Course Code: PCAB301E	Data Mining and Warehousing	Credits : 03
Hours/Week (L:T:P) :3:0:0		CIE Marks : 50
Total Hours: 03		SEE Marks : 50

Corse Objectives:

1. To gain knowledge on data mining and the need for pre-processing.
2. To characterize the kinds of patterns that can be discovered by association rule mining.
3. To get exposed to the concepts of data warehousing architecture and implementation.
4. To implement classification techniques on large datasets.
5. To analyze various clustering techniques in real world applications.

Module-1	10 Hrs.

Introduction and Data Pre-processing :Why data mining, What is data mining, What kinds of data can be mined, What kinds of patterns can be mined, Which Technologies Are used, Which kinds of Applications are targeted, Major issues in data mining .Data Pre-processing: An overview, Data cleaning, Data integration, Data reduction, Data transformation and data discretization.

Module-2	10 Hrs.

Need for Data Warehouse: Database versus Data Warehouse Multidimensional Data Model– Schemas for Multidimensional Databases– OLAP operations– OLAP versus OLTP– Data Warehouse Architecture .

Module-3	10 Hrs.

Association Analysis: Definition of Association Analysis, Frequent Item Set Generation, Rule Generation, Compact Representation of Frequent Item Sets. Alternate Method of Generating Item Sets, FP Growth Algorithms, Evaluation of Association Pattern.

Module-4	10 Hrs.

Classification: Preliminaries, General Approach To Solving Classification Problem, Decision Tree Based Classifier, Rule Based Classifier, Nearest Neighbour Classifier. **Cluster Analysis**– Partitioning Methods: k-Means and k- Mediods– Hierarchical Methods: Agglomerative and Divisive. Density–Based Method: DBSCAN Model Based Clustering Methods Clustering High Dimensional Data Outlier Analysis..

Module-5	10 Hrs.

Data mining trends and research frontiers: Mining complex data types, other methodologies of data mining, Data mining applications, Data Mining and society.

Suggested Learning resources

Text Books:

1. **JiaweiHan, Micheline Kamber, JianPei, DataMining Concepts and Techniques, Morgan Kaufmann**.

Course Outcomes:

CO1: Explain the fundamental concepts, scope, and applications of data mining and the importance of data pre-processing.

CO2: Apply suitable data pre-processing techniques such as cleaning, integration, reduction, and transformation on real-world datasets.

CO3: Describe the architecture and design of data warehouses and differentiate OLTP and OLAP operations.

CO4: Implement and analyze data mining techniques such as association rule mining, classification, and clustering on large datasets.

CO5: Summarize emerging trends, research issues, and societal impacts of data mining applications.

Course Outcomes	Programme Outcomes							
	1	2	3	4	5	6	7	8
CO1	1							
CO2		2						
CO3	1		3					
CO4	1							
CO5	1			4				

PCAB302E	Big Data Analytics	Credits:03
Hrs/Week:L:T:P:S **3:0:0:0**		CIEMarks:50
TotalHours/Week: 03 Hrs		SEEMarks:50

Course Objectives:

1. Understand and define digital data and types, business intelligence, big data and analytics.
2. Appreciate different techniques for big data analytics.
3. Comprehend HDFS architecture and programming environment.
4. Learn NoSQL and able to write statement to process unstructured data.
5. Study programming in Hive and Pig technologies.

MODULE-I	10Hrs.

Types of digital data: Classification of Digital Data, Structured Data, Semi-Structured Data, Unstructured Data. Introduction to Big Data: Characteristics, Evolution, Definitions and Challenges of big data, other characteristics of data which are not definitional traits of big data, Why big data? Are we just an information consumer or do we also produce information? Traditional Business Intelligence (BI) versus Big data, A typical Data Warehouse environment, A typical Hadoop environment, What is changing in the realms of big data? Big data analytics Where do we begin? What is big data analytics? What big data analytics isn't? Why this sudden hype around big data analytics? Classification of analytics, Top challenges facing big data, why is big data analytics important? Greatest challenges that prevent businesses from capitalizing on big data.

MODULE-II	10 Hrs.

What kind of technologies are we looking towards to help meet the challenges posed by big data? Data science, Data Scientist, Terminologies used in big data environment, BASE. The big data technology landscape: NoSQL, Where is it used? What is it? Types of NoSQL databases, Why NoSQL? Advantages of NoSQL, What we miss with NoSQL? NoSQL Vendors, SQL Versus NoSQL , NewSQL, Comparison of SQL, NoSQL and NewSQL, Hadoop: Features of Hadoop, Key advantages of Hadoop,Versions of Hadoop, Overview of Hadoop Ecosystems, Hadoop Versus, SQL, Integrated Hadoop systems offered by leading market vendors, Cloud based Hadoop solutions.

MODULE-III	10Hrs.

Hadoop: Introducing Hadoop, Why not RDBMS, Distributed Computing Challenges, History of Hadoop, Hadoop Overview, Hadoop Components, High Level Architecture of Hadoop, Hadoop Distributed File System(HDFS), HDFS Architecture, Daemons Related to HDFS, Working with HDFS Command, Special Features of Hadoop, Processing Data With Hadoop, Introduction, How Map Reduce Works? Map Reduce Example, Word Count Example using Java. Managing Resources and Applications with YARN, Introduction, Limitation of Hadoop 1.0, Hadoop 2: HDFS, Hadoop 2: YARN, Interacting with Hadoop EcoSystem, Hive,Pig, HBASE, Sqoop, Business Intelligence on Hadoop.

MODULE-IV	10 Hrs.

NoSQL - MongoDB: What is MongoDB? Why MongoDB? Using JSON, Creating or generating a unique key, Support for dynamic queries, Storing binary data, Replication, Sharding, Updating information in-place, Terms used in RDBMS and MongoDB, Data types in MongoDB, MongoDB - CRUD (Insert(), Update(), Save(),

Remove(), find()), MongoDB- Arrays, Java Scripts, Cursors, Map Reduce Programming, Aggregations. NoSQL - Cassandra: What is Cassandra? Why Cassandra? Peer to peer network, Gossip and Failure detection, Anti-Entropy & Read Repair, Writes in Cassandra, Hinted handoffs, Tunable consistency,Cassandra- CQLSH - CRUD, Counter, List, Set, Map, Tracing.

MODULE-V	10 Hrs.

Hadoop Hive: Introduction to Hive - The Problem, Solution - Hive Use Case, Data Growth, Schema Flexibility and Evolution, Extensibility, What is Hive, History of Hive and Recent Releases of Hive, Hive Features, Hive Integration and Work Flow, Hive Data Units, Hive Architecture, Hive Primitive Data Types and Collection Types, Hive File Formats, Hive Query Language - Statements, DDL , DML, Hive Partitions, Bucketing, Views, Sub Query, Joins, Hive User Defined Function, Aggregations in Hive, Group by and Having, Serialization and Deserialization, Hive Analytic Functions. Hadoop - Pig: Introducing Pig, History and Anatomy of Pig, Pig on Hadoop, Pig Features, Pig Philosophy, Word count example using Pig, Use Case for Pig, Pig Primitive Data Types, Collection Types and NULL, Pig Latin Overview, Pig Latin Grammar - Comments, Keywords, Identifiers, Case sensitivity in Pig, Common Operators in Pig, Pig Statements, LOAD, STORE, DUMP, Interactive Shell - GRUNT, FILTER, SORT, GROUP BY, ORDER BY, JOIN, LIMIT, Pig Latin Script, Local Mode, Map Reduce Mode, Running Pig Script, Working with Field, Tuple, Bag, User Defined Function, Parameters in Pig.

Reference Books*

1. Seema Acharya, Subhashini Chellappan, Big Data and Analytics, Wiley Publications, 2nd Edition, 2019, ISBN:978-81-265-5478-2.
2. Raj Kamal, Preethi Saxena, Big Data Analytics, Introduction to Hadoop, Spark and Machine Learning, McGraw hill Education.
3. Cindi Howson, Successful Business Intelligence, McGraw-Hill Publications, E-ISSN:0-07-149851-6.
4. Paul Zikopoulos, Dirk deRoos, Krishnan Parasuraman, Thomas Deutsch , James Giles, David Corrigan, "Harness the Power of Big data – The big data platform", McGraw Hill, 2012.

Course Outcomes:

CO1: Differentiation of digital data and to define big data and analytics.

CO2: Apply different techniques for big data analytics.

CO3: Comprehend HDFS architecture and programming environment.

CO4: Device NoSQL statement to process unstructured data.

CO5: Understanding programming in Hive and Pig technologies.

Course Outcomes	Programme Outcomes							
	1	2	3	4	5	6	7	8
CO1	1							
CO2		2						
CO3	1		3					
CO4	1							
CO5	1			4				

Course Code	Deep Learning Fundamentals	Credits: 03
L:T:P - 3:0:0		CIEMarks:50
Total Hours/Week: 03		SEEMarks:50

Course Learning objectives:

1. Understand the fundamentals of deep learning
2. Understanding the working of Convolution Neural Networks and RNN in decision making.
3. Illustrate the strength and weaknesses of many popular deep learning approaches.
4. Introduce major deep learning algorithms, the problem settings, and their applications to solve real world problems

MODULE-I	10 Hrs.

Introduction: What is a Neural Network? The Human Brain, Models of a Neuron, Neural Networks Viewed As Directed Graphs, Feedback, Network Architectures Rosenblatt's Perceptron: Introduction, Perceptron, The Perceptron Convergence Theorem, Relation Between the Perceptron and Bayes Classifier for a Gaussian Environment.

MODULE –II	10 Hrs.

Multilayer Perceptrons: Introduction, Some Preliminaries, Batch Learning and On-Line Learning, The Back Propagation Algorithm, XOR Problem, Heuristics for Making the Back- Propagation Algorithm Perform Better, Computer Experiment: Pattern Classification, Back Propagation and Differentiation.

MODULE –III	10 Hrs.

Regularization for Deep Learning: Parameter Norm Penalties, Norm Penalties as Constrained Optimization, Regularization and Under Constrained Problem, Dataset Augmentation, Semi-Supervised Learning. Optimization for Training Deep Models: How Learning Differs from pure Optimization, Challenges in Neural Network Optimization, Basic Algorithms, Parameter Initialization Strategies, Algorithms with Adaptive Learning Rate.

MODULE –IV	10 Hrs.

Convolution Networks: The Convolution Operation, Motivation, Pooling, Convolution and Pooling as an Infinitely Strong Prior, Variants of the Basic Convolution Function, Structured Outputs, Data Types, Efficient Convolution Algorithms, Random or Unsupervised Features, The Neuroscientific Basic for Convolutional Network, Convolutional Networks and the History of Deep Learning.

MODULE –V	10 Hrs.

Sequence Modeling: Recurrent and Recursive Nets: Unfolding Computational Graphs, Recurrent Neural Networks, Bidirectional RNNs, Encoder-Decoder Sequence-to- Sequence Architectures, Deep Recurrent Networks, Recursive Neural Networks, The Long Short-Term Memory and Other Gated RNNs.

Text Books

Suggested Learning Resources:

Textbooks

1. Simon Haykin, Neural networks and Learning Machines, Third Edition, Pearson, 2016

2. Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016.

Reference Books

Reference book

1. Bengio, Yoshua. "Learning deep architectures for AI." Foundations and trends in Machine Learning, 2009

2. N.D. Lewis, “Deep Learning Made Easy with R: A Gentle Introduction for Data Science”, January 2016

3. Nikhil Buduma, “Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms”, O’Reilly publications

Course Outcomes

CO1: Understanding Deep Learning Fundamentals

CO2: Design and Implementation of Neural Networks

CO3: Optimization and Performance Tuning

CO4: Application of Advanced Deep Learning Architectures

Course Outcomes	Programme Outcomes							
	1	2	3	4	5	6	7	8
CO1	1							
CO2		2		2				
CO3								
CO4			1,2					3
CO5								

Course Code: PCAC301E	Computer Networks	Credits: 03
L:T:P - 3:0:0		CIEMarks: 50
Total Hours/Week: 3Hrs		SEEMarks: 50

Course Objectives:

1. Understand computer networks fundamentals and issues.
2. Appreciate computer network models and role of layers in each model.
3. Understand different protocols and techniques supported at each level of network software
4. Comprehend application and challenges of computer networks.

MODULE – I — 10 Hrs.

Introduction: Uses of Computer Networks, Network Hardware. **Network Software:** Protocol Hierarchies, Design Issues for the Layers. **Reference Models:** The OSI Reference Model, The TCP/IP Reference Model, A Comparison of the OSI and TCP/IP Reference Models.

Physical Layer- Guided Transmission Media, Digital Modulation and Multiplexing.

MODULE – II — 10 Hrs.

Data Link Layer-Data link Layer Design issues, Framing, Flow Control and Error Correcting and Detection codes, Sliding Window Protocols (Stop and Wait, Go-Back-N (GBN) and Selective Repetitive (SR)), Medium Access Control-The Channel Allocation Problem, Multiple Access Protocols, and Ethernet.

Data Link Layer Switching: Uses of bridges, repeaters, hubs, switches, routers and gateways.

MODULE – III — 10 Hrs.

The Network Layer: Network Layer Design issues, Routing algorithms- The Optimality Principal, Shortest Path Algorithm, Flooding, Distance Vector Routing, Link State Routing, Hierarchical routing, Congestion Control Algorithms, Quality of Service, Internetworking.

MODULE – IV — 10 Hrs.

The Network Layerin the Internet: The Network Layer in the internet- IP version 4 Protocol(IPv4), The Main IPv6 Header, Extension Headers, Internet Control Protocols: ICMP, ARP, DHCP.**The Transport Layer -**The Transport Service: Services Provided to the Upper Layers,Berkeley Sockets, Elements of Transport Protocols, Internet transport protocols.

MODULE – V — 10 Hrs.

TCP: Introduction to TCP, The Service Model, Protocol, Segment Header, UDP.

The Application Layer- The Domain Name System, Electronic Mail, The World-Wide-Web,Streaming Audio and Video.

Reference Books

1. Andrew S. Tanenbaum, David J Wetherall, “Computer Networks”, Pearson Education, Pearson Publication, 5th Edition, 2012.

2. Behrouz A Forouzan, Firouz Mosharraf, “Computer Networks A Top-Down Approach”,

Tata McGraw-Hill Education Pvt. Ltd, 2011.

3. William Stallings, "Data and Computer Communication" ,8th edition,

Pearson Publications,2007.

Course Outcomes

After completion of the course student will be able to

CO1: To comprehend basics of data communication system.

CO2: Enumerate the layers of the OSI, TCP/IP model and demonstrate functions of each layer and comprehend the concept of data link protocols.

CO3: To exhibit the ability to apply different error detection and correction technique to solve communication problem.

CO4: To exhibit the ability to understand issues related to transport layer and protocols.

CO5: Demonstrate the concept of internetworking, routing techniques of network layer.

Course Outcomes	Programme Outcomes							
	1	2	3	4	5	6	7	8
CO1	2							
CO2		1	2					
CO3		1	2					
CO4			2					
CO5			1	2				

Course Code: PCAC304E	Cloud Essentials	Credits: 03
L:T:P - 3:0:0		CIEMarks:50
Total Hours/Week: 03		SEEMarks:50

Course Learning objectives:

1. Explain the fundamentals of cloud.
2. Analyze Business Benefits and Risks of Cloud Computing.
3. Evaluate Emerging Trends in Cloud Computing.

UNIT-I	10 Hrs.

Introduction to Cloud: Defining a cloud, Characteristics of Cloud Computing, Cloud computing reference model, Architectures for parallel and distributed computing, Elements of parallel computing and Elements of distributed computing. Cloud Service Models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Function as a Service (FaaS), Blockchain-as-a-Service (BaaS) and use cases, Cloud Deployment Models: Public Cloud, Private Cloud, Hybrid and Multi-Cloud, Community Cloud..

UNIT–II	10 Hrs.

Core Components of Cloud Architecture: Compute Services, Storage Services, Networking Services, Virtualization, Types of Virtualization, Containers vs. Virtual Machines (VMs), Load Balancing in Cloud, Auto-Scaling & Fault Tolerance, Content Delivery Networks(CDN), Bare Metal Cloud ,Cloud Orchestration and Automation..

UNIT–III	10 Hrs.

Cloud Automation, DevOps, and Future Innovations: Cloud Automation and Infrastructure as Code (IaC), DevOps and Continuous Integration/Continuous Deployment (CI/CD), Multi-Cloud and Hybrid Cloud Strategies, Sustainability and Green Cloud Computing, Cloud Innovations- AI-powered cloud automation-5G and its impact on cloud computing..

UNIT–IV	10 Hrs.

Cloud Security & Risk Management: Cloud Adoption, Advantages and Challenges of Cloud Adoption, Security Risks in Cloud- Data Breaches, Identity Theft, Network Security in Cloud. Security Solutions in Cloud: Identity and Access Management (IAM), Data Encryption Techniques, Firewalls & Intrusion Detection Systems: Compliance & Regulatory Frameworks, Disaster Recovery and Business Continuity Planning in Cloud.

UNIT–V	10 Hrs.

Emerging Trends and case study: AI, Edge Computing, Quantum Cloud, Event-driven architecture in cloud, Cloud-based AI services -AWS SageMaker, Google Vertex AI, Azure ML. Case Study: Netflix's Cloud Migration, Zoom's Cloud Scalability

Text Books

1. CompTIA Cloud Essentials+ Study Guide: Exam CLO-002 [2 ed.] .
2. Rajkumar Buyya, Christian Vecchiola, and Thamrai Selvi Mastering Cloud Computing McGraw Hill Education.
3. Handbook of Cloud Computing, Borko Furht- Armando Escalante.
4. Cloud Essentials: CompTIA Authorized Courseware for Exam CLO-001.
5. RjkumarBuyya, Christian Vecchiola, and ThamaraiSelci, Mastering Cloud Computing, Tata McGraw Hill, New Delhi, India, 2013.

Reference Books

1. Cloud Computing for Dummies by Judith Hurwitz, R.Bloor, M. Kanfman, F.Halper (Wiley India Edition)
2. Toby Velte, Anthony Velte, Cloud Computing: A Practical Approach, McGraw-Hill Osborne Media

Course Outcomes

CO1: Demonstrate the fundamental concepts of cloud computing.

CO2: Understand Cloud Security Challenges

CO3: Understand and Explain Cloud Compute Services and Analyze Cloud Networking Services

CO4: Compare different deployment and service models of cloud to develop different variety of applications with securities

Course Outcomes	Programme Outcomes							
	1	2	3	4	5	6	7	8
CO1	2							
CO2	2	2			2			
CO3	2	2			3			
CO4	2	3		3				3
CO5								

Course Code	Cyber Security	Credits: 03
L:T:P - 3:0:0		CIEMarks:50
Total Hours/Week: 03		SEEMarks:50

Course objectives:

1. To learn cybercrime and cyber law
2. To understand the cyber-attacks and tools for mitigating them.
3. To understand information gathering.
4. To learn how to detect a cyber-attack.
5. To learn how to prevent a cyber-attack.

MODULE-I — 10 Hrs.

INTRODUCTION Cyber Security – History of Internet – Impact of Internet – CIA Triad; Reason for Cyber Crime – Need for Cyber Security – History of Cyber Crime; Cyber criminals – Classification of Cyber crimes– A Global Perspective on Cyber Crimes; Cyber Laws – The Indian IT Act –Cyber crime and Punishment.

MODULE –II — 10 Hrs.

ATTACKSANDCOUNTERMEASURESOSWAP ; Malicious Attack Threats and Vulnerabilities: Scope of Cyber-Attacks – Security Breach – Types of Malicious Attacks – Malicious Software – Common Attack Vectors – Social engineering Attack – Wireless Network Attack – Web Application Attack – Attack Tools – Counter measures.

MODULE –III — 10 Hrs.

RECONNAISSANCE Harvester – Who is Netcraft–Host–Extracting Information from DNS – Extracting Information from E-mail Servers – Social Engineering Reconnaissance; Scanning – Port Scanning – Network Scanning and Vulnerability Scanning – Scanning Methodology – Ping Sweer Techniques – Nmap Command Switches.

MODULE –IV — 10 Hrs.

INTRUSION DETECTION Host -Based Intrusion Detection – Network -Based Intrusion Detection– Distributed or Hybrid Intrusion Detection – IntrusionDetection Exchange Format – Honeypots – Example System Snort.

MODULE –V — 10 Hrs.

INTRUSION PREVENTION: Firewalls and Intrusion Prevention Systems: Need for Firewalls – Firewall Characteristics and Access Policy – Types of Firewalls – Firewall Basing – Firewall Location and Configurations –Intrusion Prevention Systems – Example Unified Threat Management Products.

Text Books

1. PatrickEngebretson,—TheBasicsofHackingandPenetrationTesting:EthicalHackingand Penetration Testing Made easy‖, Elsevier, 2011. (Unit-3)
2. WilliamStallings,LawrieBrown,—ComputerSecurityPrinciplesandPractice‖,ThirdEdition, Pearson Education, 2015. (Unit-4 & 5)

3. AnandShinde, —IntroductiontoCyberSecurityGuidetotheWorldofCyberSecurity‖,Notion Press, 2021. (Unit-1 &2)

Reference Books

1.David Kim, MichaelG. Solomon, —Fundamentals ofInformationSystemsSecurity‖, Jones & Bartlett Learning Publishers, 2013.
2.NinaGodbole,SunitBelapure,—CyberSecurity:UnderstandingCyberCrimes,Computer Forensics and Legal Perspectives‖, Wiley Publishers, 2011.

Course Outcomes

CO1: Explain the basics of cyber security, cybercrime and cyber law

CO2: Classify various types of attacks and learn the tools to launch the attacks

CO3: Apply various tools to perform information gathering

CO4: Apply intrusion techniques to detect intrusion

CO5: Apply intrusion prevention techniques to prevent intrusion

Course Outcomes	Programme Outcomes							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
CO1		1						
CO2		2		3				
CO3	2	2			3			
CO4	2	2			3	3		
CO5		1		2			3	