SEMESTER III – ELECTIVES

140CP9119	BIG DATA AND ANALYTICS	L T P C M 3 0 2 4 100

AIM:

This course aims to introduce to the Big Data paradigm, the various analytics approaches, and the related methods and tools.

OBJECTIVES:

At the end of the course, the students should be able to:

- Define Big Data, identify its application in analytics, and justify the use of Big Data for use cases from selected business domains
- Describe the design and related concepts of Hadoop / HDFS
- Perform MapReduce jobs on Hadoop/HDFS
- Use Hadoop-based tools such as Pig, Hive, HBase, Zookeeper, Sqoop, Flume, Avro for data analytics
- Evaluate various analytics platforms and tools such as Mahout, BigQuery, Berkeley Data Analytics Stack, and R

UNIT I INTRODUCTION **8**

Introduction –Big Data and its History –4 V's of Big Data– Structured, Semi-structured, and Unstructured Data – Applications of Big Data Analytics and Industry Examples –Implementation of Big Data Analytics– Hadoop Ecosystem and Tools.

UNIT II HADOOP AND HDFS **9**

Hadoop: Data Storage and Analysis, Comparison with other Systems: RDBMS, Grid Computing, Volunteer Computing –Hadoop Releases –HDFS: Design and Concepts, CLI, File Systems and Interfaces, Data Flow, Hadoop Archives –Hadoop I/O.

UNIT III MAPREDUCE & YARN **10**

MapReduce: Analysis with Hadoop, Scaling Out, Hadoop Streaming, Hadoop Pipes -Classic vs YARN MapReduce: Anatomy of Job Run, Failures, Scheduling, Shuffle and Sort, Task Execution - Types and Formats, Features, Apache MRUnit.

UNIT IV HADOOP-BASED TOOLS **10**

Pig: Comparison with Databases, Pig Latin– Hive: Comparison with Traditional Databases, HiveQL, Querying, UDFs –HBase: Hbasics, Concepts, HBase versus RDBMS – Introduction to ZooKeeper, Sqoop, Flume, Avro.

UNIT V DATA ANALYTICS **8**

Machine Learning with Large Datasets: Challenges, Mahout, and MLBase –Google BigQuery: Analytics DB, Dremel, Data Analytics as a Service, Building Big Data Dashboard –Using R with Large Datasets: Data Frames and Matrices, Large Matrix Manipulation, biglm, RHadoop – Visualization Strategies: ggplot2 , matplotlib , D3.js – Berkeley Data Analytics Stack – CrowdSourcing Analytics – Inter and Trans-Firewall Analytics.

L: 45, P: 30, Total: 75

BOS CHAIRMAN

REFERENCES:

1. Arvind Sathi, "Big Data Analytics: Disruptive Technologies for Changing the Game", Mc Press, Boise, ID-USA, 2013.
2. Tom White, "Hadoop: The Definitive Guide", O'Reilly Publication and Yahoo!Press, Third Edition, 2012.
3. Michael Manoochehri, "Data Just Right: Introduction to Large-scale Data & Analytics", Addison-Wesley, First Edition, 2013.
4. Michael Minelli, Michele Chambers, and Ambiga Dhiraj, "Big data, big analytics: emerging business intelligence and analytic trends for today's businesses", John Wiley & Sons, Hoboken, NJ-USA, 2013
5. The Apache Software Foundation, "Apache™ Hadoop® URL: http://hadoop.apache.org/ ", Accessed: 06 March 2015.
6. The Apache Software Foundation, "Apache MRUnit™ URL: http://mrunit.apache.org/", Accessed: 06 March 2015.
7. AMPLab, "Software | AMPLab - UC Berkeley URL: https://amplab.cs.berkeley.edu/software/", Accessed: 06 March 2015.

BOS CHAIRMAN

140CP9120	CYBER SECURITY AND COMPUTER FORENSICS	L T P C M 3 1 0 4 100

AIM:

This course aims to introduce the concepts related to detection and control of Cyber Crime.

OBJECTIVES:

At the end of the course, the students should be able to:

- Understand about Information Security aspects and Cyber Offences
- Describe various aspects pertaining to Cyber crime
- Gain awareness about Indian Cyber law
- Comprehend about Computer Forensic mechanisms
- Understand the importance of Cyber Security and best practices to be adopted

UNIT I INTRODUCTION **8+3**

Cybercrime – Information Security – Classifications – Legal – Indian ITA – cyber offences – Social engineering – Cyberstalking – Cybercafe – Botnets – Attack vector – Cloud Computing.

UNIT II CYBER CRIME ISSUES **10+3**

Tools & Methods – Proxy servers – Phishing – Password Cracking – Keyloggers – Virus & Worms – Trojan horses & back doors – Steganography – DoS – SQL Injection – Buffer overflow – Attacks on Wireless Networks – Mobile & Wireless Devices – Mobility – Credit card frauds – Security challenges.

UNIT III CYBER LAW **7+3**

Cybercrime – legal perspectives – Need for Law – Indian IT Act – Challenges – Digital Signatures – Amendments – Punishments – Cyberlaw, Technology and Students.

UNIT IV FORENSICS **10+3**

History – Cyber forensics – Science – Need – Digital Evidence – Analysis of Email – Life cycle – Network Forensics – Investigation – Relevance of OSI layers to Computer Forensics – Challenges – Tools & Techniques – Auditing- Antiforensics- Hand held devices and Digital Forensics.

UNIT V CYBER SECURITY IMPLICATIONS **10+3**

Cyber crimes & IPR Issues – Web Threats for Organizations – Security & privacy implications – Risks in Social media marketing – Protecting privacy – Organization guidelines – Best practices – Social, political, ethical & psychological Dimensions.

L: 45, T: 15, Total: 60

REFERENCES:

1. Sunit Belapure, Nina Godbole, "Cyber Security: Understanding Cyber Crimes, Computer Forensics and Legal Perspectives", Wiley India Pvt Ltd, 2011.
2. Bill Nelson, Amelia Phillips, Frank Enfinger and Christopher Steuart, "Computer Forensics and Investigations", Cengage Learning, New Delhi, 2009.
3. Kevin Mandia, Chris Prosise, Matt Pepe, "Incident Response and Computer Forensics", Tata McGraw -Hill, New Delhi, 2006.
4. Aparna Viswanathan, "Cyber Laws: Indian and International Perspectives on Key topics including Data Security, E-commerce, Cloud Computing and Cyber Crimes", LexisNexis Butterworths Wadhwa, 2012.

BOS CHAIRMAN

140CP9121	INFORMATION RETRIEVAL TECHNIQUES	L T P C M 3 0 2 4 100

AIM

The course aims to introduce the various aspects of Text and Multimedia Information Retrieval

OBJECTIVES

At the end of the course, the students should be able to:

- Understand the principles of various Information Retrieval models
- Design query processing systems with relevance feedback mechanisms
- Apply text processing operations, construct text indices and design effective user interfaces
- Develop systems for searching and retrieving multimedia data
- Design prototype Search Engines

UNIT I INTRODUCTION **9**

Basic Concepts – Retrieval Process – Modeling – Classic Information Retrieval – Set Theoretic, Algebraic and Probabilistic Models – Structured Text Retrieval Models – Retrieval Evaluation

UNIT II QUERYING **9**

Languages – Key word based Querying – Pattern Matching – Structural Queries – Query Operations – User Relevance Feedback – Local and Global Analysis – Text and Multimedia languages

UNIT III TEXT OPERATIONS AND USER INTERFACE **9**

Document Preprocessing – Clustering – Text Compression - Indexing and Searching – Inverted files – Boolean Queries – Sequential searching – User Interface and Visualization – Access Process – Query Specification – Context – User relevance Judgment

UNIT IV MULTIMEDIA INFORMATION RETRIEVAL **9**

Data Models – Query Languages – Spatial Access Methods – Generic Approach – One Dimensional Time Series – Two Dimensional Color Images – Feature Extraction

UNIT V APPLICATIONS **9**

Searching the Web – Challenges – Characterizing the Web – Search Engines – Browsing – Meta-searchers – Parallel Information Retrieval – Distributed Information Retrieval

L: 45, P: 30, Total: 75

REFERENCES:

1. Ricardo Baeza-Yates, Berthier Ribeiro-Neto, "Modern Information Retrieval", Second Edition, ACM Press Books, 2011.
2. Daniel Jurafsky and James H. Martin, "Speech and Language Processing", Second Edition, Prentice Hall, 2009.
3. Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, "Introduction to Information Retrieval", Cambridge University Press, 2008.

BOS CHAIRMAN

140CP9122	INFORMATION VISUALIZATION	L T P C M 3 0 2 4 100

AIM:

This course aims to impart knowledge on various aspects of Information Visualization and focuses on guidelines for developing clear and effective presentation of Information.

OBJECTIVES:

At the end of the course, the students should be able to:

- Understand the foundations of Data Visualization and its models.
- Comprehend the concepts of Patterns, Visual objects and Space perception
- Explore various techniques for Interacting with Visualizations
- Design interactive visualizations
- Develop data driven documents

UNIT I FOUNDATIONS OF DATA VISUALIZATION AND VISUAL INFORMATION 12

A Model of Perceptual processing, Types of Data; Luminance, Brightness, lightness and Gamma; Color: Standards, Appearance and Applications in Visualization; Visual Attention and Information: Visual field, Iconic Buffer, Gabor Model, Texture in Visualization, glyphs and Multivariate Discrete data.

UNIT II PATTERNS AND OBJECTS 9

Static and Moving Patterns: Gestalt laws, Contours, Patterns in Motion; Visual Objects and Data Objects: Image-Based Object recognition, Structure-based Object Recognition, Geon diagram, Perceiving the Surface shapes of Objects; Space Perception and Display of data: Depth Cue Theory, Task Based Space Perception.

UNIT III INTERACTING WITH VISUALIZATIONS 6

Interacting with Visualizations: Data Selection and Manipulation loop, Exploration and Navigation loop; Memory systems, Eye movements, Problem Solving with Visualizations.

UNIT IV CODING & ALGORITHMS IN VISUALIZATION SYSTEMS 9

Coding and Algorithms: Variables, Conditionals, Loops, Functions, Objects, Arrays, Algorithms, Debugging and Libraries; Visualization Techniques and Systems: Structural Analysis, Statistical Exploration, Practical problems in conducting user studies.

UNIT V DATA DRIVEN DOCUMENTS AND CASE STUDIES 9

D3 Technology Fundamentals: DOM, CSS, JavaScript; Data: Binding data; Drawing with data; Scales, Interactivity; Case Studies: Patient Record Analysis – Business visualization analysis - Child studies – Cross Cultural studies.

L: 45, P: 30, Total: 75

REFERENCES:

1. Colin Ware, "Information Visualization: Perception for Design", Morgan Kaufmann, 2004.
2. Ben Fry, "Visualizing Data", First Edition, O'Reilly, 2008.
3. Scott Murray," Interactive Data Visualization for the Web -An Introduction to Designing with D3", O'Reilly, 2013

BOS CHAIRMAN

140CP9123	MACHINE LEARNING TECHNIQUES	L T P C M 3 0 2 4 100

AIM:

This course aims to impart knowledge on the theory and techniques associated with Machine Learning.

OBJECTIVES:

At the end of the course, the students should be able to:

- Comprehend the working of Supervised learning systems
- Apply Bayesian Classification principles and parametric estimation techniques
- Design Clustering algorithms and non-parametric classification & regression systems
- Construct Neural Network and Support Vector Machine Classifiers
- Evaluate the performance of classifiers and combine multiple learners

UNIT I SUPERVISED LEARNING **9**

Machine Learning – Paradigms and Applications – Supervised Learning: Learning from Examples – VC Dimension – Probably Approximately Correct Learning – Handling Noise – Learning Multiple Classes – Regression – Model Selection and Generalization.

UNIT II PROBABILISTIC LEARNING AND PARAMETRIC METHODS **9**

Bayesian Decision Theory: Classification – Utility Theory – Bayesian Networks – Parametric methods: Maximum Likelihood Estimation – Bayes Estimator – Parametric Classification and Regression – Dimensionality Reduction: Principal Component Analysis – Factor Analysis.

UNIT III CLUSTERING AND NON-PARAMETRIC METHODS **9**

Clustering: k-means Clustering – Expectation Maximization – Hierarchical Clustering – Non-parametric methods: Density Estimation – Generalization to Multi-variate data – Non-parametric Classification and Regression.

UNIT IV MULTILAYER PERCEPTRONS AND KERNEL MACHINES **9**

Linear Discrimination: Generalizing the Linear model – Geometry of Linear Discriminant – Multilayer Perceptrons: Perceptron Learning – Universal Approximators – Back propagation Algorithm – Training – Temporal networks – Kernel Machines: Separating Hyperplanes – Kernels – Regression.

UNIT V CLASSIFICATION ASSESSMENT AND MULTIPLE LEARNERS **9**

Assessing Classification Algorithms: Cross Validation and Re-sampling – Error measurement – Hypothesis Testing – Performance Assessment – Comparison of Algorithms – Combining Multiple Learners: Voting – Bagging – Boosting.

L: 45, P: 30, Total: 75

REFERENCES:

1. Ethem Alpaydin, "Introduction to Machine Learning", Prentice Hall of India, Second Edition, 2010.
2. Peter Harrington, "Machine Learning in Action", Wiley India Pvt Ltd, 2012.
3. Brett Lantz, "Machine Learning with R", Packt Publishing, 2013.

BOS CHAIRMAN

140CP9124	SOCIAL NETWORK ANALYSIS	L T P C M 3 0 2 4 100

AIM:

The course aims to provide foundation knowledge about social networks along with various models and algorithms suitable for mining real world social networks. This course also aims in providing research directions in mining social networks.

OBJECTIVES:

At the end of the course, the students should be able to:

- Analyze the statistical properties of social networks and its data sets
- Gain expertise in various modelling techniques along with link prediction methods
- Visualize the social data and categorize mining methods suitable for real time applications.
- Explore multimedia networks, tagging and privacy preservation mechanisms in social mining.
- Mine various real world social web applications.

UNIT I PROPERTIES OF SOCIAL NETWORKS **9**

Statistical Properties: Static and Dynamic properties – Random walks in Social networks – Evaluation and datasets – Community discovery – Node classification in social networks.

UNIT II MODELS AND ALGORITHMS **9**

Evolution in social networks – Incremental mining for community tracing – Social influence analysis – Expert location in social networks – Link prediction – Bayesian probabilistic models – Probabilistic relational models – Linear algebraic methods

UNIT III MINING AND VISUALIZATION **9**

Visualizing social networks – Interaction and analytics – Data mining methods for social media – Text mining in social networks – Keyword Search – Classification Algorithms – Integrating Sensors and Social Networks.

UNIT IV TAGGING AND PRIVACY **9**

Multimedia Information Networks - Links from Semantics and Community media - Network of Personal Photo Albums and Geographical Information - Inference Methods – Social tagging: Generation models and System design – Visualization, recommendations and applications – Tagging problems - Privacy in social networks – privacy breaches and preservation mechanisms.

UNIT V APPLICATIONS **9**

Influence and Homophily - Recommendation in Social Media - Recommendation Using Social Context - Behavior Analytics: Individual and Collective Behavior Analysis, Modelling and Prediction. Case Studies: Mining Facebook, LinkedIn and Web Pages.

L: 45, P: 30, Total: 75

REFERENCES:

1. Charu. C.Aggarwal, "Social Network Data Analytics", Springer, 2011
2. Reza Zafarani, Mohammad Ali Abbasi, Huan Liu, "Social Media Mining", Cambridge University Press, 2014
3. Matthew A. Russell, "Mining the Social Web", Second Edition, O'Reilly Media Inc., 2014
4. Mrutyunjaya Panda, Satchidananda Dehuri, Gi-Nam, "Social Networking – Mining, Visualization and Security", Springer, 2014
5. Borko Furht, "Handbook of Social Network Technologies and Applications" Springer, First Edition, 2010.

BOS CHAIRMAN

140CP9125	SOFT COMPUTING	L T P C M 3 0 2 4 100

AIM:

To provide a conceptual understanding of various Soft Computing techniques such as Genetic Algorithms, Neural Networks, Fuzzy Systems and Neuro-Fuzzy hybrid systems.

OBJECTIVES:

At the end of the course, the students should be able to:

- Understand the concept of Soft Computing
- Apply Genetic Algorithms for solving optimization problems
- Develop Neural Network architectures for solving problems
- Apply Fuzzy Logic for Inference ,Decision making and Expert system
- Design hybrid Neuro-Fuzzy systems

UNIT I INTRODUCTION TO SOFT COMPUTING **8**

Evolution of Computing - Soft Computing Constituents – From Conventional AI to Computational Intelligence - Machine Learning Basics.

UNIT II GENETIC ALGORITHMS **10**

Introduction – Biological Background – Operators and Techniques in GA – Classification of GA – Applications.

UNIT III NEURAL NETWORKS **9**

Machine Learning Using Neural Network, Adaptive Networks – Feed forward Networks – Supervised Learning Neural Networks – Radial Basis Function Networks – Reinforcement Learning – Unsupervised Learning Neural Networks – Adaptive Resonance architectures – Advances in Neural networks.

UNIT IV FUZZY LOGIC **9**

Fuzzy Sets – Operations on Fuzzy Sets – Fuzzy Relations – Membership Functions- Fuzzy Rules and Fuzzy Reasoning – Fuzzy Inference Systems – Fuzzy Expert Systems – Fuzzy Decision Making.

UNIT V NEURO-FUZZY MODELING **9**

Adaptive Neuro-Fuzzy Inference Systems – Coactive Neuro-Fuzzy Modeling – Classification and Regression Trees – Data Clustering Algorithms – Rule base Structure Identification – Neuro-Fuzzy Control – Case studies.

L: 45, P: 30, Total: 75

REFERENCE BOOKS:

1. Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani, "Neuro-Fuzzy and Soft Computing", Prentice-Hall of India, 2003.
2. Timothy J. Ross, "Fuzzy Logic with Engineering Applications", Third Edition, Wiley, 2010
3. Simon Haykin, "Neural Networks and Learning Machines", Third Edition, Prentice Hall, 2009
4. Mitchell Melanie, "An Introduction to Genetic Algorithm", Prentice Hall, 1998.
5. Sivanandam. S.N., Deepa .S.N., "Introduction to Genetic Algorithms", Springer, 2008.
6. Rajasekaran. S., Vijayalakshmi Pai.G.A., "Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications", PHI Learning Pvt. Ltd., 2004.

BOS CHAIRMAN

140CP9126	SOFTWARE PROJECT MANAGEMENT	L T P C M 3 1 0 4 100

Aim:
To introduce software project management techniques to successfully initiate, evaluate, plan, manage and control Information Technology projects.

OBJECTIVES:
The students will be able to:

- Plan a newly proposed project using different techniques.
- Evaluate a proposed project.
- Prepare a realistic estimation of a proposed project.
- Elucidate planning of project activities based on effort & duration estimation techniques and effectively manage project risks using different approaches
- Schedule, monitor and control a project successfully.

UNIT I PROJECT EVALUATION AND MANAGEMENT **9+3**

Importance of Software project management – types of project – contract and Technical project management – activities – plans, methods and methodologies- categorizing software projects – stakeholders – setting objectives- project success and failure –management control and practices – Portfolio management – evaluation techniques – Risk evaluation – Programme management.

UNIT II PROJECT PLANNING AND SELECTION OF APPROACHES **9+3**

Stepwise project planning–build or buy- Choosing methodologies and technologies – software process and models – prototyping – categorizing prototypes – Incremental delivery – Atern/Dynamic System development –RAD – Agile methods – Extreme programming(XP) - SCRUM – Managing iterative process – Selecting appropriate process model.

UNIT III SOFTWARE EFFORT ESTIMATION **8+3**

Estimation – problems in Estimation – Basis for estimation – Software effort estimation techniques – Bottom up estimating – Top down approach and parametric models – Expert judgment – Estimating by analogy – Albercht Function point analysis – Function points Mark II – COSMIC Full Function points – parametric productivity model – Capers Jones estimating rules of thumb.

UNIT IV ACTIVITY PLANNING AND RISK MANAGEMENT **10+3**

Objectives of activity planning– Project Schedules – project and activities - Sequencing and Scheduling Activities – Network planning model – Forward pass – The Backward Pass – Activity Float – project duration – critical activities – activity on arrow networks - Risk – categories of risk – Identification – Assessment – planning – Management – evaluating risk – applying PERT – Monte carlo Simulation – Critical chain concepts.

UNIT V RESOURCE ALLOCATION, MONITORING AND CONTROL **9+3**

Identifying resource requirements – Scheduling resources – Creating critical paths – Counting the cost – Resource schedule publishing – Cost schedules – Scheduling sequence – Creating framework – Collecting data – Visualizing progress – Cost monitoring – Earned value Analysis–Managing contracts.
Case Study: PRINCE2 – Project Management tools

L: 45, T: 15, Total: 60

References:

1. Bob Hughes, Mike Cotterell, Rajib Mall, "Software Project Management", TMH, Fifth Edition, 2014.
2. Gopalswamy Ramesh, "Managing global software projects", TMH, 2007.
3. Watts S Humphery, "Managing Software Process", Pearson 8th Impression, 2010.
4. Walker Royce, "Software Project Management, A Unified framework", Pearson Education, 2006.

BOS CHAIRMAN

140CP9127	WIRELESS SECURITY	L T P C M 3 1 0 4 100

AIM:

The course aims to provide technology fundamentals of wireless network security along with designing, installation and deployment issues in wireless networks. This course also aims in providing research directions in security aspects of wireless networks.

OBJECTIVES:

At the end of the course, the students should be able to:

- Characterize the security features of various wireless technologies
- Analyze various issues in designing security for wireless LAN & WAN
- Plan, install and deploy security mechanisms and evaluate its performance
- Gain expertise in management of wireless security and its countermeasures
- Provide security solutions for real world scenarios.

UNIT I TECHNOLOGY FUNDAMENTALS 9 +3

Wireless Network Security Fundamentals - Types of Security technologies – Policies – Standards: WEP, IEEE 802.11b, IEEE 802.11i, IEEE 802.1X, Bluetooth, SSL, WTLS, WPA, WPA2.

UNIT II DESIGNING WIRELESS SECURITY 8 +3

Security for Wireless LANs and WANs - Handling Wireless Private Information - Design Issues - Cost Justification and Consideration - Architectural design issues – WMTIP.

UNIT III PLANNING, INSTALLATION AND DEPLOYMENT 9 +3

Implementation Plan Development - Planning Techniques - Testing Techniques – Internetworking - Installation and Deployment - Securing and Certification of Performance.

UNIT IV MAINTENANCE AND COUNTERMEASURES 12 +3

Configuring secure access - Management of Wireless network security - Ongoing Maintenance - Standards Development - Ensuring Site Security - Security for Government and Industry groups - Security tactics of the Military - Security tactics of Private Enterprises.

UNIT V ENHANCING SECURITY 7 +3

Security solutions for ISP intranet, Internet and E-Commerce - Enhancing Wireless Web Security - Security Solutions for Consideration: WiFi Wireless LAN, Solutions for enterprises and residential users.

L: 45, T: 15, Total: 60

REFERENCES:

1. John R.Vacca, "Guide to Wireless Network Security", Springer, 2006
2. Randall K. Nichols, Panos C. Lekkas, "Wireless Security: Models, Threats, and Solutions", McGraw Hill, 2002
3. Yang Xiao, Xuemin Shen, Ding-Zhu Du, "Wireless Network Security", Springer, 2007
4. Tyler Wrightson, "Wireless Network Security: A Beginners' Guide", McGraw-Hill, 2012

BOS CHAIRMAN