

**Faculty:ScienceandTechnology****Syllabus Prescribedfor3YearB.Sc. Data Science andAnalytics UG Programme [CBCS]****SEMESTER:III****Programme: B.Sc. Data Science and AnalyticsTitle:****Introduction to Natural Language Processing****Type:AEC****Credits:3**

<b>TotalMarks-50</b>		<b>CourseCode:3DA1</b>		<b>(TotalNumberofPeriods)Hrs</b>	
<b>TheoryExamMarks:40</b>	<b>InternalMarks:10</b>	<b>MinPassing:20</b>		<b>45</b>	

**CourseOutcome(CO):**

Uponcompletionofthiscourse,thestudentsshouldbeableto:

1. Understand the origins, challenges, and applications of natural language processing (NLP), with a focus on Indian languages.
2. Differentiate and implement various grammar-based and statistical language models in practical NLP tasks.
3. Utilize techniques like regular expressions, finite-state automata, and probabilistic parsing for effective text analysis.
4. Apply first-order logic and description logics for semantic representation, along with word sense disambiguation techniques.
5. Explain articulatory and acoustic phonetics concepts and apply digital signal processing methods in speech analysis.

<b>Unit</b>	<b>Content</b>
<b>Unit-1 Overview and Language Modelling (9HRS)</b>	Overview: Origins and challenges of NLP, Language and Grammar, ProcessingIndian Languages, NLP Applications, Information Retrieval. LanguageModelling:VariousGrammarbasedLanguageModels,Statistical LanguageModel.
<b>Unit-2WordLevel And Syntactic Analysis(9HRS)</b>	Word Level Analysis: Regular Expressions, Finite-State Automata,Morphological Parsing, Spelling Error Detection and correction, Words andWord classes, Part-of Speech Tagging. SyntacticAnalysis:Context-freeGrammar,Constituency,Parsing-Probabilistic Parsing.
<b>Unit-3 Semantics and Pragmatics (9HRS)</b>	Requirementsfor representation,First-OrderLogic,DescriptionLogics– Syntax-Driven Semantic analysis, Semanticattachments:WordSenses,RelationsbetweenSenses,Thematic Roles, selectional restrictions – Word Sense Disambiguation, WSD using Supervised,Dictionary&Thesaurus,Bootstrappingmethods–WordSimilarity usingThesaurusandDistributionalmethods.

<b>UNIT-4 Extracting Relations from Text (9HRS)</b>	From Word Sequences to Dependency Paths: Introduction, Subsequence Kernels for Relation Extraction, A Dependency-Path Kernel for Relation Extraction and Experimental Evaluation. Mining Diagnostic Text Reports by Learning to Annotate Knowledge Roles: Introduction, Domain Knowledge and Knowledge Roles, Learning to Annotate Cases with Knowledge Roles and Evaluations.
<b>UNIT-5 Basic Concepts of Speech Processing (9HRS)</b>	Speech Fundamentals: Articulatory Phonetics – Production And Classification Of Speech Sounds, Acoustic Phonetics – Acoustics Of Speech Production, Review Of Digital Signal Processing Concepts, Short-Time Fourier Transform, Filter Bank And LPC Methods.

**Text Book:**

1. “Speech and Language Processing” by Daniel Jurafsky and James H. Martin
2. “Natural Language Processing with Python” by Steven Bird, Ewan Klein, and Edward Loper
3. Nikhil Buduma, Nicholas Locascio, “Fundamentals of Deep Learning: Designing Next Generation Machine Intelligence Algorithms”, O’Reilly Media, 2017.

**Reference Book:**

1. Ian Goodfellow, “Deep Learning”, MIT Press, 2017.
2. Lawrence Rabiner and Bing-Hwang Juang, “Fundamentals Of Speech Recognition”, Pearson Education, 2003.
3. Nitin Indurkha and Fred J. Damerau, —Handbook of Natural Language Processing, Second Edition, Chapman and Hall/CRC Press, 2010.
4. Tanveer Siddiqui, U.S. Tiwary, —Natural Language Processing and Information Retrieval, Oxford University Press, 2008.

**Programme: B.Sc.Data Science and Analytics****Title: Big Data Technology****Type: DSE****Credits: 4**

<b>Total Marks-100</b>		<b>Course Code: 3DA2(E1)</b>		<b>(Total Number of Periods) Hrs</b>	
<b>Theory Exam Marks: 80</b>	<b>Internal Marks: 20</b>	<b>Min Passing: 40</b>	<b>60</b>		

**Course Outcome (CO):**

Upon completion of this course, the students should be able to:

1. Identify the various sources of Big Data
2. Able to collect and store Big Data from various sources
3. Able to write Pig Scripts-Extract, Transform and Load the data on HDFS
4. Able to write Hive Scripts-Extract, Transform, Load and Analyze the data present in HDFS
5. Able to write scripts to extract data from structured and un-structured data for analytics
6. Able to extract and process semi and un-structured data using HBase

Unit	Content
<b>Unit-1Introduction To Big Data Acquisition (12 HRS)</b>	Big data framework, fundamental concepts of Big Data Management and analytics, Current challenges and trends in Big Data Acquisition. Map Reduce Algorithm, Hadoop Storage [HDFS], Common Hadoop Shell commands - Anatomy of File Write and Read, NameNode, Secondary NameNode, and DataNode,HadoopConfiguration,PigConfiguration,HiveConfiguration, HBaseConfiguration.
<b>Unit-2 Data CollectionAnd Transmission(11 HRS)</b>	Big data collection , Strategies , Types of Data Sources , Structured Vs Unstructureddata,ELTvsETL,storageinfrastructurerequirements, Collection methods , Log files , sensors , Methods for acquiring network data (Libcap-based and zero-copy packet capture technology) , Specialized network monitoringsoftwares(Wireshark,SmartsniffandWinnetcap),Mobile equipments,Transmissionmethods,Issues.
<b>Unit-3Apache Pig (12 HRS)</b>	Introduction - Pig features - Pig Architecture - Pig Execution modes, Pig Grunt shellandShellcommands.PigLatinBasics:Datamodel,DataTypes,Operators - Pig Latin Commands - Load & Store, Diagnostic Operators, Grouping,Cogroup, Joining, Filtering, Sorting, Splitting - Built-In Functions, User define functions.PigExecutionModes:BatchMode,EmbeddedMode,Pig ExecutioninBatchMode,Usecases-MapReduceprogramswithPig,PigVs SQL
<b>UNIT-4Hive(15HRS)</b>	Hive Features - Hive architecture -Hive Meta store - Hive data types -56 Hive Tables - Table types - Creating database, Altering database, Create table, alter table, Drop table, Built-In Functions - Built-In Operators, User defined functions(UDFs), View, Pig Vs Hive. HiveQL, Introduction, HiveQL Select, HiveQL , MapReduce using HiveQL OrderBy, Group By Joins, LIMIT,Distribute By, Cluster By - Sorting And Aggregation , Partitioning: Static & Dynamic partitioning , Index Creation - Bucketing , Analysis of MapReduce execution,HiveOptimization,SettingHiiivngParameters.Comparison betweenMapReduce,HiveQL andSQL.
<b>UNIT-5 Hbase (10 HRS)</b>	HBasics, Features of HBase, Concepts, Clients, Example, Hbase VersusRDBMS, Limitations of HBase Big Data Privacy And Applications: Data Masking , Privately identified Information (PII) , Privacy preservation in BigData,PopularBigDataTechniquesandtools,Applications-SocialMedia Analytics,FraudDetection.

### ReferenceBook:

1. Bart Baesens, "Analytics in a Big Data World: The Essential Guide to Data Science and its Applications", John Wiley & Sons, 2014.
2. TomWhite"Hadoop:TheDefinitiveGuide"ThirdEditon,O'reillyMedia,2012.
3. SeemaAcharya,SubhasiniChellappan,"BigDataAnalytics"Wiley2015.
4. Min Chen. Shiwen Mao, Yin Zhang. Victor CM Leung, Big Data: Related Technologies, Challenges and Future Prospects, Springer, 2014.
5. Michael Minelli, Michele Chambers Ambiga Dhiraj, "Big Data, BigAnalytics : Emerging Business Intelligence and Analytic Trends", John Wiley & Sons, 2013.
6. Raj.Pethuru"HandbookofResearchonCloudInfrastructuresforBigDataAnalytics",IGIGlobal.

**Programme:B.Sc. DataScienceandAnalytics****Title:CloudComputing Type:****DSE****Credits:4**

<b>TotalMarks-100</b>		<b>CourseCode:3DA2(E2)</b>		<b>(TotalNumberofPeriods)Hrs</b>	
<b>TheoryExamMarks:80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>60</b>		

**CourseOutcome(CO):**

Uponcompletionofthiscourse,thestudentsshouldbeableto:

1. Understand different computing paradigms and potential of the paradigms and specifically cloud computing
2. Understand cloud service types, cloud deployment models and technologies supporting and driving the cloud
3. Acquire the knowledge of programming models for cloud and development of software application that runs the cloud and various services available from major cloud providers
4. Understandthesecurityconcernsandissuesincloudcomputing
5. Acquiretheknowledgeofadvancesincloudcomputing.

<b>Unit</b>	<b>Content</b>
<b>Unit-1Cloud Computing Fundamentals (10 HRS)</b>	ComputingParadigms,CloudComputingFundamentals,CloudComputing Architecture and Management
<b>Unit-2 Models (14 HRS)</b>	CloudDeploymentModels,CloudServiceModels,TechnologicalDriversfor CloudComputing:SOA andCloud,MulticoreTechnology,Web2.0andWeb 3.0, Pervasive Computing, Operating System, Application Environment
<b>Unit-3Virtualization (12HRS)</b>	Virtualization,ProgrammingModelsforCloudComputing:MapReduce,Cloud Haskell,SoftwareDevelopmentinCloud
<b>UNIT-4Networking for Cloud Computing (12 HRS)</b>	NetworkingforCloudComputing:Introduction,OverviewofDataCenter Environment,NetworkingIssuesinDataCenters,TransportLayerIssuesin DCNs, Cloud Service Providers
<b>UNIT-5Securityin Cloud Computing (12 HRS)</b>	SecurityinCloudComputing,andAdvancedConceptsinCloudComputing

**TextBook:**

1. Chandrasekaran,K.Essentialsofcloudcomputing.CRCPress,2014.

**ReferenceBook:**

1. Cloud Computing: Principles and Paradigms, Editors: Rajkumar Buyya, James Broberg,Andrzej M. Goscinski, Wiley, 2011
2. Enterprise Cloud Computing - Technology,Architecture, Applications, Gautam Shroff, Cambridge University Press, 2010
3. CloudComputingBible,BarrieSosinsky,Wiley-India,2010.

**Programme:B.Sc. DataScienceandAnalytics****Title:DataAnalyticsusing R****Type: Core Skill****Credits:4**

<b>TotalMarks-100</b>		<b>CourseCode:3DA3</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>TheoryExamMarks:80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>60</b>	

**CourseOutcome(CO):**

After successful completion of the course, students will be able to:

1. Understand the basic concepts and environment of R programming.
2. Perform data manipulation and cleaning using R packages.
3. Create and customize various types of data visualizations.
4. Conduct statistical analysis and interpret the results.
5. Utilize advanced R programming techniques including parallel computing and packaged development.
6. Conduct reproducible research and document analysis using R Markdown.

<b>Unit</b>	<b>Content</b>
<b>Unit-1 Introduction to R (14 HRS)</b>	Overview of R and R Studio, Installing R and R Studio, Basic R Syntax, Variables and Data Types, Vectors, Lists, Matrices, Arrays, and Data Frames, Basic Operations in R
<b>Unit-2 Data Manipulation in R (13 HRS)</b>	Importing and Exporting Data, Data Cleaning and Preparation, dplyr and tidyr for Data Manipulation, String Manipulation with stringr, Handling Missing Values
<b>Unit-3 Data Visualization in R (12 HRS)</b>	Introduction to Data Visualization, Basic Plotting with Base R, Advanced Plotting with ggplot2, Customizing Plots, Creating Interactive Visualizations with plotly and shiny
<b>UNIT-4 Statistical Analysis in R (11 HRS)</b>	Descriptive Statistics, Probability Distributions, Hypothesis Testing, ANOVA and Regression Analysis, Time Series Analysis.
<b>UNIT-5 Advanced Topics in R (10 HRS)</b>	Introduction to Parallel Computing in R, Writing Functions and Scripts, Package Development in R, Version Control with Git and GitHub, Reproducible Research with R Markdown

**Text Book:**

1. "R for Data Science" by Hadley Wickham and Garrett Grolemund ISBN-13: 978-1491910399

**Reference Book:**

1. "Advanced R" by Hadley Wickham ISBN-13: 978-1498759809
2. "The Art of R Programming" by Norman Matloff ISBN-13: 978-1593273842
3. "R in Action" by Robert I. Kabacoff ISBN-13: 978-1617291388

**Programme:B.Sc. DataScienceandAnalytics****Title:BusinessAnalytics****Type: Core Skill****Credits: 3**

<b>TotalMarks-100</b>		<b>CourseCode:3DA4</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>TheoryExamMarks:80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>45</b>	

**CourseOutcome(CO):**

After successful completion of the course, students will be able to:

1. Analyze and visualize data in different industries such as manufacturing, service, retail, software, banking and finance, sports etc.
2. Use technical skills in descriptive modeling to support business decision-making.
3. Use technical skills in predictive modeling to support business decision-making.
4. Use technical skills in prescriptive modeling to support business decision-making.
5. Demonstrated decision making with and without Risk for solving problems in different industries.

<b>Unit</b>	<b>Content</b>
<b>Unit-1 Overview of Business analytics(10 HRS)</b>	Overview of Business analytics, Scope of Business analytics, Business Analytics Process, Relationship of Business Analytics Process and organization, competitive advantages of Business Analytics. Organization Structures of Business analytics, Team management, Management Issues, Designing Information Policy, Outsourcing, Ensuring Data Quality, Measuring contribution of Business analytics, Managing Changes.
<b>Unit-2 Descriptive Analytics(8HRS)</b>	Statistical Tools: Statistical Notation, Descriptive Statistical methods, Review of probability distribution and data modeling, sampling and estimation methods overview. Important Resources, Business Analytics Personnel, Data and models for Business analytics, problem solving, Visualizing and Exploring Data, Business Analytics Technology.
<b>Unit-3 Predictive Analytics(8 HRS)</b>	Modeling Relationships and Trends in Data, Simple Linear Regression, Residual Analysis and Regression Assumptions, Multiple Linear Regression, Building Good Regression Models, Forecasting Methods: Forecasting Techniques: Qualitative and Judgmental Forecasting, Statistical Forecasting Models, Forecasting Models for Stationary Time Series, Selecting Appropriate Forecasting Models.
<b>UNIT-4 Prescriptive Analytics (10 HRS)</b>	Linear Optimization : Optimization Models, Linear Optimization in Bank Financial Planning, Analytics in Practice: Using Optimization Models for Sales Planning at NBC, Developing Linear Optimization Models, Identifying Decision Variables, the Objective, and Constraints, Developing a Mathematical Model, Implementing Linear Optimization Models, Solving Linear Optimization Models
<b>UNIT-5 Simulation and Risk Analysis (9 HRS)</b>	Model-Building Strategies: Building Models Using Logic and Business Principles, Building Models Using Influence Diagrams, Building Models Using Historical Data. Introduction of simulation and Risk Analysis, Types of simulations, Risk Management, Risk Assessment, Impact Analysis, Monte

CarloSimulationUsingAnalyticSolverPlatform,New-ProductDevelopment Model,NewsvendorModel,OverbookingModel,CashBudgetModel.
---

**TextBook:**

1. Business analytics Principles, Concepts, andApplications by Marc J. Schniederjans, Dara G. Schniederjans, Christopher M. Starkey, Pearson FT Press.
2. BusinessAnalytics:Methods,ModelsandDecisionsbyJamesEvans,personsEducation,3rdEdition

**ReferenceBook:**

1. R. Sharda, D. Delen, and E. Turban, Business Intelligence and Analytics. Systems for Decision Support,10th Edition. Pearson/Prentice Hall, 2015. ISBN-13: 978-0-13-305090-5, ISBN-10: 0-13-305090-4;
2. Carlo Vercellis, “Business Intelligence - Data Mining and Optimization for Decision Making”, Wiley Publications, ISBN: 9780470753866
3. Business Analytics A Practitioner’s Guide by Rahul Saxena, AnandSrinivasanInternational Series in Operations Research & Management Science, Springer New York, December 5, 2012,ISBN: 9781461460800.

**Programme: B.Sc.Data Science andAnalytics****Title: R Programming- LAB****Type:SEC/LAB****Credits: 2**

<b>TotalMarks-50</b>		<b>CourseCode:3DA5</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>ExternalMarks:25</b>	<b>InternalMarks:25</b>	<b>MinPassing:20</b>	<b>60</b>	

**ListofPractical’s:**

**NOTE:** The listsuggestssampleprogramset.Hence,theconcernedstaffmay modify thelistasneeded (Minimum 15).

1. WriteaRprogramtocretealistcontainingstrings,numbers,vectorsandalogicalvalues.
2. WriteaRprogramtolistcontainingavector,amatrixandalistandgivenamestotheelements in the list.
3. WriteaRprogramtocretealistcontainingavector,amatrixandalistandgivenamestothe elements in the list. Access the first and second element of the list.
4. WriteaRprogramtocretealistcontainingavector,amatrixandalistandaddelementattheend of the list.
5. WriteaRprogramtoselectsecondelementofagivennestedlist.
6. WriteaRprogramtocretealistcontainingavector,amatrixandalistandremovethesecond element.
7. WriteaRprogramtocretealistcontainingavector,amatrixandalistandupdatethelastelement.

**Programme: B.Sc.Data Science andAnalytics****Title: Business Analytics - LAB****Type:SEC/LAB****Credits: 2**

<b>TotalMarks-50</b>		<b>CourseCode:3DA6</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>ExternalMarks:25</b>	<b>InternalMarks:25</b>	<b>MinPassing:20</b>	<b>60</b>	

**ListofPractical's:**

**NOTE:** The listsuggestssampleprogramset.Hence,theconcernedstaffmay modify thelistasneeded (Minimum 15).

1. ImporttheLegacydatafromdifferentsourcesandloadinthetargetsystem
2. PerformExtractTransformandLoading(ETL)processtoconstructthedatabaseinSQLServ
3. ExecutetheMDX(MultiDimensionalExpression)queriestoExtractthedatafromDatawarehouse.
4. Apply the What-If Analysis for data visualization .Design and generate necessary reports based on Data warehouse.
5. PerformTimeSeriesAnalysis.
6. PerformLinearRegressiononthegivenDataWarehouseData.
7. PerformtheDataClusteringusingClusteringAlgorithm.

**Programme: B.Sc.Data Science andAnalytics****Title: LAB basedon 3DA2****Type:SEC/LAB****Credits: 2**

<b>TotalMarks-50</b>		<b>CourseCode:3DA7</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>ExternalMarks:25</b>	<b>InternalMarks:25</b>	<b>MinPassing:20</b>	<b>60</b>	

Minimum 15 experiments / programming assignments must be completed based on the respective syllabus (3CS2E1/ 3CS2 E2).

**Programme: B.Sc.CyberSecurity Title:**

**Environment Studies**

**Type:VEC[CollegeLevelTheory]**

**Credits: 2**

<b>TotalMarks-50</b>	<b>CourseCode:3DA8</b>	<b>(TotalNumberofPeriods)Hrs</b>
<b>InternalMarks:50</b>	<b>MinPassing:20</b>	<b>45</b>

**UnitI:(a)TheMultidisciplinarynatureofenvironmentalstudies:**

Definition,Principles,Scopeandimportance,ManandEnvironment,Needforpublicawareness.

(b)NaturalResources:Renewableandnonrenewableresources:

Availability, use, overexploitation and associated environmental problems related to following Natural resource:

- Forestresources:
- Waterresources:.
- MineralResources:
- FoodResources:
- EnergyResources:
- LandResources:
- Roleofindividualinconservationofnaturalresources'

**UnitII:Ecosystems:**

- Conceptandcomponentsofanecosystem.
- Typesofecosystem
- Structureandfunctionofforestandpondecosystem.
- Energyflowintheecosystems.
- Foodchains,foodwebsandecologicalpyramids.
- Ecologicalsuccession:Generalmechanism

**UNITIII:Biodiversityandit'sConservation:**

- Introduction,definitionandtypesofbiodiversity.
- Bio-geographicalclassificationofIndia.
- Indiaasamega-diversitynation.
- Hot-spotsofbiodiversity.
- Threatstobiodiversity:habitatsloss,poachingofwildlife,manwildlifeconflicts.

- Endangered and endemic species of India.
- Conservation of biodiversity: in-situ and ex-situ conservation of biodiversity.

**UNIT IV: Environmental Pollution:**

- Definition, Causes, effects and control measures of: a. Air pollution b. Water pollution c. Soil pollution d. Noise pollution e. Nuclear hazards.
- Solid waste Management: Principles, methods and significance
- Disaster management: Floods, earth quake, cyclone and landslides.

**Unit V: Social issues and the Environment:**

- From unsustainable to sustainable development
- Urban problems related to energy
- Water conservation: rainwater harvesting, watershed management
- Environmental ethics: issues and possible solutions
- Climate change, global warming, acid rain, ozone layer depletion and nuclear accidents
- Wasteland reclamation
- Environmental Legislation: Environment protection Act (1986); Air (prevention and control of pollution) Act (1981-82); Water (prevention and control of pollution) Act (1974); Wildlife protection act (1972); Forest conservation act (1980), Issues involved in enforcement of environmental legislation

**Recommended Books:**

1. Text Book of Environmental Studies, Erach Bharucha, UGC.
2. Fundamental concepts in Environmental Studies, D.D. Mishra, S. Chand & Co Ltd.
3. Ecology and Environment, P.D. Sharma.
4. Ecology, M.P. Arora, Himalaya Publishing House.

**Faculty:ScienceandTechnology****Syllabus Prescribedfor3YearB.Sc. Data ScienceandAnalytics UG Programme [CBCS]****SEMESTER:IV****Programme: B.Sc.Data Science andAnalytics****Title: Cloud Services for IOT****Type:AEC****Credits: 3**

<b>TotalMarks-100</b>		<b>CourseCode:4DA1</b>		<b>(TotalNumberofPeriods)Hrs</b>	
<b>ExternalMarks(T):80</b>		<b>InternalMarks:20</b>		<b>MinPassing:40</b>	
				<b>45</b>	

**CourseOutcome(CO):**

Uponcompletionofthiscourse,thestudentsshouldbeableto:

1. BuildanappropriatecloudarchitectureandidentifythecLOUDservices
2. Handlevarioussensorsandthetechnologies
3. DevelopIoTapplicationsusingcloudplatforms
4. IntegrateIoTapplicationsintothecloudservices
5. Accessthesecurityissuesinapplicationsandnetworks

<b>Unit</b>	<b>Content</b>
<b>Unit 1 CloudPlatformArchite cture and Services (9 HRS)</b>	Cloud computing and service models: Public, Private and Hybrid clouds, Infrastructure as a service(IaaS), Platform as a service(PaaS), Software as a service(SaaS) Architectural design of compute and storage clouds: Layered cloud architectural development, Architectural design challenges, Publiccloudplatforms:GAE,AWSandAzure
<b>Unit-2 Programming IOT Devices for Cloud Interface (9HRS)</b>	Basics of Sensors and actuators, examples and working principles of sensors and actuators, Cloud computing and IOT, Arduino/Equivalent Microcontroller platform. IoTCommunicationTechnologies,RFID,Bluetooth,Zigbee,Wifi,Wired Communication
<b>Unit-3 Cloud PlatformsforIOT (9HRS)</b>	Thinkspeak IoT Cloud Platform, Kaa Open Source Iot Cloud Platform, AWS IoT Cloud Platform,AWS IoT Device SDK,ArduinoAWS IoT development, RaspberryPi3,AWSIoTdevelopment
<b>UNIT-4Cloud Services forIOT (9HRS)</b>	ServiceManagementinCloudComputing,ServiceLevelAgreements(SLAs), Managing IoT Data, Looking at Data, Scalability & Cloud Services, Database & Data Stores in Cloud, Large Scale Data Processing.
<b>UNIT-5 Security and Applications(9HRS)</b>	Application Safety and Service Vulnerability in Cloud Network, IoT Securityand Privacy Preservation, Security and Challenges in Mobile Cloud Computing, ThevitalroleofFogcomputinginInternetofThings

**TextBook:**

1. Kai Hwang, Geoffrey C Fox, Jack G Dongarra, “Distributed and Cloud Computing, From parallel processing to the Internet of Things”, Morgan Kaufmann Publishers, 2012.
2. Raj Kamal, “Internet of Things:Architecture and Design Principles”, McGraw-Hill Education Pvt. Ltd., 2018.
3. CharalamposDoukas,“BuildingInternetofThingswiththeArduino”,CreateSpace, April2002.
4. AgusKurniawan“LearningAWSIoT”PacktPublishing(January29,2018)

**ReferenceBook:**

1. Dac-NhuongLe,ChintanBhatt,ManiMadhukar“SecurityDesignsfortheCloud,IoT,andSocialNetworking” John Wiley & Sons (11 October 2019).
2. HonboZhou,“TheInternetofThingsintheCloud:AMiddlewarePerspective”,CRCPress,2013.
3. MarcoSchwartz,“InternetofThingswithArduinoCookbook”,PacktPublications,2016.
4. RajkumarBuyya,ChristianVecchiola.S.ThamaraiSelvi, “Mastering Cloud Computing”,McGrawHill Education, 2013.
5. NickAntonopoulos and Lee Gillam, “Cloud Computing: Principles, Systems andApplications”, Second Edition, Springer, 2017.

**Programme: B.Sc.Data Science andAnalytics****Title: Machine Learning****Type:CoreSkill Credits:**

3

<b>TotalMarks-100</b>		<b>CourseCode:4DA2</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>ExternalMarks(T):80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>45</b>	

**CourseOutcome(CO):**

Uponcompletionofthiscourse,thestudentsshouldbeableto:

1. Identify various machine learning algorithms and terminologies and perform data pre-processing using standard ML library.
2. Designa predictive model using appropriate supervised learning algorithms to solve any given problem.
3. Develop an application using appropriate unsupervised learning algorithms for performing clustering and dimensionality reduction.
4. Solvecomplexproblemsusingartificialneuralnetworksandkernelmachines.
5. Implementprobabilisticgraphicalmodelsforsuitableapplications

<b>Unit</b>	<b>Content</b>
<b>UnitIIntroduction ToArtificialNeural Networks (9 Hrs)</b>	FundamentalsOfNeuralNetworks,ModelofArtificialNeuron,Neural Network Architectures, Learning Methods, Taxonomy Of Neural Network Architectures,Applications

<b>Unit-2 Feed Forward Neural Networks (9hrs)</b>	Perceptron Models: Discrete, Continuous and Multi, Category –Training Algorithms: Discrete and Continuous Perceptron Networks, Limitations of the Perceptron Model. Credit Assignment Problem, Generalized Delta Rule, Derivation of Back propagation (BP) Training, and Summary of Back propagationAlgorithm,KolmogorovTheorem
<b>Unit-3 Machine Learning(9hrs)</b>	Machine Learning Fundamentals –Types of Machine Learning , Supervised, Unsupervised, Reinforcement, The Machine Learning process, Terminologies in ML. Testing ML algorithms: Overfitting, Training, Testing and Validation Sets, Confusion matrix ,Accuracy metrics, ROC Curve,Basic Statistics: Averages, Variance and Covariance, The Gaussian, The Bias, Variance trade off, ApplicationsofMachineLearning.
<b>Unit-4 Supervised Learning (9hrs)</b>	Regression:LinearRegression,MultivariateRegression, Classification:LinearDiscriminant Analysis,LogisticRegression,K,Nearest Neighbor classifier. Decision Tree based methods for classification and Regression, Ensemble methods.
<b>Unit-5 Unsupervised Learning (9hrs)</b>	Clustering, K, Means clustering, Hierarchical clustering , The Curse of Dimensionality , Dimensionality Reduction , Principal ComponentAnalysis , ProbabilisticPCA,IndependentComponentsanalysis

**TextBook:**

1. CharuC.Aggarwal“NeuralNetworksandDeeplearning”SpringerInternationalPublishing,2018
2. SatishKumar,“NeuralNetworks,AClassroomApproach”,TataMcGraw-Hill,2007.
3. KevinP.Murphy,“MachineLearning:AProbabilisticPerspective”,MITPress,2012.
4. StephenMarsland,“MachineLearning–AnAlgorithmicPerspective”,CRCPress,2009.
5. Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das, “Machine Learning”, Pearson Education, 2018.
6. ChristopherBishop,“PatternRecognitionandMachineLearning”Springer,2011.

**ReferenceBook:**

1. Andreas C. Muller, “Introduction to Machine Learning with Python:AGuide for Data Scientists”, O’Reilly,2016.
2. SebastianRaschka,“PythonMachineLearning”,PacktPublishing,2015.
3. Hastie, Tibshirani, Friedman, “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”,2nd Edition, Springer, 2017.
4. EthemAlpaydin,“IntroductiontoMachineLearning”,2ndRevisededition,MITPress,2010

**Programme: B.Sc.Data Science andAnalytics****Title: Computer Vision****Type:DSE****Credits: 3**

<b>TotalMarks-100</b>		<b>CourseCode:4DA3(E1)</b>		<b>(TotalNumberofPeriods)Hrs</b>	
<b>ExternalMarks(T):80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>45</b>		

**CourseOutcome(CO):**

Upon completion of this course, the student should be able to:

1. Acquire knowledge on Image retrieval and processing.
2. Design and deploy various image formation models.
3. Apply the techniques of motion estimation and object recognition.

Unit	Content
<b>Unit 1 Introduction (9Hrs)</b>	Image Processing, Computer Vision: Definition and Overview, computer vision applications: Document Image Analysis, Biometrics, Object Recognition, Tracking, Medical Image Analysis, Content Based Image Retrieval, Video data Processing, Multimedia, Virtual Reality and Augmented Reality
<b>Unit-2 Image Formation Models (9hrs)</b>	Monocular Imaging System, Camera Model and Camera calibration, Binocular Imaging System, Multiple views geometry, Structure Determination, Shape from Shading-Construction of 3D model from images. Image Processing and Feature Extraction-Image representation, Edge detection
<b>Unit-3 Motion Estimation (9hrs)</b>	Optical Computation, Structure from motion. Shape Representation and Segmentation Contour based representation, Region based representation, Deformable curves and surfaces, multi-resolution analysis
<b>Unit-4 Object Recognition (9hrs)</b>	Hough transforms and other simple object recognition methods, shape correspondence and shape matching, Principal component analysis, Shape priors for recognition. Image Understanding-Pattern recognition methods- HMM, GMM and EM
<b>Unit-5 Applications (9hrs)</b>	Photo album-face detection, Face recognition, Eigen faces, Surveillance-foreground background separation, particle filters, chamfer matching, tracking-occlusion combining views from multiple cameras-locating roadway-road markings-identifying road signs-locating pedestrians

**Text Book:**

1. Computer Vision – A modern approach, by D. Forsyth and J. Ponce Prentice Hall Robot Vision, by B.K.P. Horn. McGraw-Hill
2. Introductory Techniques by 3D Computer Vision, by E. Trucco and A. Verri, Publisher: Prentice Hall.
3. R.C. Gonzalez, R.E. Woods. Digital Image Processing. Addison Wesley Longman, Inc., 1992.
4. Richard Szeliski, Computer Vision: Algorithms and Applications. Springer, 2010.

**Reference Book:**

1. R.H. Ballard, C.M. Brown, Computer Vision, Prentice Hall, Englewood Cliffs, 1982.

**Programme: B.Sc. Data Science and Analytics**

**Title: Computer Networks**

**Type: DSE**

**Credits: 3**

<b>Total Marks-100</b>	<b>Course Code: 4DA3(E2)</b>	<b>(Total Number of Periods) Hrs</b>
------------------------	------------------------------	--------------------------------------

**CourseOutcome(CO):**

Uponcompletionofthiscourse,thestudentsshouldbeableto:

<b>ExternalMarks(T):80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>45</b>
----------------------------	-------------------------	----------------------	-----------

**CourseOutcome(CO):**

Upon completion of this course, the student should be able to:

1. Understand the fundamental concepts of Computer Networks and their applications.
2. Develop problem-solving skills related to network design, implementation, and troubleshooting.
3. Implement network protocols and configure network devices.

Unit	Content
<b>Unit 1 Introduction to Computer Networks (9Hrs)</b>	Overview of Computer Networks, Fundamental Concepts of Network, Network Topologies, Network models. Networking Devices: Routers, Switches, Hubs, Bridges, Gateways, Functions and Configuration of Each Device.
<b>Unit-2 Data Communication Fundamentals and Techniques(9hrs)</b>	Analog and Digital Signal, Data-Rate Limits, Digital to Digital Line Encoding Schemes, Pulse Code Modulation, Parallel and Serial Transmission, Digital to Analog Modulation - Multiplexing Techniques- FDM, TDM, Transmission Media. Network Switching, Dial-Up Modems, Digital Subscriber Line.
<b>Unit-3 : Data Link Layer and Networking Protocols (9hrs)</b>	Data Link Layer Fundamentals- Functions, Framing, Error Detection, and Error Correction, Flow Control Mechanisms. Ethernet- Standards and Frame Structure, MAC Addressing and ARP, Ethernet Switching Network Protocols: Introduction to TCP/IP Protocol Suite, IP Addressing: IPv4 and IPv6 Subnetting, CIDR Notation Address Resolution Protocol (ARP): ARP Operation and Table, ARP Spoofing and Security Considerations, Virtual LANs (VLANs).
<b>Unit-4 Network Layer and Transport Layer (9hrs)</b>	Network Layer: IP Routing, Routing Protocols: RIP, OSPF, BGP, Network Address Translation (NAT) Transport Layer: TCP vs. UDP, TCP Handshake and Connection Management, Flow Control and Congestion Control in TCP, Congestion Control Algorithms: Slow Start, Congestion Avoidance, Fast Retransmit, Fast Recovery, TCP Variants: TCP Reno, TCP Vegas, QoS Principles and Mechanisms, Differentiated Services (DiffServ) and Integrated Services (IntServ) Network Security Fundamentals: Threats and Vulnerabilities, Basic Security Mechanisms: Firewalls, VPNs, Encryption
<b>Unit-5 Application Layer and Emerging Technologies Applications(9hrs)</b>	Application Layer Protocols: HTTP/HTTPS: Structure and Operation, FTP, SMTP, POP3, IMAP: Protocols and Uses, DNS: Domain Name System and Resolution Network Applications: Web Browsing, Email Communication, File Transfer, Voice over IP (VoIP) and Streaming. Emerging Technologies: Software-Defined Networking (SDN), Network Function Virtualization (NFV), Internet of Things (IoT) and Its Impact on Networking Network Management: SNMP: Simple Network Management Protocol, Network Monitoring Tools and Techniques. Future Trends in Networking: 5G and Beyond, Network Automation and Artificial Intelligence in Networking.

**TextBook:**

1. Andrew S. Tanenbaum, "Computer Networks", 5th Edition, Pearson Education, 2011.
2. James F. Kurose and Keith W. Ross, "Computer Networking: A Top-Down Approach", 8th Edition, Pearson, 2021.

**ReferenceBook:**

1. Behrouz A. Forouzan, "Data Communications and Networking", 5th Edition, McGraw-Hill Education, 2012.
2. Larry L. Peterson and Bruce S. Davie, "Computer Networks: A Systems Approach", 6th Edition, Morgan Kaufmann, 2019.
3. Bhavneet Sidhu, An Integrated Approach to Computer Networks, Khanna Publishing House, 2023.
4. Mastering PC Hardware & Networking, Khanna Publishing House, 2024.

**Programme: B.Sc.Data Science and Analytics****Title: Data Handling and Visualisation****Type: Core Skill Credits:**

3

<b>Total Marks-100</b>		<b>Course Code:4DA4</b>		<b>(Total Number of Periods)Hrs</b>	
<b>External Marks(T):80</b>	<b>Internal Marks:20</b>	<b>Min Passing:40</b>	<b>45</b>		

**Course Outcome(CO):**

Upon completion of this course, the students should be able to:

1. Understand basics of Data Visualization.
2. Implement visualization of distributions.
3. Write programs on visualization of time series, proportions & associations.
4. Apply visualization on Trends and uncertainty.
5. Explain principles of proportions.

<b>Unit</b>	<b>Content</b>
<b>Unit-1 Introduction To Visualization (9Hrs)</b>	Visualizing Data-Mapping Data onto Aesthetics, Aesthetics and Types of Data, Scales Map Data Values onto Aesthetics, Coordinate Systems and Axes- Cartesian Coordinates, Nonlinear Axes, Coordinate Systems with Curved Axes, Color Scales- Color as a Tool to Distinguish, Color to Represent Data Values ,Color as a Tool to Highlight, Directory of Visualizations- Amounts, Distributions, Proportions, x-y relationships, Geospatial Data
<b>Unit-2 Visualizing Distributions (9hrs)</b>	Visualizing Amounts-Bar Plots, Grouped and Stacked Bars, Dot Plots and Heatmaps, Visualizing Distributions: Histograms and Density Plots- Visualizing a Single Distribution, Visualizing Multiple Distributions at the Same Time, Visualizing Distributions: Empirical Cumulative Distribution Functions and Q-Q Plots-Empirical Cumulative Distribution Functions, Highly Skewed Distributions, Quantile- Quantile Plots, Visualizing Many Distributions at Once-Visualizing Distributions Along the Vertical Axis, Visualizing Distributions Along the Horizontal Axis
<b>Unit-3 Visualizing Associations &amp; Time Series (9hrs)</b>	Visualizing Proportions-A Case for Pie Charts, A Case for Side-by-Side Bars, A Case for Stacked Bars and Stacked Densities, Visualizing Proportions Separately as Parts of the Total , Visualizing Nested Proportions- Nested Proportions Gone Wrong, Mosaic Plots and Treemaps, Nested Pies ,Parallel Sets. Visualizing Associations Among Two or More Quantitative Variables- Scatterplots, Correlograms, Dimension

	Reduction, Paired Data. Visualizing Time Series and Other Functions of an Independent Variable-Individual Time Series , Multiple Time Series and Dose–ResponseCurves,TimeSeriesofTwoorMoreResponseVariables
<b>Unit-4 Visualizing Uncertainty(9hrs)</b>	Visualizing Trends-Smoothing, Showing Trends with a Defined Functional Form, Detrending and Time-Series Decomposition, Visualizing Geospatial Data-Projections, Layers, Choropleth Mapping, Cartograms, Visualizing Uncertainty-Framing Probabilities as Frequencies, Visualizing the Uncertainty of Point Estimates,VisualizingtheUncertaintyofCurveFits,HypotheticalOutcomePlots
<b>Unit-5 Principle Of Proportional Ink (9hrs)</b>	The Principle of Proportional Ink-Visualizations Along Linear Axes, Visualizations Along LogarithmicAxes, Direct Area Visualizations, Handling Overlapping Points-Partial Transparency and Jittering, 2D Histograms, Contour Lines, Common Pitfalls of Color Use Encoding Too Much or Irrelevant Information ,Using Nonmonotonic ColorScalestoEncodeDataValues,NotDesigningforColor-VisionDeficiency.

**TextBook:**

1. Claus Wilke, “Fundamentals of Data Visualization:A Primer on Making Informative and Compelling Figures”, 1st edition, O’Reilly Media Inc, 2019.
2. James F. Kurose and Keith W. Ross, "Computer Networking:ATop-DownApproach", 8th Edition, Pearson, 2021.

**ReferenceBook:**

1. .TonyFischetti,BrettLantz,R:DataAnalysisandVisualization,O’Reilly,2016.
2. Ossama Embarak, Data Analysis and Visualization Using Python:Analyze Data to Create Visualizations for BI Systems,Apress, 2018.

**Programme: B.Sc.Data Science andAnalytics****Title: Operating System****Type:CoreSkill Credits:**

3

<b>TotalMarks-100</b>	<b>CourseCode:4DA7</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>ExternalMarks(T):80</b>	<b>InternalMarks:20</b>	<b>MinPassing:40</b>	<b>45</b>

**CourseOutcome(CO):**

Uponcompletionofthiscourse,thestudentsshouldbeableto:

1. Characterizethebasicfunctionsofoperatingsystems.
2. Designtheconceptsofprocessmanagement.
3. Implementtheconceptsofdeadlocks.
4. Describevirtualmemoryandfilesystem.
5. AnalyzetheFilesystemimplementationanddiskI/Otechnique.

Unit	Content
<b>Unit 1 - Introduction (9 Hrs)</b>	Introduction, Computer System Organization, Computer System Architecture, Computer System Structure, Operating System Operations, Process Management, Memory Management, Storage Management, Distributed Systems, Operating System Services, User Operating System Interface, System Calls, Types of System calls, System Programs, Process Concept, Process Scheduling, Operations on Processes, Inter-process Communication
<b>UNIT 2 - SCHEDULING (9hrs)</b>	Threads, Overview, Multithreading Models, CPU Scheduling, Basic Concepts, Scheduling Criteria, Scheduling Algorithms, Thread Scheduling, Multiple-Processor Scheduling, The Critical-Section Problem, Peterson's Solution, Synchronization Hardware, Semaphores
<b>Unit3 - Deadlocks (9hrs)</b>	System Model, Deadlock Characterization, Methods for handling Deadlocks, Deadlock Prevention- Deadlock avoidance- Deadlock detection- Recovery from Deadlock Storage Management, Swapping- Contiguous Memory allocation
<b>Unit 4 - Paging And file System (9hrs)</b>	Paging- Demand Paging, Copy-on Write, Page Replacement, Allocation of frames – Thrashing- Virtual Memory, File Concept, Access Methods, Directory and Disk Structure
<b>Unit 5 - File Management (9hrs)</b>	File System Structure, File System Implementation, Directory Implementation, Allocation Methods, Free-space Management – Disk Structure – Disk Attachment, Disk Scheduling Disk Management, Swap-Space Management, RAID Structure

**TextBook:**

1. .Abraham Silberschatz, Peter Baer Galvin and Greg Gagne, "Operating System Concepts", Eighth Edition, John Wiley & Sons (ASIA) Pvt. Ltd, 2009.

**ReferenceBook:**

1. Harvey M. Deitel, "Operating Systems", Second Edition, Pearson Education, 2002.
2. William Stallings, "Operating System", Prentice Hall of India, 4th Edition, 2003.
3. Andrew S. Tanenbaum, "Modern Operating Systems", Prentice Hall of India, 2003.

**Programme: B.Sc.Data Science and Analytics****Title: Ethics and Etiquettes of Social Media****Type: VEC****Credits:3**

Total Marks-100		Course Code:4DA8		(Total Number of Periods)Hrs
College level theory	Marks:50	Min Passing:20	45	

**Course Outcome(CO):**

1. Understand the ethical principles underlying social media use.
2. Analyze ethical dilemmas related to privacy, authenticity, and transparency on social media platforms.
3. Demonstrate proficiency in applying proper etiquette and professional conduct in online interactions.
4. Develop strategies for promoting digital citizenship and responsible social media engagement.
5. Critically evaluate the impact of social media on individuals, communities, and society as a whole.

**Unit1:IntroductiontoEthicsandSocialMedia(09Hours)**

Ethical Theories and Principles , Online Communication Ethics, Evolution of Social Media Platforms, Ethical Case Studies, Algorithms and Online Behavior, Cultural Perspectives on Social Media Ethics

**Unit2:EthicalDilemmasinSocialMedia(09Hours)**

Privacy Concerns and Data Protection, Authenticity vs. Anonymity, Misinformation and Fake News, Ethical Advertising Practices, Transparency in Marketing, Platform Moderation Ethics

**Unit3:ProfessionalConductandEtiquette(09Hours)**

Building a Professional Online Presence, Social Media Communication Etiquette, Handling Criticism and Conflicts, Ethics of Content Creation and Sharing, Personal vs. Professional Identities, Ethical Guidelines for Influencers.

**Unit4:DigitalCitizenshipandResponsibleEngagement(09Hours)**

Promoting Digital Literacy, Understanding Online Actions' Consequences, Advocating for Digital Rights, Social Media for Social Good-Bridging the Digital Divide, Ethics in Online Communities.

**Unit5:ImpactofSocialMediaonSociety(09Hours)**

Social Media and Mental Health, Addressing Cyberbullying and Harassment, Influence on Public Discourse, Algorithmic Bias and Discrimination, Social Media and Political Polarization, Promoting Empathy and Civility Online.

**ReferenceBooks:**

1. "The Social Media Handbook" by Jeremy Harris Lipschultz
2. "Ethics for the Information Age" by Michael J. Quinn
3. "The Ethics of Influence: Government in the Age of Social Media" by Cass R. Sunstein
4. "Social Media Ethics Made Easy: How to Comply with FTC Guidelines" by Patricia Barnes

**Programme: B.Sc.Data Science And Analytics****Title: Data Handling and Visualization- LAB****Type: SEC/LAB****Credits:2**

<b>Total Marks-50</b>		<b>Course Code:4DA5</b>		<b>(Total Number of Periods)Hrs</b>
<b>External Marks:25</b>	<b>Internal Marks:25</b>	<b>Min Passing:20</b>		<b>60</b>

**List of Practical's:**

**NOTE:** The list suggests sample program set. Hence, the concerned staff may modify the list as needed (Minimum 15).

1. Introduction to various Data Visualization tools
2. Basic Visualization in Python
3. Basic Visualization in R
4. Introduction to Tableau and Installation
5. Connecting to Data and preparing data for visualization in Tableau

6. DataAggregationandStatisticalfunctionsinTableau
7. DataVisualizationsinTableau
8. BasicDashboardsinTableau

**Programme: B.Sc.Data Science andAnalytics**

**Title: Machine Learning LAB**

**Type:SEC/LAB**

**Credits: 2**

<b>TotalMarks-50</b>		<b>CourseCode:4DA6</b>		<b>(TotalNumberofPeriods)Hrs</b>
<b>ExternalMarks:25</b>	<b>InternalMarks:25</b>	<b>MinPassing:20</b>	<b>60</b>	

**NOTE:** The listsuggestssampleprogramset.Hence,theconcernedstaffmay modify thelistasneeded (Minimum 15).

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 probabilitythat it is Fridayis 20 %. What is theprobability that a student isabsentgiven that today is Friday? Apply Baye's rule in python to get the result. (Ans: 15%)
2. Implementlinearregressionusingpython.
3. Extractthedatafromdatabaseusingpython
4. Implementk-nearestneighboursclassificationusingpython
5. ImplementNaïveBayestheoremtoclassifytheEnglishtext
6. Implementanalgorithmtodemonstratethesignificanceofgeneticalgorithm