Sixth Semester					
Theory					
Sl. No.	Category	Subject Code	Subject Name	L-T-P	Credit
1	BSC	22CM6BS01T	Optimization Engineering	4-0-0	4
2	PCC-9	22CS6PC01T	Machine Learning	3-0-0	3
3	PCC-10	22CS6PC02T	Software Engineering	3-0-0	3
4	PEC-3	Professional Elective-3:			
		22CS6PE01T 22CS6PE02T 22CS6PE03T 22CS6PE04T	Computer Network Security / Internet of Things Computer Vision Cloud Computing	3-0-0	3
		Professional Elective-4:			3
5	PEC-4	22CS6PE05T 22CS6PE06T 22CS6PE07T 22CS6PE08T 22CS6PE09T	Natural Language Processing Big Data Analytics Pattern Recognition Software Testing and Quality Assurance Fundamentals of 5G Communications	3-0-0	
6	OEC-3	Open Elective-3:(For CSE Students)			
			Courses from Other Departments	3-0-0	3
		Open Elective-3: (For Other Branch Students)			
		22CS6OE01T	Data Analytics		
Total Credit (Theory)					19
Practical					
1	PCC-9	22CS6PC01L	Machine learning Lab	0-0-2	1
2	PCC-10	22CS6PC02L	Software Engineering Lab	0-0-2	1
3	PSI	22CM6PS01L 22CM6PS02L	Technical writing using LATEX Simulation-based Laboratory using NS2/ NS3	0-0-3	2
4	HSMC	22CM6HS01L	Business Communication and Interview Skills	0-0-3	2
Total Semester Credit					25

22CM6BS01T	**Optimization Engineering (4-0-0)**	**4 Credits**

Course objectives:

1. To understand the theory of optimization methods and algorithms developed for solving various types of optimization problems.
2. To develop and promote research interest in applying optimization techniques in problems of Engineering and Technology.
3. To apply the mathematical results and numerical of optimization theory to different Engineering problems.

Module-1:

[8 Hrs]

Idea of Engineering optimization, Classification of optimization Problems, Optimization Problem and Model Formulation. **Linear programming:** Formulation of LPP, Simplex method, Big-M method, Two-phase Method, Dual Simplex method, Sensitivity analysis in linear programming.

Module-2: [8 Hrs]

Transportation problems: Finding an initial basic feasible solution by Northwest Corner rule, Vogel's approximation method, Degeneracy, Optimality test, MODI method, Least common cell method.

Assignment problems: Hungarian method for solution of Assignment Problems

Integer Programming: Integer Programming, Mixed Integer programming, Branch and Bound

method

Module-3:

[10 Hrs]

Non-linear programming: Introduction to non-linear programming. Constrained optimization, Multivariable optimization: Method of Lagrange Multipliers, Kuhn-Tucker conditions.

Unconstraint optimization: Powell's Method, Steepest Descent (Cauchy) Method, Conjugate Gradient (Fletcher-Reeves) Method, Newton's method.

Module-4:

[8 Hrs]

Game Theory: Concept, Game models, Two persons zero sum games and their solution, Pure & Mixed Strategy, solution of 2Xn and mX2 games by graphical approach.

Decision Theory: Concept, Decision under risk (EMV) & uncertainty**.**

Module-5:

[8 Hrs]

Queuing models: General characteristics, Markovian queuing model, M/M/1model, Limited queue capacity, multiple server, Finite sources, Queue discipline.

Course outcomes:

1. Understand importance of optimization of industrial process management.
2. Apply basic concepts of mathematics to formulate an optimization problem.
3. Analyses and appreciate variety of performance measures for various optimization problems

TEXT BOOKS:

1. S. S. Rao, *Engineering Optimization*, New Age International Publications.
2. Ravindran, D. T. Philips, J. Solberg, *Operations Research- Principle and Practice*, Second edition, Wiley India Pvt Ltd.
3. H.A.Taha, A.M.Natarajan, P.Balasubramanie, A.Tamilarasi, *Operations Research*, Pearson Education, Eighth Edition.

Reference Books:

1. S.D.Sharma, *Operations Research*, Kedarnath Publications.
2. F.S.Hiller, G.J.Lieberman, *Operations Research*, Tata McGraw Hill.
3. P.C.Biswal,*Optimization Engineering*, Scitech Publications
4. Prem Kumar Gupta, D.S.Hira, *Operations Research*, S.Chand Publications.

Digital Learning Resources

Course Name	Optimization from fundamentals
Course Link	https://nptel.ac.in/courses/105/103/105103210/
Course Instructor	Prof. Ankur A. Kulkarni IIT Bombay

Course Name	Introduction to Operating Systems
Course Link	https://nptel.ac.in/courses/106/106/106106144/
Course Instructor	Prof. Chester Rebeiro IIT Madras

22CS6PC01T	Machine Learning (3-0-0)	3 Credits

Course Objective:

1. Develop a concise knowledge on understanding of the fundamental concept of machine learning.
2. Understand the different learning algorithms and implement them.
3. Gain experience in applying machine learning algorithms to real world problem.

Module-1: [6 Hrs.]

Introduction to Machine Learning, Model Preparation, Modelling and Evaluation, Human learning versus machine learning, types of machine learning, applications of machine learning, tools for machine learning, Machine Learning Activities, Data structures for machine learning, Data Pre- processing, selecting a model, training a model, model representation and interpretability, evaluating performance of a model, improving performance of a model, Learning theory, Hypothesis and target class, Hilbert space, Inductive bias and bias-variance tradeoff.

Module-2: [6 Hrs.]

Feature Engineering, Bayesian Concept Learning, Introduction to feature engineering, feature transformation, feature subset selection, Importance of Bayesian methods, Bayes' theorem, concept learning through Bayes' theorem, Bayesian Belief Network

Module-3: [12 Hrs.]

Supervised Learning –Classification, Regression, Example of supervised learning, classification model, classification learning steps, common classification algorithms – KNN, Decission trees random forest, SVM, example of regression, common regression algorithms,

Module-4: [12 Hrs.]

UnsupervisedLearning –Clustering, pattern finding using association rules, Unsupervised learning versus supervised learning, applications of unsupervised learning, clustering and its types, Apriori algorithm for association rule learning

Module-5: [4 Hrs.]

Neural Network: Understanding the biological neuron, exploring artificial neuron, types of activation functions, early implementation of artificial neural network, architectures of neural network, learning process in artificial neural network, backpropagation, Overview of Deep Learning

Course Outcome:

1. Equip students with knowledge of fundamentals concepts in machine learning.
2. Ability to analyse and validate different learning algorithms.
3. fine tune machine learning algorithms and evaluate models generated from data.

Textbooks:

1: Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das, “Machine Learning”, Pearson Education 2: C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2010..

Reference Books:

J. Friedman, T. Hastie, and R. Tibshirani.The elements of statistical learning.Vol. 1, no. 10. New York: Springer series in statistics, 2001.

S. Shalev-Shwartz, and S. Ben-David. Understanding machine learning: From theory to algorithms. Cambridge university press, 2014.

Digital Learning Resources

Course Name	Introduction to Machine Learning
Course Link	https://nptel.ac.in/courses/106/106/106106139/
Course Instructor	Dr. Balaraman Ravindran, IIT Madras

22CS6PC02T	Software Engineering(3-0-0)	3 Credits

Course Objectives:

1. To understand the phases in a software project
2. To understand fundamental concepts of requirements engineering and Analysis Modeling.
3. To understand the various software design methodologies
4. To learn various testing and maintenance measures

Module-1: SOFTWARE PROCESS AND AGILE DEVELOPMENT [10 Hrs]

Introduction to Software Engineering, Overview of software development activities, Software Process, Perspective and Specialized Process Models –Introduction to Agility-Agile process- Extreme programming-XP Process.

Module-2: SOFTWARE REQUIREMENTS ENGINEERING [10 hrs]

Software Requirements: Functional and Non-Functional, User requirements, System requirements, Software Requirements Document – Requirement Engineering Process: Feasibility Studies, Requirements elicitation and analysis, requirements validation, requirements management. Classical analysis: Structured system Analysis, Data Dictionary.

Module-3: SOFTWARE ANALYSIS & DESIGN [10 hr]

Overview of design process: High-level and detailed design, Cohesion and coupling, Modularity and layering, Design process – Design Concepts-Design Model– Design Heuristic – Architectural Design.

Basic concepts of Object Oriented Analysis & Design.Structured Analysis using DFD -Structured Design using Structure Chart, User interface design, Command language, menu and iconic interfaces.

Module- 4: TESTING AND MAINTENANCE [10 hr]

Software testing fundamentals-Internal and external views of Testing-white box testing – basis path testing-control structure testing-black box testing- Regression Testing – Unit Testing – Integration Testing – Validation Testing – System Testing And Debugging –Software Implementation Techniques: Coding practices-Refactoring-Maintenance and Reengineering-BPR model-Reengineering process model-Reverse and Forward Engineering. **Software Project Management:** Introduction to software Estimation techniques, COCOMO I & II Model – Project Scheduling – Earned Value Analysis Planning – Project Plan, Planning Process

Course Outcome:

1. To apply the software engineering lifecycle by demonstrating competence in communication, planning, analysis, design, construction, and deployment
2. An ability to work in one or more significant application domains
3. Work as an individual and as part of a multidisciplinary team to develop and deliver quality software
4. Demonstrate an understanding of and apply current theories, models, and techniques that provide a basis for the software lifecycle
5. Demonstrate an ability to use the techniques and tools necessary for engineering practice

Text Book:

1. Rajib Mall, Fundamentals of Software Engineering, Fifth Edition, PHI, 2018

Reference Books:

1. Software Engineering, A Practitioner's Approach, Roger S. Pressman, TMG Hill.
2. Fundamentals of Software Engineering, Rajib Mall, PHI, 2014.
3. Software Engineering, I. Sommerville, 9th Ed., Pearson Education.

Digital Learning Resources

Course Name	Software Engineering
Course Link	NPTEL :: Computer Science and Engineering - NOC:Software Engineering
Course Instructor	Prof. Rajib Mall, IIT Kharagpur

22CS6PE01T	Computer Network Security (3-0-0)	3 Credits

Course Objective:

1. Exhibit knowledge to secure corrupted systems, protect personal data, and secure computer networks in an Organization.
2. Practice with an expertise in academics to design and implement security solutions.
3. Understand key terms and concepts in Cryptography, Governance and Compliance.
4. Develop cyber security strategies and policies
5. Understand principles of web security and to guarantee a secure network by monitoring and analyzing the nature of attacks through cyber/computer forensics software/tools.

Module:-I **[**10 Hrs]

Introduction to Security: What is security? Why we need Security? Security concerns, Security Goals: Confidentiality, Integrity, Availability, Authenticity and Accountability, Computer security challenges, Security Breach Impact levels: Low, Moderate and High, Security threats/attacks: passive and active, Security Policy, Security issues, Brief History of Malware, Types of Malware, Network Security Audit, The Orange Book, Legal Issues.

TCP/IP Security Attacks: TCP Segment Format, TCP Connection Setup, TCP Disconnection, IP Address Spoofing, Covert Channel, IP Fragment Attacks, TCP Flags, Syn Flood, Ping of Death, Smurf, Fin, UDP Flood Attack, Connection Hijacking, ARP Spoofing, DNS Spoofing, E-Mail Spoofing, Web Spoofing.

Module:-2 **[**8 Hrs]

Introduction to Cryptography, Symmetric-Key Cryptography: Traditional Ciphers, Simple Modern Ciphers, Modern Round Ciphers, Mode of Operations. Asymmetric-key Cryptography: RSA and Diffie-Hellman.

Network Security: Security Services, Message Confidentiality, Message Integrity, Message Authentication: MAC and HMAC, Digital Signature, Key Management: Symmetric-key Distribution: KDC, Session Keys, Kerberos, Public-key Distribution: Certification Authority, X.509, PKI.

Module:-3 **[**10 Hrs]

Authentication, Authentication methods, Passwords, Challenge-Response, Biometrics, Something you have, Two-factor authentication., Single Sign-On and Web Cookies.

Authorization, A brief history of authorization, Access control matrix, Compartments, Covert Channel, Inference Control, CAPTCHA, Firewalls and Proxies, Defense in depth, Computer Networks security zones, Concept of Demilitarized Zones (DMZ) in designing Corporate Networks, Analysis of Network Infrastructure, DMZ: Mail server, WWW Server, DNS Server. Network flooding, Anticipating attacks, IDS.

Module:-4 **[**6 Hrs]

Simple Security Protocols, Authentication Protocols: authentication using symmetric keys, authentication using public keys, session keys, perfect forward secrecy, mutual authentication, session keys, and PFS, Timestamps, Authentication and TCP, Zero knowledge proofs.

Module:-5 **[**6 Hrs]

SSH, SSL/TSL: SSL and Man-in-the-Middle, SSL connections, SSL Versus IPSec, , IPSec: IKE Phase I:

Digital Signature, Symmetric Key, Public Key Encryption, IPSec Cookies, IKE Phase II, IPSec and IP

Datagrams, Transport and Tunnel Modes, ESP and AH, Application Layer Security: Pretty Good Privacy (PGP).

Course Outcome:

1. Analyze and evaluate the cyber security needs of an organization.
2. Determine and analyze software vulnerabilities and security solutions to reduce the risk of exploitation. Measure the performance and troubleshoot cyber security systems.
3. Implement cyber security solutions and use of cyber security, information assurance, and cyber/computer forensics software/tools.
4. Comprehend and execute risk management processes, risk treatment methods, and key risk and performance indicators
5. Design and develop security architecture for an organization.
6. Design operational and strategic cyber security strategies and policies.

Textbooks:

1. Mark Stamp, Information Security: Principles and Practices, John Wiley & Sons, Hoboken, NJ, 2011. Chapters 1, 7, 8, 9, 10, 11, 13
2. Behrouza Forouzan, Data Communications and Networking, McGraw-Hill, 2006. Chapters 30, 31, 32.
3. Matt Bishop, Introduction to Computer Security, Addison-Wesley, 2005. Chapters 9, 10.4.2, 11, 22, 23.
4. Gert De Laet and Gert Schauwere, Network Security Fundamentals, Cisco Press, Indiana, 2004. Chapters 1, 2, 9. 10.

Reference books:

Richard Bejtlich, The Tao of Network Security Monitoring: Beyond Intrusion Detection, Addison- Wesley. Use this book for Practical.

Digital Learning Resources

Course Name	Cryptography And Network Security
Course Link	NPTEL :: Computer Science and Engineering - NOC:Cryptography And NetworkSecurity
Course Instructor	Prof. Sourav Mukhopadhyay, IIT Kharagpur

Course Name	Cryptography And Network Security
Course Link	NPTEL :: Computer Science and Engineering - Cryptography and Network Security
Course Instructor	Dr. Debdeep Mukhopadhyay , IIT Kharagpur

22CS6PE02T	Internet of Things (3-0-0)	3 Credits

Course Objective

1. To learn the basic issues, policy and challenges in the Internet.
2. To get an idea of the application areas where Internet of Things can be applied.
3. To understand the cloud and internet environment and various modes of communications with Internet.
4. To understand the various modes of communications with Internet.

Module-1: [10 Hrs]

Introduction & Concepts: Introduction to Internet of Things, Physical Design of IOT, Logical Design of IOT, IOT Enabling Technologies, IOT Levels. Domain Specific IOTs: Home Automation, Cities, Environment, Energy, Retail, Logistics, Agriculture, Industry, Health & Life Style, Challenges and Issues.

Module-2: [10 Hrs]

M2M & System Management with NETCONF-YANG: M2M, Difference between IOT and M2M, SDN and NFV for IOT, Software defined Networking, Network Function Virtualization, Need for IOT Systems Management, Simple Network Management Protocol, Limitations of SNMP, Network Operator Requirements, NETCONF, YANG, IOT Systems management with NETCONF-YANG

Module-3: [10Hrs]

IoT Protocols: Protocol Standardization for IoT and WSN Protocols-SCADA and RFID Protocols-Issues with IoT Standardization Protocols IEEE802.15.4-BACNet Protocol- , Architecture

- Network layer – APS Layer – Security.

Module-4: [10 Hrs]

Case Study and IoT Application Development: IoT applications in home- infrastructuressecurity- Industries- IoT electronic equipments. Use of Big Data and Visualization in IoT Industry 4.0 concepts - Sensors and sensor Node – Interfacing using Raspberry Pi/Arduino- Web Enabled Constrained Devices.

Course Outcome:

1. Able to understand the application areas of IOT
2. Able to realize the revolution of Internet in Mobile Devices, Cloud & Sensor Networks
3. Able to understand building blocks of Internet of Things and characteristics.

Text Book:

1. VijayMadisetti, Arshdeep Bahga," Internet of ThingsA Hands-On- Approach",2014, ISBN:978 0996025515
2. Honbo Zhou, "The Internet of Things in the Cloud:A Middleware Perspective" — CRCPress- 2012.
3. Arshdeep Bahga, Vijay Madisetti, "Internet of Things (A Hands-On-Approach)", VPT, 2014.
4. Luigi Atzor et.al, "The Internet of Things: A survey, ", Journal on Networks, Elsevier Publications,

October 2010.
5. Olivier Hersent, David Boswarthick, Omar Elloumi, “The Internet of Things – Key applications and Protocols”, Wiley, 2012..
6. Adrian McEwen, “Designing the Internet of Things”, Wiley Publishers, 2013.

Digital Learning Resources

Course Name	Introduction to internet of things
Course Link	NPTEL :: Computer Science and Engineering - NOC:Introduction to internet ofthings
Course Instructor	Prof. Sudip Misra, IIT Kharagpur

22CS6PE03T	Computer Vision (3-0-0)	3 Credits

Course Objective:

1. To Gain an understanding of the fundamental of extracting information from digital imagery.
2. To learn different task and algorithms related to computer vision
3. To learn to build to integrated computer vision model for solving real world problem

Module-1: [8 Hrs]

Introduction to computer vision, Machine vision systems, optics and lenses, image sensors, human vision and Neurovisual model; Marr's paradigm; Imaging geometry— world co-ordinate system and camera coordinate system, co-ordinate transformations, projection geometry, camera calibration, radiometry,Image brightness and radiometry, image formation and surface reflectance under different conditions, reflectance map and bidirectional reflectance distribution function, photometric stereo recovering albedo and surface orientation, shape from shading.

Module-2: [12 Hrs]

Feature detection, interest point and corner, harris corner detection, shift, surf, ransac and transformation , Local image features, level curve curvature, histogram of oriented gradients, blob detection, difference of gaussian, difference of Laplacian, Gestalt law, saliency detection. Texture - gray level co-ocurrence feature, Local binary pattern. Shape from Texture for Planes, Shape from Texture for Curved Surfaces. Segmentation- Clustering and Segmentation with Mean Shift, Graph cut segmentation.

Module-3: [12 Hrs]

Binocular technique— stereo pair, epipolar line and plane, Stereo matching, photogrammetry, monocular technique— texture processing and shape from texture, depth from focusing and symmetry, different range finder (active) -laser range finder, light-stripe method. Understanding line drawings, gradient and dual space, generalized cylinder, volumetric representation, edge and junction labelling; Labelling and recognition of scene objects; Construction of model-base and visual learning, model based recognition system— Acronym, model based recognition from sparse range data, 3D model based vision system, scene understanding, virtual reality.

Module-4: [10 Hrs]

Motion field, optical flow— smoothness, boundary conditions, discontinuities of optical flow, block based method, pre-recursive method, Bayesian method, motion segmentation method, motion from points and lines, token tracking, stereo and motion tracking, use of Kalman filter, focus of expansion, structure from motion, motion compensated filtering and restoration, video compression, active and passive surveillance

Course Outcome:

1. Identify and define basic concepts in computer vision, including camera parameters, camera calibration, edge detection, line detection, stereo algorithms, motion detection, structure from motion, image mosaic, face detection, and object recognition
2. Evaluate the correctness and generality of a computer vision method
3. Build integrated model computer vision model for solving real life problem.

Text Books:

1. Forsyth, David A., and Jean Ponce. *Computer vision: a modern approach*. Prentice Hall Professional Technical Reference, 2002.
2. Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press..

3. Szeliski, Richard. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

Digital Learning Resources

Course Name	Computer Vision and Image Processing - Fundamentals and Applications
Course Link	NPTEL :: Electrical Engineering - NOC:Computer Vision and Image Processing - Fundamentals and Applications
Course Instructor	Prof. M. K. Bhuyan, IIT Guwahati

Course Name	Computer Vision
Course Link	NPTEL :: Computer Science and Engineering - NOC:Computer Vision
Course Instructor	Prof. Jayanta Mukhopadhyay, IIT Kharagpur

19CS5PE02T	Cloud Computing (3-0-0)	3 Credits

Course Objectives

1. To provide students with the fundamentals and essentials of Cloud Computing.
2. To provide students a sound foundation of the Cloud Computing so that they are able to start using and adopting Cloud Computing services and tools in their real life scenarios.
3. To enable students exploring some important cloud computing driven commercial systems and applications.
4. To expose the students to frontier areas of Cloud Computing and information systems, while providing sufficient foundations to enable further study and research.

Module- 1: **[10 Hrs.]**

Evolution of Computing Paradigms - Overview of Existing Hosting Platforms, Grid Computing, Utility Computing, Autonomic Computing, Dynamic Datacenter Alliance, Hosting / Outsourcing, Introduction to Cloud Computing, Workload Patterns for the Cloud, "Big Data", IT as a Service, Technology Behind Cloud Computing,

Module- 2: **[10 Hrs.]**

A Classification of Cloud Implementations- Amazon Web Services - IaaS, The Elastic Compute Cloud (EC2), The Simple Storage Service (S3), The Simple Queuing Services (SQS), VMware vCloud - IaaS, vCloud Express, Google AppEngine - PaaS, The Java Runtime Environment,

Module- 3: **[10 Hrs.]**

The Python Runtime Environment- The Datastore, Development Workflow, Windows Azure Platform - PaaS, Windows Azure, SQL Azure, Windows Azure AppFabric, Salesforce.com - SaaS / PaaS, Force.com, Force Database - the persistency layer, Data Security, Microsoft Office Live - SaaS, LiveMesh.com, Google Apps - SaaS, A Comparison of Cloud Computing Platforms, Common Building Blocks.

Module- 4: **[10 Hrs.]**

Cloud Security – Infrastructure security – Data security – Identity and access management Privacy- Audit and Compliance.

Course Outcome:

1. Explain the core concepts of the cloud computing paradigm: how and why this paradigm shift came about, the characteristics, advantages and challenges brought about by the various models and services in cloud computing.
2. Apply the fundamental concepts in datacenters to understand the tradeoffs in power, efficiency and cost.
3. Identify resource management fundamentals, i.e. resource abstraction, sharing and sandboxing and outline their role in managing infrastructure in cloud computing.
4. Analyze various cloud programming models and apply them to solve problems on the cloud.

Text Book:

1. Kai Hwang, Geoffrey C. Fox and Jack J. Dongarra, "Distributed and Cloud Computing from Parallel Processing to the Internet of Things", Morgan Kaufmann, Elsevier, 2012
2. Barrie Sosinsky, "Cloud Computing Bible" John Wiley & Sons, 2010

Reference Books:

Tim Mather, Subra Kumaraswamy, and Shahed Latif, "Cloud Security and Privacy An Enterprise Perspective on Risks and Compliance", O'Reilly 2009

Digital Learning Resources

Course Name	Cloud computing
Course Link	https://nptel.ac.in/courses/106/105/106105167/
Course Instructor	Prof. Soumya Kanti Ghosh, IIT Kharagpur

Course Name	Cloud Computing and Distributed Systems
Course Link	https://nptel.ac.in/courses/106/104/106104182/
Course Instructor	Dr.Rajiv Misra, IIT Patna

22CS6PE05T	Natural Language Processing (3-0-0)	3 Credits

Course Objectives

1. Developing ability to syntactically, semantically and pragmatically understand Natural Language data for rule based processing.
2. Learn how Natural Language can also be analyzed statistically.
3. And to learn nuances of Natural Language based machine learning.

Module:I **[**10 Hrs]

Introduction, Regular Expression, Text Normalization, Edit Distance, N-gram Language Model, Data Generalization and Smoothing, Kneser-Ney Smoothing.

Lexical Semantics, Vector Semantics, Words and Vectors, Similarity Metrics Measures, Term Frequency Inverse Document Frequency, Word Embedding and its Semantic properties, Word2vec Models, Parts-of-Speech, HMM based POS Tagging, Sequence Processing, Context Free Grammar, Treebanks, Lexicalized Grammars.

Module:II **[10 Hrs]**

Parsing: Ambiguity Resolution, Statistical/Probabilistic Parsing, PCFG, Evaluating Parser,

Dependency Parsing, Dependency Relations and Formalisms, Transition-Based Dependency Parsing, Graph-Based Dependency Parsing, Representation of Meaning, Model-Theoretic Semantics, First-Order Logic, Event and State Representations, Description Logics.

Module III [10 Hrs]

Information Extraction, Named Entity Recognition, Relation Extraction, Time Extraction, Event Extraction, Template Filling, Semantic Role Labeling, Diathesis Alterations, The Proposition Bank, FramNet, Selection Restrictions, Sentiment Analysis: Defining Emotions, Creating Affect Lexicons, Semi-supervised Induction of Affect Lexcons, Sentiment Recognition, Affect Recognition, Connotation Frames.

Module: IV [10 Hrs]

Extraction based Text Summarization, Abstraction based Text Summarization, Coreference Resolution, Discourse Analysis, Machine translation, Information Retrieval based Question Answering, Knowledge based Question Answering, Dialog Systems, Chatbots.

Course Outcome:

1. Provides a modern and statistical perspective on natural language processing.
2. Enable the student to: acquire fundamentals of language technology; understand, implement, and apply state-of-the-art techniques to novel problems involving natural language data
3. Able to read and understand current research literature.

Text Books:

1. Dan Jurafsky and James H. Martin. *Speech and Language Processing*, Prentice-Hall. (3rd Edition)
2. James Allen. *Natural Language Understanding*, Pearson.
3. Chris Manning and Hinrich Schuetze. *Foundations of Statistical Natural Language Processing*, MIT Press.

4. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*, MIT Press.

Online Materials

1. Natural Language Processing, Skills gain - NLP, Tenserflow, Dialog Systems, Deep Learning.
2. Natural Language Processing using Python, Skills gain - NLP, Machine Learning specific NLP models.
3. Advanced NLP using Deep Learning, Skills gain - Deep Learning, Advanced NLP.
4. NLP Notes by Jacob Eisenstein.

Digital Learning Resources

Course Name	**Natural Language Processing**
Course Link	NPTEL :: Computer Science and Engineering - NOC:Natural Language Processing
Course Instructor	Prof. Pawan Goyal, IIT Kharagpur

Course Name	**Natural Language Processing**
Course Link	NPTEL :: Computer Science and Engineering - Natural Language Processing
Course Instructor	Prof. Pushpak Bhattacharyya, IIT Bombay

22CS6PE06T	Big Data Analytics (3-0-0)	3 Credits

Course Objective :

1. Understand the Big Data Platform and its Use cases
2. Real- time Analytics Platform(RTAP) and IBM Info sphere
3. Provide HDFS Concepts and Interfacing with HDFS
4. Understand Map Reduce Jobs
5. Provide hands on Hodoop Eco System and apply analytics on Structured, Unstructured Data.

MODULE -1: [8 Hrs]

Introduction: Big Data Overview, BI Versus Data Science, Current Analytical Architecture, Drivers of Big Data.

Data Analytics Lifecycle - Overview, Phases - Discovery, Data Preparation and Model planning, Model building, Communicate Results and Operationalize.

Industry examples of Big Data

MODULE -2: [12 Hrs]

Big Data Technology: Hadoop's parallel world,Data Discovery, Cloud and Big data, Predictive analytics, crowd sourcing analytics, Inter and Trans firewall analytics.

Information management: Big data foundation, Big data computing platforms, Big data computation, more on Big data storage, Big data computational limitaitons.

MODULE -3: [10 Hrs]

Estimating moments, Counting oneness in a window , Decaying window - Real- time Analytics Platform(RTAP) applications, IBM Info sphere , Big data at rest , Info sphere streams ,Data stage , Statistical analysis , Intelligent scheduler , Info sphere Streams, Predictive Analytics , Supervised , Unsupervised learning , Neural networks, Mining Frequent item sets , Market based model

MODULE -4: [10 Hrs]

Apriori Algorithm , Handling large data sets in Main memory , Limited Pass algorithm , Counting frequent item sets in a stream , Clustering Techniques , Hierarchical –KMeans, Clustering high dimensional data Visualizations, Visual data analysis techniques, interaction techniques, Systems and applications IBM for Big Data , Map Reduce Framework, Hadoop , Hive Sharding ,No SQL Databases , Hadoop Distributed file systems , Hbase, Impala , Analyzing big data with twitter , Big data for E-Commerce , Big data for blogs.

Course Outcome:

1. Identify Big Data and its Business Implications.
2. List the components of Hadoop and Hadoop Eco-System
3. Access and Process Data on Distributed File System
4. understand Data stage , Statistical analysis , Intelligent scheduler , Info sphere Streams
5. Develop Big Data Solutions using Hadoop Eco System
6. Analyze Infosphere BigInsights Big Data Recommendations.
7. Apply Machine Learning Techniques using R.

Recommended Books:

1. Data Science and Big data Analytics - EMC Education Services , 2015- WILEY

2. Michael Minelli, Michehe Chambers, “Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today’s Business”, 1st Edition, Ambiga Dhiraj, Wiely CIO Series
3. Big Data Analytics - Pyne , Rao and Rao, Springer 4: 2. Arvind Sathi, “Big Data Analytics: Disruptive Technologies for Changing the Game”, 1st Edition, IBM Corporation, 2012.

Seema Acharya, Subhasini Chellappan, "Big Data Analytics" Wiley 2015.

Digital Learning Resources

Course Name	Data Analytics with Python
Course Link	NPTEL :: Computer Science and Engineering - NOC:Data Analytics with Python
Course Instructor	Prof. A. Ramesh, IIT Roorkee

22CS6PE07T	Pattern Recognition(3-0-0)	3 Credits

Course Objective:

1. To introduce the fundamental principles of pattern recognition
2. To learn the algorithms involved in design and construction of various pattern recognition system component.
3. To analyse different components of pattern recognition system and learn to integrate them to form a unified system.

Module-1: **[6 Hrs]**

Metric, positive definite matrix, Eigen values and eigen vectors, SVD, statistical measures- mean, median, mode, correlation, dispersion matrix, binomial distribution, normal distribution, multi- variate normal distribution, basic concepts in probability theory-Bayes' theorem, Chebyshev's inequality, Laws of large numbers, central limit theorem. Unbiased estimation, maximum- likelihood, linear discriminant function, Introduction to pattern, recognition and its application, feature, and its characteristics, supervised and unsupervised approach to pattern recognition, pattern recognition system

Module-2: **[16 Hrs]**

Classification systems and its characteristics- overfitting, underfitting, training and test sets, standardization, normalization, distance function- Euclidean, Mahalanobis, Bayes decision classifier, error in bayes decision classifier, minimum distance classifier, Naive-Bayes rule, perceptron, K-NN classifier, SVM classifier, Multi-layer perceptron, decision tree, ensemble method of classifiers- bagging, boosting, random forest, cross-validation, Selection criteria of classification model- No Free Lunch Theorem, Bias-Variance problem, assessment of classification system- confusion matrix, precision, recall, F-score, sensitivity, specificity, ROC analysis AUC analysis. Class imbalance problem.

Module-3: **[10 Hrs]**

Clustering, clustering criterion, similarity measures, types of clustering, partition-based K-means, K- medoid, PAM, affinity based, Fuzzy C-means. Hierarchical clustering- agglomerative, divisive, Density based clustering- DBSCAN, probability based- mean-shift clustering, mixture model, spectral clustering, ensemble of clustering, subspace clustering for high dimensional data, cluster validation- external and internal criteria.

Module-4: **[8 Hrs]**

Dimensionality reduction, curse of dimensionality, feature selection, feature selection criterion, class separability measures- divergence, Chernoff bound, Bhattacharyya distance, scatter matrices, Filters- mutual information, the pointwise mutual information, Pearson product-moment correlation coefficient, Relief-based algorithms, branch and bound, wrappers, feature reduction- Karhunen-loeve transform, independent component analysis, principle component analysis, linear discriminant analysis, Kernel PCA. Application of Pattern recognition for image analysis.

Course Outcome:

1. Equip students with knowledge of different algorithms for classification, clustering and dimensionality reduction.
2. Ability to analyse and validate different algorithms used for pattern recognition.
3. Ability to integrate algorithm to build a unified pattern recognition model for solving real world problem.

Suggested Books:

1. Duda, R.O., Hart, P.E., and Stork, D.G. *PatternClassification*. Wiley-Interscience. 2nd Edition. 2001
2. Theodoridis, S. and Koutroumbas, K. *Pattern Recognition*. Edition 4. Academic Press, 2008.
3. Hastie, T., Tibshirani, R. and Friedman, J. *The Elements of Statistical Learning*. Springer. 2001.
4. Bishop, C. M. *Pattern Recognition and Machine Learning*. Springer. 2007

Digital Learning Resources

Course Name	**Pattern Recognition**
Course Link	NPTEL :: Computer Science and Engineering - Pattern Recognition
Course Instructor	Prof. Sukhendu Das, Prof. C.A. Murthy

22CS6PE08T	Software Testing & Quality Assurance (3-0-0)	3 Credits

Course Objective:

1. To learn the evolution of software testing techniques, myths and facts of software testing, models for testing processes, various types of software testing.
2. To design test cases using black-box and white-box testing techniques
3. To understand basic concepts of regression testing, problems of regression testing and types of regression testing techniques.
4. To learn the strategies for testing of object-oriented applications and web-based applications
5. To learn about test automation and present day practices

Unit-1

Quality Revolution, Software Quality, Role of Testing, Verification and Validation, Failure, Error, Fault and Defect, Notion of Software Reliability, Objective of Testing, What is a Test Case?, Expected Outcome, Concept of Complete Testing, Testing Activities, Test Oracle, Testing Levels, Regression Testing, White-Box and Black Testing, Test Planning and Design, Monitoring and Measuring Test Execution, Test Tools and Automation

Unit Testing: Concept of Unit Testing, Static and Dynamic unit Testing, Mutation Testing, Debugging, Unit Testing in eXtreme Programming.

Unit-2

Control Flow Testing: Outline of Control Flow Testing, Control Flow Graph, Path in a CFG, Path selection Criteria, All-Path Coverage Criterion, Statement Coverage Criterion, Branch Coverage Criterion, Generation of Test Input, Example of Test Data Selection.

Data Flow Testing: Data Flow Anomaly,. Overview of Dynamic Data Flow Testing, Data Flow Graph, Data Flow Terms, Data Flow Testing Criteria, Comparison of Data Flow Test Selection Criteria, Feasible Paths and Test Selection Criteria, Comparison of Testing Techniques.

System Integration Testing: Concept of Integration Testing, Different Types of Interfaces and Interface Errors, Granularity of System Integration Testing, System Integration Techniques, Software and Hardware Integration, Test Plan for System Integration, Off-the-Shelf Component Integration, Off-the-Shelf Component Testing, Built-in Testing

Unit – 3

System Test Categories: Basic Tests, Functionality Tests, Robustness Tests, Interoperability Tests, Performance Tests, Scalability Tests, Stress Tests, Load and Stability Tests, Reliability Tests, Regression Tests, Documentation Tests.

Functional Testing: Equivalence Class Partitioning, Boundary Value Analysis, Decision Tables, Random Testing, Error Guessing, Category Partition.

System Test Planning And Automation: Structure of a System Test Plan, Introduction and Feature Description, Assumptions, Test Approach, Test Suite Structure, Test Environment, Test Execution Strategy, Test Effort Estimation, Scheduling and Test Milestones, System Test Automation, Evaluation and Selection of Test Automation Tools, Test Selection Guidelines for Automation, Characteristics of Automated Test Cases, Structure of an Automated Test Case, Test Automation Infrastructure.

Acceptance Testing: Types of Acceptance Testing, Acceptance Criteria, Selection of Acceptance Criteria, Acceptance Test Plan, Acceptance Test Execution, Acceptance Test Report, Acceptance Testing in eXtreme Programming.

Unit-4

Software Reliability: Definition, Factors Influencing Software Reliability, Application of Software Reliability, Operational Profiles.

Software Quality: Five Views of Software Quality, McCall's Quality Factors and Criteria, Quality Factors Quality Criteria, Relationship between Quality Factors and Criteria, Quality Metrics, ISO 9126 Quality Characteristics, ISO 9000:2000 Software Quality Standard ISO 9000:2000 Fundamentals, ISO 9001:2000 Requirements.

Maturity Models: Basic Idea in Software Process, Capability Model(CMM) Model, Architecture, Five Levels of Maturity and Key Process Areas, Common Features of Key Practices, Application of CMM, CMMI, Test Process Improvement (TPI), Testing Maturity Model (TMM).

Textbook:

Software Testing and Quality Assurance: Theory and Practice, Kshirasagar (Sagar) Naik, University of Waterloo, Priyadarshi (Piyu) Tripathy, NEC, Wiley , 2008.

Reference Book:

Software Quality Assurance, Daniel Galin, Pearson Education

MOOC course:

1. https://www.edx.org/course/software-testing-fundamentals-usmx-umuc-stv1-1x-4 by-Dr. Michael Scott Brown, Program Chair of the Software Engineering Master's

University of Maryland University CollegeDr. Renata Rand McFadden, Adjunct Professor University of Maryland University College

Digital Learning Resources

Course Name	**Software Testing**
Course Link	NPTEL :: Computer Science and Engineering - NOC:Software testing
Course Instructor	Prof. Meenakshi D'Souza, IIIT Bangalore

22CS6OE01T	Data Analytics (3-0-0)	3 Credits

Course Objective:

To optimize business decisions and create competitive advantage with Big Data analytics

- To explore the concepts regression and classification.
- To learn to analyze the complexity of different techniques.
- To understand the various additive models and boosting techniques.
- To understand the Neural Networks, Support Vector Machines,and K-nearest Neighbor.

Module: 01 **[8 Hours]**

Linear Methods for Regression and Classification: Overview of supervised learning, Linear regression models and least squares, Multiple regression, Multiple outputs, Subset selection , Ridge regression, Lasso regression , Linear Discriminant Analysis , Logistic regression , Perceptron learning algorithm.

Module: 02 **[8 Hours]**

Model Assesment and Selection : Bias,Variance,and model complexity,Bias-variance trade off, Optimisim of the training error rate ,Esimate of In-sample prediction error,Effective number of parameters, Bayesian approach and BIC, Cross- validation ,Boot strap methods, conditional or expected test error.

Module: 03 **[8 Hours]**

Additive Models,Trees,and Boosting: Generalized additive models, Regression and classification trees , Boosting methods-exponential loss and AdaBoost, Numerical Optimization via gradient boosting ,Examples (Spam data, California housing , NewZealand fish, Demographic data)

Module: 04 **[8 Hours]**

Neural Networks(NN) , Support Vector Machines(SVM),and K-nearest Neighbor: Fitting neural networks, Back propagation, Issues in training NN, SVM for classification, Reproducing Kernels, SVM for regression, K-nearest – Neighbour classifiers(Image Scene Classification)

Module: 05 **[8 Hours]**

Unsupervised Learning and Random forests: Association rules, Cluster analysis, Principal Components, Random forests and analysis. (II) Inferential Statistics and Prescriptive analytics.

Text Books:

Trevor Hastie, Robert Tibshirani, Jerome Friedman , The Elements of Statistical Learning-Data Mining, Inference, and Prediction ,Second Edition , Springer Verlag, 2009.

G.James,D.Witten,T.Hastie,R.Tibshirani-An introduction to statistical learning with applications in R,Springer,2013.

E.Alpaydin, Introduction to Machine Learning, Prentice Hall Of India,2010,(Chapter-19) References

C.M.Bishop –Pattern Recognition and Machine Learning,Springer,2006

L.Wasserman-All of statistics.

Digital Learning Resources

Course Name	Data Analytics with Python
Course Link	NPTEL :: Computer Science and Engineering - NOC:Data Analytics with Python
Course Instructor	Prof. A. Ramesh, IIT Roorkee

22CS6PC02L	Software Engineering Lab (0-0-2)	1 Credits

Course Objectives:

1. To understand the software engineering methodologies involved in the phases for project development.
2. To gain knowledge about open source tools used for implementing software engineering methods.
3. To exercise developing product-startups implementing software engineering methods.
4. Open source Tools: StarUML / UMLGraph / Topcased

Laboratory Experiments

Prepare the following documents and develop the software project startup, prototype model, using software engineering methodology for at least two real time scenarios.

Problem Analysis and Project Planning & Software Requirement Analysis-Thorough study of the problem – Identify Project scope, Objectives, Describe the individual Phases/modules of the project and Identify deliverables. Identify functional and non-functional requirements. **[3hr]**

Data Modeling – Use work products – data dictionary. **[1hr 30 min]**

Software Designing - Develop DFD, use case diagrams and activity diagrams, class diagrams, sequence diagram, interaction diagram. **[8hr 30 min]**

Software Testing – Manual Testing process with real time example / sample experiments.

[2hr]

Sample problem statement for Experiments:

Title: Library Information System:

- The library has 10,000 books each book is assigned a unique identification number. The library clerk

should be able to enter the details of the books into the library information system (LIS) through suitable interface.

- There are four categories of members i.e., undergraduate students, postgraduate student, research scholar and faculty members.
- Each library member is assigned a unique library member code.
- Each undergraduate student can issue up to two books for one month duration.
- Each postgraduate student can issue up to four books for one month duration.
- Each research scholar student can issue up to six books for three months duration.
- Each faculty member can issue up to ten books for six months duration.
- The LIS should answer query regarding the availability of the book. If available, it also shows the rack number and number of copies available.
- LIS issue book to each register member. When the member returns the book, update the member's account and makes books available for future use.
- Member can reserve books, which have been issued. When the book is returned, the LIS must issue the book to the member who has reserve the book.
- When the member return the book, LIS checks for the issue date and return date, if it is greater than the authorized days than issue a fine to them.
- LIS issues reminder to the members who issue the book and it is overdue.
- LIS should allow the librarian to create and delete member records.

Course Outcome:

1. Build a fully functional, interactive, layered, distributed, database-backed software system from the ground-up as part of a small, agile, development team in a laboratory setting
2. Become acquainted with historical and modern software methodologies
3. Understand the phases of software projects and practice the activities of each phase
4. Practice clean coding
5. Take part in project management
6. Become adept at such skills as distributed version control, unit testing, integration testing, build management, and deployment

Text Books:

Robert C. Martin, Clean Code: A Handbook of Agile Software Craftsmanship, Prentice Hall, 2008.

22CS6PC02L	Machine Learning Lab **(0-0-2)**	**1 Credits**

1. Make use of Data sets in implementing the machine learning algorithms
2. Implement the machine learning concepts and algorithms in any suitable language of choice.

Laboratory Experiments

1. Build a multivariate logistic regression model to classify glass type of glass given different glass mixture features using the Glass Identification Dataset from UCI Machine Learning Repository.
2. Implement supervised machine learning algorithm (Classification – K Nearest Neighbourhood) in python to classify breast tumour data into malignant breast tumour or benign breast tumour (use breast tumour dataset) and obtain its accuracy level.
3. Implement supervised machine learning algorithm (Classification – K Nearest Neighbourhood) in python to classify iris data into setosa, virginica, versicolor using iris dataset and obtain its accuracy level.
4. Implement supervised machine learning algorithm (Classification – Support Vector Machine) in python to classify breast tumour data into malignant breast tumour or benign breast tumour (use breast tumour dataset) and obtain its accuracy level.
5. Write a python program to build an email spam classifier using support vector machines for the Spam base dataset from UCI machine learning repository.
6. Implement unsupervised machine learning algorithm (Clustering – K Means) in python on Titanic dataset to cluster data (use Titanic dataset) by removing the class label.
7. Implement unsupervised machine learning algorithm (Clustering – K Means) in python on Breast Tumour dataset to cluster data (use Breast Tumour dataset) by removing the class label.
8. Implement unsupervised machine learning algorithm (Clustering – Hierarchical) in python on Titanic dataset to cluster data (use Titanic dataset).
9. Implement Apriori algorithm in python to find rules which explain association between different products for given transactions at a retail store. (The data is available at https://drive.google.com/file/d/1NUXoptUlHY8z4KcFKpFA6sQN5KnWzk3p/view?usp=sharing)
10. Implement text classification using neural network in python/R on Twenty Newsgroupdataset from UCI machine learning repository.
11. Implement supervised machine learning algorithm (Classification - Naïve Bayes algorithm) in python/R on Pima Indians Diabetes dataset and obtain its accuracy level.
12. classification and prediction algorithms on UCI dataset using Python's

scikit-learnlibrary

References:

1) Peter Harrington, “Machine Learning in Action”, DreamTech
2) Michael Bowles, “Machine Learning in Python”, Wiley
3) Gavin Hackeling, Mastering Machine Learning with scikit-learn, Packt
4) Giuseppe Bonaccorso, Machine Learning Algorithms - Second Edition, Packt

Course Outcome:

1. Understand the implementation procedures for the machine learning algorithms
2. Design Java/Python programs for various Learning algorithms.
3. Apply appropriate data sets to the Machine Learning algorithms
4. Identify and apply Machine Learning algorithms to solve real world problems