JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B.Tech in COMPUTER ENGG. (SOFTWARE ENGINEERING)
III & IV YEAR COURSE STRUCTURE & TENTATIVE SYLLABUS (R18)

Applicable From 2020-21 Admitted Batch

III YEAR I SEMESTER

S. No.	Course Code	Course Title	L	T	P	Credits
1		Design and Analysis of Algorithms	3	0	0	3
2		Computer Networks	3	0	0	3
3		Database Management Systems	3	0	0	3
4		Software Requirements & Estimation	3	0	0	3
5		Professional Elective – I	3	0	0	3
6		Professional Elective – II	3	0	0	3
7		Computer Networks Lab	0	0	3	1.5
8		Database Management Systems Lab	0	0	3	1.5
9		Advanced Communication Skills Lab	0	0	2	1
10		Intellectual Property Rights	3	0	0	0
		Total Credits	**21**	**0**	**8**	**22**

III YEAR II SEMESTER

S. No.	Course Code	Course Title	L	T	P	Credits
1		Automata Theory and Compiler Design	3	1	0	4
2		Software Testing Methodologies	3	1	0	4
3		Software Architecture and Design Patterns	3	1	0	4
4		Professional Elective – III	3	0	0	3
5		Open Elective – I	3	0	0	3
6		Compiler Design Lab	0	0	3	1.5
7		Software Testing Lab	0	0	3	1.5
8		Professional Elective - III Lab	0	0	2	1
9		Environmental Science	3	0	0	0
		Total Credits	**18**	**3**	**8**	**22**

IV YEAR I SEMESTER

S. No.	Course Code	Course Title	L	T	P	Credits
1		Agile Software Development	3	0	0	3
2		Machine Learning	2	0	0	2
3		Professional Elective – IV	3	0	0	3
4		Professional Elective – V	3	0	0	3
5		Open Elective – II	3	0	0	3
6		Machine Learning Lab	0	0	2	1
7		Industrial Oriented Mini Project/ Summer Internship	0	0	0	2*
8		Seminar	0	0	2	1
9		Project Stage – I	0	0	6	3
		Total Credits	**14**	**0**	**10**	**21**

IV YEAR II SEMESTER

S. No.	Course Code	Course Title	L	T	P	Credits
1		Organizational Behaviour	3	0	0	3
2		Professional Elective – VI	3	0	0	3
3		Open Elective – III	3	0	0	3
4		Project Stage – II	0	0	14	7
		Total Credits	**9**	**0**	**14**	**16**

*Note: Industrial Oriented Mini Project/ Summer Internship is to be carried out during the summer vacation between 6th and 7th semesters. Students should submit report of Industrial Oriented Mini Project/ Summer Internship for evaluation.
MC - Environmental Science – Should be Registered by Lateral Entry Students Only.
MC – Satisfactory/Unsatisfactory

Professional Elective-I

	Data Warehousing and Business Intelligence
	Artificial Intelligence
	Web Programming
	Image Processing
	Computer Graphics

Professional Elective - II

	Mining Massive Datasets
	Information Retrieval Systems
	Internet of Things
	DevOps
	Software Design Methodologies

Professional Elective - III

	Object Oriented Analysis & Design
	Introduction to Data Science
	Scripting Languages
	Mobile Application Development
	Cryptography and Network Security

 # Courses in PE - III and PE - III Lab must be in 1-1 correspondence.

Professional Elective - IV

	Quantum Computing
	Data Visualization Techniques
	Natural Language Processing
	Information Storage Management
	Software Project Management

Professional Elective - V

	Privacy Preserving in Data Mining
	Cloud Computing
	Data Stream Mining
	Software Quality Assurance
	Exploratory Data Analysis

Professional Elective – VI

	Software Metrics
	Web security
	Computational Complexity
	Blockchain Technology
	Parallel and Distributed Computing

AGILE SOFTWARE DEVELOPMENT

B.Tech. IV Year I Sem.   L   T   P   C
                                                                           3   0   0   3

Course Objectives: The objectives of this subject are to:
1. Organize Agile Software Development, Extreme Programming and Software Development Rhythms.
2. Describe their unique features relative to traditional software practices.
3. Examine their applications in the real world and address their impacts on developing software.

Course Outcomes: Upon completion of the subject, students will be able to:
1. Summarize the agile methodologies: extreme programming, scrum, and feature driven programming.
2. Apply the Twelve XP Practices and Illustrate pair programming and its characteristics.
3. Apply XP to a small project.
4. Examine Feature-Driven Development and Regaining Control.
5. Relate Agile Modeling and RUP and Choose Tools to help with Agile Development.

UNIT - I
Introduction: Agile Methods, Agile Manifesto, and Agile Modeling Introduction, What Is Agile, The Agile Manifesto, Agile Methods, XP: Extreme Programming, DSDM, SCRUM, Feature-Driven Development, Modeling Misconceptions, Agile Modeling, Tools of Misconceptions, Updating Agile Models

UNIT - II
Extreme Programming: Introduction, Core XP Values, The Twelve XP Practices, About Extreme Programming, Planning XP Projects, Test First Coding, Making Pair Programming Work

UNIT - III
Agile Modeling and XP: Introduction, The Fit, Common Practices, Modeling Specific Practices, XP Objections to Agile Modeling, Agile Modeling and Planning XP Projects, XP Implementation Phase

UNIT - IV
Feature-Driven Development: Introduction, Incremental Software Development, Regaining Control: The Motivation behind FDD, Planning an Iterative Project, Architecture Centric, FDD and XP

UNIT - V
Agile Methods with RUP and PRINCE2 and Tools and Obstacles: Agile Modeling and RUP, FDD and RUP, Agile Methods and Prince2, Tools to Help with Agile Development, Eclipse: An Agile IDE, Obstacles to Agile Software Development, Management Intransigence, The Failed Project Syndrome, Contractual Difficulties, Familiarity with Agility.

TEXT BOOKS:
1. Agile software construction, 1/e, John hunt, springer, 2005.
2. Agile and Iterative Development: a manager's guide, Addison-Wesley Craig Larman, [Pearson Education] - 2004.

REFERENCE BOOKS:
1. The Art of Agile Development, Pearson, Robert C. Martin, Juli, James Shore, Chromatic, 2013, O'Reilly Media.
2. Agile Testing, Elisabeth Hendrickson, Quality Tree Software Inc 2008.

MACHINE LEARNING

B.Tech. IV Year I Sem.

L	T	P	C
2	0	0	2

Pre-requisites:
1. Data Structures.
2. Knowledge on statistical methods.

Course Objectives:
1. This course explains machine learning techniques such as decision tree learning, Bayesian learning etc.
2. To understand computational learning theory.
3. To study the pattern comparison techniques.

Course Outcomes:
1. Understand the concepts of computational intelligence like machine learning
2. Ability to get the skill to apply machine learning techniques to address the real time problems in different areas
3. Understand the Neural Networks and its usage in machine learning application.

UNIT - I
Introduction - Well-posed learning problems, designing a learning system, Perspectives and issues in machine learning. Concept learning and the general to specific ordering – introduction, a concept learning task, concept learning as search, find-S: finding a maximally specific hypothesis, version spaces and the candidate elimination algorithm, remarks on version spaces and candidate elimination, inductive bias**.**
Decision Tree Learning – Introduction, decision tree representation, appropriate problems for decision tree learning, the basic decision tree learning algorithm, hypothesis space search in decision tree learning, inductive bias in decision tree learning, issues in decision tree learning.

UNIT - II
Artificial Neural Networks-1– Introduction, neural network representation, appropriate problems for neural network learning, perceptions, multilayer networks and the back-propagation algorithm.
Artificial Neural Networks-2- Remarks on the Back-Propagation algorithm, An illustrative example: face recognition, advanced topics in artificial neural networks.
Evaluation Hypotheses – Motivation, estimation hypothesis accuracy, basics of sampling theory, a general approach for deriving confidence intervals, difference in error of two hypotheses, comparing learning algorithms.

UNIT - III
Bayesian learning – Introduction, Bayes theorem, Bayes theorem and concept learning, Maximum Likelihood and least squared error hypotheses, maximum likelihood hypotheses for predicting probabilities, minimum description length principle, Bayes optimal classifier, Gibs algorithm, Naïve Bayes classifier, an example: learning to classify text, Bayesian belief networks, the EM algorithm.
Computational learning theory – Introduction, probably learning an approximately correct hypothesis, sample complexity for finite hypothesis space, sample complexity for infinite hypothesis spaces, the mistake bound model of learning.
Instance-Based Learning- Introduction, *k*-nearest neighbour algorithm, locally weighted regression, radial basis functions, case-based reasoning, remarks on lazy and eager learning.

UNIT- IV
Genetic Algorithms – Motivation, Genetic algorithms, an illustrative example, hypothesis space search, genetic programming, models of evolution and learning, parallelizing genetic algorithms.

Learning Sets of Rules – Introduction, sequential covering algorithms, learning rule sets: summary, learning First-Order rules, learning sets of First-Order rules: FOIL, Induction as inverted deduction, inverting resolution.

Reinforcement Learning – Introduction, the learning task, Q–learning, non-deterministic, rewards and actions, temporal difference learning, generalizing from examples, relationship to dynamic programming.

UNIT - V

Analytical Learning-1- Introduction, learning with perfect domain theories: PROLOG-EBG, remarks on explanation-based learning, explanation-based learning of search control knowledge.

Analytical Learning-2-Using prior knowledge to alter the search objective, using prior knowledge to augment search operators.

Combining Inductive and Analytical Learning – Motivation, inductive-analytical approaches to learning, using prior knowledge to initialize the hypothesis.

TEXT BOOK:
1. Machine Learning – Tom M. Mitchell, - MGH.

REFERENCE BOOK:
1. Machine Learning: An Algorithmic Perspective, Stephen Marshland, Taylor & Francis.

QUANTUM COMPUTING (Professional Elective – IV)

B.Tech. IV Year I Sem.   L T P C
                                                              3 0 0 3

Course Objectives:
1. To introduce the fundamentals of quantum computing
2. The problem-solving approach using finite dimensional mathematics

Course Outcomes:
1. Understand basics of quantum computing
2. Understand physical implementation of Qubit
3. Understand Quantum algorithms and their implementation
4. Understand the Impact of Quantum Computing on Cryptography

UNIT - I
Introduction to Essential Linear Algebra: Some Basic Algebra, Matrix Math, Vectors and Vector Spaces, Set Theory. **Complex Numbers:** Definition of Complex Numbers, Algebra of Complex Numbers, Complex Numbers Graphically, Vector Representations of Complex Numbers, Pauli Matrice, Transcendental Numbers.

UNIT - II
Basic Physics for Quantum Computing: The Journey to Quantum, Quantum Physics Essentials, Basic Atomic Structure, Hilbert Spaces, Uncertainty, Quantum States, Entanglement.
Basic Quantum Theory: Further with Quantum Mechanics, Quantum Decoherence, Quantum Electrodynamics, Quantum Chromodynamics, Feynman Diagram Quantum Entanglement and QKD, Quantum Entanglement, Interpretation, QKE.

UNIT - III
Quantum Architecture: Further with Qubits, Quantum Gates, More with Gates, Quantum Circuits, The D-Wave Quantum Architecture. **Quantum Hardware:** Qubits, How Many Qubits Are Needed? Addressing Decoherence, Topological Quantum Computing, Quantum Essentials.

UNIT - IV
Quantum Algorithms: What Is an Algorithm? Deutsch's Algorithm, Deutsch-Jozsa Algorithm, Bernstein-Vazirani Algorithm, Simon's Algorithm, Shor's Algorithm, Grover's Algorithm.

UNIT - V
Current Asymmetric Algorithms: RSA, Diffie-Hellman, Elliptic Curve. **The Impact of Quantum Computing on Cryptography:** Asymmetric Cryptography, Specific Algorithms, Specific Applications.

TEXT BOOKS:
1. Nielsen M. A., Quantum Computation and Quantum Information, Cambridge University Press
2. Dr. Chuck Easttom, Quantum Computing Fundamentals, Pearson

REFERENCE BOOKS:
1. Quantum Computing for Computer Scientists by Noson S. Yanofsky and Mirco A. Mannucci
2. Benenti G., Casati G. and Strini G., Principles of Quantum Computation and Information, Vol. Basic Concepts. Vol. Basic Tools and Special Topics, World Scientific.
3. Pittenger A. O., An Introduction to Quantum Computing Algorithms.

DATA VISUALIZATION TECHNIQUES (Professional Elective – IV)

B.Tech. IV Year I Sem.   **L T P C**
                                                                                            3 0 0 3
Course Objective: To understand various data visualization techniques.

Course Outcomes:
1. Visualize the objects in different dimensions.
2. Design and process the data for Virtualization.
3. Apply the visualization techniques in physical sciences, computer science, applied mathematics and medical science.
4. Apply the virtualization techniques for research projects. (K1,K3).

UNIT - I
Introduction and Data Foundation: Basics - Relationship between Visualization and Other Fields - The Visualization Process - Pseudo code Conventions - The Scatter plot. Data Foundation - Types of Data - Structure within and between Records - Data Preprocessing - Data Sets

UNIT - II
Foundations for Visualization: Visualization stages - Semiology of Graphical Symbols - The Eight Visual Variables - Historical Perspective - Taxonomies - Experimental Semiotics based on Perception Gibson's Affordance theory – A Model of Perceptual Processing.

UNIT - III
Visualization Techniques: Spatial Data: One-Dimensional Data - Two-Dimensional Data – Three-Dimensional Data - Dynamic Data - Combining Techniques. **Geospatial Data:** Visualizing Spatial Data - Visualization of Point Data -Visualization of Line Data - Visualization of Area Data - Other Issues in Geospatial Data Visualization **Multivariate Data:** Point-Based Techniques - Line- Based Techniques - Region-Based Techniques - Combinations of Techniques – Trees Displaying Hierarchical Structures – Graphics and Networks- Displaying Arbitrary Graphs/Networks.

UNIT - IV
Interaction Concepts and Techniques: Text and Document Visualization: Introduction - Levels of Text Representations - The Vector Space Model - Single Document Visualizations -Document Collection Visualizations - Extended Text Visualizations **Interaction Concepts**: Interaction Operators - Interaction Operands and Spaces - A Unified Framework. **Interaction Techniques**: Screen Space - Object-Space -Data Space -Attribute Space- Data Structure Space - Visualization Structure - Animating Transformations -Interaction Control

UNIT - V
Research Directions in Virtualizations: Steps in designing Visualizations – Problems in designing effective Visualizations- Issues of Data. Issues of Cognition, Perception, and Reasoning. Issues of System Design Evaluation, Hardware and Applications.

TEXT BOOKS:
1. Matthew Ward, Georges Grinstein and Daniel Keim, "Interactive Data Visualization Foundations, Techniques, Applications", 2010.
2. Colin Ware, "Information Visualization Perception for Design", 2nd edition, Margon Kaufmann Publishers, 2004.

REFERENCE BOOKS:
1. Robert Spence "Information visualization – Design for interaction", Pearson Education, 2nd Edition, 2007.
2. Alexandru C. Telea, "Data Visualization: Principles and Practice," A. K. Peters Ltd, 2008.

NATURAL LANGUAGE PROCESSING (Professional Elective – IV)

B.Tech. IV Year I Sem.   **L T P C**
                                                                        3 0 0 3

Prerequisites: Data structures, finite automata and probability theory

Course Objective:
1. Introduce to some of the problems and solutions of NLP and their relation to linguistics and statistics.

Course Outcomes:
1. Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
2. Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
3. Able to manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.
4. Able to design, implement, and analyze NLP algorithms
5. Able to design different language modeling Techniques.

UNIT - I
Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models.
Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches.

UNIT - II
Syntax Analysis: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms, Models for Ambiguity Resolution in Parsing, Multilingual Issues.

UNIT - III
Semantic Parsing: Introduction, Semantic Interpretation, System Paradigms, Word Sense Systems, Software.

UNIT - IV
Predicate-Argument Structure, Meaning Representation Systems, Software.

UNIT - V
Discourse Processing: Cohension, Reference Resolution, Discourse Cohension and Structure.
Language Modeling: Introduction, N-Gram Models, Language Model Evaluation, Parameter Estimation, Language Model Adaptation, Types of Language Models, Language-Specific Modeling Problems, Multilingual and Cross lingual Language Modeling.

TEXT BOOKS:
1. Multilingual natural Language Processing Applications: From Theory to Practice –   Daniel M. Bikel and Imed Zitouni, Pearson Publication.
2. Natural Language Processing and Information Retrieval: Tanvier Siddiqui, U.S. Tiwary.

REFERENCE BOOK:
1. Speech and Natural Language Processing - Daniel Jurafsky & James H Martin, Pearson Publications.

INFORMATION STORAGE MANAGEMENT (Professional Elective – IV)

B.Tech. IV Year I Sem.   **L T P C**
                                                                      3 0 0 3

Course Objectives:
1. To understand the basic components of Storage System Environment.
2. To understand the Storage Area Network Characteristics and Components.
3. To examine emerging technologies including IP-SAN.
4. To describe the different backup and recovery topologies and their role in providing disaster recovery and business continuity capabilities.
5. To understand the local and remote replication technologies.

Course Outcomes:
1. Understand the logical and physical components of a Storage infrastructure.
2. Evaluate storage architectures, including storage subsystems, DAS, SAN, NAS, and CAS.
3. Understand the various forms and types of Storage Virtualization.
4. Describe the different roles in providing disaster recovery and business continuity capabilities.
5. Distinguish different remote replication technologies.

UNIT - I
Introduction to Storage Technology: Data proliferation and the varying value of data with time & usage, Sources of data and states of data creation, Data center requirements and evolution to accommodate storage needs, Overview of basic storage management skills and activities, The five pillars of technology, Overview of storage infrastructure components, Evolution of storage, Information Lifecycle Management concept, Data categorization within an enterprise, Storage and Regulations.

UNIT - II
Storage Systems Architecture: Intelligent disk subsystems overview, Contrast of integrated vs. Modular arrays, Component architecture of intelligent disk subsystems, Disk physical structure-components, properties, performance, and specifications, Logical partitioning of disks, RAID & parity algorithms, hot sparing, Physical vs. logical disk organization, protection, and back end management, Array caching properties and algorithms, Front end connectivity and queuing properties, Front end to host storage provisioning, mapping, and operation, Interaction of file systems with storage, Storage system connectivity protocols.

UNIT - III
Introduction to Networked Storage: JBOD, DAS, SAN, NAS, & CAS evolution, Direct Attached Storage (DAS) environments: elements, connectivity, & management, Storage Area Networks (SAN): elements & connectivity, Fibre Channel principles, standards, & network management principles, SAN management principles, Network Attached Storage (NAS): elements, connectivity options, connectivity protocols (NFS, CIFS, ftp), & management principles, IP SAN elements, standards (iSCSI, FCIP, iFCP), connectivity principles, security, and management principles, Content Addressable Storage (CAS): elements, connectivity options, standards, and management principles, Hybrid Storage - solutions overview including technologies like virtualization & appliances.

UNIT - IV
Introductions to Information Availability: Business Continuity and Disaster Recovery Basics, Local business continuity techniques, Remote business continuity techniques, Disaster Recovery principles & techniques. **Managing & Monitoring:** Management philosophies (holistic vs. system & component), Industry management standards (SNMP, SMI-S, CIM), Standard framework applications, Key management metrics (thresholds, availability, capacity, security, performance), Metric analysis methodologies & trend analysis, Reactive and proactive management best practices, Provisioning & configuration change planning, Problem reporting, prioritization, and handling techniques, Management

tools overview.

UNIT - V
Securing Storage and Storage Virtualization: Define storage security. List the critical security attributes for information systems, describe the elements of a shared storage model and security extensions, Define storage security domains, List and analyze the common threats in each domain, Identify different virtualization technologies, describe block-level and file level virtualization technologies and processes.

TEXT BOOKS:
1. Marc Farley Osborne, "Building Storage Networks'', Tata McGraw Hill, 2001.
2. Robert Spalding and Robert Spalding, "Storage Networks: The Complete Reference", Tata McGraw Hill, 2003.
3. Meeta Gupta, "Storage Area Network Fundamentals", Pearson Education Ltd., 2002.

REFERENCE BOOKS:
1. Gerald J Kowalski and Mark T Maybury," Information Storage Retrieval Systems theory & Implementation", BS Publications, 2000.
2. Thejendra BS, "Disaster Recovery & Business continuity", Shroff Publishers & Distributors, 2006.

SOFTWARE PROJECT MANAGEMENT (Professional Elective – IV)

B.Tech. IV Year I Sem.   **L T P C**
                                                                      3 0 0 3

Prerequisites: A course on "Software Engineering".

Course Objectives
1. To develop skills in software project management
2. The topics include-software economics; software development life cycle; artifacts of the process; workflows; checkpoints; project organization and responsibilities; project control and process instrumentation.

Course Outcomes
1. Gain knowledge of software economics, phases in the life cycle of software development, project organization, project control and process instrumentation.
2. Analyze the major and minor milestones, artifacts and metrics from management and technical perspective.
3. Design and develop software product using conventional and modern principles of software project management

UNIT - I
Conventional Software Management: The waterfall model, conventional software Management performance. **Evolution of Software Economics:** Software economics, pragmatic software cost estimation.

UNIT - II
Improving Software Economics: Reducing Software product size, improving software processes, improving team effectiveness, improving automation, Achieving required quality, peer inspections.
The old way and the new: The principles of conventional software Engineering, principles of modern software management, transitioning to an iterative process.

UNIT - III
Life cycle phases: Engineering and production stages, inception, Elaboration, construction, transition phases. **Artifacts of the process:** The artifact sets, Management artifacts, Engineering artifacts, programmatic artifacts. **Model based software architectures:** A Management perspective and technical perspective. **Work Flows of the process:** Software process workflows, Iteration workflows.

UNIT - IV
Checkpoints of the process: Major milestones, Minor Milestones, Periodic status assessments. Iterative Process Planning**:** work breakdown structures, planning guidelines, cost and schedule estimating, Iteration planning process, Pragmatic planning. **Project Organizations and Responsibilities:** Line-of-Business Organizations, Project Organizations, evolution of Organizations. Process Automation**:** Automation building blocks, The Project Environment.

UNIT - V
Project Control and Process instrumentation: The seven core Metrics, Management indicators, quality indicators, life cycle expectations, pragmatic Software Metrics, Metrics automation. Tailoring the Process**:** Process discriminates. **Future Software Project Management:** modern Project Profiles, Next generation Software economics, modern process transitions. **Case Study:** The command Center Processing and Display system- Replacement (CCPDS-R).

TEXT BOOKS:
1. Software Project Management, Walker Royce: Pearson Education, 2005.

REFERENCE BOOKS:
1. Software Project Management, Bob Hughes and Mike Cotterell: Tata McGraw-Hill Edition.
2. Software Project Management, Joel Henry, Pearson Education.
3. Software Project Management in practice, Pankaj Jalote, Pearson Education. 2005.

PRIVACY PRESERVING IN DATA MINING (Professional Elective – V)

B.Tech. IV Year I Sem.   L T P C
                                                                            3 0 0 3

Prerequisites: A course on "Data Mining".

Course Objectives:
1. The aim of the course is to introduce the fundamentals of Privacy Preserving Data Mining Methods
2. The course gives an overview of - Anonymity and its Measures, Multiplicative Perturbation for Privacy-Preserving Data Mining, techniques for Utility-based Privacy Preserving Data

Course Outcomes:
1. Understand the concepts of Privacy Preserving Data Mining Models and Algorithms.
2. Demonstrate a comprehensive understanding of different tasks associated in Inference Control Methods for Privacy-Preserving Data Mining.
3. Understand the concepts of Data Anonymization Methods and its Measures.
4. Evaluate and Appraise the solution designed for Multiplicative Perturbation.
5. Formulate, Design and Implement the solutions for Utility-based Privacy Preserving Data.

UNIT - I
Introduction, Privacy-Preserving Data Mining Algorithms, The Randomization Method, Group Based Anonymization, Distributed Privacy-Preserving Data Mining

UNIT - II
Interface Control Methods
Introduction, A Classification of Microdata Protection Methods, Perturbative Masking Methods, Non-Perturbative Masking Methods, Synthetic Microdata Generation, Trading off Information Loss and Disclosure Risk.

UNIT - III
Measure of Anonymity
Data Anonymization Methods, A Classification of Methods, Statistical Measure of Anonymous, Probabilistic Measure of Anonymity, Computational Measure of Anonymity, reconstruction Methods for Randomization, Application of Randomization

UNIT - IV
Multiplicative Perturbation
Definition of Multiplicative Perturbation, Transformation Invariant Data Mining Models, Privacy Evaluation for Multiplicative Perturbation, Attack Resilient Multiplicative Perturbation, Metrics for Quantifying Privacy Level, Metrics for Quantifying Hiding Failure, Metrics for Quantifying Data Quality.

UNIT - V
Utility-Based Privacy-Preserving Data
Types of Utility-Based Privacy Preserving Methods, Utility-Based Anonymization Using Local Recording, The Utility-Based Privacy Preserving Methods in Classification Problems, Anonymization Merginal: Injection Utility into Anonymization Data Sets.

TEXT BOOK:
1. Privacy – Preserving Data Mining: Models and Algorithms Edited by Charu C. Aggarwal and S. Yu, Springer.

REFERENCE BOOKS:
1. Charu C. Agarwal, Data Mining: The Textbook, 1st Edition, Springer.
2. J. Han and M. Kamber, Data Mining: Concepts and Techniques, 3rd Edition, Elsevier.
3. Privacy Preserving Data Mining by Jaideep Vaidya, Yu Michael Zhu and Chirstopher W. Clifton, Springer.

CLOUD COMPUTING (Professional Elective – V)

B.Tech. IV Year I Sem.　　　　　　　　　　　　　　　　　　　　**L　T　P　C**
　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　3　0　0　3

Pre-requisites: Courses on Computer Networks, Operating Systems, Distributed Systems.

Course Objectives:
1. This course provides an insight into cloud computing.
2. Topics covered include- distributed system models, different cloud service models, service-oriented architectures, cloud programming and software environments, resource management.

Course Outcomes:
1. Ability to understand various service delivery models of a cloud computing architecture.
2. Ability to understand the ways in which the cloud can be programmed and deployed.
3. Understanding cloud service providers.

UNIT - I
Computing Paradigms: High-Performance Computing, Parallel Computing, Distributed Computing, Cluster Computing, Grid Computing, Cloud Computing, Bio computing, Mobile Computing, Quantum Computing, Optical Computing, Nano computing.

UNIT - II
Cloud Computing Fundamentals: Motivation for Cloud Computing, The Need   for Cloud Computing, Defining Cloud Computing, Definition of Cloud computing, Cloud Computing Is a Service, Cloud Computing Is a Platform, Principles of Cloud computing, Five Essential Characteristics, Four Cloud Deployment Models

UNIT - III
Cloud Computing Architecture and Management: Cloud architecture, Layer, Anatomy of the Cloud, Network Connectivity in Cloud Computing, Applications, on the Cloud, Managing the Cloud, Managing the Cloud Infrastructure Managing the Cloud application, Migrating Application to Cloud, Phases of Cloud Migration Approaches for Cloud Migration.

UNIT - IV
Cloud Service Models: Infrastructure as a Service, Characteristics of IaaS. Suitability of IaaS, Pros and Cons of IaaS, Summary of IaaS Providers, Platform as a Service, Characteristics of PaaS, Suitability of PaaS, Pros and Cons of PaaS, Summary of PaaS Providers, Software as a Service, Characteristics of SaaS, Suitability of SaaS, Pros and Cons of SaaS, Summary of SaaS Providers, Other Cloud Service Models.

UNIT V
Cloud Service Providers: EMC, EMC IT, Captiva Cloud Toolkit, Google, Cloud Platform, Cloud Storage, Google Cloud Connect, Google Cloud Print, Google App Engine, Amazon Web Services, Amazon Elastic Compute Cloud, Amazon Simple Storage Service, Amazon Simple Queue ,service, Microsoft, Windows Azure, Microsoft Assessment and Planning Toolkit, SharePoint, IBM, Cloud Models, IBM Smart Cloud, SAP Labs, SAP HANA Cloud Platform, Virtualization Services Provided by SAP, Sales force, Sales Cloud, Service Cloud: Knowledge as a Service, Rack space, VMware, Manjra soft, Aneka Platform

TEXT BOOK:
1. Essentials of cloud Computing: K. Chandrasekhran, CRC press, 2014

REFERENCE BOOKS:
1. Cloud Computing: Principles and Paradigms by Rajkumar Buyya, James Broberg and Andrzej M. Goscinski, Wiley, 2011.
2. Distributed and Cloud Computing, Kai Hwang, Geoffery C. Fox, Jack J. Dongarra, Elsevier, 2012.
3. Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance, Tim Mather, Subra Kumaraswamy, Shahed Latif, O'Reilly, SPD, rp 2011.

DATA STREAM MINING (Professional Elective – V)

B.Tech. IV Year I Sem.

L	T	P	C
3	0	0	3

Prerequisites
1. A basic knowledge of "Data Mining"

Course Objectives:
1. The aim of the course is to introduce the fundamentals of Data Stream Mining.
2. The course gives an overview of – Mining Strategies, methods and algorithms for data stream mining.

Course Outcomes:
1. Understand how to formulate a knowledge extraction problem from data streams.
2. Ability to apply methods / algorithms to new data stream analysis problems.
3. Evaluate the results and understand the functioning of the methods studied.
4. Demonstrate decision tree and adaptive Hoeffding Tree concepts

UNIT - I
MOA Stream Mining, Assumptions, Requirements, Mining Strategies, Change Detection Strategies, MOA Experimental Settings, Previous Evaluation Practices, Evaluation Procedures for Data Streams, Testing Framework, Environments, Data Sources, Generation Speed and Data Size, Evolving Stream Experimental Setting.

UNIT - II
Hoeffding Trees, The Hoeffding Bound for Tree Induction, The Basic Algorithm, Memory Management, Numeric Attributes, Batch Setting Approaches, Data Stream Approaches.

UNIT - III
Prediction Strategies, Majority Class, Naïve Bayes Leaves, Adaptive Hybrid, Hoeffding Tree Ensembles, Data Stream Setting, Realistic Ensemble Sizes.

UNIT - IV
Evolving Data Streams, Algorithms for Mining with Change, A Methodology for Adaptive Stream Mining, Optimal Change Detector and Predictor, Adaptive Sliding Windows, Introduction, Maintaining Updated Windows of Varying Length.

UNIT - V
Adaptive Hoeffding Trees, Introduction, Decision Trees on Sliding Windows, Hoeffding Adaptive Trees, Adaptive Ensemble Methods, New methods of Bagging using trees of different size, New method of bagging using ADWIN, Adaptive Hoeffding Option Trees, Method performance.

TEXT BOOK:
1. DATA STREAM MINING: A Practical Approach by Albert Bifet and Richard Kirkby.

REFERENCE BOOKS:
1. Knowledge discovery from data streams by Gama João. ISBN: 978-1-4398-2611-9.
2. Machine Learning for Data Streams by Albert Bifet, Ricard Gavalda; MIT Press, 2017.

SOFTWARE QUALITY ASSURANCE (Professional Elective – V)

B.Tech. IV Year I Sem.   **L   T   P   C**
                                                                                     3   0   0   3

Course Objectives: Assurance of software quality using metrics, factors and process standards.

Course Outcomes:
1. Understand software quality challenge
2. Classify software requirements
3. Understand the models for cost of software quality
4. Implement software quality standards

UNIT - I
The software quality challenge: The uniqueness of software quality assurance, The environments for which SQA methods are developed.
What is software quality? What is software? Software errors, faults and failures, Classification of the causes of software errors, Software quality – definition, Software quality assurance – definition and objectives, Software quality assurance and software engineering

UNIT - II
Software quality factors: The need for comprehensive software quality Requirements, Classifications of software requirements into software quality factors, Product operation software quality factors, Product revision software quality factors, Product transition software quality factors, Alternative models of software quality factors, Who is interested in the definition of quality Requirements? Software compliance with quality factors.

UNIT - III
Software quality metrics: Objectives of quality measurement, Classification of software quality metrics, Process metrics, Product metrics, Implementation of software quality metrics, Limitations of software metrics. **Costs of software quality:** Objectives of cost of software quality metrics, The classic model of cost of software quality, An extended model for cost of software quality, Application of a cost of software quality system, Problems in the application of cost of software, quality metrics.

UNIT - IV
SQA project process standards – IEEE software engineering standards: Structure and content of IEEE software engineering standards, IEEE/EIA Std 12207 – software life cycle processes, IEEE Std 1012 – verification and validation, IEEE Std 1028 – reviews.
Management and its role in software quality assurance: Top management's quality assurance activities, Department management responsibilities for quality assurance, Project management responsibilities for quality assurance.

UNIT - V
The SQA unit and other actors in the SQA system
The SQA unit, SQA trustees and their tasks, SQA committees and their tasks, SQA forums – tasks and methods of operation

TEXT BOOK:
1. Software Quality Assurance From theory to implementation, DANIEL GALIN, Pearson Education.

REFERENCE BOOK:
1. Metrics and Models In Software Quality Engineering, Kan Pearson Education.

EXPLORATORY DATA ANALYSIS (Professional Elective – V)

B.Tech. IV Year I Sem.   L T P C
                                                                                        3 0 0 3

Course Objectives:
1. This course introduces the methods for data preparation and data understanding.
2. It covers essential exploratory techniques for understanding multivariate data by summarizing it through statistical methods and graphical methods.
3. Supports to Summarize the insurers use of predictive analytics, data science and Data Visualization.

Course Outcomes:
1. Handle missing data in the real-world data sets by choosing appropriate methods.
2. Summarize the data using basic statistics. Visualize the data using basic graphs and plots.
3. Identify the outliers if any in the data set.
4. Choose appropriate feature selection and dimensionality reduction.
5. Techniques for handling multi-dimensional data.

UNIT - I:
Introduction to Exploratory Data Analysis: Data Analytics lifecycle, Exploratory Data Analysis (EDA)– Definition, Motivation, Steps in data exploration, The basic data types Data Type Portability.

UNIT - II:
Preprocessing - Traditional Methods and Maximum Likelihood Estimation: Introduction to Missing data, Traditional methods for dealing with missing data, Maximum Likelihood Estimation – Basics, Missing data handling, Improving the accuracy of analysis. **Preprocessing Bayesian Estimation:** Introduction to Bayesian Estimation, Multiple Imputation-Imputation Phase, Analysis and Pooling Phase, Practical Issues in Multiple Imputation, Models for Missing Notation Random Data.

UNIT - III:
Data Summarization & Visualization: Statistical data elaboration, 1-D Statistical data analysis, 2-D Statistical data Analysis, N-D Statistical data analysis.

UNIT - IV:
Outlier Analysis: Introduction, Extreme Value Analysis, Clustering based, Distance Based and Density Based outlier analysis, Outlier Detection in Categorical Data. **Feature Subset Selection:** Feature selection algorithms: filter methods, wrapper methods and embedded methods, Forward selection backward elimination, Relief, greedy selection, genetic algorithms for features selection.

UNIT - V
Dimensionality Reduction: Introduction, Principal Component Analysis (PCA), Kernel PCA, Canonical Correlation Analysis, Factor Analysis, Multidimensional scaling, Correspondence Analysis.

TEXT BOOKS:
1. Making sense of Data: A practical Guide to Exploratory Data Analysis and Data Mining, by Glenn J. Myatt.

REFERENCE BOOKS:
1. Charu C. Aggarwal, "Data Mining The Text book", Springer, 2015.
2. Craig K. Enders, "Applied Missing Data Analysis", The Guilford Press, 2010.
3. Inge Koch, "Analysis of Multivariate and High dimensional data", Cambridge University Press, 2014.
4. Michael Jambu, "Exploratory and multivariate data analysis", Academic Press Inc., 1990.
5. Charu C. Aggarwal, "Data Classification Algorithms and Applications", CRC press, 2015.

MACHINE LEARNING LAB

B.Tech. IV Year I Sem.   **L T P C**
                                                                              0 0 2 1

Course Objective: The objective of this lab is to get an overview of the various machine learning techniques and can able to demonstrate them using python.

Course Outcomes: After the completion of the course the student can able to:
1. understand complexity of Machine Learning algorithms and their limitations;
2. understand modern notions in data analysis-oriented computing;
3. be capable of confidently applying common Machine Learning algorithms in practice and implementing their own;
4. Be capable of performing experiments in Machine Learning using real-world data.

List of Experiments
1. The probability that it is Friday and that a student is absent is 3 %. Since there are 5 school days in a week, the probability that it is Friday is 20 %. What is theprobability that a student is absent given that today is Friday? Apply Baye's rule in python to get the result. (Ans: 15%)
2. Extract the data from database using python
3. Implement k-nearest neighbours classification using python
4. Given the following data, which specify classifications for nine combinations of VAR1 and VAR2 predict a classification for a case where VAR1=0.906 and VAR2=0.606, using the result of k-means clustering with 3 means (i.e., 3 centroids)

VAR1	VAR2	CLASS
1.713	1.586	0
0.180	1.786	1
0.353	1.240	1
0.940	1.566	0
1.486	0.759	1
1.266	1.106	0
1.540	0.419	1
0.459	1.799	1
0.773	0.186	1

5. The following training examples map descriptions of individuals onto high, medium and low credit-worthiness.

   medium skiing design   single twenties no -> highRisk
   high   golf   trading   married forties yes -> lowRisk
   low   speedway transport married thirties yes -> medRisk
   medium football banking   single thirties yes -> lowRisk
   high   flying media   married fifties yes -> highRisk
   low   football security   single twenties no -> medRisk
   medium golf   media   single thirties yes -> medRisk
   medium golf   transport married forties yes -> lowRisk
   high   skiing banking   single thirties yes -> highRisk
   low   golf   unemployed married forties yes -> highRisk

Input attributes are (from left to right) income, recreation, job, status, age-group, home-owner. Find the unconditional probability of `golf' and the conditional probability of `single' given `medRisk' in the dataset?

6. Implement linear regression using python.
7. Implement Naïve Bayes theorem to classify the English text
8. Implement an algorithm to demonstrate the significance of genetic algorithm
9. Implement the finite words classification system using Back-propagation algorithm