

B.Tech. CSE Sem – VI [January-June 2021]
(CSE-Core/CSE-IOT/CSE-AI/CSE-CTIS/CSE-DS)

Sr. No.	Course Code	Courses	L	T	P	Credit
1	CS3CO13	Design and Analysis of Algorithms	3	1	2	5
2	CS3CO15	Object Oriented Analysis and Design	3	1	2	5
3	CS3CO27	Compiler Design	3	0	2	4
4	CS3EXXX	Elective-3	3	0	0	3
5	CS3EXXX	Elective-4	3	0	0	3
6	CS3ES14	Software Workshop-III	0	0	2	1
7	OE000XX	Open Elective-2	3	0	0	3
		Total	18	2	8	24
		Total Contact Hours	28			

Elective-3: *(Choose any one)*

[Track: Web Technology]: Programming with XML (CS3EL08)

[Track: AI]: Soft Computing (CS3EA03)

[Track: Data Engineering]: Big Data Engineering (CS3ED04)

[Track: CTIS (inerture)]: Ethical Hacking (CS3EY05)

[Track: Data Science(inerture)]: Big Data Analytics (CS3ET05)

Elective-4: *(Choose any one)*

[Track: Web Technology]: Internet of Things (CS3EL06)

[Track: AI]: Natural Language Processing (CS3EA06)

[Track: Data Engineering]: Database Application and Tools (CS3ED01)

[Track: CTIS (inerture)]: Principles of Virtualization (CS3EY04)

[Track: Data Science(inerture)]: Machine Learning (CS3ET07)

Open Elective-2(CSE-Core): *(Choose any one)*

R Programming (OE00051)

Agile Development (OE00015)

Blockchain Architecture (OE00016)

Open Elective-2(CSE-CTIS):

Cloud Web Services (OE00074)

Open Elective-2(CSE-DS):

Exploratory Data Analytics (OE00075)

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
CS3CO13	Design and Analysis of Algorithms	3	1	2	6	5

Unit-I: Introduction to Algorithms

Algorithms, Analysis, Performance issues: Time and Space complexity; Asymptotic Notations.

Mathematical preliminaries: functions & their growth rates; Recurrence relations, Methods for solving recurrences. Elementary Sorting techniques and its analysis: Selection, Bubble, Insertion sort

Unit-II: Sorting and Divide & Conquer

Advance sorting techniques and its analysis: Heap sort, Radix sort and Bucket sort, Divide and Conquer techniques and its analysis - Binary search, Merge Sort, Quick sort, Strassen's Matrix multiplication.

Unit-III: Greedy Algorithms

Greedy problems and its complexity analysis: Optimal merge patterns, Huffman coding, Minimum spanning trees, Knapsack problem, Job sequencing with deadlines, Single source shortest path problem - Dijkstra's Algorithm

Unit-IV: Dynamic Programming

Dynamic programming problems and its complexity analysis: 0/1 Knapsack, Multistage graph, Bellman Ford Algorithm, Reliability design, Floyd-Warshall algorithm, Matrix Chain Multiplication, Longest Common subsequence.

Unit-V: Backtracking and Branch & Bound

Backtracking Approach: N-Queen's problem, Hamiltonian cycle, Graph coloring problem, Sum of Subset problem. Introduction to branch & bound method, examples of branch and bound method like 15 puzzle traveling salesman problem, 0/1 knapsack. An introduction to P, NP, NP Complete and NP hard problems.

Text Books:

1. Thomas H Cormen, Charles E Lieserson, Ronald L Rivest and Clifford Stein, Introduction to Algorithms, Second Edition, MIT Press/McGraw-Hill
2. E. Horowitz, S. Sahni, S Rajasekaran, Computer Algorithms, Galgotia Publications

Reference Books:

1. Saara Base, Computer Algorithms: Introduction to Design and Analysis, Addison Wesley.
2. A V Aho, J E Hopcroft & J D Ullman, The Design and Analysis of Computer Algorithms, Addison Wesley.

Course Code	Course Name	Hours Per Week			
		L	T	P	Credits
CS3CO15	Object Oriented Analysis and Design	3	1	2	5

Unit I

Structure of Complex Systems, Object Oriented Development Methods, Characteristics of Objects, Fundamental Concepts of Object orientation, UML- Overview, RUP and its Phases

Unit II

Models, Concepts in UML, Structural and Behavioral Models, Use Cases and functional Requirements, Use Case Descriptions, Classes, Relationships, Association, Generalization, Realization, Dependencies, Constraints

Unit III

State Machine View, Activity View, Interaction View, Sequence Diagram, Collaboration Diagram, Interaction Diagrams

Unit IV

Physical View, Component Diagram, Deployment Diagram, Package, Dependencies on Packages, Modelling System and Subsystems, Patterns and Types of Patterns, Applying Patterns

Unit V

Object Oriented Testing, Types of Testing, Quality Assurance Methods, Reusability, Reverse Engineering, Case Studies

Text Book:

1. Grady Booch, Object Oriented Analysis and Design with Applications, Addison Wesley
2. James Rumbaugh, Ivar Jacobson, Grady Booch, The Unified Modelling Language Reference Manual, Addison Wesley

Reference Book:

1. Design Patterns - Elements of Reusable Object-Oriented Software, Gamma, et. al., Addison-Wesley.
2. Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development, by Craig Larman, Pearson Education.

Course Code	Course Name	Hours Per Week			
		L	T	P	Credits
CS3CO27	Compiler Design	3	0	2	4

Unit I

Compiler structure: Pass Structure of compiler, Translators, Interpreter, Assembler, Phases of Compilers, Symbol Table, Error Handling, Lexical Analyzer: Role of Lexical Analyzer, Specification of tokens, Recognition of tokens and input Buffering, The Syntactic Specification of Programming Languages, Cross Compiler, bootstrap Compiler.

Unit II

Ambiguous Grammar, LL(0) and LL(1) grammar, Parsing, Basic Parsing Techniques: Top Down parsers, Recursive Descent Parsers, First() and Follow(), Recursive and Non- Recursive Predictive Parsers.

Unit III

LR Grammar, Operator Grammar, Bottom Up Parsing: Operator precedence parsing, LR(0) parsers, Construction of SLR, Canonical LR and LALR parsing tables.

Unit IV

Syntax Directed Definition, Translation Scheme, Synthesized and inherited attributes, dependency graph, Construction of syntax trees, S-attributed and L-attributed definitions, Three address codes, quadruples, triples and indirect triples, Translation of assignment statements.

Unit V

Storage organization, activation trees, activation records, allocation strategies, Parameter passing symbol table, dynamic storage allocation, Basic blocks and flow graphs, Optimization of basic blocks, Loop optimization, Global data flow analysis, Loop invariant computations.

Text Book:

1. Alfred V. Aho, and J.D. Ullman, Principle of Compiler Design, Narosa Publication.
2. A.Barret William and R.M. Bates, Compiler construction (Theory and Practice), Galgotia Publication.

Reference Book:

1. A.C. Holub, Compiler design in C, PHI.
2. O.G. Kakde, Compiler Design, Laxmi Publications

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
CS3EL08	Programming with XML	3	0	0	3	3

Unit-I: Introduction to XML

XML overview, Markup languages, Comparison with HTML, Usage, Rules for writing XML, XML syntax, creating notebook XML, Tree structure of XML, Elements, Attributes and values, Root element, Child element, nesting of elements, Empty elements, adding attributes, Elements and Attributes uses, Writing comments, Predefined entities, XML tools, XML validation.

Unit-II: XML-DTDs (Document Type Definitions)

Document Type Definition, DTD syntax, creating a DTD for notebook XML, defining elements with children, Empty element, Number of occurrences, defining choices, Attribute definitions, Internal and external DTD's, Validating XML with DTD, Pros and cons of using DTD.

Unit-III: XML-Schema

Introduction to Schema, Namespace, Schema definition, Data types, Simple and complex data types, Attributes definition, Restrictions on values, creating schema definition for notebook XML, Link and Validate XML with schema.

Unit-IV: XSLT

Introduction to XSL, Layout of an XSL Document and Templates, Linking XSL to your XML Source, Transforming XML with XSLT, xsl:output, xsl:template, xsl:apply-templates, Looping over nodes using xsl:for-each, Apply conditions using xsl:if, Processing and output using xsl:value-of, Sorting nodes, Create a XSLT for notebook and XML file and generate output in different conditions.

Unit-V: XPath and Project

Introduction to XPath, Using XPath to navigate an XML document, Predicates.

Sample Project: Store the information of students in XML file, validate it using XML schema and display the information of students in HTML using XSLT with proper formatting and conditions like having enrollment number, name start with, having CGPA between, in sorted order, etc.

Text Books:

1. Introduction to XML V.1, O'Reilly Publication.
2. Deitel H.M., XML How to Program, Pearson Publication.
3. Uttam K. Roy, Web Technologies, Oxford University Press.

References

1. Michael J. Young, XML Step by Step, Microsoft Press; 2nd edition
2. Elliotte Rusty Harold, XML Bible Second Edition, Hungry Minds Publication.

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
CS3EA03	Soft Computing	3	0	0	3	3

Unit- I: Concept of Computing Systems, Introduction to Soft Computing, Soft Computing vs. Hard Computing, Components of Soft Computing, Neural Networks: Structure and function of Biological Neuron and Artificial Neuron, Definition and characteristics of ANN, Training techniques in different ANNs, Activation functions, Different ANNs architectures, Introduction to basic ANN architectures: McCulloch & Pitts model, Perceptron Model, Linear separability, ADALINE and MADALINE.

Unit- II: Neural Network Architectures: Supervised Learning: Backpropagation Network architecture, Backpropagation algorithm, Limitation, Characteristics and Application of EBPA, Bidirectional associative memories (BAM), Unsupervised Learning: Hebbian Learning, Generalized Hebbian learning algorithm, Competitive learning, Self- Organizing Computational Maps: Kohonen Network, Applications of ANN to solve real world's problems.

Unit- III: Fuzzy Logic: Introduction to Fuzzy logic, Fuzzy sets and membership functions, Operations on Fuzzy sets, Fuzzy relations, rules, propositions, implications and inferences, Defuzzification techniques, Fuzzy logic controller design, Some applications of Fuzzy logic.

Unit- IV: Genetic Algorithms: Concept of "Genetics" and "Evolution" and its application to probabilistic search techniques, Basic GA framework and different GA architectures, GA operators: Encoding, Crossover, Selection, Mutation, etc, solving single-objective optimization problems using GAs.

Unit- V: Hybrid Systems: Genetic Algorithm based Backpropagation Network, Fuzzy – Backpropagation, Fuzzy Logic Controlled Genetic Algorithms. Case studies. Case studies in Engineering

Text Book:

1. Sinha, N.K. and Gupta, M. M.: “Soft Computing and Intelligent Systems - Theory and Applications”, Academic Press.
2. S. Rajasekaran and G. A. Vijayalakshmi Pai, “Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis, and Applications”, Prentice Hall of India, 2007.

Reference Books:

1. D. K. Pratihar, “Soft Computing”, Narosa, 2008.
2. Jang, J-S. R., Sun,C-T, Mizutani, E.: “Neuro–Fuzzy and Soft Computing”, Prentice Hall of India.

Course Code	Course Name	Hours Per Week			
		L	T	P	Credits
CS3EA06/IT3EA06	Natural Language Processing	3	0	0	3

Unit-1

Introduction: Human languages, Main approach of NLP, Knowledge in speech and language processing, Ambiguity, Models and algorithms, Formal language and Natural Language, Regular Expression and automata.

Unit-2

Text Pre-processing, Tokenization, Feature Extraction from text, Morphology: Inflectional and Derivational, Finite state morphological parsing, Finite state transducer

Part of Speech Tagging: Rule based, Stochastic POS, Transformation based tagging.

Unit-3

Speech Processing: Speech and phonetics, Vocal organ, Phonological rules and Transducer, Probabilistic models: Spelling error, Bayesian method to spelling, Minimum edit distance, Bayesian method of pronunciation variation.

Unit-4

N-Grams: Simple N-Gram, perplexity, Smoothing, Backoff, Entropy, Parsing: Statistical Parsing, Probabilistic parsing, TreeBank.

Unit-5

Application: Sentiment analysis, spelling correction, Word sense disambiguation, Machine translation, Text Classification, Question answering system.

Text Book:

1. Daniel Jurafsky and James H. Martin, "Speech and Language Processing", Pearson Education.
2. James Allen, "Natural Language Understanding", Pearson Education.

Reference book:

1. Christopher D. Manning and Hinrich Schutze, "Foundation of statistical Natural Language Processing", MIT Press.
2. Mary Dee Harris "Introduction to Natural language Processing", Reston.

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
CS3ED04	Big Data Engineering	3	0	0	3	3

Unit 1: Foundations of Big Data Systems

Introduction to Big Data and its Applications Data Abstraction Linear data structures like Hashtables, Hashmaps, Bloom Filters Non-linear data structures like Binary Search Trees, KD Trees Distributed Algorithm Design Algorithm Design using MapReduce

Unit 2: Platforms for Big Data

Distributed Computing Environment for Big Data NoSQL databases for Big Data Storage Applications (HBase) Distributed Processing of data using MapReduce & Pig In-memory distributed processing using Apache Spark Data Storage on Cloud (Amazon S3 & Dynamo DB)

UNIT 3: Processing Big Data – ETL & Batch Processing

Performing ETL Operations Concepts in Data Warehousing and its relevance for Big Data Ingesting data into Big Data Platforms using Sqoop & Flume Workflow management for Hadoop using OOZIE Batch Processing on Cloud

UNIT 4: Processing of Real Time Data & Streaming Data

Applications of Streaming Data in Industry Sourcing Streaming data using Apache Flume Building real-time data pipelines using Apache Storm Streaming on Apache Spark

UNIT 5: Big Data Analytics

Regression, Clustering & Classification using Spark MLlib Building visualizations using Big Data Case Studies on applications of Big Data Analytics

Text Books

1. Mayank Bhushan, Big Data and Hadoop- Learn by Example, BPB Publications
2. Erl/Khattak/Buhler, Big Data Fundamentals: Concepts Drivers and Techniques, Prentice Hall

References

1. Jeffrey Aven , Hadoop in 24 Hours, Sams Teach Yourself, SAMS Publications.
2. DT Editorial Services, Big Data, Black Book: Covers Hadoop 2, MapReduce, Hive, YARN, Pig, R and Data Visualization, Dream Tech Publications

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
CS3ED01	Database Application and Tools	3	0	0	3	3

UNIT I

Database Environment: Data versus information, traditional file processing, disadvantages, database approach, range of database application, advantages of database approach. Cost and risk factors, components of database environment, evolution of database system.

Database Development Process: Information engineering, information architecture, enterprise data model, planning, SDLC, CASE etc. Steps of planning, strategic planning factors, corporate planning objects. Developing preliminary data model, and use of planning matrices, SDLC steps, CASE role, people in database development, three-schema architecture for database development. Examples to demonstrate the development process.

UNIT II

Modeling Data in the Organization: Modeling of the rules of organization, data names and definitions, ER model constructs entities and its types, attributes, relationships, degree, unary, binary, ternary, n-ary, cardinalities constraints, ER modeling examples.

Enhanced ER modeling: supertype, subtypes, specialization, generalization, specifying constraints in EER models, completeness, Disjointness, discriminators, defining super/sub type hierarchies, EER modeling examples, live demos modelling for few scenarios.

UNIT II

Logical database design: and relational model development, Relational model properties, keys, primary, secondary, composite, properties of relations. Codd's rules, integrity constraints, creating relational tables, Transform EER diagrams into relations, seven different steps for mapping EER model into relations.

UNIT IV

Introduction to normalization: steps, functional dependencies, basic normal forms, definition of first, second, third normal form and removing anomalies from the relations. De-normalization and merging relations.

UNIT V

Special Topics (Overview): Data Warehousing, Data Mining, Distributed Databases, Object oriented modeling, definitions, activities in phases of model development, advantages of OOM, UML class diagrams, Example of a model development.

Text Book:

1. Hoffer, Prescott, "Modern Database Management", Seventh Edition, McFadden Pearson Education.

Reference Book(s):

1. Thomas M. Connolly, Carolyn E. Begg, "Database Systems", Pearson Education.
2. Raghu R and Johannes G., "Database management Systems", Mc Hill 3rd Edition, 2002.
3. Elmasri R, Navathe S, "Fundamentals of Database Systems", Addison Wesley 4th Edition.

Course Code	Course Name	Hours Per Week			
		L	T	P	Credits
CS3EL06/IT3EL06	Internet of Things	3	0	0	3

Unit-I

Introduction: Definition, Characteristics of IoT, IoT Architectural view, Physical design of IoT, IoT Protocols, Communication Models of IoT, IoT Communication APIs, IoT Enabling Technologies.

Unit-II

IoT and M2M: Machine-to-Machine (M2M), Difference between M2M and IoT, SDN (Software Defined Networking) and NFV (Network Function Virtualization) for IoT, Data Storage in IoT, IoT Cloud Based Services.

Unit –III

IoT Platform Design Methodology: Specifications of Purpose and Requirement, Process, Domain Model, Information Model, Service, IoT Level, Functional View, Operational View, Device and Component Integration, Application Development.

Unit –IV

Security issues in IoT: Introduction, Vulnerabilities, Security requirements and threat analysis, IoT security Tomography, layered attacker model, identity management and establishment, access control.

Unit-V

Application areas of IoT: Home Automation, smart lighting, home intrusion detection, smart cities, smart parking, environment, weather monitoring system, agriculture.

Text Books:

1. Arshdeep Bahga, Vijay Madiseti, “Internet of Things – A hands-on approach”, Universities Press.
2. Rajkamal, ” Internet of Things”, Tata McGraw Hill publication

Reference Books:

1. Olivier Hersent, David Boswarthick, Omar Elloumi , “The Internet of Things – Key applications and Protocols”, Wiley
2. Donald Norris “The Internet of Things: Do-It-Yourself at Home Projects for Arduino, Raspberry Pi and BeagleBone Black”, McGraw Hill publication.

Open Learning Source:

1. <https://nptel.ac.in/courses/106105166/>
2. <https://github.com/connectIOT/iottoolkit>

Course Code	Course Name	Hours per Week			Total
		L	T	P	Credits
CS3EY04	Principle of Virtualization	3	0	0	3

Unit I: Introduction Introduction to Virtualization - Types of virtualization - Difference between cloud and virtualization - Physical infrastructure and virtual infrastructure - Virtualization approaches - Partitioning - Hosting - Isolation - Hardware independence - Virtual machine - Hypervisor - Types of hypervisor - Virtual machine manager - Types of hypervisor - Introduction to datacenter virtualization Esxi - Difference between Esxi and Esx - Versions of Esxi - Installation and configuration of Esxi 6.0 - vSphere 6.0

Unit II: Components of vSphere 6.0 Components of VMware vSphere - vSphere 6.0: Overview and Architecture - Topology of vSphere 6.0 Data Center - vSphere 6.0 Configuration MaximumsvCenter Server - vCenter Server Features - Certificate Management - Alarms and Alerts - Monitoring Features - Template Management - Linked Mode Deployment - Storage Features in vSphere - Shared Storage - Storage Protocols - Datastores - Virtual SAN - Virtual Volumes - Networking Features in vSphere - Virtual Networking - Virtual Switches and its types

Unit III: Features of vSphere and NSXvSphere Resource Management Features - vMotion - Distributed Resource Scheduler (DRS) - Distributed Power Management (DPM) - Storage vMotion - Storage DRS - Storage I/O Control - Network I/O Control - vSphere Availability Features - vSphere Data Protection - High Availability - Fault Tolerance - vSphere Replication - Introduction to NSX.

Unit VI: vSphere Solutions to Data Center Challenges and vSphere Security.Challenges - Availability Challenges - Scalability Challenges - Management Challenges - Optimization Challenges - Application Upgrade Challenges - Cloud Challenges - Security - Describe the features and benefits of VMware Platform Services Controller - Configure ESXi host access and authorization - Secure ESXi - vCenter Server - and virtual machines - Upgrade ESXi and vCenter Server instances

Unit V: Resource optimization and resource management Network Optimization - Configure and manage vSphere distributed switches - Migrate virtual machines from standard switches to distributed switches - Explain distributed switch features such as port mirroring - LACP - QoS tagging - and NetFlow - CPU Optimization - Explain the CPU scheduler operation - NUMA support - and other features that affect CPU performance - Monitor key CPU performance metrics - Memory Optimization - Explain ballooning - memory compression - and host swapping techniques for memory reclamation when memory is overcommitted - Monitor key memory performance metrics - Storage Optimization - Diagnose storage access problems - Configure VMware vSphere Flash Read Cache - Monitor key storage performance metrics

Text Books:

1. Virtualization Essentials Paperback – 26 Apr 2012 by Matthew Portnoy - wiley publications
2. VMware Cookbook Paperback – 17 Jul 2012 by Troy - Shroff/O'Reilly; Second edition (17 July 2012)

Reference Book:

1. Mastering VMware vSphere 5.5 (SYBEX) Paperback – 2014 by Scott Lowe, Nick Marshall, Forbes Guthrie, Matt Liebowitz, Josh Atwell - Wiley (2014) edition.

Course Code	Course Name	Hours per Week			Total
		L	T	P	Credits
CS3EY05	Ethical Hacking	3	0	0	3

Unit I: Introduction to Ethical Hacking Ethical Hacking, why is it necessary, scope and limitations, skills required, phases of ethical hacking, tools and techniques, Black Box, Gray Box and White Box techniques, differences between vulnerability assessment, ethical hacking and penetration testing, Reverse engineering, Ethical hacking terminology, Exploit, Vulnerability – Zero-day, manual PT, Case Studies on data breaches and cybercrimes involving hacking

Unit II: Ethical Hacking through Attacks and Exploits EH methodology, attacks, exploits, Denial of Service, Sniffers, malware, Session Hijacking and ethical hacking of Web Servers and applications, Password Cracking, Key Logger, Hash Injection attack, replay and man-in-the-middle attacks, rainbow table attack, distributed network attack, spoofing, phishing, spyware, rootkits, hiding files, counter measures

Unit III: Web and Network Hacking Enumeration and scanning, host discovery, type of scanning – TCP SYN, ACK, XMAS & UDP Port scanning, SQL Injection, Social Engineering, Buffer Overflows, Input data validation, physical penetration attacks, Hacking Wireless Networking, hacking mobile platforms, Windows and Linux Hacking, Evading IDS, Firewalls and Honeypots, DDoS attacks, using metasploit, counter measures.

Unit VI: Report Writing & Mitigation Introduction to Report Writing & Mitigation, requirements for low level reporting & high-level reporting of Penetration testing results, Demonstration of vulnerabilities and mitigation of issues identified including tracking, CVSS scoring for vulnerabilities, rating and prioritization, impact of these in reporting.

Unit V: Ethical Hacking and Legal System Overview of India's Information Technology Amendment Act 2008 (IT Act 2008), hacker vs cracker, liabilities – civil and penal, cyber theft and IPC sec 378, IT Act 2008 – sections 43, 65 and 66, how to file a complaint of suspected hacking, Case Studies, understanding how hacking is legally dealt with among BRICS countries

Text Books:

1. Gray Hat Hacking the Ethical Hackers Handbook, 3rd Edition Paperback – 1 Jul 2017 by Allen Harper, Shon Harris, Jonathan Ness, Chris Eagle, McGraw Hill Education; 3 ed (1 July 2017)
2. CEH v9: Certified Ethical Hacker Version 9 Study Guide by Sean-Philip Oriyano, Sybex; Stg edition (17 June 2016)
3. Hacking for Beginners: Ultimate 7 Hour Hacking Course for Beginners. Learn Wireless Hacking, Basic Security, Penetration Testing by Anthony Reynolds, CreateSpace Independent Publishing Platform (10 April 2017)
4. An Ethical Guide To WI-FI Hacking and Security by Swaroop Yermalkar, BecomeShakespeare.com; First edition (15 August 2014)
5. Hands-On Ethical Hacking and Network Defense by Michael T. Simpson | Kent Backman | James Corley, Cengage India 1st edition (2016)

Reference Book:

1. The Basics of Hacking and Penetration Testing: Ethical Hacking and Penetration Testing Made Easy by Patrick Egebretonson, Syngress; 2 editions (12 September 2013)
2. Hacking with Python: The Complete Guide to Ethical Hacking, Basic Security, Botnet Attack, Python hacking and Penetration Testing Kindle Edition by John C. Smalls.

Course Code	Course Name	Hours per Week			Total
		L	T	P	Credits
CS3ET07	Machine Learning	3	0	0	3

Unit I: Introduction to Machine Learning Algorithms

Introduction to Machine learning – Statistical Learning – types of Machine Learning –learning models: geometric, probabilistic and logistic models, introduction to supervised, unsupervised and reinforcement learning – model evaluation – model implementation – model accuracy indicators.

Unit II: Supervised Learning – Simple Linear Regression Analysis

Introduction to parametric machine learning method, assumptions of parametric machine learning methods, linear model and its assumptions, simple linear regression, scatter diagram, Simple linear Regression parameter estimation, properties of regression parameters, testing the significance of regression parameters using ANOVA and t test, estimation of σ^2 , Interval Estimation of the Mean Response, R Square, Adjusted R Square, Normality of response variable, prediction of new observations, Confidence interval for β_0 , β_1 and σ^2 .

Unit III: Supervised Learning – Multiple Linear Regression Analysis I

Multiple linear regression model, assumptions of Multiple linear regression variables – multicollinearity, homoscedasticity, autocorrelation, effects of multicollinearity, effect of homoscedasticity and auto autocorrelation in parameter estimation, Least - Squares Estimation of the Regression Coefficients, Geometrical Interpretation of Least Squares, Properties of the Least - Squares Estimators, Estimation of σ^2 , Inadequacy of Scatter Diagrams in Multiple Regression.

Unit IV:Supervised Learning – Multiple Linear Regression Analysis II

testing the general linear hypothesis, Test for Significance of Regression, Tests on Individual Regression Coefficients and Subsets of Coefficients, Special Case of Orthogonal Columns in X, Confidence Intervals on the Regression Coefficients, CI Estimation of the Mean Response, Simultaneous Confidence Intervals on Regression Coefficients, predicting new observations, residual analysis, model adequacy and validation.

Unit V: Supervised Learning – Non-Linear Regression Analysis

Introduction to non-linear regression models, non-linear least square method to estimating the regression parameters, transformation of non-linear model to linear model, linearization, other parameter estimation methods, starting values, statistical inference in non-linear regression models

Text Books

1. Introduction to Linear Regression Analysis, Fifth Edition - DOUGLAS C. MONTGOMERY, ELIZABETH A. PECK, G. GEOFFREY VINING, A JOHN WILEY & SONS, INC., PUBLICATION
2. Introduction to Machine Learning - EthemAlpaydm, The MIT Press

Reference Books

1. Python Machine Learning - Sebastian Raschka, PACKT Publishing
2. Using Multivariate Statistics - Barbara G. Tabachnick, Linda S. Fidell, Pearson Education Inc

Course Code	Course Name	Hours per Week			Total
		L	T	P	Credits
CS3ET05	Big Data Analytics	3	0	0	3

Unit I: Understanding BigData Defining Data, Types of Data, Structured Data, Semi Structured Data, Unstructured Data, how data being Generated, Different source of Data Generation, Rate at which Data is being generated, Different V's, Volume, Variety, Velocity, Veracity, Value, How single person is contributing towards BigData, Significance for BigData, Reason for BigData, Understanding RDBMS and why it is failing to store BigData. Future of BigData, BigData use cases for major IT Industries.

Unit II : Introduction to Hadoop What is Hadoop, Apache Community, Cluster, Node, Commodity Hardware, Rack Awareness, History of Hadoop, need for Hadoop, how is Hadoop Important, Apache Hadoop Ecosystem, Different Hadoop offering, Hadoop 1.x Architecture, Apache Hadoop Framework, Master- Slave Architecture, Advantages of Hadoop.

Unit III:Storage Unit Hadoop Distributed File System, Design of HDFS, HDFS Concept, How files are stored in HDFS, Hadoop File system, Replication factor, Name Node, Secondary Name Node, Job Tracker, Task tracker, Data Node, FS Image, Edit-logs, Check-pointing Concept, HDFS federation, HDFS High availability, Architectural description for Hadoop Cluster, When to use or not to use HDFS, Block Allocation in Hadoop Cluster, Read operation in HDFS, Write operation in HDFS, Hadoop Archives, Data Integrity in HDFS, Compression & Input Splits.

Unit IV: Processing Unit What is MapReduce, History of MapReduce, how does MapReduce work, Input files, Input Format types Output Format Types, Text Input Format, Key Value Input Format, Sequence File Input Format, Input split, Record Reader, MapReduce overview, Mapper Phase, Reducer Phase, Sort and Shuffle Phase, Importance of MapReduce

Data Flow, Counters, Combiner Function, Partition Function, Joins, Map Side Join, Reduce Side Join, MapReduce Web UI, Job Scheduling, Task Scheduling, Fault Tolerance, Writing MapReduce Application, Driver Class, Mapper Class, Reducer Class, Serialization, File Based Data Structure, writing a simple MapReduce program to Count Number of words, MapReduce Work Flows

Unit V: YARN &Hadoop Cluster YARN, YARN Architecture, YARN Components, Resource Manager, Node Manager, Application Master, Concept of Container, Difference between Hadoop 1.x and 2.x Architecture, Execution of Job in Yarn Cluster, Comparing and Contrasting Hadoop with Relational Databases Cluster Specification, Cluster Setup and Installation, Creating Hadoop user, Installing Hadoop, SSH Configuration, Hadoop Configuration, Hadoop daemon properties, Different modes of Hadoop, Standalone Mode, Pseudo Distributed Mode, Fully Distributed Modes,

Text Book:

1. Hadoop: The Definitive Guide, By: Tom White, O'REILLY

Reference Books:

1. Hadoop for Dummies, By: Dirk deRoos, Paul C. Zikopoulos, Bruce Brown, Rafael Coss, and Roman B. Melnyk, A Wiley brand
2. Hadoop in Action, Writer: [Chuck Lam](#) Published By: [Manning Publications](#)

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
OE00051	R Programming	3	0	0	3	3

Unit 1 - R basics

Introduction: Basic features of R, advantages of using R, Limitations, R resources, Arithmetic and objects, Math, Variables, and Strings, Vectors and Factors, Vector operations.

Unit 2 - Data structures in R

Data types, Arrays, Tables, Matrices: operations, Lists: operations, Data frames: creation, factors, reading.

Unit 3 - R programming fundamentals

Conditions and loops, Functions in R, Objects and Classes, Recursion, Debugging

Unit 4 - Working with data in R

Reading CSV and Excel Files, reading text files, Writing and saving data objects to file in R, reading in larger, Datasets, Exporting data. Interface to outside world.

Unit 5 – String & Dates in R, Graphics

String operations in R, Regular Expressions, Dates in R, Time in R, Graphics: one dimension plot, legends, function plot, box plot.

Text Books

1. Andrie de Vries, Joris Meys, R Programming for Dummies, Wiley Publications.
2. Roger D. Peng, R Programming for Data Science, Leanpub.

References

1. Emmanuel Paradis, R for Beginners, CRAN Publications.
2. Michael J. Crawley, The R Book, Wiley Publications.

Course Code	Course Name	Hours Per Week			
		L	T	P	Credits
OE00015	Agile Development	3	0	0	3

Unit-I

Understanding Agile: Introduction to Agile Project Management, Agile Manifesto, Agile Principles, Agile Benefits: Product Development and customers, Development teams etc.

Unit-II

Agile Frameworks: Agile approaches, reviewing the big three: Lean, Extreme programming and Scrum. Putting Agile in action: Environment, Behaviors- Agile roles, New values, Team philosophy.

Unit-III

Working in Agile: Planning in Agile, product vision, creating the product roadmap, refining requirement and estimates, release planning and Sprint planning.

Unit-IV

Managing in Agile: Managing Scope and procurement, managing time and cost, team dynamics and communication, managing quality and risk

Unit-V

Ensuring Agile Success: Building a foundation- Commitment, choosing the right project team members-Development team, scrum master etc. Being a change agent, Key benefits and key resources for agile project management.

Text Books:

1. Mark C. Layton, Agile Project Management for Dummies, Wiley publishers
2. Jim Robert Highsmith, Agile Project Management: Creating Innovative Products, Pearson education
3. Hitzler, Markus, Rudolph, Foundations of Semantic Web Technologies, Chapman & Hall/CRC
4. Allemang, Hendler, Semantic Web for the working Ontologist, Elsevier Pub

Reference Books:

1. Charles G. Cobb, Making Sense of Agile Project Management: Balancing Control and Agility, Wiley
2. Mike Cohn, Agile Estimating and Planning, Pearson
3. Liz Sedley and Rachel Davies, Agile Coaching, The Pragmatic Bookshelf

Course Code	Course Name	Hours Per Week				
		L	T	P	Hrs.	Credits
OE00016	Blockchain Architecture	3	0	0	3	3

Unit I: Cryptocurrency: History, electronic cash, double spending problem, Bitcoin protocols, Mining strategy and rewards, Types of crypto currency wallets, Legal aspects of crypto currency, Crypto currency exchanges.

Unit II: Introduction to Blockchain: History of blockchain, Hash functions, SHA-256, Symmetric cryptography, Asymmetric cryptography, Keys & Digital signatures, benefits and limitation of block chain, features of blockchain.

Unit III: Consensus: Nakamoto consensus, Proof of work, Proof of stake, Proof of burn, Difficulty Level, Sybil attack, Energy utilization, collision of energy utilization, Introduction to ethereum.

Unit IV: Blockchain Architectures: Blockchain network, Merkle patricia Tree, Soft & hard fork, Private and public blockchain, Tokenized blockchain.

Unit V: Blockchain Applications: Financial Sector, Medical record management system, domain name service and future of block chain, case study: Government on blockchain. Introduction to hashgraph and tangle.

Text Books:

1. Andreas Antonopoulos “Mastering Bitcoin Unlocking Digital Cryptocurrencies” O’Reilly publication.
2. Imran Bashir “Mastering Blockchain: Distributed ledger technology, decentralization, Packt publishing”.

Reference Books:

1. Wattenhofer, The Science of the Blockchain
2. Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller and Steven Goldfeder, “Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction” Princeton University.

Course Code	Course Name	Hours per Week			Total
		L	T	P	Credits
OE00074	Cloud Web Services	3	0	0	3

UNIT I CLOUD COMPUTING AS A SERVICE

Cloud Computing, Software-as-a-service: SaaS, Platform-as-a-service: PaaS, Hardware-as-a-service: HaaS, Infrastructure-as-a-service: IaaS, Google Cloud Infrastructure, Google File System, Search engine, Map Reduce, Grid Computing, Amazon Web Services, REST APIs, SOAP API, Query API, User Authentication, Connecting to the Cloud, Open SSH Keys, Tunneling / Port Forwarding, Simple Storage Service – S3, Overview,Buckets,Objects,ACL,Logging,Signed URI,S3 Applications, Elastic Cloud Compute - EC2.

UNIT II NETWORKING BASICS

Overview, Keypairs, Network Types, LAN, Gateways and Router, IP Classes and Subnets, CIDR, Utilities, Instances Management, Image Management, Security groups, Amazon Elastic Block Storage - EBS, Ubuntu in the Cloud, Installation, Utilities, File system, Shell.

UNIT III PROGRAMMING AND CONTROL STRUCTURES

Programming, Control Structures, Event based Init Daemon, Apache Instances in EC2, Introduction, Installation and Running, Testing server and content, Configuring Apache, Directives, Virtual hosts, Amazon Simple Queue Service, Amazon Simple Notification Service, Amazon Simple DB, Amazon Relational Database Service, Mysql Server Replication in Cloud, Mysql Database, ▪ Batch mode, Mysql Apache Integration, Storage Engines, Replication Basics, Availability and scalability, Caching, Proxy.

UNIT IV BACKUP AND RECOVERY

Backup and Recovery, Database Sharding, EC2 Applications, Web application design, Focus on Search Engine, Security, Firewall, Data, Network and Host, AWS EC2 Capacity Planning, Apache Servers, Mysql Servers.

UNIT V AMAZON CLOUD

Amazon Cloud Watch, Monitoring Tools, Amazon Cloud Front, Youtube, Amazon Elastic Load Balancing, Cluster Balancing, Amazon Auto Scaling, Apache Scaling, Mysql Scaling, Amazon Virtual Private Cloud, DHCP, DNS, NFS, NIS, Virtualization, Private Cloud for Enterprise, Hybrid Cloud for Enterprise.

TEXT BOOKS:

1. Cloud Computing: Principles and Paradigms, Editors: RajkumarBuyya, James Broberg, Andrzej M. Goscinski, Wiley, 2011
2. Visible Ops Private Cloud: From Virtualization to Private Cloud in 4 Practical Steps, Andi Mann, Kurt Milne, Jeanne Morain, IT Process Institute, Inc.; first edition (April 8, 2011)

REFERENCE BOOK:

1. Cloud Computing Explained: Implementation Handbook for Enterprises, John Roton, Recursive Press (November 2, 2009)

Course Code	Course Name	Hours per Week			Total
		L	T	P	Credits
OE00075	Exploratory Data Analytics	3	0	0	3

Unit I: Introduction to Data and its types

Definition and importance of data, classification of data : based on observation – Cross Sectional, times series and panel data, based on measurement – ratio, interval, ordinal and nominal, based on availability – primary, secondary, tertiary, based on structural form – structured, semi structured and unstructured, based on inherent nature – quantitative and qualitative, concepts on sample data and population, small sample and large sample, statistic and parameter, types of statistics and its application in different business scenarios, frequency distribution of data.

Unit II: Introduction to Exploratory Data Analysis (EDA)

Definition of EDA, difference between EDA with classical and Bayesian Analysis, comparison of EDA with Classical data summary measures, goals of EDA, Underlying assumptions in EDA, importance of EDA in data exploration techniques, introduction to different techniques to test the assumptions involved in EDA, role of graphics in data exploration, introduction to unidimensional, bidimensional and multidimensional graphical representation of data.

Unit III: Data Preparation

Introduction to data exploration process for data preparation, data discovery, issues related with data access, characterization of data, consistency and pollution of data, duplicate or redundant variables, outliers and leverage data, noisy data, missing values, imputation of missing and empty places, with different techniques, missing pattern and its importance, handling non numerical data in missing places.

Unit IV: Univariate Data Analysis

Description and summary of data set, measure of central tendency – mean: Arithmetic, geometric and harmonic mean – Raw and grouped data, confidence limit of mean, median, mode, quartile and percentile, interpretation of quartile and percentile values, measure of dispersion, concepts on error, range, variance, standard deviation, confidence limit of variance and standard deviation, coefficient of variation, mean absolute deviation, mean deviation, quartile deviation, interquartile range, concepts on symmetry of data, skewness and kurtosis, robustness of parameters, measures of concentration.

Unit V: Bivariate Data Analysis

Introduction to bivariate distributions, association between two nominal variables, contingency tables, Chi-Square calculations, Phi Coefficient, scatter plot and its causal interpretations, correlation coefficient, regression coefficient, relationship between two ordinal variables – Spearman Rank correlation, Kendall's Tau Coefficients, measuring association between mixed combination of numerical, ordinal and nominal variables.

Text Books

1. Exploratory Data Analysis – John W Tukey, Addison Wesley Publishing Company
2. Exploratory Data Analysis in Business and Economics - An Introduction Using SPSS, Stata and Excel – Thomas Cleff, Springer Publication

Reference Books

1. Graphical Exploratory Data Analysis - S.H.C. du Toit A.G.W. Steyn R.H. Stumpf, Springer Publication
2. Hand book of Data Visualization – Chun-houh Chen, Wolfgang Härdle, Antony Unwin, Springer Publication.