Department of Computer Applications
DIT University Dehradun

Course Structure & Syllabus
for
Pre Ph.D. (Computer Science) Course Work
Session: 2018-19

Pre Ph.D (Computer Science)

Course Category	Course Code	Course Name	Periods			Credits
			L	T	P	
UC	MB901	Research Methodology	4	0	0	4
DE		Elective 1	4	0	0	4
DE		Elective 2	4	0	0	4
DE		Elective3	4	0	0	4
DC	DS001	Seminar	0	0	2	1

List of Electives

S.No.	Subject Code	Course
Elective 1		
1	CA941	Modeling and Simulation
2	CA942	Cloud Technologies
3	CA943	Fuzzy Logic & Genetic Algorithms
4	CA944	Big Data Analytics
5	CA945	Digital Image Processing
Elective 2		
6	CA946	Cryptography
7	CA947	Advanced Computer Networks
8	CA948	Neural Networks & Neuro Fuzzy Systems
9	CA949	Machine Learning
10	CA951	Wireless Sensor Networks
Elective 3		
11	CA952	Advanced Software Engineering
12	CA953	Artificial Intelligence
13	CA954	Distributed Systems
14	CA955	Information & Coding Theory
15	CA956	Mobile and Ad-Hoc Networks
16	CA957	Advanced Data Warehousing and Mining

Note : Apart from above listed Elective courses, Research Scholar may choose any course across departments being offered at PG level, if it is required/suggested by the Research Committee.

Pre Ph.D (Computer Science)

Subject Code	MB901	Subject Title	Research Methodology						
LTP	4 0 0	Credit	4	Subject Category	UC	Year	1st	Semester	I / II

UNIT – I
Fundamentals of Research: Defining research, Objectives of research, types, research process, deductive and inductive reasoning;
Identifying and formulating a research problem, Literature review: Search for existing literature (World Wide Web, Online data bases), Review the literature selected (Case studies, review articles and Meta-analysis), Develop a theoretical and conceptual framework, Writing up the review,
Definition of variables: Concepts, indicators and variables, Types of variables, Types of measurement scales, Constructing the Hypothesis- Null(Research) and alternative, one-tailed and two-tailed testing, errors in testing. Ethical and Moral Issues in Research, Plagiarism, tools to avoid plagiarism – Intellectual Property Rights – Copy right laws – Patent rights

UNIT – II
Research Design: Design of Experiments: Research Designs -Exploratory, Descriptive and Experimental, Experimental designs- Types of Experimental Designs

UNIT – III
Sampling, Sampling distribution, and Data Collection: Sampling distribution, Normal and binomial distribution, Reasons for sampling, sampling technique, sampling errors.Sources of Data-Primary Data, Secondary Data, Data Collection methods

UNIT – IV
Statistical Data Analysis: Descriptive and inferential statistical analysis. Testing of hypothesis with Z-test, T-test and its variants, Chi-square test, ANOVA, Correlation, Regression Analysis, Introduction to data analysis data using SPSS20.0

UNIT – V
Research Report: Writing a research report- Developing an outline, Formats of Report writing, Key elements- Objective, Introduction, Design or Rationale of work, Experimental Methods, Procedures, Measurements, Results, Discussion, Conclusion, Referencing and various formats for reference writing of books and research papers, Writing a Research Proposal.

Books Recommended:
1. Ganesan R, Research Methodology for Engineers , MJP Publishers, Chennai. 2011
2. C.R.Kothari, "Research Methodology", 5th edition, New Age Publication,
3. Cooper, "Business Research Methods", 9th edition, Tata McGraw hills publication
4. Walpole R.A., Myers R.H., Myers S.L. and Ye, King: Probability & Statistics for Engineers and Scientists, Pearson Prentice Hall, Pearson Education, Inc. 2007.
5. Anderson B.H., Dursaton, and Poole M.: Thesis and assignment writing, Wiley Eastern 1997.
6. Bordens K.S. and Abbott, B.b.: Research Design and Methods, McGraw Hill, 2008.
7. Morris R Cohen: An Introduction to logic and Scientific Method (Allied Publishers) – P 197-222; 391–403

Pre Ph.D (Computer Science)

Subject Code	CA941	Subject Title	Modelling & Simulation						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives:-The goal is to introduce students to basic methods and tools for modeling and simulation of continuous, discrete and combined systems in time and space dimension.

Unit I (8 L)
Modelling: Definition of System, system concepts, types of system, continuous and discrete systems, modelling process, verification and validation. Simulation: Introduction, classification of simulation models, advantages and disadvantages of simulation.

UnitII (8 L)
Discrete system simulation: Monte Carlo method, Random Number Generation: Congruence generators, long period generators, uniformity and independence testing. Random Variate Generation: Location, scale and shape parameters, discrete and continuous probability distributions; Inverse transformation method.

Unit III (8 L)
Queuing Theory: Introduction, notation and assumption, Little's theorem, queuing model with poison input, exponential service and arbitrary service times, simulation of queuing system, simulation of single-server queue, Simulation of two server queuing system.

Unit IV (8 L)
Inventory Control: Elements of Inventory Theory, more complex inventory models, finite and infinite delivery rate model with and without back ordering, simulation of inventory systems.

Unit V (10 L)
Evaluation of simulation: length of simulation runs, variance reduction techniques. Project Management: PERT/CPM techniques, simulation of PERT networks. Model as components of information systems, modelling for decision support. Virtual Reality: the ultimate interactive model.

Learning Outcomes: - Students will acquire knowledge of system definition, classification and basic knowledge of simulation system principles. They will be able to create simulation models of various types.

Text Books:
1. **"Discrete-Event Modeling and Simulation: A Practitioners Approach",**Wainer, A. G, CRC Press, Boca Raton, FL,2009

2. **"System Simulation Techniques with MATLAB and Simulink**",DingyüXue, Yang Quan Chen, John Wiley & Sons,UK,2014

References Books:
1. **"Discrete-Event System Simulation Banks"**, J., J.S. Carson, B.L. Nelson, and D.M. Nicol,Prentice-Hall, Upper Saddle River, NJ,4/e, 2005

2. **"Simulation Modeling and Analysis,Law",** A.M. and W.D. Kelton, McGraw-Hill, New York, NY,, 3/e 2000

Pre Ph.D (Computer Science)

Subject Code	CA942	Subject Title	Cloud Technologies						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -The objective of the course is to make the students able to learn about cloud computing and the technologies working behind it. The students will learn the technical foundation of cloud such as Virtualization and Service oriented architecture. They will learn the basic cloud applications and their benefits and why it is the wave of future.

Unit I (8 L)

Fundamentals of Cloud Computing: What it is & what it is not?, The old IT infrastructure versus the cloud, www, Internet, Cloud and cloud computing, Motivation for Cloud Computing, A comparison of IT infrastructure options, Historical developments: Client-Server computing, Peer to Peer Computing, Distributed Computing, Cluster & Grid Computing, Principals of parallel and distributed computing- Eras of computing, parallel vs. Distributed computing, Elements of parallel of computing, Elements of Distributed of computing, Technologies of distributed computing: Service oriented computing and architecture.

Unit II (9 L)

Introduction to High performance computing, Technical foundations of Cloud Computing, Goals of Cloud Computing, Resource Sharing at various levels,Cloud Architecture, NIST USA Cloud Computing Model: Essential Characteristics, Delivery Models, IaaS, PaaS, Saas, Deployment Models, Public, private, Hybrid & Community, Pricing Model of Cloud Computing, Advantages & Disadvantages of Cloud Computing,Open challenges: Cloud definition, cloud interoperability and standards, scalability and fault tolerance, Security, trust and privacy, Organizational aspects.

Unit III (8 L)

Companies in Cloud Computing, Cloud computing Engines- GAE, EC2, Microsoft Azure, Virtualization: Characteristics of virtualized environment, taxonomy of virtualization techniques, Virtualization and cloud computing, pros and cons of virtualization, Technology examples, Flex Tenancy architecture, Pros & Cons of Cloud service development, Cloud application development, Cloud computing applications, Cloud Computing for everyone, Computing for Community & corporate.

Unit IV (8 L)

Security: Securing the cloud-The security boundary, Security service boundary, security mapping, Data security- Brokered cloud storage access, storage location and tenancy, encryption, Auditing and compliance, Establishing identity and presence, Network security, Host security, compromise response, High-throughput computing: Task programming.

Unit V (9 L)

Advanced topics in cloud computing: energy efficiency in clouds, market based management of clouds- market oriented cloud computing; A reference model of MOCC technologies and initiatives supporting MOCC, Federated clouds / InterClouds, Third party cloud services- Meta CDN; SpotCloud, Using the mobile cloud.

Learning Outcomes:-Having successfully completed this course, the student will demonstrate:

1. An ability to make a cloud application.
2. An ability to share online calendars and to-do-list.
3. An ability to collaborate online for a big project.
4. An ability of hardware and software virtualization.
5. An ability of private cloud development.

Text Books:
1. **"Handbook of Cloud Computing"**,Furht, Borko, Escalante, Armando,Springer,USA, 1/e,2010
2. **"Cloud Computing: Concepts, Technology &Architecture",**ThomasErl et al,Prentice Hall,USA,1/e 2013

Reference Books:
1. **"A Road to Cloud Computing:A Beginner's Perspective"** HarjotDhawan,LAP Lambert Academic Publishing,USA,2012
2. **"Cloud Computing for Dummies",** Judith Hurwitz,, ,John Wiley &Sons,USA,1/e 2010

Pre Ph.D (Computer Science)

Subject Code	CA943	Subject Title	Fuzzy Logic & Genetic Algorithms						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: - The course has been designed with the purpose to teach the students the fundamentals of Reasoning under uncertain circumstances (covered through Probabilistic Reasoning) and the Reasoning with imprecise knowledge (covered through Fuzzy Sets and Fuzzy Logic). The objective of the course is also to teach the fundamentals and applications of Genetic Algorithms.

Unit I (8 L)

Fuzzy Sets (Introduction)

Basic concepts of fuzzy logic, Fuzzy sets and Crisp sets, Fuzzy set theory, Basic operations, Properties of fuzzy sets, Fuzzy and Crisp relations, Fuzzy to Crisp conversion.

Unit II (8 L)

Fuzzy Logic (Fuzzy Membership, Rules)

Membership functions, Propositional logic and predicate logic, Resolution, Inference in fuzzy logic, Fuzzy if-then rules, Fuzzy mapping rules, Fuzzy implications, Min-Max Theorem, Resolution Rule under Fuzzy environment, Refutation method for theorem proving, Defuzzification, Fuzzy Classification

Unit III (10 L)

Reasoning with uncertain and incomplete information: The statistical approach to uncertainty, Introduction, Uncertain & incomplete knowledge. Review of Probability theory, Conditional Probabilities, Bayes Theorem with Evidence variable, Numerical Problems.

Unit IV (8 L)

Bayesian Networks, Bayesian reasoning. Decision Making, Joint Probabilities, Relationships, Chain Rule, Polytrees.,Dempster-Shafer theory of evidence, Certainty Theory, Non-monotonic systems.

Unit V (8 L)

Theoretical Foundation of Genetic Algorithms

Introduction: Basic Operators: Reproduction, Crossover & Mutation. Fitness function. Search Space, Schemas & Two-Armed and k-armed problem, exact mathematical models, Applications of Genetic Algorithms.

Learning Outcomes:- At the end of the course, the students shall have learnt the fundamentals and gained the requisite knowledge on the following:

 (i) The Fuzzy Sets Fuzzy Systems and the reasoning using the Fuzzy Logic
(ii) How to develop a decision support system using the fuzzy knowledge
(iii) The Reasoning Techniques under uncertain circumstances
(iv) The Bayesian Networks and the Bayesian Reasoning
(v) Dempsters- Shafers theory and Certainty theory as alternative to Bayesian reasoning uncer uncertain circumstances.
(vi) Genetic Algorithms provide one of the best optimization techniques when the serch space is large. Starting from the very fundamentals the students shall learn as to how to find solutions to problems using Genetic Algorithms.

Text Books:
1. "**Fuzzy Logic with Engineering Applications"** Timothy J. Ross, John Wiley &Sons,Englnd,2/e, 2004
2. **Neural Networks,Fuzzy Logic and genetic Algorithm**,SRajsekaran et al,PHI,New Delhi,1/e,2006
3. **"Genetic Algorithms in search ,optimization & Machine Learning"**, Goldberg, Pearson Education, Indi,4(Reprint) ,2009
4. **Neural Networks and Learning Machines**, Simon O. Haykin,Pearson Education,USA,3/e,2009

Reference Books:
1. **Machine Learning: Neural Networks, Genetic Algorithms, and Fuzzy Systems**, HojjatAdeli et al,John Wiley &Sons,USA,1/e,1995
2. **Machine Learning: The Art and Science of Algorithms that Make Sense of Data,**PeterFlach,Cambridge University Press,UK,1/e, 2012

Pre Ph.D (Computer Science)

Subject Code	CA944	Subject Title	Big Data Analytics						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -:- This course on Big Data Analytics consists of coherent body of ideas and methods to acquaint the student with the basic programs in the computational and human Intelligence field and their underlying theory. Students will be capable to quickly adapt to new technology in the field of Big Data, assimilate new information, and solve real world problems.

Unit 1 Introduction   (8L)
Examples, data science articulated, history and context, technology landscape.

Unit 2 Data Manipulation at Scale   (8 L)
Databases and the relational algebra ,Parallel databases, parallel query processing, in-database analytics ,MapReduce, Hadoop, relationship to databases, algorithms, extensions, languages ,Key-value stores and NoSQL; tradeoffs of SQL and NoSQL

Unit 3 Analytics   (8 L)
Topics in statistical modeling: basic concepts, experiment design, pitfalls, Topics in machine learning: supervised learning (rules, trees, forests, nearest neighbor, regression), optimization (gradient descent and variants), unsupervised learning.

Unit 4 Communicating Results   (7 L)
Visualization, data products, visual data analytics, Provenance, privacy, ethics, governance.

Unit 5 Special Topics   (9L)
Graph Analytics: structure, traversals, analytics, PageRank, community detection, recursive queries semantic web.

Learning Outcomes: - On completion of the programme the candidates will be able to apply the knowledge of computing tools and techniques in the field of Big Data for solving real world problems encountered in the Software Industries. Candidate will be able to analyze the various technologies & tools associated with Big Data and will be able to identify the challenges in Big Data with respect to IT Industry and pursue quality research in this field with social relevance.

Text Books:
1. **Intelligent Data Analysis**,Michael Berthold, David J. Hand,2/e, Springer, 2007.
2. **Taming the Big Data Tidal Wave:Finding Opportunities in Huge Data Streams with Advanced Analytics, Bill Franks**", 1/e,John Wiley & sons, 2012.

Reference Books:
1. **Making Sense of Data I**,Glenn J. Myatt, 2/e, John Wiley & Sons, 2014
2. **Data Mining Concepts and Techniques**, Jiawei Han, Micheline kamber, 2/e, Elsevier, Reprinted 2008.

Pre Ph.D (Computer Science)

Subject Code	CS945	Subject Title	Digital Image Processing						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -To learn and understand the fundamentals of digital image processing, and various image Transforms, Image Enhancement Techniques, Image restoration Techniques and methods, image compression and Segmentation used in digital image processing.

Unit I (8 L)

Introduction and Fundamentals

Motivation and Perspective, Applications, Components of Image Processing System, Element of Visual Perception, A Simple Image Model, Sampling and Quantization.

Image Enhancement in Spatial Domain

Introduction; Basic Gray Level Functions – Piecewise-Linear Transformation Functions:Contrast Stretching; Histogram Specification; Histogram Equalization; Local Enhancement; Enhancement using Arithmetic/Logic Operations – Image Subtraction, Image Averaging; Basics of Spatial Filtering; Smoothing - Mean filter, Ordered Statistic Filter; Sharpening – The Laplacian.

Unit II (8 L)

Image Enhancement in Frequency Domain

Fourier Transform and the Frequency Domain, Basis of Filtering in Frequency Domain, Filters – Low-pass, High-pass; Correspondence Between Filtering in Spatial and Frequency Domain; Smoothing Frequency Domain Filters – Gaussian Lowpass Filters; Sharpening Frequency Domain Filters – Gaussian Highpass Filters; Homomorphic Filtering.

Image Restoration

A Model of Restoration Process, Noise Models, Restoration in the presence of Noise only-Spatial Filtering – Mean Filters: Arithmetic Mean filter, Geometric Mean Filter, Order Statistic Filters – Median Filter, Max and Min filters; Periodic Noise Reduction by Frequency Domain Filtering – Bandpass Filters; Minimum Mean-square Error Restoration.

Unit III (8 L)

Color Image Processing

Color Fundamentals, Color Models, Converting Colors to different models, Color Transformation, Smoothing and Sharpening, Color Segmentation.

Morphological Image Processing: Introduction, Logic Operations involving Binary Images, Dilation and Erosion, Opening and Closing, Morphological Algorithms – Boundary Extraction, Region Filling, Extraction of Connected Components, Convex Hull, Thinning, Thickening

Unit IV (8 L)

Registration:

Introduction, Geometric Transformation – Plane to Plane transformation, Mapping, Stereo Imaging – Algorithms to Establish Correspondence, Algorithms to Recover Depth

Segmentation: Introduction, Region Extraction, Pixel-Based Approach, Multi-level Thresholding, Local Thresholding, Region-based Approach, Edge and Line Detection: Edge Detection, Edge Operators, Pattern Fitting Approach, Edge Linking and Edge Following, Edge Elements Extraction by Thresholding, Edge Detector Performance, Line Detection, Corner Detection.

Unit V (8 L)

Feature Extraction: Representation, Topological Attributes, Geometric Attributes Description, Boundary-based Description, Region-based Description and Relationship.

Object Recognition: Deterministic Methods, Clustering, Statistical Classification, Syntactic Recognition, Tree Search, Graph Matching

Learning Outcomes: -To understand and gain complete knowledge about

Pre Ph.D (Computer Science)

1. The fundamentals of digital image processing

2. Image transform used in digital image processing

3. Image enhancement techniques used in digital image processing

4. Image restoration techniques and methods used in digital image processing

Text Books:
1. **Digital Signal Processing**,S. Salivahanan et al, TMH, 3rd edition, 2000.
2. **Theory and Application of Digital Signal Processing**,L. R. Rabiner& B. Gold, PHI, NJ, 3rd edition, 2009

Reference Books:
1. **Digital Image Processing and Computer Vision**,R.J. Schalkoff , John Wiley and Sons, NY,3rd Edition, 1996.

2. **Fundamentals of Digital Image Processing**,A. K. Jain, Prentice Hall, Upper Saddle River, NJ,2nd edition, 2004.

3. **Digital Signal Processing** - A Computer based approach,S. K. Mitra, TMH, 2nd edition, 2002.

Pre Ph.D (Computer Science)

Subject Code	CA946	Subject Title	Cryptography						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -
Security attacks, mechanisms, and services.Network security and access security models.Overview of secret-key and public-key cryptography.Authentication protocols and key management.Network security practice. Email security. IP security and web security.Intrusion detection and prevention systems.Firewalls and virtual private networks. Wireless network security.

UNIT I **(8 L)**
Introduction to Security in Networks, Characteristics of Networks, Intrusion, Kinds of security breaches, Plan of attack, Points of vulnerability, Methods of defense, Control measures, Effectiveness of controls.

UNIT II **(9 L)**
Basic encryption and decryption, Encryption techniques, Characteristics of good encryption Systems, Secret key cryptography, Data Encryption Standard, International Data Encryption Algorithm, Advanced Encryption Standard.

UNIT III **(8 L)**
Public Key encryptions, Introduction to number theory, RSA algorithm, Diffie-Hellman, Digital Signature standard, Elliptic Curve cryptography, Digital signatures and authentication trusted intermediaries, Security handshake pitfalls, Hash and MAC algorithms, Security standards, Kerberos.X.509AuthenticationService

UNIT IV **(8 L)**
Secure sockets, IPsec overview, IP security architecture, IPsec-Internet Key Exchanging (IKE) , IKE phases ,encoding , Internet security ,Threats to privacy ,Packet sniffing, Spoofing , Web security requirements , Real Time communication security .

UNIT V **(7 L)**
Secret Sharing Schemes, The Shamir Threshold Scheme, Access Structure and General Secret key sharing, Information Rate and Construction of Efficient Schemes, Multicast Security and Copyright production-Multicast Security, Broadcast Encryption ,Multicast Rekeying, Copyright Protection ,Tracing Illegally Redistribution keys.

Learning Outcomes: -Upon successful completion of this course, students will:

* Have internalized the fundamental notions of threat, vulnerability, attack and countermeasure.

* Able to identify the security goals of an information system, point out contradictory goals and suggest compromises.

* Have a theoretical understanding of the principles underlying cryptography and cryptanalysis and have a technical understanding of the main cryptographic concepts and technologies available today, including symmetric and asymmetric encryption, hashing, and digital signatures.

* Understand the purpose of security protocols and be witness to the difficulties of their verification.

* Understand how malicious code functions (e.g., viruses), what the vulnerabilities that make propagation possible (e.g., buffer overflows), and what methods and practices are available for mitigation (e.g., the Common Criteria).

* Understand the threats and vulnerabilities that are specific of a networked environment, and explain countermeasures including firewalls and intrusion detection systems.

* Have an understanding for the vulnerabilities brought about by modern web-based application and services, and discuss countermeasures.

* Have an appreciation for the concerns of privacy and some of the approaches to fend them off.

* 10. Understand and model the economics of cybersecurity.

Text Books:
1. **"Cryptogrphy and Network Security"**, Behrouz A Frorouzan, ,TMH, India, Special Edition,2008

Reference Books:
1. **"Introduction to cryptography"**,JohannesA.Buchmann ,Springer-Verlag,USA,2/e,2004
2. **"Cryptography and Network Security",**AtulKahate,,TMH, New Delhi,2/e 2005

Pre Ph.D (Computer Science)

Subject Code	CA947	Subject Title	Advanced Computer Networks						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -
- To uncover and understand the current directions of computer networks from literature readings.
- To expose students to the "full span" of the computer network's frontier.
- To encourage a performance perspective towards analysis of computer and communications networks.
- To "fill-in" gaps in students' networking knowledge.

Unit I (8L)
Network Layer design Issues, IPv4, IPv6, Shortest Path Routing, Distance Vector Routing, Flooding, Hierarchical Routing, Broadcast Routing, and Multicast Routing.

Unit II (8L)
Wireless Networks, GSM Architecture, CDMA, Mobility in networks, Handoffs. Mobile IP- IP Packet Delivery, Agent Discovery, Registration, Tunneling and Encapsulation.

Unit III (8L)
Mobile TCP- Traditional TCP (Congestion Control, Slow Start, Fast Retransmit/Fast Recovery), Indirect TCP, Snooping TCP, Mobile TCP, Selective Retransmission, Transaction Oriented TCP.

Unit IV (8L)
Wireless LAN- Infrared Vs Radio Transmission, Infrastructure and Ad-hoc Network, IEEE 802.11-System Architecture, Protocol Architecture, Physical Layer, Bluetooth.

Unit V (8L)
IP Security: Architecture, Authentication header, Encapsulating security payloads, combining security associations, key management, SSL.

Learning Outcomes: -
1. To master the terminology and concepts of the OSI reference model and the TCP-IP Reference model.
2. To master the concepts of protocols, network interfaces, and Design/performance issues in local area networks and wide area networks,
3. To be familiar with wireless networking concepts,
4. To be familiar with contemporary issues in networking technologies,
5. To be familiar with network tools and network programming

Text Books:
1. "**Mobile Communication**",JSchiller, Pearson Education Ltd.,India,2/e,2003
2. "**Computer Network**", Andrew S. Tanenbaum, Pearson Education Ltd.,India,5/e,2010

Reference Books:
1. "**AdhocNetworking**",Charles E. Perkins,Addison-Wesley,USA,1/e,2001
2. "**TCP/IP Protocol Suite**",BehrouzForouzan, TMH,India,4/e,2009
3. "**Internetworking with TCP/IP Volume One**", Douglas E. Comer, Addison-Wesley,USA,6/e,2013

Pre Ph.D (Computer Science)

Subject Code	CS948	Subject Title	Neural Networks & Neuro Fuzzy Systems						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -This course aims at introducing the fundamental theory and concepts of computational intelligence methods, in particular neural networks, fuzzy systems and their applications in the area of machine intelligence.

Unit I   **(8L)**
Introduction: Biological neural system, Artificial Intelligent Systems, Modeling human performance. Uncertain & incomplete knowledge, Expert Systems Vs Neural Networks.

Unit II   **(8L)**
Foundations for connectionist networks. Architecture, Activation functions, Characteristics of Neural Networks, McCulloch-Pitts Neurons, Liner Separability

Unit III   **(8 L)**
Perceptron learning, Pattern Classification. Hebb Rule. Adaline.Madaline.Delta Rule.Back propagation learning, Competitive learning, Hebbian learning, BAMs.

Unit IV   **(8L)**
Supervised and Unsupervised learning. Reinforcement learning, Kohonen Self Organizing Maps, Applications of SOMs, Adaptive Resonance Theory,   Neural Network Applications.

Unit V   **(8 L)**
Neuro-Fuzzy Systems: Types of Fuzzy Neural Nets, Neural components in a Fuzzy System Fuzzy-ANN Controller, Support Vector Machines, Applications of SVMs.

Learning Outcomes: -
1. To understand the fundamental theory and concepts of neural networks, neuro-modeling, several neural network paradigms and its applications.
2. To understand the concepts of fuzzy sets, knowledge representation using fuzzy rules, approximate reasoning, fuzzy inference systems, and fuzzy logic control and other machine intelligence applications of fuzzy logic.

Text Book:
1. Simon Haykin, "Neural Networks: A Comprehensive Foundation", Prentice Hall
2. Nils J. Nilsson, "Artificial Intelligence - A New Synthesis", Morgan Kaufmann Publishers.
3. Robert J. Scholkoof "Artificial Neural Networks"; McGraw Hill Eductaion, 2011
Reference Book:
1. "Fuzzy Logic with Engineering Applications", Timothy J. Ross, Wiley India.
2."Artificial Intelligence A Modern Approach", Stuart Russel, Peter Norvig, Pearson (3rd Ed.)

Pre Ph.D (Computer Science)

Subject Code	CA949	Subject Title	Machine Learning						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: -This course will introduce the field of Machine Learning, in particular focusing on the core concepts of supervised and unsupervised learning.

Unit 1 (8L)

Overview and Introduction to Bayes Decision Theory: Machine intelligence and applications, pattern recognition concepts classification, regression, feature selection, supervised learning class conditional probability distributions, Examples of classifiers bayes optimal classifier and error, learning classification approaches.

Unit 2 (8L)

Linear machines: General and linear discriminants, decision regions, single layer neural network, linear separability, general gradient descent, perceptron learning algorithm, mean square criterion and widrow-Hoff learning algorithm; multi-Layer perceptrons: two-layers universal approximators, backpropagation learning, on-line, off-line error surface, important parameters.

Unit 3 (8L)

Learning decision trees: Inference model, general domains, symbolic decision trees, consistency, learning trees from training examples entropy, mutual information, ID3 algorithm criterion, C4.5 algorithm continuous test nodes, confidence, pruning, learning with incomplete data.

Unit 4 (8L)

Instance-based Learning: Nearest neighbor classification, k-nearest neighbor, nearest neighbor error probability. Machine learning concepts and limitations: Learning theory, formal model of the learnable, sample complexity, learning in zero-bayes and realizable case, VC-dimension, fundamental algorithm independent concepts, hypothesis class, target class, inductive bias, occam's razor, empirical risk, limitations of inference machines, approximation and estimation errors, Tradeoff.

Unit 5 (8L)

Machine learning assessment and Improvement: Statistical model selection, structural risk minimization, bootstrapping, bagging, boosting. Support Vector Machines: Margin of a classifier, dual perceptron algorithm, learning nonlinear hypotheses with perceptron kernel functions, implicit non-linear feature space, theory, zero-Bayes, realizable infinite hypothesis class, finite covering, margin-based bounds on risk, maximal margin classifier.

Learning outcomes

On completion of the course students will be expected to:

1. Have a good understanding of the fundamental issues and challenges of machine learning: data, model selection, model complexity, etc.
2. Have an understanding of the strengths and weaknesses of many popular machine learning approaches.
3. Appreciate the underlying mathematical relationships within and across Machine Learning algorithms and the paradigms of supervised and un-supervised learning.4
4. Be able to design and implement various machine learning algorithms in a range of real-world applications.

Text Books

1. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2006.
2**.** T. M. Mitchell, Machine Learning, McGraw-Hill, 1997.
3. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
4. R. O. Duda, P. E. Hart, and D.G. Stork, Pattern Classification, John Wiley and Sons, 2001.

Reference Book:

1. Vladimir N. Vapnik, Statistical Learning Theory, John Wiley and Sons, 1998.
2. J. Shawe-Taylor and N. Cristianini, Cambridge, Introduction to Support Vector Machines, University Press, 2000.

Pre Ph.D (Computer Science)

Subject Code	CA951	Subject Title	Wireless Sensor Networks						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives:-
This course deals with the comprehensive knowledge about wireless sensor networks. It provides an insight into different layers and their design considerations. A thorough knowledge of infrastructure establishment and sensor network platform is provided.

UNIT-1 (8L)
Characteristics Of WSN: Characteristic requirements for WSN - Challenges for WSNs – WSN vs Adhoc Networks - Sensor node architecture – Commercially available sensor nodes –Imote, IRIS, Mica Mote, EYES nodes, BTnodes, TelosB, Sunspot -Physical layer and transceiver design considerations in WSNs, Energy usage profile, Choice of modulation scheme, Dynamic modulation scaling, Antenna considerations.

UNIT-2 (7L)
Medium Access Control Protocols: Fundamentals of MAC protocols - Low duty cycle protocols and wakeup concepts - Contentionbased protocols - Schedule-based protocols - SMAC - BMAC - Traffic-adaptive medium access protocol (TRAMA) - The IEEE 802.15.4 MAC protocol.

UNIT-3 (8L)
Routing And Data Gathering Protocols Routing Challenges and Design Issues in Wireless Sensor Networks, Flooding and gossiping – Data centric Routing – SPIN – Directed Diffusion – Energy aware routing - Gradient-based routing - Rumor Routing – COUGAR – ACQUIRE – Hierarchical Routing - LEACH, PEGASIS – Location Based Routing – GF, GAF, GEAR, GPSR – Real Time routing Protocols – TEEN, APTEEN, SPEED, RAP - Data aggregation - data aggregation operations - Aggregate Queries in Sensor Networks - Aggregation Techniques – TAG, Tiny DB.

UNIT-4 (8L)
Embedded Operating Systems: Operating Systems for Wireless Sensor Networks – Introduction - Operating System Design Issues - Examples of Operating Systems – TinyOS – Mate – MagnetOS – MANTIS - OSPM - EYES OS – SenOS – EMERALDS – PicOS – Introduction to Tiny OS – NesC – Interfaces and Modules- Configurations and Wiring - Generic Components -Programming in Tiny OS using NesC, Emulator TOSSIM.

UNIT-5 (8L)
Applications Of WSN: WSN Applications - Home Control - Building Automation - Industrial Automation - Medical Applications - Reconfigurable Sensor Networks - Highway Monitoring - Military Applications - Civil and Environmental Engineering Applications - Wildfire Instrumentation - Habitat Monitoring - Nanoscopic Sensor Applications – Case Study: IEEE 802.15.4 LR-WPANs Standard - Target detection and tracking - Contour/edge detection - Field sampling.

Learning Outcomes: On completion of this course you should be able to:
- Apply knowledge of wireless sensor networks (WSN) to various application areas.
- Design and implement WSN.
- Conduct performance analysis of WSN and manage WSN.
- Formulate and solve problems creatively in the area of WSN.

Text Books:
1. **"Fundamentals of Wireless Sensor Networks: Theory and Practice"**,WaltenegusDargie, Christian Poellabauer, ,Wiley,India,1st edition ,2010
2. **"Multihop Wireless Networks: Opportunistic Routing**",Kai Zeng, Wenjing Lou, Ming Li ,Wiley,India, 1st edition, 2011

References:
1. **"Protocols and Architectures for Wireless Sensor Network"**,HolgerKerl, Andreas Willig, Wiley,India,1st edition,,2005
2. **"Wireless Sensor Network"**, Raghavendra, Cauligi S, Sivalingam, Krishna M., ZantiTaieb, Springer,USA, 1st edition, 2004
3. **"Sensor Network Wireless"**Feng Zhao, Leonidas Guibas,Elsevier,USA, 1st edition, 2004

Pre Ph.D (Computer Science)

Subject Code	CA952	Subject Title	Advanced Software Engineering						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives:-

In this course the student will learn about some of the most advanced topics on Software Engineering. The objective of this course is to teach students the methodology to design and write secure code applying the Secure Software Engineering life Cycle.

Unit-I   (8L)

Introduction: Software Crisis, Software Processes & Characteristics, Software life cycle models, Waterfall, Prototype, Evolutionary and Spiral Models, Overview of Quality Standards like ISO 9001, SEI – CMM. Software Requirements analysis & specifications: Requirement engineering, requirement elicitation techniques like FAST, QFD & Use case approach, requirements analysis using DFD, Data dMEionaries& ER Diagrams, Requirements documentation, Nature of SRS, Characteristics & organization of SRS.

Unit-II   (8L)

Software Architecture: Role of Software Architecture, Architecture views, Component and Connector view: Components, Connectors, Architecture style for C and C view: pipe and filter, shared data style, client server style, Evaluating Architecture. Software Project Planning:Size Estimation like lines of Code & Function Count, Cost Estimation Models, Static single & Multivariable Models, COCOMO, COCOMO-II, Putnam resource allocation model, Risk Management.

Unit-III   (8L)

Software Design: Cohesion & Coupling, Classification of Cohesiveness & Coupling, Function Oriented Design, Object Oriented Design, User Interface Design. Software Maintenance: Management of Maintenance, Maintenance Process, Maintenance Models, Reverse Engineering, Software Re-engineering, Configuration Management, Documentation. Software Metrics:Software measurements: What & Why, Token Count, Halstead Software Science Measures, Design Metrics, Data Structure Metrics, Information Flow Metrics

Unit-IV   (8L)

Software Testing:Testing process, Design of test cases, functional testing: Boundary value analysis, Equivalence class testing, Decision table testing, Cause effect graphing, Structural testing, Path Testing, Data flow and mutation testing, Unit Testing, Integration and System Testing, Debugging, Alpha & Beta Testing, Regression Testing, Testing Tools & Standards. Software Reliability:Importance, Hardware Reliability & Software Reliability, Failure and Faults, Reliability Models, Basic Model, Logarithmic Poisson Model, Calender time Component.

Learning Outcomes:

After completing this course, the student should be able to: Have a higher-level understanding of how to write secure code by using the Secure Software Engineering Life Cycle. 1. Comprehend, apply, and implement Secure Software Requirements 2. Comprehend, apply, and implement Secure Software Design 3. Comprehend and apply Secure Software Implementation 4. Comprehend, apply, and implement Secure Software Testing

Test Books:

1. K. K. Aggarwal & Yogesh Singh, "Software Engineering", New Age International, 2001.
2. R. S. Pressman, "Software Engineering – A practitioner's approach", 5th Ed., McGraw Hill Int. Ed., 2001.
3. P. Jalote, "An Integrated approach to Software Engineering", Springer Publications, 2005.

Reference Books:

1. R. Fairley, "Software Engineering Concepts", Tata McGraw Hill, 1997.
2. Yogesh Singh, "Software Testing", Cambridge University Press, New York, 2012.
3. Stephen R. Schach, "Classical & Object Oriented Software Engineering", IRWIN, 1996.

Pre Ph.D (Computer Science)

Subject Code	CA953	Subject Title	Artificial Intelligence						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives: - This course covers theoretical and practical approaches to AI, with topics to include search, logic, knowledge representation, uncertainty, and different aspects of the performance of AI techniques.

Unit I (7L)
Overview of Artificial Intelligence: Definition & Importance of AI.
Knowledge: General Concepts: Introduction, Definition and Importance of Knowledge, Knowledge-Based Systems, And Representation of Knowledge, Knowledge Organization, Knowledge Manipulation, And Acquisition of Knowledge.

Unit II (7L)
Knowledge Representation: Introduction, Syntax and Semantics for Propositional logic, Syntax and Semantics for FOPL, Properties of Wffs, Conversion to Clausal Form, Inference Rules, The Resolution Principle, No deductive Inference Methods, Representations Using Rules.

Unit III (8L)
Dealing with Inconsistencies and Uncertainties: Introduction, Truth Maintenance Systems, Default Reasoning and the Closed World Assumption, Predicate Completion and Circumscription, Modal and Temporal Logics.
Probabilistic Reasoning: Introduction, Bayesian Probabilistic Inference, Possible World Representations, Dumpster-Shafer Theory, Ad-Hoc Methods.

Unit IV (8L)
Structured Knowledge: Graphs, Frames and Related Structures: Introduction, Associative Networks, Frame Structures, Conceptual Dependencies and Scripts. search and control strategies: Introduction, Preliminary Concepts, Examples of Search Problems, Uninformed or Blind Search, Informed Search, Searching And-Or Graphs.
Matching Techniques: Introduction, Structures Used in Matching, Measures for Matching, Matching Like Patterns, Partial Matching.

Unit V (7L)
Knowledge Organization and Management: Introduction, Indexing and Retrieval Techniques, Integrating Knowledge in Memory, Memory Organization Systems.
Expert Systems Architectures: Introduction, Rule Based System Architecture, Non-Production System Architecture, Dealing with uncertainty, Knowledge Acquisition and Validation, Knowledge System Building Tools.

Learning outcomes
1. Students are expected to apply AI techniques to different complex problems using programs.
2. Students will be able to create knowledge base and serve their applications in different fields.
3. Students are expected to implement perfect and better heuristics for different applications.
4. Students are able to create Expert systems.

Text Books:
1. "**Artificial Intelligence**", Elaine Rich, Kevin Knight and ShivashankarB.Nair,Tata McGraw-Hill,Third edition, 2009.

2. "**Artificial Intelligence: A Modern Approach**", Stuart J. Russell and Peter Norvig,Pearson Education Asia, Second edition, 2003.

3. "**Artificial Intelligence and Intelligent System**",N. P. Padhy, Oxford University Press, Second edition, 2005.

Reference Books:
1. "**Introduction to Artificial Intelligence**",RajendraAkerkar, Prentice-Hall of India, 2005.

2. "**Artificial Intelligence**", Patrick Henry Winston, Pearson Education Inc., Third edition, 2001.

3. "**Introduction to Artificial Intelligence**",EugeneCharniak and Drew Mc Dermott, Addison-Wesley,ISE Reprint, 1998

Pre Ph.D (Computer Science)

Subject Code	CA954	Subject Title	Distributed Systems						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives:-This course covers the foundations of distributed systems including models of computing, different types of communication (Layered Protocols, Remote Procedure Calls, Remote Objects, messages, streams). The course also makes students aware about some key concept such as Distributed mutual exclusion, Distributed file system, Distributed shared memory. In the last unit course covers some of the applications areas of distributed systems such as Fault tolerance computing, Grid computing and High performance computing.

Unit I (8L)
Fundamentals of Distributed Computing: Architectural models for distributed computing systems, Issues and challenges in Distributed systems, Basic concepts in distributed computing such as clocks, message ordering, consistent global states.
Distributed Environments
Current systems and developments (DCE, CORBA, JAVA).

Unit II (8L)
Coordination & Synchronization: Introduction, Distributed Mutual Exclusion, Elections, Multicast Communication, Atomic transactions, Deadlocks in Distributed systems.
Message Passing & Remote Procedure Calls: Features of a good message-passing system, RPC model. Implementing RPC mechanism, Stub Generation, RPC Messages, Marshaling Arguments and Results, Server Management, Communication protocols for RPCs.

Unit III (8L)
Distributed File Systems: Features of Good DFS, File Models, File-Accessing models, File Service Architecture, File-sharing semantics, File Caching schemes, File replications.

Unit IV (8L)
Distributed Shared Memory: Shared memory consistency models, Page based distributed shared memory, Shared variable distributed shared memory, Object based distributed shared memory.
Replication: Introduction, System Model & Group Communication, Fault Tolerant Services, Transactions with Replicated Data.

Unit V (8L)
Advanced Topics in Distributed Computing: High Performance Computing-HPF, Distributed and mobile multimedia systems.Adaptability in Mobile Computing.Grid Computing and applications. Fault tolerant Computing Systems.

Learning Outcomes: - During this course, students will learn a range of fundamental and applied techniques in distributed systems. The learning objectives for Distributed Systems are:

1. To develop and apply knowledge of distributed systems techniques and methodologies.

2. To gain experience in the design and development of distributed systems and distributed systems applications.

3. To gain experience in the application of fundamental Computer Science methods and algorithms in the development of distributed systems and distributed systems applications.

4. To gain knowledge about the key concepts of distributed systems such as distributed mutual exclusion, distributed file system, distributed shared memory.

Text Books:
1. "**Distributed Databases - Principles and Systems**"; Stefano Ceri, GuiseppePelagatti; Tata McGraw Hill; 1985.
2. "**Database Systems- Design, Implementation and Management**"; Peter Rob, Carlos Coronnel;
3. Course Technology; 2000.

Reference Books:
1. "**Principles of Distributed database systems**" by M.T. Ozsu/S. Sridhar, Pearson education, 2006.
2. "**Database Management Systems**" by Raghu RamaKrishnan, JohnaasGehrke; Tata McGrawHill; 2000.
3. "**Fundamentals of Database Systems**" Elmasri, Navathe; Addison-Wesley;Third Edition;2002.

Pre Ph.D (Computer Science)

Subject Code	CA955	Subject Title	Information & Coding Theory						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives:-
- To deeply understand the mathematics of Information Theory and its physical meaning
- Relate the joint, conditional, and marginal entropies of variables in terms of their coupled probabilities
- To understand various Source coding techniques
- Construct efficient codes for data on imperfect communication channels
- To provide an overview on some basic tools as well as several advanced topics in lossy and lossless source coding. From both the theoretic and practical aspects students are encouraged to propose research oriented projects and investigate them.
- Understand and explain limits in human perception that are exploited by lossy compression techniques
- Provide a good overview of the principles and characteristics of several widely-used compression techniques and standards for image and video signals
- To understand various channel coding techniques
- Can apply the knowledge to real problems in communication application

Unit I Information Theory (8 L)

Information-Entropy, Information rate, classification of codes, Kraft McMillan inequality, Source coding theorem, Shannon-Fano coding, Huffman coding, Extended Huffman coding -Joint and conditional entropies, Mutual information - Discrete memory less channels – BSC,BEC – Channel capacity, Shannon limit.

Unit II Source Coding: Text, Audio and Speech (8 L)

Text: Adaptive Huffman Coding, Arithmetic Coding, LZW algorithm Audio: Perceptual coding, Masking techniques, Psychoacoustic model, MEG Audio layers I, II, III, Dolby AC3 Speech: Channel Vocoder, Linear Predictive Coding

Unit III Source Coding: Image and Video (8 L)

Image and Video Formats – GIF, TIFF, SIF, CIF, QCIF Image compression: READ, JPEG.Video Compression: Principles-I, B, P frames, Motion estimation, Motion compensation,H.261, MPEG standard.

Unit IV Error Control Coding: Block Codes (8 L)

Definitions and Principles: Hamming weight, Hamming distance, Minimum distance decoding - Single parity codes, Hamming codes, Repetition codes - Linear block codes, Cyclic codes - Syndrome calculation, Encoder and decoder - CRC

Unit V   Error Control Coding: Convolutional Codes (8 L)

Convolutional codes– code tree, trellis, and state diagram - Encoding – Decoding: Sequential search and Viterbi algorithm – Principle of Turbo coding

Learning Outcomes: - The participants must at the end of the course be able to:
- Understand and explain the basic concepts of information theory, source coding, channel and channel capacity, channel coding and relation among them.
- Describe the real life applications based on the fundamental theory.
- Calculate entropy, channel capacity, bit error rate, code rate, steady-state probability and so on.
- Implement the encoder and decoder of one block code or convolutional code using any program language.

Text Books:
1. R Bose, "Information Theory, Coding and Cryptography", TMH 2007

Reference Books:
1. FredHalsall ,"Multimedia Communications: Applications, Networks, protocols and Standards", Perason Education Asia, 2002 .
2. Steven Roman, Introduction to Coding and Information Theory.
3. T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed. Wiley interscience, 2006 .

Pre Ph.D (Computer Science)

Subject Code	CA956	Subject Title	Mobile and Ad-Hoc Networks						
LTP	4 0 0	Credit	4	Subject Category	DE	Year	1	Semester	-

Course Objectives:-This course aims to provide basic understanding to realize the vision of "Optimally Connected Anywhere, Anytime". The course is an introduction to the fundamentals of mobile computing, to learn about different architectures of mobile application development, to study the routing algorithm, security mechanism and transport protocols used in mobile adhoc networks, and to know recent and future trends in mobile computing.

Unit I (8 L)

Issues in Mobile Computing, Wireless Telephony: Frequency Reuse, Cell Design, Cellular Architecture. Second Generation Cellular System- IS-95, GSM, IS-41, Data Operations- CDPD, HCSD, & GPRS. Third Generation Cellular System- 3G Concept, 3G Spectrum Allocation, 3G Service Classes & Application, UMTS, Introduction to Future Trends- 4G and Beyond.

Unit II (8 L)

WLAN Overview, Infrared LAN, Spread-Spectrum LAN,Narrowband Microwave LAN, MAC issues, Wireless Multiple Access Protocols- MACA & MACAW, IEEE 802.11 variants, HIPERLAN, PAN Technologies- Bluetooth and HomeRF, Wireless Metropolitan Area Networks (Wireless Local Loop).

Unit III (8 L)

Wireless Internet: IP Limitations, Mobile IP &its working, Issues in Mobile IP. TCP over Wireless, Wireless Access Protocol (WAP): Architecture & Protocol Stack. Security in Wireless Systems- WEP and Virtual Private Networks, Mobile Agents Computing, Transaction Processing in Mobile Computing Environment

Unit IV (8 L)

Ad Hoc Networks: Ad Hoc Networks vs. Cellular Networks, Issues in Ad Hoc Wireless Networks, Routing Protocols for Ad Hoc Wireless Networks-Design Issues and Classification, Vehicular AdHoc Networks-Introduction and Routing in VANETs, Delay Tolerant Networks- Introduction and Routing in DTNs.

Unit V (8 L)

Wireless Sensor Networks: Overview, Application Areas, Sensor Nodes' Architecture,
Data Aggregation, Routing, and Query Processing in WSN. Intermittently Connected Delay Tolerant WSN-Introduction, End-to-End Reliability, Routing, and Link Layer Protocols in ICDT-WSN.

Learning Outcomes: - After completion of this course student will be able to:
- Describe wireless and mobile communications systems and be able to choose an appropriate mobile system from a set of requirements.
- Be able to avoid or work around the weaknesses of mobile computing, or to reject mobile computing as a solution.
- Understand the routing concept of mobile adhoc networks.
- Understand the reliable and unreliable communication and provide solutions to improve the quality of service in mobile adhoc networks.
- Program applications on a mobile computing system and interact with servers and database systems.

Text Book:
1. Jochen H. Schiller, Mobile Communication, Addison-Wesley, Pearson Education
2. Sipra Das Bit, Biplab K. Sikdar, Mobile Computing, PHI Learning
3. Vijay Garg, Elsevier,Wireless Communication & Networking , Morgan Kaufmann Publisher

Reference Book:
a. Subir K Sarkar, T G Basavaraju, C Puttamadappa, Ad Hoc Mobile Wireless Networks, AuerbachPublication .
b. A. Ananda, MunChoon Chan, Mobile, Wireless & Sensor Networks,Wei Tsang Ooi , IEEE press, Willy Intersciences.
c. ShambhuUpadhyaya et al,"Mobile Computing: Implementing Pervasive Information and Communications Technologies" Springer US,USA, illustrated edition, 2013

Pre Ph.D (Computer Science)

Subject Code	CA957	Subject Title	Advanced Data Warehousing and Mining						
LTP	4 0 0	**Credit**	4	**Subject Category**	DE	**Year**	1	**Semester**	-

Course Objectives:-This course aims to provide a basic understanding of Data Mining and Data Warehouse concepts and technologies. This course introduces students to enterprise data and the process and technologies to integrate data from a variety of sources.

Unit I (8L)

Evolution of Database System Technology, Architecture of Data Mining System, Data Warehouse, Advanced Data and Information Systems and Advanced Applications, Data Mining Functionalities, Data Mining Task Primitives

Unit II (8L)

Data Preprocessing, Descriptive Data Summarization, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation

Unit III (8L)

Data Warehouse Architecture, Data Warehouse Schemas, Multidimensional Data Model, OLTP, OLAP, OLAP operations, ROLAP, MOLAP, HOLAP

Unit IV (8L)

Market Basket Analysis, Apriori Algorithm, Mining Multilevel Association Rules, Classification and Prediction, Classification by Decision Tree Induction, Bayesian Classification, and Classification by Back propagation

Unit V (8L)

Cluster Analysis, Types of Data in Cluster Analysis, Partitioning Methods, K-Means Method, K-Medoids Method, DBSCAN, OPTICS. Mining Sequence pattern in Biological Data, Mining Multidimensional, Multilevel Sequential Pattern

Learning Outcomes: - At the end of the course, the students should be able to:
- Describe the components of an Enterprise data warehouse architectures.
- Model the relational database required for an enterprise data warehouse.
- Extract, cleanse, consolidated, and transform heterogeneous data into a single enterprise data warehouse.
- Analyze data to generate information and knowledge that lead to informed decisions for businesses.
- Differentiate between Transaction Processing and Analytical applications and describe the need for data preprocessing and data mining functionalities.
- Demonstrate understanding of predictive analysis of vast amount of data.
- Understanding of classification and clustering algorithms for data analysis.

Text Book:
1. **"Data Mining Concepts and Techniques",** Jiawei Han and Micheline Kamber, Elsevier, Third Edition, 2012.
Reference Books:
1. "**Data-Mining. Introductory & Advanced Topics**", Margaret H.Dunham, Pearson Education, India, 3/e, 2008.
2**. "Data Warehousing System",**Mallach, McGraw –Hill Pearson Education, 2000