. Categorical: Stacked bar chart - Grouped bar chart - Segmented bar chart - Improving the color and labeling - Other plots - Quantitative vs. Quantitative: Scatterplot - Line plot- Categorical vs. Quantitative: Bar chart (on summary statistics) - Grouped kernel density plots - Box plots	7Hr

Text Books	
1	K.P. Soman, Shyam Diwakar and V. Ajay, —Insight into Data Mining Theory and Practice, Eastern Economy Edition, Prentice Hall of India, 2006.
2	Rob Kabacoff , Data Visualization with R, Bookdown, 2018. Chapters: 1-13 https://rkabacoff.github.io/datavis/
Reference Books	
1	Alex Berson and Stephen J.Smith, —Data Warehousing, Data Mining & OLAP‖, Tata McGraw – Hill Edition, 35th Reprint 2016.
2	Kirthi Raman - Mastering Python Data Visualization -Packt Publishing -2015

Subject Name: ELECTIVE IV: Deep Learning

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 (Theory)	Hrs	3	-	3	30	70	100

Course Objectives:	
1	To understand the evolution of deep architectures
2	To apply deep learning principles to Natural Language Processing
	To assess the challenges of multimodality and reinforcement learning

Course Outcomes	
At the end of this course students are able to:	
CO1	Understand the evolution of deep architectures
CO2	Understand the Deep Learning in Computer Vision
CO3	Apply deep learning principles to Natural Language Processing
CO4	Assess the challenges of multimodality learning
CO5	Analyze the Deep Learning Hacks.

Unit 1	Deep Architecture: need, applications, Hyper-parameters in Deep Neural Networks Encoding, Layers, Loss function, Learning Rate, Momentum and Optimization, Regularization and dropout, Batch Norms, vanishing gradient problem, and ways to mitigate it	7Hr
Unit 2	Deep Learning in Computer Vision: Origins of CNNs- Convolutional Neural Networks -Fine-tuning CNNs - Popular CNN architectures	7Hr
Unit 3	NLP - Vector Representation: Traditional NLP - Deep learning NLP – Applications. Advanced Natural Language Processing: Deep learning for text - Recurrent neural networks - Long short-term memory network – Applications	7Hr
Unit 4	Multimodality: What is multimodality learning? - Challenges of multimodality learning- Image captioning - Visual question answering - Multi-source based self-driving Deep Reinforcement Learning: What is reinforcement learning (RL)? - Deep reinforcement learning - Implementing reinforcement learning	8Hr
Unit 5	Deep Learning Hacks: Massaging your data - Tricks in training - Fine-tuning - Model compression Deep Learning Trends: Recent models for deep learning - Novel applications	7Hr

Text Books	
1	Deep Learning Author Ian Goodfellow and Yoshua Bengio and Aaron Courville. Publisher MIT Press Edition 2016
2	Anurag Bhardwaj, Wei Di, JianingWei,"**Deep Learning Essentials**", Packt Publishing, 2018. Chapters:1-10
Reference Books	
1	Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016.
2	Bishop, C. ,M., Pattern Recognition and Machine Learning, Springer, 2006.

Subject Name: Elective-IV Information Extraction and Retrieval

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 (Theory)	Hrs	3	-	3	30	70	100

Course Objectives:	
1	To understand the basics of the information retrieval process.
2	To evaluate the performance of the IR system.

Course Outcomes	
At the end of this course students are able to:	
CO1	Recognize the basic Concepts of information retrieval, and the distributed and multimedia IR
CO2	Apply various storage and searching techniques
CO3	Evaluate performance of information retrieval and ontology
CO4	Analyze Distributed and Multimedia IR
CO5	To understand the various applications of Information Retrieval giving emphasis to multimedia and distributed IR - web Search.

Unit 1	Introduction: Basic Concepts of IR - Data Retrieval & Information Retrieval - IR system block diagram. Automatic Text Analysis: Luhn's ideas - Conflation Algorithm - Indexing and Index Term Weighing - Probabilistic Indexing - Automatic Classification. Measures of Association - Different Matching Coefficient - Classification Methods - Cluster Hypothesis - Clustering Algorithms	8Hr
Unit 2	Storage and Searching Techniques: Storage: Inverted file - Suffix trees & suffix arrays - Signature Files - Scatter storage or hash addressing - Clustered files.IR Models: Basic concepts - Boolean Model - Vector Model Searching strategies: Boolean Search - Serial search - cluster-based retrieval - Query languages - Types of queries - Patterns matching - structural queries.	7Hr
Unit 3	Retrieval Performance Evaluation and Ontology: Performance evaluation: Precision and recall - alternative measures Ontology: Ontology-based information sharing - Ontology languages for semantic web - Ontology creation.	7Hr
Unit 4	Distributed and Multimedia IR: Distributed IR: Introduction - Collection Partitioning - Source Selection - Query Processing - web issues. MULTIMEDIA IR: Introduction - Data Modeling - Query languages - Generic multimedia indexing approach - One-dimensional time series - two-dimensional color images - Automatic feature extraction.	7Hr
Unit 5	Web Searching and Web Recommendation: Searching the Web: Challenges - Characterizing the Web - Search Engines - Browsing - Meta-searchers - Web crawlers - Meta-crawler - Web data mining - Finding a needle in the Haystack - Searching using Hyperlinks - Page ranking algorithms. Collaborative Filtering and Content-Based Recommendation of Documents and Products – Information Extraction and Integration: Extracting Data from Text. Semantic Web - Collecting and Integrating Specialized Information on the web.	7Hr

Text Books	
1	Yates &Neto - "Modern Information Retrieval" - Pearson Education - ISBN 81-297-0274-6.
2	Heiner Stuckenschmidt - Frank van Harmelen - "Information Sharing on the Semantic Web" - Springer International Edition-ISBN 3-540-20594-2.
Reference Books	
1	Christopher D. Manning - PrabhakarRaghavan and HinrichSchutze "Introduction to Information Retrieval" - Cambridge University Press - ISBN 978-0-521-86571-5
2	V. S. Subrahamanian - Satish K. Tripathi "Multimedia information System" - Kulwer Academic Publisher.

Subject Name: Elective-IV Time Series Analysis and Forecasting

Load	Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)	3	-	3	30	70	100

Course Objectives:	
1	Present time series in an informative way, both graphically and with summary statistics,
2	Model time series to analyses the underlying structure(s) in both the time and frequency domains.

Course Outcomes	
At the end of this course students are able to:	
CO1	Understand Time Series Components and Patterns
CO2	Apply Descriptive Techniques to Time Series Data
CO3	Develop and Estimate Time Series Models
CO4	Utilize Software Tools for Time Series Forecasting
CO5	Implement Advanced Time Series Techniques

Unit		
Unit 1	**Introduction to Time Series Analysis:** Definition and Importance of Time Series, Types of Time Series Data, Components of Time Series: Trend, Seasonality, Cyclic, and Irregular Variations, Time Series Plots and Graphical Techniques, Applications of Time Series Analysis	7Hr
Unit 2	**Descriptive Techniques:** Measures of Central Tendency and Dispersion, Moving Averages, Exponential Smoothing, Decomposition of Time Series, Identifying Patterns in Time Series Data	7Hr
Unit 3	**Time Series Modeling:** Autoregressive (AR) Models; Moving Average (MA) , Models, Autoregressive Moving Average (ARMA) Models, Autoregressive Integrated Moving Average (ARIMA) Models, Seasonal ARIMA (SARIMA) Models	8Hr
Unit 4	**Forecasting Techniques:** Qualitative vs Quantitative Forecasting, Judgmental Forecasting Methods, Exponential Smoothing State Space Models (ETS), Machine Learning Approaches to Time Series Forecasting, Combining Forecasts	7Hr
Unit 5	**Recent Trends and Research in Time Series Analysis:** Big Data and Time Series Analysis, Real-time Forecasting and Streaming ,Data Advances in Machine Learning for Time Series, Future Directions and Research Opportunities	8Hr 7Hr

Text Books	
1	Brockwell, P.J. and Daris, R. A. (2002). Introduction to time Series and Forecasting, 2nd ed., Springer-Verlag.
2	Montgomery, D.C. and Johnson, L.A. (1976). Forecasting and Time Series Analysis, McGraw Hill.
Reference Books	
1	Chatfield C. (1980): The Analysis of Time Series –An Introduction, Chapman & Hall.
2	Mukhopadhyay P. (2011): Applied Statistics, 2nd ed. Revised reprint, Books and Allied

Subject Name: Elective-V Big Data Anaytics

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 (Theory)	Hrs	3	-	3	30	70	100

Course Objectives:	
1	To get familiar with the Concepts, Terminology, Characteristics and Life Cycle of Big Data Analytics
2	To Analyze the Benefits and Challenges of core components of Hadoop and apply File System Shell Commands for HDFS.

Course Outcomes	
At the end of this course students are able to:	
CO1	Recognize the Big Data Analytics Concepts, Terminology, Characteristics and Life Cycle of Big Data Analytics
CO2	Evaluate the different core components of Hadoop and apply File System Shell Commands in HDFS.
CO3	Analyze the Mapper and Reducer Key/Value Types for MapReduce Jobs and perform the Mapreduce job in MR1Framework and Yarn Framework
CO4	Apply Query Statements for Create Table, Insert Data and Alter Data in the Apache Hive and HBase Environment .
CO5	Apply Pig Latin Scripts for Writing, Evaluation and Filter Functions

Unit 1	Understanding Big Data: Concepts and Terminology – Big Data Characteristics – Types of Data – Case Study Background – Drivers for Big Data Adoption: Information and Communication Technology – Big Data Analytics Lifecycle	7Hr
Unit 2	Core components of Hadoop- Apache Hadoop – HDFS Daemons – MapReduce Daemons – HDFS High Availability Daemons – Benefits and Challenges of HDFS – File Sizes, Block Sizes and Block Abstraction in HDFS – Data Replication – How does HDFS Store, Read, and Write Files? – Data Serialization Options – File System Shell Commands for HDFS	8Hr
Unit 3	**ADVANCED MAPREDUCE TECHNIQUES** Simple, advanced, and in-between Joins, Graph algorithms, using language-independent data structures. Hadoop configuration properties - Setting up a cluster, Cluster access control, managing the NameNode, Managing HDFS, MapReduce management, Scaling.	~~8Hr~~ 7Hr
Unit 4	**HADOOP STREAMING** Hadoop Streaming - Streaming Command Options - Specifying a Java Class as the Mapper/Reducer - Packaging Files With Job Submissions - Specifying Other Plug-ins for Jobs.	7Hr
Unit 5	**HIVE & PIG :** Architecture, Installation, Configuration, Hive vs RDBMS, Tables, DDL & DML, Partitioning & Bucketing, Hive Web Interface, Pig, Use case of Pig, Pig Components, Data Model, Pig Latin.	7Hr

Text Books	
1	Deepak Vohra, "Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-related Frameworks and Tools", Apress, 2016.
2	Thomas Erl, WajidKhattak, Paul Buhler, "Big Data Fundamentals Concepts, Drivers & Techniques", Service Tech Press, 2015
Reference Books	
1	Noreen Burlingame , "The little book on Big Data", New Street publishers, 2012.
2	Pethuru Raj, Anupama Raman, DhivyaNagaraj and Siddhartha Duggirala, High-Performance Big-Data Analytics: Computing Systems and Approaches, Springer, 2015.

Subject Name: Elective-V Data Modeling and Simulation

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)		3	-	3	30	70	100

Course Objectives:	
1	Provide a thorough understanding of the principles, techniques, and methodologies used in data modeling and simulation.
2	Equip students with the ability to collect, preprocess, and analyze data, and to build and validate models using appropriate software tools.

Course Outcomes	
At the end of this course students are able to:	
CO1	Understand Data Modeling Concepts
CO2	Perform Data Collection and Preparation:
CO3	Build and Validate Models
CO4	Implement Simulation Techniques
CO5	Analyze and Interpret Simulation Results

Unit 1	**Introduction to Data Modeling and Simulation:** Definition and Importance of Data Modeling, Types of Data Models: Conceptual, Logical, and Physical, Introduction to Simulation: Definitions and Applications, Differences between Modeling and Simulation, Case Studies Highlighting the Importance of Data Modeling and Simulation	8Hr
Unit 2	**Data Collection and Preparation:** Data Sources: Primary and Secondary, Data Cleaning and Preprocessing, Handling Missing Data and Outliers, Data Normalization and Transformation, Exploratory Data Analysis (EDA)	7Hr
Unit 3	**Simulation Techniques and Methods:** Types of Simulation: Discrete-Event, Continuous, and Hybrid; Monte Carlo Simulation, Agent-Based Modeling, System Dynamics; Queuing Theory and Network Simulations	7Hr
Unit 4	**Model Building and Validation:** Steps in Building a Model: Conceptualization to Implementation, Data Fitting and Parameter Estimation, Model Verification and Validation Techniques, Sensitivity Analysis, Model Calibration	7Hr
Unit 5	**Applications of Data Modeling and Simulation:** Business and Economic Simulations, Engineering and Manufacturing Simulations, Healthcare Simulations, environmental and Ecological Simulations, Social and Behavioral Simulations	7Hr

Text Books	
1	"Simulation Modeling and Analysis" by Averill M. Law **ISBN:** 978-0073401324
2	"Data Modeling and Database Design" by Narayan S. Umanath and Richard W. Scamell **ISBN:** 978-1285073980
Reference Books	
1	"The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness, and Determinism" by Foster Morrison ☐ **ISBN:** 978-0486449049
2	"Principles of Modeling and Simulation: A Multidisciplinary Approach" by John A. Sokolowski and Catherine M. Banks

Subject Name: Elective-V Data Warehousing

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 (Theory)	Hrs	3	-	3	30	70	100

Course Objectives:	
1	To understand data warehouse concepts, architecture, business analysis and tools
2	To understand data pre-processing and data visualization techniques
3	To study algorithms for finding hidden and interesting patterns in data

Course Outcomes	
At the end of this course students are able to:	
CO1	Design a Data warehouse system and perform business analysis with OLAP tools.
CO2	Apply suitable pre-processing and visualization techniques for data analysis
CO3	Apply frequent pattern and association rule techniques for data analysis
CO4	Apply appropriate classification and clustering techniques for data analysis
CO5	Analyze Meta data Management in Data Warehouse

Unit 1	Data Warehouse Fundamentals Introduction to Data Warehouse, OLTP Systems; Differences between OLTP Systems and Data Warehouse: Characteristics of Data Warehouse; Functionality of Data Warehouse: Advantages and Applications of Data Warehouse; Advantages, Applications: Top Down and Bottom Up Development Methodology: Tools for Data warehouse development: Data Warehouse Types	8Hr
Unit 2	Planning and Requirements Introduction: Planning Data Warehouse and Key Issues: Planning and Project Management in constructing Data warehouse: Data Warehouse Project; Data Warehouse Development Life Cycle, Kimball Lifecycle Diagram, Requirements Gathering Approaches: Team organization, Roles, and Responsibilities:	7Hr
Unit 3	Data Warehouse Architecture Introductions, Components of Data Warehouse Architecture: Technical Architectures; Data warehouse architectures Tool selection: Federated Data Warehouse Architecture:	7Hr
Unit 4	**Dimensional Modeling** : Introduction: E-R Modeling: Dimensional Modeling: E-R Modeling VS Dimensional Modeling: Data Warehouse Schemas; Star Schema, Inside Dimensional Table, Inside Fact Table, Fact Less Fact Table, Granularity, Star Schema Keys: Snowflake Schema: Fact Constellation Schema: **Data Warehouse & OLAP** : Introduction: What is OLAP?; Characteristics of OLAP, Steps in the OLAP Creation Process, Advantageous of OLAP: What is Multidimensional Data: OLAP Architectures; MOLAP,	7Hr
Unit 5	**Meta data Management in Data Warehouse** : Introductions to Metadata: Categorizing Meta data: Meta data management in practice; Meta data requirements gathering, Meta data classification, Meta data collection strategies: Meta Data Management in Oracle and SAS: Tools for Meta data management:	7Hr

Text Books	
1	Alex Berson and Stephen J.Smith, ―Data Warehousing, Data Mining & OLAP‖, Tata McGraw – Hill Edition, 35th Reprint 2016.
2	K.P. Soman, Shyam Diwakar and V. Ajay, ―Insight into Data Mining Theory and Practice, Eastern Economy Edition, Prentice Hall of India, 2006.
Reference Books	
1	Ian H.Witten and Eibe Frank, ―Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, Second Edition

Subject Name: Open Elective-II Machine Learning

Load		Lecture	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)		3	-	3	30	70	100

Course Objectives:	
1	To understand the need for machine learning for various problem solving..
2	To study the various supervised, semi – supervised and unsupervised learning algorithms in machine learning.
	To understand the latest trend in machine learning.

Course Outcomes	
At the end of this course students are able to:	
CO1	Differentiate between supervised, unsupervised, semi supervised machine learning approaches
CO2	Discuss the decision tree algorithm and identity and overcome the problems.
CO3	Discuss and apply the back propagation algorithm and genetic algorithms to various problems.
CO4	Apply the Bayesian concepts to machine learning
CO5	Analyzes and suggest appropriate machine learning approaches for various types of problems.

Unit 1	Introduction: ML Techniques and overview, Version Space and Candidate Eliminations, Inductive Bias, Decision Tree Learning - Representation & Algorithm, Hypothesis Evaluation, Heuristic Space Search	7Hr
Unit 2	Neural Networks and Genetic Algorithms Neural Network Representation, Problems, Perceptron, Multilayer Networks and Back Propagation Algorithms, Genetic Algorithms, Model Evaluation and Learning	7Hr
Unit 3	Bayesian and Computational Learning Bayes Theorem, Concept Learning, Bayes Optimal Classifier, Gibbs Algorithm, Naive Bayes Classifier, Bayesian Belief Network, Probability Learning, Mistake bond Model, Support Vector Machine	7Hr
Unit 4	Instance Based Learning K- Nearest Neighbor Learning, Locally Weighted Regression, Radial Basis Functions, Case Based Learning	7Hr
Unit 5	Unsupervised Learning: K means Clustering Algorithm, Reinforcement Learning, Elements of Reinforcement Learning, Exploration vs. Exploitation dilemma, Q – Learning.	8Hr

Text Books	
1	Tom M Mitchell, Machine Learning, McGraw – Hill Education (India) Private Limited, 2013.
2	Machine Learning: An Algorithmic Prespective, CRC Press, 2009, by Stephen Marsland
Reference Books	
1	Ethem Adpaydin, Introduction to Machine Learning (Adaptive Computation and Machine Learning), The MIT Press 2004
2	Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies by John D. Kelleher, Brian Mac Namee, and Aoife D'Arcy

Lab Name: Machine Learning and data analytics Lab

Load	Practical	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
2 Hrs (Practical)	2	-	1	25	25	50

1	Implement and analyze Linear Regression in Python (Single Variable and Multi Variable)
2	Implement and analyze the Logistic Regression in Python
3	Implement and analyze the Decision tree algorithm in Python
4	Implement and analyze the Random Forest algorithm in Python
5	Implementation of two samples T –Test and paired two sample T-Test in excel.
6	Implementation of one way and two way ANOVA in Excel
7	Write the steps of installing the Hadoop in Windows 10
8	Working with Hadoop Commands
9	Implementation of word count example using MapReduce
10	Implementation of MapReduce program to count unique number of times a song is played based on userid and trackid

Lab Name: Data Mining and Visualization Lab

Load	Practical	Tutorial	Credits	College Assessment Marks	University Evaluation	Total Marks
2 Hrs (Practical)	2	-	1	25	25	50

1	Explore WEKA Data Mining/Machine Learning Toolkit
2	Perform data preprocessing tasks on i. Add attribute ii. Add expression iii. Copy attribute iv. Remove attribute
3	Demonstrate performing classification on data sets
4	Demonstrate performing association rule mining on data sets
5	Demonstrate performing SVM classification on data sets
6	Demonstrate performing clustering on data sets
7	Connecting to Data and preparing data for visualization in Tableau
8	Data Aggregation and Statistical functions in Tableau
9	Data Visualizations in Tableau
10	Basic Dashboards in Tableau

Sr. No.	Subject	Teaching Scheme			Evaluation Schemes			Credits	Category
	B. Tech. Eight Semester CSE(DS)								
		L	T	P	CA	UE	TOTAL		
1	Industry Project	-	-	16	75	75	150	8	PROJ-CS
2	MOOCS Course-1	-	-	-	-	-	-	3	OEC
3	MOOCS Course-2	-	-	-	-	-	-	3	OEC
	Total	0	0	16	75	75	150	14	

MOOCS Course-1:

i) Social Networks

ii) Reinforcement Learning iii) Introduction to Internet of Things

MOOCS Course-2

i) Cyber Security and Privacy   ii) Block Chain and its Applications   iii) Ethical Hacking

Subject:	Social Networks				
Load	Credits	Total Marks	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)	03	100	30	70	100

Aim: To understand social networks and use of tools for social network analysis.

Prerequisite(s): Computer Networks

Course Objective/Learning Objective:

1	To understand highly interconnected and hence more complex social networks
2	To represent connected social networks in form of graph
3	To apply graph theory, sociology, game theory
4	To use tools and extract statistics from social networks

Course Outcome:

At the end of this course Student are able to:

CO1	Learn social networks , its types and representation
CO2	Understand weak ties, strong and weak relationships , homophily and calculate
CO3	Analyze various links concern with social Networks
CO4	Understand Power Laws and Rich-Get-Richer Phenomena
CO5	Understand Small World Phenomenon

Course layout

Week 1: Introduction

Week 2: Handling Real-world Network Datasets

Week 3: Strength of Weak Ties

Week 4: Strong and Weak Relationships (Continued) & Homophily

Week 5: Homophily Continued and +Ve / -Ve Relationships

Week 6: Link Analysis

Week 7: Cascading Behaviour in Networks

Week 8: Link Analysis (Continued)

Week 9: Power Laws and Rich-Get-Richer Phenomena

Week 10: Power law (contd..) and Epidemics

Week 11: Small World Phenomenon

Week 12: Pseudocore (How to go viral on web)

Books and references

1. https://onlinecourses.nptel.ac.in/noc24_cs120/preview

2. Networks, Crowds and Markets by David Easley and Jon Kleinberg, Cambridge University Press, 2010 (available for free download).

3. Social and Economic Networks by Matthew O. Jackson, Princeton University Press, 2010.

Subject:	Reinforcement Learning				
Load	**Credits**	**Total Marks**	**College Assessment Marks**	**University Evaluation**	**Total Marks**
3 Hrs (Theory)	03	100	30	70	100

Aim: The goal of the course is to introduce the basic mathematical foundations of reinforcement learning, as well as highlight some of the recent directions of research.

Prerequisite(s): Learnings & Neural Networks

Course Objective/Learning Objective:

1	It aims to model the trial-and-error learning process that is needed in many problem situations where explicit instructive signals are not available.
2	It has roots in operations research, behavioral psychology and AI.
3	The goal of the course is to introduce the basic mathematical foundations of reinforcement learning.
4	It highlight some of the recent directions of research

Course Outcome:

At the end of this course Student are able to:

CO1	Understand Bandit algorithm and its mathematical formulation.
CO2	Use dynamic programming for reinforcement learning
CO3	Perform function approximation and apply LSM
CO4	Analyze Fit Q, DQN & Policy Gradient for Full RL
CO5	understand combinatorial models for complex problems

Course layout

Week 1 : Introduction

Week 2 : Bandit algorithms – UCB, PAC

Week 3 : Bandit algorithms –Median Elimination, Policy Gradient

Week 4 : Full RL & MDPs

Week 5 : Bellman Optimality

Week 6 : Dynamic Programming & TD Methods

Week 7 : Eligibility Traces

Week 8 : Function Approximation

Week 9 : Least Squares Methods

Week 10 : Fitted Q, DQN & Policy Gradient for Full RL

Week 11 : Hierarchical RL

Week 12 : POMDPs

Books and references

1. https://onlinecourses.nptel.ac.in/noc24_cs102/preview
2. R. S. Sutton and A. G. Barto. Reinforcement Learning - An Introduction. MIT Press. 1998.

Subject:	Introduction to Internet of Things				
Load	Credits	Total Marks	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)	03	100	30	70	100

Aim: The goal of the course is to envision pervasive connectivity, storage, and computation, which, in turn, gives rise to building different IoT solutions

Prerequisite(s): Basic programming knowledge

Course Objective/Learning Objective:

1	To provide a solid foundation in the basic concepts, architecture, and communication protocols of the Internet of Things (IoT).
2	To equip with practical skills through the design, development, and implementation of IoT systems using various sensors, actuators, and microcontrollers.
3	To impart knowledge about the security and privacy challenges in IoT systems
4	To expose to a variety of real-world IoT applications and case studies across different industries such as healthcare, agriculture, smart cities, and manufacturing.

Course Outcome:

At the end of this course Student are able to:

CO1	understand the key components and architecture of IoT systems, including sensors, actuators, communication protocols, and cloud integration.
CO2	design and implement IoT solutions using microcontrollers
CO3	develop skills in collecting, storing, processing, and analyzing IoT data using relevant tools and technologies
CO4	Propose and implement basic security measures to protect IoT systems.
CO5	Analyze different IoT applications across various industries, evaluate their impact,

Course layout

Week 1: Introduction to IoT: Part I, Part II, Sensing, Actuation, Basics of Networking: Part-I

Week 2: Basics of Networking: Part-II, Part III, Part IV, Communication Protocols: Part I, Part II

Week 3: Communication Protocols: Part III, Part IV, Part V, Sensor Networks: Part I, Part II

Week 4: Sensor Networks: Part III, Part IV, Part V, Part VI, Machine-to-Machine Communications

Week 5: Interoperability in IoT, Introduction to Arduino Programming: Part I, Part II, Integration of Sensors and Actuators with Arduino: Part I, Part II

Week 6: Introduction to Python programming, Introduction to Raspberry Pi, Implementation of IoT with Raspberry Pi

Week 7: Implementation of IoT with Raspberry Pi (contd), Introduction to SDN, SDN for IoT

Week 8: SDN for IoT (contd), Data Handling and Analytics, Cloud Computing

Week 9: Cloud Computing(contd), Sensor-Cloud

Week 10: Fog Computing, Smart Cities and Smart Homes

Week 11: Connected Vehicles, Smart Grid, Industrial IoT

Week 12: Industrial IoT (contd), Case Study: Agriculture, Healthcare, Activity Monitoring

Books and references

1) https://onlinecourses.nptel.ac.in/noc24_cs115/preview
2) S. Misra, A. Mukherjee, and A. Roy, 2020. *Introduction to IoT*. Cambridge University Press. Availability: https://www.amazon.in/Introduction-IoT-Sudip-Misra/dp/1108959741/ref=sr_1_1?dchild=1&keywords=sudip+misra&qid=1627359928&sr=8-1
3) S. Misra, C. Roy, and A. Mukherjee, 2020. *Introduction to Industrial Internet of Things and Industry 4.0*. CRC Press. Availability: https://www.amazon.in/dp/1032146753/ref=sr_1_3?dchild=1&keywords=sudip+misra&qid=1627359971&sr=8-3
4) Research Papers

Subject:	Cyber Security and Privacy				
Load	Credits	Total Marks	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)	03	100	30	70	100

Aim: The goal of the course is to to develop a wholesome understanding about cyber security and privacy risks to businesses covering governance, compliance and risk mitigation and closely study certain business domains.

Prerequisite(s): A core course on Management Information Systems desirable

Course Objective/Learning Objective:

1	To provide a comprehensive understanding of the principles, concepts, and terminology of cyber security, including threat modeling and risk management.
2	To impart knowledge about privacy laws, regulations, and frameworks
3	To teach methods for ensuring data privacy in various contexts.
4	To develop an awareness of the ethical, legal, and societal implications of cyber security and privacy.

Course Outcome:

At the end of this course Student are able to:

CO1	Apply fundamental concepts and techniques in cyber security, such as cryptography, network security, and risk management.
CO2	Implement and manage security measures, perform security assessments, and develop secure software solutions.
CO3	Apply privacy principles, regulations, and frameworks to protect personal and sensitive information in various scenarios.
CO4	Identify security threats, performing forensic analysis
CO5	Execute effective incident response and recovery strategies.

Course layout

Week 1:
Introduction - Introduction to cyber security, Confidentiality, integrity, and availability.
Week 2:
Foundations - Fundamental concepts, CIA, CIA triangle, data breach at target.

Week 3:
Security management, Governance, risk, and compliance (GRC)- GRC framework, security standards.

Week 4:
Contingency planning - Incidence response, Disaster Recovery, BCP.

Week 5:
Cyber security policy - ESSP, ISSP, SYSSP.
Week 6:
Risk Management - Cyber Risk Identification, Assessment, and Control.
Week 7:
Cyber security: Industry perspective - Defense Technologies, Attack, Exploits
Week 8:
Cyber security technologies - Access control, Encryption, Standards.

Week 9:

Foundations of privacy - Information privacy, Measurement, Theories.

Week 10:

Privacy regulation - Privacy, Anonymity, Regulation, Data Breach.

Week 11:

Privacy regulation in Europe, Privacy: The Indian Way - Data Protection, GDPR, DPDP, Aadhar.

Week 12:

Information privacy: Economics and strategy, Economic value of privacy, privacy valuation, WTA and WTC, Business strategy and privacy, espionage, Privacy vs safety. Books and references

Text Book:
1) Michael E. Whitman, Herbert J. Mattord, (2018). Principles of Information Security, 6th edition, Cenage Learning, N. Delhi.

References
1) https://onlinecourses.nptel.ac.in/noc24_cs121/preview
2) Darktrace, "Technology" https://www.darktrace.com/en/technology/#machine-learning, accessed November 2018.
3) Van Kessel, P. Is cyber security about more than protection? EY Global Information Security Survey 2018-2019.
4) Johnston, A.C. and Warkentin, M. Fear appeals and information security behaviors: An empirical study. MIS Quarterly, 2010.
5) Arce I. et al. Avoiding the top 10 software security design flaws. IEEE Computer Society Center for Secure Design (CSD), 2014.
6) Smith, H. J., Dinev, T., & Xu, H. Information privacy research: an interdisciplinary review. MIS Quarterly, 2011.
7) Subramanian R. Security, privacy and politics in India: a historical review. Journal of Information Systems Security (JISSec), 2010.
8) Acquisti, A., John, L. K., & Loewenstein, G. What is privacy worth? The Journal of Legal Studies, 2013
9) Xu H., Luo X.R., Carroll J.M., Rosson M.B. The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. Decision Support Systems, 2011.

Subject:	Block Chain and its Applications				
Load	Credits	Total Marks	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)	03	100	30	70	100

Aim: This subject will cover the basic design principles of Blockchain technology and its applications over different sectors. Additionally, the course also provides tutorials on setting up blockchain applications using one of the well- adopted permission less blockchain platforms - Ethereum, and one permissioned blockchain platform - Hyperledger.

Prerequisite(s): Computer Networks; Operating Systems; Cryptography and Network Security.

Course Objective/Learning Objective:

1	Learn its capability of providing a transparent, secured, tamper-proof solution for interconnecting different stakeholders in a trustless setup.
2	This subject will cover the basic design principles of Blockchain technology and its applications over different sectors.
3	Additionally, the course also provides tutorials on setting up blockchain applications using one of the well- adopted permissionless blockchain platforms - Ethereum, and one permissioned blockchain platform - Hyperledger.
4	Provide its applications.

Course Outcome:

At the end of this course Student are able to:

CO1	Understand basic crypto premitives
CO2	Understand elements and evolution of blockchain
CO3	Understand consensus in permissionless and permissioned models
CO4	Hands on ethereum smart contracts and hyperledgers
CO5	Perform decentralized identity management, interoperability.

Course layout

Week 1: Introduction to Blockchain Technology and its Importance

Week 2: Basic Crypto Primitives I – Cryptographic Hash

Week 3: Basic Crypto Primitives II – Digital Signature

Week 4: Evolution of the Blockchain Technology

Week 5: Elements of a Blockchain

Week 6: Blockchain Consensus I – Permissionless Models

Week 7: Blockchain Consensus II – Permissioned Models

Week 8: Smart Contract Hands On I – Ethereum Smart Contracts (Permissionless Model)

Week 9: Smart Contract Hand On II – Hyperledger Fabric (Permissioned Model)

Week 10: Decentralized Identity Management

Week 11: Blockchain Interoperability

Week 12: Blockchain Applications

References

1. Mastering Blockchain: A deep dive into distributed ledgers, consensus protocols, smart contracts, DApps, cryptocurrencies, Ethereum, and more, 3rd Edition, Imran Bashir, Packt Publishing, 2020, ISBN: 9781839213199, book website: https://www.packtpub.com/product/mastering-blockchain-third-edition/9781839213199
2. Hyperledger Tutorials - https://www.hyperledger.org/use/tutorials
3. Ethereum Development Resources - https://ethereum.org/en/developers
4. Online materials and case studies

Subject:	Ethical Hacking				
Load	Credits	Total Marks	College Assessment Marks	University Evaluation	Total Marks
3 Hrs (Theory)	03	100	30	70	100

Aim: The goal of the course is to adopt safe practices and usage of their IT infrastructure. Starting from the basic topics like networking, network security and cryptography, the course will cover various attacks and vulnerabilities and ways to secure them.

Prerequisite(s): Basic concepts in programming and networking

Course Objective/Learning Objective:

1	To provide a solid foundation in the principles, techniques, and methodologies of ethical hacking and penetration testing.
2	To educate about various network and system vulnerabilities, including how to identify, assess, and exploit them ethically to improve security.
3	To equip with practical skills in using various hacking tools and technologies for penetration testing, vulnerability assessment, and security auditing.
4	To foster an understanding of the legal, ethical, and professional issues surrounding ethical hacking, promoting responsible and lawful conduct in the field.

Course Outcome:

At the end of this course Student are able to:

CO1	Understand the fundamental concepts and methodologies of ethical hacking, including reconnaissance, scanning, enumeration, exploitation, and reporting.
CO2	Identify and exploit various network and system vulnerabilities ethically, using a range of tools and techniques.
CO3	Perform penetration tests and vulnerability assessments.
CO4	Design and implement effective security measures and countermeasures to protect systems and networks against potential threats and attacks.
CO5	Understand the legal, ethical, and professional responsibilities of an ethical hacker,

Course layout

Week 1: Introduction to ethical hacking. Fundamentals of computer networking. TCP/IP protocol stack.

Week 2: IP addressing and routing. TCP and UDP. IP subnets.

Week 3: Routing protocols. IP version 6.

Week 4: Installation of attacker and victim system.
Information gathering using advanced google search, archive.org, netcraft, whois, host, dig, dnsenum and NMAP tool.

Week 5: Vulnerability scanning using NMAP and Nessus. Creating a secure hacking environment. System Hacking: password cracking, privilege escalation, application execution. Malware and Virus. ARP spoofing and MAC attack.

Week 6: Introduction to cryptography, private-key encryption, public-key encryption.

Week 7: Cryptographic hash functions, digital signature and certificate, applications.

Week 8: Steganography, biometric authentication, network-based attacks, DNS and Email security.

Week 9: Packet sniffing using wireshark and burpsuite, password attack using burp suite. Social engineering attacks and Denial of service attacks.

Week 10: Elements of hardware security: side-channel attacks, physical inclinable functions, hardware trojans.

Week 11: Different types of attacks using Metasploit framework: password cracking, privilege escalation, remote code execution, etc.Attack on web servers: password attack, SQL injection, cross site scripting.

Week 12: Case studies: various attacks scenarios and their remedies.

Books and references
1. https://onlinecourses.nptel.ac.in/noc24_cs94/preview
2. Data and Computer Communications -- W. Stallings.
3. Data Communication and Networking -- B. A. Forouzan
4. TCP/IP Protocol Suite -- B. A. Forouzan
5. UNIX Network Programming -- W. R. Stallings
6. Introduction to Computer Networks and Cybersecurity -- C-H. Wu and J. D. Irwin
7. Cryptography and Network Security: Principles and Practice -- W. Stallings