Data Science and Big Data Analytics

314454 : DATA SCIENCE AND BIG DATA ANALYTICS

CREDITS - 04

UNIT – I INTRODUCTION: DATA SCIENCE AND BIG DATA 08 hours

Introduction to Data science and Big Data, Defining Data science and Big Data, Big Data examples, Data explosion, Data volume, Data Velocity, Big data infrastructure and challenges, Big Data Processing Architectures, Data Warehouse, Re-Engineering the Data Warehouse, Shared everything and shared nothing architecture, Big data learning approaches.

UNIT – II MATHEMATICAL FOUNDATION OF BIG DATA 08 Hours

Probability theory, Tail bounds with applications, Markov chains and random walks, Pair wise independence and universal hashing, Approximate counting, Approximate median, The streaming models, Flajolet Martin Distance sampling, Bloom filters, Local search and testing connectivity, Enforce test techniques, Random walks and testing, Boolean functions, BLR test for linearity.

UNIT - III BIG DATA PROCESSING 08 Hours

Big Data technologies, Introduction to Google file system, Hadoop Architecture, Hadoop Storage: HDFS, Common Hadoop Shell commands, Anatomy of File Write and Read, NameNode, Secondary NameNode, and DataNode, Hadoop MapReduce paradigm, Map Reduce tasks, Job, Task trackers - Cluster Setup – SSH & Hadoop Configuration, Introduction to: NOSQL, Textual ETL processing.

UNIT – IV BIG DATA ANALYTICS 08 Hours

Data analytics life cycle, Data cleaning , Data transformation, Comparing reporting and analysis, Types of analysis, Analytical approaches, Data analytics using R, Exploring basic features of R, Exploring R GUI, Reading data sets, Manipulating and processing data in R, Functions and packages in R, Performing graphical analysis in R, Integrating R and Hadoop, Hive, Data analytics.

UNIT – V Big Data Visualization 08 Hours

Introduction to Data visualization, Challenges to Big data visualization, Conventional data visualization tools, Techniques for visual data representations, Types of data visualization, Visualizing Big Data, Tools used in data visualization, Propriety Data Visualization tools, Open –source data visualization tools, Analytical techniques used in Big data visualization, Data visualization with Tableau, Introduction to: Pentaho, Flare, Jasper Reports, Dygraphs, Datameer Analytics Solution and Cloudera, Platfora, NodeBox, Gephi, Google Chart API, Flot, D3, and Visually.

UNIT – VI BIG DATA TECHNOLOGIES APPLICATION AND IMPACT 08 Hours

Social media analytics, Text mining, Mogile analytics , Roles and responsibilities of Big data person, Organizational impact, Data analytics life cycle, Data Scientist roles and responsibility, Understanding decision theory, creating big data strategy, big data value creation drivers, Michael Porter's valuation creation models, Big data user experience ramifications, Identifying big data use cases.

Text Books
1. Krish Krishnan, Data warehousing in the age of Big Data, Elsevier, ISBN: 9780124058910, 1st Edition.
2. DT Editorial Services, Big Data, Black Book, DT Editorial Services, ISBN: 9789351197577, 2016 Edition.

Reference Books

1. Mitzenmacher and Upfal, Probability and Computing: Randomized Algorithms and Probabilistic Analysis, Cambridge University press, ISBN :521835402 hardback.
2. Dana Ron, Algorithmic and Analysis Techniques in Property Testing, School of EE.
3. Graham Cormode, Minos Garofalakis, Peter J. Haas and Chris Jermaine, Synopses for Massive Data: Samples, Histograms,

Wavelets, Sketches, Foundation and trends in databases, ISBN:10.1561/1900000004.
4. A.Ohri, R for Business Analytics, Springer, ISBN:978-1-4614-4343-8.
5. Alex Holmes, Hadoop in practice, Dreamtech press, ISBN:9781617292224.
6. AmbigaDhiraj, Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business, Wiely CIO Series.
7. Arvind Sathi, Big Data Analytics: Disruptive Technologies for Changing the Game, IBM Corporation, ISBN:978-1-58347-380-1.
8. EMC Education Services, Data Science and Big Data Analytics-Discovering, analyzing Visualizing and Presenting Data.
9. Li Chen, Zhixun Su, Bo Jiang, Mathematical Problems in Data Science, Springer, ISBN :978-3-319-25127-1.
10. Philip Kromer and Russell Jurney, Big Data for chips, O'Reilly, ISBN :9789352132447.
11. EMC Education services, Data Science and Big Data Analytics, EMC2 Wiley, ISBN :978812655653-3.
12. Mueller Massaron, Python for Data science, Wiley, ISBN :9788126557394.
13. EMC Education Services, Data Science and Big Data Analytics, Wiley India, ISBN:9788126556533
14. Benoy Antony, Konstantin Boudnik, Cheryl Adams,,Professional Hadoop, Wiley India, ISBN :9788126563029
15. Mark Gardener, Beginning R: The Statistical Programming Language ,Wiley India, ISBN:9788126541201
16. Mark Gardener, The Essential R Reference ,Wiley India, ISBN : 9788126546015
17. Judith Hurwitz, Alan Nugent, Big Data For Dummies, Wiley India, ISBN : 9788126543281

Computer Network Technology

314450 : COMPUTER NETWORK TECHNOLOGY

CREDITS - 03

UNIT – I NETWORK LAYER 06 Hours
Network Layer Services, IPv4 Addresses: Classful and Classless Addressing, Special Addresses, NAT, Subnetting, Supernetting, Delivery and Forwarding of IP Packet, Structure of Router, IPv4:

Fragmentation, Options, Checksum, ARP: Address Mapping, ARP Protocol, RARP, DHCP, ICMPv4, Unicast Distance Vector Routing, Link State Routing, Unicast Routing Protocols: RIP,EIGRP,OSPF,BGP, IPv6 Addressing.

UNIT – II TRANSPORT LAYER 06 Hours

Transport Layer Services, UDP: Datagram, Services, Applications, TCP: Services, Features, Segment, TCP Connection, Window in TCP, Flow control, Congestion Control, Congestion Control Algorithms, Leaky Bucket, Token Bucket and QoS, TCP Timers, Options, TCP Package, Applications, SCTP: Features, Services, Packet Format,

Socket: TCP and UDP Socket, Applications.

UNIT – III APPLICATION LAYER 06 Hours

Client Server Paradigm: Communication using TCP and UDP, Peer to Peer Paradigm, Application Layer
Protocols: DNS, FTP, TFTP, HTTP, SMTP, POP, IMAP, MIME, Network Management: SNMP.

UNIT – IV WIRELESS STANDARDS 06 Hours

Electromagnetic Spectrum: Spectrum Allocation, Radio Propagation Mechanism, Characteristics of Wireless Channel, Wireless LANs: Architectural Comparison, Characteristics, Access Control,

IEEE 802.11: Architecture, MAC Sub Layer, Addressing Mechanism, Physical Layer, Bluetooth: Architecture, Layers,

IEEE 802.16/WiMax: Services, Architecture, Layers, Differences between Bluetooth, IEEE 802.11 and IEEE 802.16.

UNIT – V ADHOC WIRELESS NEWTORK 06 Hours

Infrastructure Network and Infrastructure-less Wireless Networks, Issues in Adhoc Wireless Network, Adhoc Network MAC Layer: Design Issues, Design Goal, Classification, MACAW, Adhoc Network Routing Layer: Issues in Designing a Routing Protocol for Ad-hoc Wireless Networks – Classifications of Routing Protocols, DSDV, AODV, DSR,

Adhoc Transport Layer: Issues in Designing a Transport Layer

Protocol for Ad hoc Wireless Networks – Design Goals of a Transport Layer Protocol for Ad hoc Wireless Networks –Classification of Transport Layer Solutions, TCP over Adhoc Wireless Networks.

UNIT – VI RECENT TRENDS IN COMMUNICATION NETWORKS 06 Hours

Satellite Network: Operation, GEO Satellites, MEO Satellites, LEO Satellites, Wireless Sensor Network:Functioning, Characteristics, Operation, Cluster Management

Computational Grid: Design, Issues

Internet of Things: Vision, Trends, Significance, Technical Building Blocks, Issues and Challenges, Applications, IoE.
Software Defined Network: SDN Implication for research and innovation, Genesis of SDN, Characteristics of SDN, SDN Operations, SDN Devices, SDN Controllers, SDN Application, OpeFLow Overview, Network Function
Virtualization: Introduction, Applications

Network Neutrality: Need, Requirements (e Reference from research papers and web)

Text Books
1. Behrouz A. Forouzan, TCP/IP Protocol Suite, McGraw Hill Education, ISBN: 978-0-07-070652-1, 4th Edition.
2. C. Siva Ram Murthy, B. S. Manoj, Adhoc Wireless Networks: Architecture and Protocols, Pearson Education, ISBN: 978-81-317-0688-6, 1st Edition.
3. Behrouz A. Forouzan, Data Communication and Networking, McGraw Hill Education, ISBN: 978-1-25-906475-3, 5th Edition.

Reference Books

1. Andrew S. Tanenbaum, David J. Wethrall, Computer Network, Pearson Education, ISBN: 978-0-13-212695-3.
2. Kurose Ross, Computer Networking: A Top Down Approach Featuring the Internet, Pearson Education, ISBN: 978-81-7758-878-1.
3. Charles E. Perkins, Adhoc Networking, Pearson Education, 978-81-317-2096-7.
4. Andrea Goldsmith, Wireless Communication, Cambridge University

Press, ISBN:978-0-521-83716-3.

5. Mayank Dave, Computer Network, Cengage Learning, ISBN: 978-81-315-0986-9.

6. C. K. Toh, Ad Hoc Mobile Wireless Networks Protocols and Systems, Prentice Hall, ISBN: 978-01-324-42046.

7. Paul Goransson, Chuck Black, Software Defined Networks: A Comprehensive Approach, Morgan Kaufmann, ISBN: 978-0124166752.

8. Natalia Olifer, Victor Olifer, Computer Networks: Principles, Technologies and Protocols for Network Design, Wiley India, ISBN: 9788126509171

9. Kazem Sohraby, Daniel Minoli, Taieb Znati, Wireless Sensor Networks: Technology, Protocols and Applications, Wiley India, ISBN: 9788126527304

10. P. Nicopolitidis, M.S. Obaidat, G.I. Papadimitriou, A.S. Pomportsis, Wireless Networks, Wiley India, ISBN : 9788126522200

Design and Analysis of Algorithms

314452 : DESIGN AND ANALYSIS OF ALGORITHMS

CREDITS - 04

UNIT – I INTRODUCTION 08 Hours
Brute Force method: Introduction to Brute Force method & Exhaustive search, Brute Force solution to 8 queens' problem.
Proof Techniques: Minimum 2 examples of each: Contradiction, Mathematical Induction, Direct proofs, Proof by counterexample, Proof by contraposition.
Analysis of Algorithm: Efficiency- Analysis framework, asymptotic notations – big O, theta and omega.
Amortized Analysis: Aggregate, Accounting & Potential method with the example of stack operations. Analysis of Non-recursive and recursive algorithms: Solving Recurrence Equations (Homogeneous and nonhomogeneous).

UNIT – II DIVIDE AND CONQUER AND GREEDYMETHOD 08 Hours
Divide & Conquer: General method, Control abstraction, Merge sort, Quick Sort – Worst, Best and average case. Binary search, Finding Max-Min, Large integer Multiplication (for all above algorithms analysis to be done with recurrence).
Greedy Method: General method and characteristics, Prim's method

for MST , Kruskal's method for MST (using nlogn complexity), Dijkstra's Algorithm, Optimal storage on tapes, Fractional Knapsack problem, Job Sequencing.

UNIT - III DYNAMIC PROGRAMMING 08 Hours
General strategy, Principle of optimality, 0/1 knapsack Problem, Bellman-Ford Algorithm , Multistage Graph problem, Optimal Binary Search Trees, Travelling Salesman Problem.

UNIT – IV BACKTRACKING 08 Hours
General method, Recursive backtracking algorithm, Iterative backtracking method. 8-Queen problem, Sum of subsets, Graph coloring, Hamiltonian Cycle , 0/1 Knapsack Problem.

UNIT – V BRANCH AND BOUND 08 Hours
The method, Control abstractions for Least Cost Search, Bounding, FIFO branch and bound, LC branch and bound, 0/1 Knapsack problem – LC branch and bound and FIFO branch and bound solution, Traveling sales person problem

UNIT – VI COMPUTATIONAL COMPLEXITY AND PARALLEL ALGORITHMS 08 Hours
Computational Complexity: Non Deterministic algorithms, The classes: P, NP, NP Complete, NP Hard, Satisfiability problem, Proofs for NP Complete Problems: Clique, Vertex Cover.
Parallel Algorithms: Introduction, models for parallel computing, computing with complete binary tree, Pointer doubling algorithm.

Text Books
1. Horowitz and Sahani, Fundamentals of computer Algorithms, Galgotia, ISBN 81-7371-612-9.
2. S. Sridhar, Design and Analysis of Algorithms, Oxford, ISBN 10 : 0-19-809369-1.

Reference Books
1. Thomas H Cormen and Charles E.L Leiserson, Introduction to Algorithm, PHI, ISBN:81-203-2141-3.
2. R. C. T. Lee, SS Tseng, R C Chang, Y T Tsai, Introduction to Design and Analysis of Algorithms, A Strategic approach, Tata McGraw Hill,

ISBN-13: 978-1-25-902582-2. ISBN-10: 1-25-902582-9.
3. Anany Levitin, Introduction to the Design & Analysis of Algorithm, Pearson, ISBN 81- 7758-835-4.
4. Steven S Skiena, The Algorithm Design Manual, Springer, ISBN 978-81-8489-865-1.
5. George T. Heineman, Gary Pollice, Stanley Selkow, Algorithms in a Nutshell, A Desktop Quick Reference, O'Reilly, ISBN: 9789352133611.
6. Gilles Brassard, Paul Bratle, Fundamentals of Algorithms, Pearson, ISBN 978-81-317-1244-3.
7. Michael T. Goodrich, Roberto Tamassia, Algorithm Design: Foundations, Analysis and Internet Examples, Wiley India, ISBN: 9788126509867
8. Rod Stephens, Essential Algorithms: A Practical Approach to Computer Algorithms, Wiley India, ISBN: 9788126546138

Cloud Computing

314453 : CLOUD COMPUTING

CREDITS - 03

UNIT – I FUNDAMENTALS OF CLOUD COMPUTING 06 Hours

Origins and Influences, Basic Concepts and Terminology, Goals and Benefits, Risks and Challenges, Roles and Boundaries, Cloud Characteristics, Cloud Delivery Models, Cloud Deployment Models, Federated Cloud/Intercloud, Types of Clouds.
Cloud-Enabling Technology: Broadband Networks and Internet Architecture, Data Center Technology, Virtualization Technology, Web Technology, Multitenant Technology, Service Technology.

UNIT – II VIRTUALIZATION AND COMMON STANDARDS IN CLOUD COMPUTING 06 Hours

Implementation Levels of Virtualization, Virtualization Structures/Tools and Mechanisms, Types of Hypervisors, Virtualization of CPU, Memory, and I/O Devices, Virtual Clusters and Resource Management, Virtualization for Data-Center Automation.
Common Standards: The Open Cloud Consortium, Open Virtualization Format, Standards for Application
Developers: Browsers (Ajax), Data (XML, JSON), Solution Stacks (LAMP and LAPP),Syndication (Atom, Atom Publishing Protocol, and RSS), Standards for Security.

UNIT - III CLOUD PROGRAMMING, ENVIRONMENTS AND APPLICATIONS 06 Hours

Features of Cloud and Grid Platforms, Programming Support of Google App Engine, Programming on Amazon AWS and Microsoft Azure, Emerging Cloud Software Environments, Understanding Core OpenStack Ecosystem.
Applications: Moving application to cloud, Microsoft Cloud Services, Google Cloud Applications, Amazon Cloud Services, Cloud Applications (Social Networking, E-mail, Office Services, Google Apps, Customer Relationship Management).

UNIT –IV CLOUD SECURITY AND ISSUES 06 Hours

Basic Terms and Concepts, Threat Agents, Cloud Security Threats and Attacks, Additional Considerations. Cloud Security Mechanisms: Encryption, Hashing, Digital Signature, Public Key Infrastructure (PKI), Identity and Access Management (IAM), Single Sign-On (SSO), Hardened Virtual Server Images.
Cloud Issues: Stability, Partner Quality, Longevity, Business Continuity, Service-Level Agreements, Agreeing on the Service of Clouds, Solving Problems, Quality of Service, Regulatory Issues and Accountability.

UNIT – V UBIQUITOUS CLOUDS AND THE INTERNET OF THINGS 06 Hours

Cloud Trends in Supporting Ubiquitous Computing, Performance of Distributed Systems and the Cloud, Enabling Technologies for the Internet of Things (RFID, Sensor Networks and ZigBee Technology, GPS), Innovative Applications of the Internet of Things (Smart Buildings and Smart Power Grid, Retailing and Supply-Chain Management, Cyber-Physical System), Online Social and Professional Networking.

UNIT – VI FUTURE OF CLOUD COMPUTING 06 Hours

How the Cloud Will Change Operating Systems, Location-Aware Applications, Intelligent Fabrics, Paints, and More, The Future of Cloud TV, Future of Cloud-Based Smart Devices, Faster Time to Market for Software Applications, Home-Based Cloud Computing, Mobile Cloud, Autonomic Cloud Engine, Multimedia Cloud, Energy Aware Cloud Computing, Jungle Computing.
Docker at a Glance: Process Simplification, Broad Support and Adoption, Architecture, Getting the Most from
Docker, The Docker Workflow.

Text Books

1. Jack J. Dongarra, Kai Hwang, Geoffrey C. Fox, Distributed and Cloud Computing: From Parallel
Processing to the Internet of Things, Elsevier, ISBN :9789381269237, 9381269238, 1st Edition.
2. Thomas Erl, Zaigham Mahmood and Ricardo Puttini, Cloud Computing: Concepts, Technology &
Architecture, Pearson, ISBN :978 9332535923, 9332535922, 1st Edition.

Reference Books

1. Srinivasan, J. Suresh, Cloud Computing: A practical approach for learning and implementation, Pearson, ISBN :9788131776513.
2. Brian J.S. Chee and Curtis Franklin, Jr., Cloud Computing: Technologies and Strategies of the Ubiquitous Data Center, CRC Press, ISBN :9781439806128.
3. Kris Jamsa, Cloud Computing: Saas, Paas, Iaas, Virtualization, Business Models, Mobile, Security, and More, Jones and Bartlett, ISBN :9789380853772.
4. John W. Ritting house, James F. Ransome, Cloud Computing Implementation, Management, and Security, CRC Press, ISBN : 978 1439806807, 1439806802.
5. Karl Matthias, Sean P. Kane, Docker: Up and Running, OReilly, ISBN:9781491917572, 1491917571.
6. Rajkumar Buyya, Christian Vecchiola, S. ThamaraiSelvi, Mastering Cloud Computing: Foundations and Applications Programming, McGraw Hill, ISBN: 978 1259029950, 1259029956.
7. Barrie Sosinsky, Cloud Computing Bible, Wiley, ISBN: 978 8126529803.
8. Gautham Shroff, Enterprise Cloud Computing, Cambridge, ISBN: 9781107648890.
9. Ronald L. Krutz and Russell D. Vines, Cloud Security: A Comprehensive guide to Secure Cloud Computing, Wiley, ISBN: 9788126528097.
10. Scott Adkins, John Belamaric, Vincent Giersch, Denys Makogon, Jason E. Robinson, OpenStack: Cloud Application Development, Wrox, ISBN :9781119194316.
11. Rajkumar Buyya, James Broberg, Andrzej Goscinski, Cloud Computing: Principles and Paradigms, Wiley India, ISBN: 9788126541256
12. Kailash Jayaswal, Jagannath Kallakurchi, Donald J. Houde, Cloud Computing Black Book ,Wiley Dreamtech,ISBN:9789351194187

13. Barrie Sosinsky, Cloud Computing Bible Wiley India, ISBN :9788126529803

Systems Programming

314451 : SYSTEMS PROGRAMMING

CREDITS - 04

UNIT – I INTRODUCTION TO SYSTEMS PROGRAMMING AND ASSEMBLERS 08 Hours

Introduction: Need of System Software, Components of System Software, Language Processing Activities, Fundamentals of Language Processing.
Assemblers: Elements of Assembly Language Programming, A simple Assembly Scheme, Pass structure of Assemblers, Design of Two Pass Assembler, Single pass assembler.

UNIT – II MACROPROCESSORS, LOADERS AND LINKERS 08 Hours

Macro Processor: Macro Definition and call, Macro Expansion, Nested Macro Calls and definition, Advanced Macro Facilities, Design of two-pass Macro Processor.
Loaders: Loader Schemes, Compile and Go, General Loader Scheme, Absolute Loader Scheme, Subroutine Linkages, Relocation and linking concepts, Self-relocating programs, Relocating Loaders, Direct Linking Loaders, Overlay Structure.

UNIT - III INTRODUCTION TO COMPILERS 08 Hours

Phase structure of Compiler and entire compilation process.
Lexical Analyzer: The Role of the Lexical Analyzer, Input Buffering.
Specification of Tokens, Recognition of Tokens, Design of Lexical Analyzer using Uniform Symbol Table, Lexical Errors.
LEX: LEX Specification, Generation of Lexical Analyzer by LEX.

UNIT – IV PARSERS 08 Hours

Role of parsers, Classification of Parsers: Top down parsers- recursive descent parser and predictive parser.Bottom up Parsers – Shift Reduce: SLR, CLR and LALR parsers. Error Detection and Recovery in Parser. YACC
specification and Automatic construction of Parser (YACC).

UNIT – V SEMANTIC ANALYSIS AND STORAGE ALLOCATION 08 Hours

Need, Syntax Directed Translation, Syntax Directed Definitions, Translation of assignment Statements, iterative statements, Boolean expressions, conditional statements, Type Checking and Type conversion.

Intermediate Code Formats: Postfix notation, Parse and syntax tress, Three address code, quadruples and triples.

Storage Allocation: Storage organization and allocation strategies.

UNIT – VI CODE GENERATION AND OPTIMIZATION 08 Hours

Code Generation: Code generation Issues. Basic blocks and flow graphs, A Simple Code Generator.

Code Optimization: Machine Independent: Peephole optimizations: Common Sub-expression elimination, Removing of loop invariants, Induction variables and Reduction in strengths, use of machine idioms, Dynamic Programming Code Generation.

Machine dependent Issues: Assignment and use of registers, Rearrangement of Quadruples for code optimization.

Text Books

1. D. M. Dhamdhere, Systems Programming and Operating Systems, Tata McGraw-Hill, ISBN 13:978-0-07-463579-7, Second Revised Edition.
2. Alfred V. Aho, Ravi Sethi, Jeffrey D. Ullman, Compilers Principles, Techniques and Tools, Addison Wesley, ISBN:981–235–885 - 4, Low Price Edition.
3. J. J. Donovan, Systems Programming, McGraw-Hill, ISBN 13:978-0-07-460482-3, Indian Edition.

Reference Books

1. Leland L. Beck, "System Software An introduction to Systems Programming", Pearson Education, ISBN13: 9788177585551.