|  |
| --- |
|  |
| M.sC., DATA ANALYTICS |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
| **SYLLABUS****FROM THE ACADEMIC YEAR****2023 - 2024** |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
| **TAMILNADU STATE COUNCIL FOR HIGHER EDUCATION, CHENNAI – 600 005** |
|  |

**contents**

1. PO and PSO Description
2. PG – Template
3. Methods of Evaluation & Methods of Assessment
4. Semester Index.
5. Subjects – Core, Elective, Nonmajor, Skill Enhanced, Ability Enhanced, Extension Activity, Environment, Professional Competency
6. Course Lesson Box
7. Course Objectives
8. Units
9. Learning Outcome
10. Refence and Text Books
11. Web Sources
12. PO & PSO Mapping tables

|  |
| --- |
| **TANSCHE REGULATIONS ON LEARNING OUTCOMES-BASED CURRICULUM FRAMEWORK FOR POSTGRADUATE EDUCATION** |
| **Programme** | **M.Sc., Data Analytics**  |
| **Programme Code** |  |
| **Duration** | **PG - Two Years** |
| **Programme Outcomes (Pos)** | **PO1: Problem Solving Skill**Apply knowledge of Management theories and Human Resource practices to solve business problems through research in Global context.**PO2: Decision Making Skill**Foster analytical and critical thinking abilities for data-based decision-making.**PO3: Ethical Value**Ability to incorporate quality, ethical and legal value-based perspectives to all organizational activities.**PO4: Communication Skill**Ability to develop communication, managerial and interpersonal skills.**PO5: Individual and Team Leadership Skill**Capability to lead themselves and the team to achieve organizational goals.**PO6: Employability Skill**Inculcate contemporary business practices to enhance employability skills in the competitive environment.**PO7: Entrepreneurial Skill**Equip with skills and competencies to become an entrepreneur.**PO8: Contribution to Society** Succeed in career endeavors and contribute significantly to society.**PO 9 Multicultural competence** Possess knowledge of the values and beliefs of multiple cultures and a global perspective.**PO 10: Moral and ethical awareness/reasoning**Ability to embrace moral/ethical values in conducting one’s life.  |
| **Programme Specific Outcomes****(PSOs)** | **PSO1 – Placement**To prepare the students who will demonstrate respectful engagement with others’ ideas, behaviors, beliefs and apply diverse frames of reference to decisions and actions.**PSO 2 - Entrepreneur**To create effective entrepreneurs by enhancing their critical thinking, problem solving, decision making and leadership skill that will facilitate startups and high potential organizations.**PSO3 – Research and Development**Design and implement HR systems and practices grounded in research that comply with employment laws, leading the organization towards growth and development.**PSO4 – Contribution to Business World**To produce employable, ethical and innovative professionals to sustain in the dynamic business world.**PSO 5 – Contribution to the Society**To contribute to the development of the society by collaborating with stakeholders for mutual benefit. |

**Template for P.G., Programmes**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Semester–I** | **Credit** | **Hours** | **Semester-II** | **Credit** | **Hours** | **Semester-III** | **Credit** | **Hours** | **Semester–IV** | **Credit** | **Hours** |
| 1.1. Core-I  | 5 | 7 | 2.1. Core-IV  | 5 | 6 | 3.1. Core-VII | 5 | 6 | 4.1. Core-XI  | 5 | 6 |
| 1.2 Core-II  | 5 | 7 | 2.2 Core-V  | 5 | 6 | 3.2 Core-VII  | 5 | 6 | 4.2 Core-XII | 5 | 6 |
| 1.3 Core – III  | 4 | 6 | 2.3 Core – VI | 4 | 6 | 3.3 Core – IX | 5 | 6 | 4.3 Project with viva voce | 7 | 10 |
| 1.4 Discipline Centric Elective -I | 3 | 5 | 2.4 Discipline Centric Elective – III | 3 | 4 | 3.4 Core – X  | 4 | 6 | 4.4Elective - VI (Industry / Entrepreneurship) 20% Theory80% Practical  | 3 | 4 |
| 1.5 Generic Elective-II:  | 3 | 5 | 2.5 Generic Elective -IV:  | 3 | 4 | 3.5 Discipline Centric Elective - V  | 3 | 3 | 4.5 Skill Enhancement course / Professional Competency Skill  | 2 | 4 |
|  |  |  | 2.6 NME I | 2 | 4 | 3.6 NME II | 2 | 3 | 4.6 Extension Activity | 1 |  |
|  |  |  |  |  |  | 3.7 Internship/ Industrial Activity | 2 | - |  |  |  |
|  | **20** | **30** |  | **22** | **30** |  | **26** | **30** |  | **23** | **30** |
| **Total Credit Points -91** |

**Choice Based Credit System (CBCS), Learning Outcomes Based Curriculum Framework (LOCF) Guideline Based Credits and Hours Distribution System**

**for all Post – Graduate Courses including Lab Hours**

**First Year – Semester – I**

|  |  |  |  |
| --- | --- | --- | --- |
| **Part** | **List of Courses** | **Credits** | **No. of Hours** |
|  | Core – I | 5 | 7 |
| Core – II | 5 | 7 |
| Core – III | 4 | 6 |
| Elective – I | 3 | 5 |
| Elective – II | 3 | 5 |
|  |  | **20** | **30** |

**Semester-II**

|  |  |  |  |
| --- | --- | --- | --- |
| **Part** | **List of Courses** | **Credits** | **No. of Hours** |
|  | Core – IV | 5 | 6 |
| Core – V | 5 | 6 |
| Core – VI | 4 | 6 |
| Elective – III | 3 | 4 |
| Elective – IV | 3 | 4 |
| Skill Enhancement Course [SEC] - I | 2 | 4 |
|  |  | **22** | **30** |

**Second Year – Semester – III**

|  |  |  |  |
| --- | --- | --- | --- |
| **Part** | **List of Courses** | **Credits** | **No. of Hours** |
|  | Core – VII | 5 | 6 |
| Core – VIII | 5 | 6 |
| Core – IX | 5 | 6 |
| Core (Industry Module) – X | 4 | 6 |
| Elective – V | 3 | 3 |
| Skill Enhancement Course - II | 2 | 3 |
|  | Internship / Industrial Activity [Credits] | 2 | - |
|  |  | **26** | **30** |

**Semester-IV**

|  |  |  |  |
| --- | --- | --- | --- |
| **Part** | **List of Courses** | **Credits** | **No. of Hours** |
|  | Core – XI | 5 | 6 |
| Core – XII | 5 | 6 |
| Project with VIVA VOCE | 7 | 10 |
| Elective – VI (Industry Entrepreneurship)  | 3 | 4 |
| Skill Enhancement Course – III / Professional Competency Skill | 2 | 4 |
| Extension Activity | 1 | - |
|  |  | **23** | **30** |

**Total 91 Credits for PG Courses**

|  |
| --- |
| **METHODS OF EVALUATION** |
| **Internal Evaluation** | Continuous Internal Assessment Test  | **25 Marks** |
| Assignments / Snap Test / Quiz |
| Seminars  |
| Attendance and Class Participation |
| **External Evaluation** | End Semester Examination | **75 Marks** |
| **Total** | **100 Marks** |
| **METHODS OF ASSESSMENT** |
| **Remembering (K1)** | * Thelowestlevelofquestionsrequirestudentstorecallinformationfromthecoursecontent
* Knowledgequestionsusuallyrequirestudentstoidentifyinformationinthetextbook.
 |
| **Understanding (K2)**  | * Understandingoffactsandideasbycomprehendingorganizing,comparing,translating,interpolatingandinterpretingintheirownwords.
* Thequestionsgobeyondsimplerecallandrequirestudentstocombinedatatogether
 |
| **Application (K3)** | * Studentshavetosolveproblemsbyusing/applyingaconceptlearnedintheclassroom.
* Studentsmust usetheir knowledgetodetermineaexactresponse.
 |
| **Analyze (K4)**  | * Analyzingthequestionisonethatasksthestudentstobreakdownsomethingintoitscomponentparts.
* Analyzingrequiresstudentstoidentifyreasonscausesormotivesandreachconclusionsorgeneralizations.
 |
| **Evaluate (K5)** | * Evaluationrequiresanindividualtomakejudgmentonsomething.
* Questionstobeaskedtojudgethevalueofanidea,acharacter,aworkofart,orasolutiontoaproblem.
* Studentsareengagedindecision-makingandproblem–solving.
* Evaluationquestionsdonothavesinglerightanswers.
 |
| **Create (K6)** | * Thequestionsofthiscategorychallengestudentstogetengagedincreativeandoriginalthinking.
* Developingoriginalideasandproblemsolvingskills
 |

**PROGRAMME OUTCOMES (PO) - PROGRAMME SPECIFIC OUTCOMES (PSO) MAPPING**

|  |
| --- |
| **PROGRAMME SPECIFIC OUTCOMES (PSO)** |
|  | **PO1** | **PO2** | **PO3** | **PO4** | **PO5** |
| **PSO1** | **3** | **3** | **3** | **3** | **3** |
| **PSO2** | **3** | **3** | **3** | **3** | **3** |
| **PSO3** | **3** | **3** | **3** | **3** | **3** |
| **PSO4** | **3** | **3** | **3** | **3** | **3** |
| **PSO5** | **3** | **3** | **3** | **3** | **3** |

**Level of Correlation between PO’s and PSO’s**

*(Suggested by UGC as per Six Sigma Tool – Cause and Effect Matrix)*

Assign the value

**1 – Low**

**2 – Medium**

**3 – High**

**0 – No Correlation**

**M.Sc., Data Analytics**

|  |  |  |  |
| --- | --- | --- | --- |
| **Course Code** | **Title of the Course** | **Credits** | **Hours**  |
| Core 1  | Descriptive Statistics | 5 | 7 |
| Core 2 | Foundations of Data Science | 5 | 7 |
| Core 3 | Linear Algebra | 4 | 6 |
| Elective 1 | Data Structures | 3 | 5 |
| Elective 2  | RDBMS and SQL | 3 | 5 |
|  | Oracle and SQL Lab |  |  |
|  | Data Analytics Lab I (R, SPSS, SciLab) |  |  |
|  | Total | 20 | 30 |

|  |
| --- |
| **II SEMESTER** |
| Core 4 | Probability and Distributions | 5 | 6 |
| Core 5 | Machine Learning | 5 | 6 |
| Core 6 | Big data framework | 4 | 6 |
| Elective 3 | Data Science with Python | 3 | 4 |
| Elective 4 |  | 3 | 4 |
|  | Data Analytics Lab II (Hadoop, MapReduce &R, SPSS) |  |  |
|  | Machine learning and Python Lab |  |  |
|  | NME / Skill Enhancement Course [SEC]  | 2 | 4 |
|  | Total | **22** | **30** |

|  |
| --- |
| **III SEMESTER** |
| Core 7 | Big Data with Spark and Hive | 5 | 6 |
| Core 8 | Applied Statistics | 5 | 6 |
| Core 9 | Data mining | 5 | 6 |
| Core 10 | Research Methodology | 4 | 6 |
| Elective 5 | Business Analytics | 3 | 3 |
|  | Data analytics Lab III (Spark, Hive) |  |  |
|  | Data mining and Visualization Lab |  |  |
|  | NME / Skill Enhancement Course [SEC] | 2 | 3 |
|  | Internship / Industrial Activity | 2 | - |
|  | **Total** | **26** | **30** |
| **IV SEMESTER** |
| Core 11 | Cloud Computing | 5 | 6 |
| Core 12 | Cyber Security | 5 | 6 |
|  | Project Work and Viva- Voce | 7 | 10 |
| Elective 6 |  | 3 | 4 |
|  | Skill Enhancement Course – III / Professional Competency Skill | 2 | 4 |
|  | Extension Activity | 1 | - |
|  | Total | **23** | **30** |

**Total Credits -91**

## SEMESTER I

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DESCRIPTIVESTATISTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| Theory | 56 | 4 | - | 4 |

**Preamble** The course introduces the measures of central tendency and dispersion. It also provides the students with systematic knowledge in correlation, regression and outlier analysis.

# CourseOutcomes

Uponthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| CO1 | Demonstratethebasicconceptsofstatistics | **K2** |
| CO2 | Applytheconceptsofcorrelation,regression | **K3** |
| CO3 | Identifythemethodsfordifferent measuresofcentraltendency, dispersion | **K4** |
| CO4 | Evaluatethemethodsforrepresentationofdata. | **K5** |
| CO5 | Constructvariousplots,outliersforregressiondiagnostics | **K6** |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **COs** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | S | S |
| CO2 | S | S | S | M |
| CO3 | S | S | M | S |
| CO4 | M | S | S | S |
| CO5 | S | S | M | S |

S-Strong;M-Medium;L-Low

**Syllabus**

# UNITI (12 hours)

Introduction: Nature and scope of Statistics, limitations of statistics - Types of data: Concept of population and sample, primary and secondary data, quantitative and qualitative data, discrete and continuous data, cross-sectional and time series data. Scales of measurement: Nominal, Ordinal, Ratio and Interval.

# UNITII (11 hours)

Presentationofdata: ConstructionofTableswithoneor more factorsofclassification, Diagrammatic representations:-Line diagram,bardiagram,piediagram andsub-dividedbardiagram,Frequency

distribution and cumulative frequency distribution and their graphical representations, Frequency polygon, histogram, ogive, frequency curves, stem and leaf displays.

# UNIT–III (11hours)

Univariate data: Different measures of location, dispersion, relative dispersion, skewness and kurtosis, Moments, Quantiles and measures based on them – comparison with moment Measures - Box-plot and detection of outliers. Trimmed mean and Winsorised mean – Simple problems.

# UNITIV (11 hours)

Bivariate data – scatter diagram, correlation coefficient and its properties, Correlationratio, Intraclass correlation, Rank correlation – Spearman’s and Kendall’s measures

# UNITV (11 hours)

Regression: Introduction– Usesofregressionanalysis – regression lines – regression equationsofX on Y and Y on X – regression equation in terms of correlation table – standard error of estimate-

UseCases

# TextBook

1. StatisticalMethods, S.P.Gupta,SultanChand and sons(Unit I:Chapter1,3,5(uptopgno.108) UnitII: Chapter 5 (108-126), Chapter 6 Unit V: Chapter 11)
2. BasicStatistics,R.Wilcox,OxfordUniversityPress,2009(Unit III:2.2-2.5,5.5,7.3,9.3)
3. FundamentalsofStatistics:VolumeI, GoonA.M,GuptaM.K.,DasguptaB,Worldpress, 1998(Unit IV: Chapter 10:10.1 – 10.6,Chapter 12: 12.1 – 12.5)

# ReferenceBooks

1. MurrayRSpiegelandLarryJStephens:Statistics,Schaum’sOutline,Fourthedition,2008
2. R.S.N.Pillai,Statistics,S.ChandPublishingCompanyPvtLtd,1992
3. https:/[/www.indi](http://www.indiabix.com/data-interpretation/questions-and-answers/)a[bix.com/data-interpretation/questions-and-answers/](http://www.indiabix.com/data-interpretation/questions-and-answers/)
4. https:/[/www.mathsisfun.com/data/pictogr](http://www.mathsisfun.com/data/pictographs.html)a[phs.html](http://www.mathsisfun.com/data/pictographs.html)

# Pedagogy

Lectures,Simulationexercises,Demonstration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **FOUNDATIONSOFDATASCIENCE** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course provides the fundamental concepts in data science. It includes Data Classification, Sources of Data, Data Science user- roles and skills, Process of big data technology, Security and Intelligence, Basics of R and statistical measures.

# Prerequisite

Mathematics/StatisticsandJava

# CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1. | Understanddataclassification,processofbigdatatechnology, userrolesandskillsindata science. | K2 |
| CO2. | Applythefundamentalconceptsandtechniquesofdatasciencein360viewofCustomer | K3 |
| CO3. | Analyzethemethodologiesofdatascience | K4 |
| CO4. | Detect mythsinbig data | K5 |
| CO5. | DesignthecodefortheproblemsrelatedtodatascienceusingR | K6 |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **COs** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | S | M | M |
| CO3. | S | S | S | S |
| CO4. | S | S | S | S |
| CO5. | S | S | S | S |

S-Strong;M-Medium;L-Low

# Syllabus

## UNITI (11Hours)

**DataEvolution**:DataDevelopment TimeLine - ICT Advancement - aPerspective–DataGrowth-a Perspective - IT Components - Business Process – Landscape - Data to Data Science - **Understanding data**: Introduction - Types of Data: Numeric - Categorical – Graphical - High Dimensional Data - Data Classification - Sources of Data: Time Series - Transactional Data - Biological Data - Spatial Data - Social Network Data- Data Evolution - Data Sources.

# UNITII (12Hours)

**Data Science:** Data Science - A Discipline –Data Science vs Statistics, Data Sciencevs Mathematics,Data Science vs Programming Language, Data Science vs Database, Data Sciencevs

Machine Learning. Data Analytics- Relation: Data Science, Analytics and Big Data Analytics. Data Science Components – Big data technology – Data Science user- roles and skills- Data Science use cases

# UNITIII (11Hours)

Digital Data-an Imprint: Evolution of Big Data –What is Big Data –Sources of Big Data. Characteristics of Big Data 6Vs –Big Data Myths –Data Discovery-Traditional Approach, Big data Technology: Big Data TechnologyProcess –Big DataExploration -Data Augmentation –Operational Analysis –360 View of Customers –Security and Intelligence

# UNITIV (11 Hours) R

Basics: Introduction- Packages and Library – Data types – Basic operators – R objects- Vectors – Lists- Arrays – Matrix- Factors – Data frame- R file formats- Importing and exporting files – Data Visualization in R: Lattice package- Box plot- bar chart – scatter plot- GGplot2

# UNITV (11Hours)

Statistical Measures in R: Measures of central tendency – Range- inter quartile range – Mean – Median – variance- Standard deviation – Sampling distribution – probabilitydistributions- hypothesis tests

Usecases:Insurancepolicyoffers,Discounttargetinginonlineshopping

# TextBook

1. V.Bhuvaneswari,T.Devi,(2016).BigDataAnalytics: APractitioner’sApproach,Bharathiar

University

1. V.Bhuvaneswari(2016).DataAnalyticswithR,Bharathiar University.

# ReferenceBooks

1. NinaZumal,JohnMount(2014).PracticalDatascience inR,ManagingPublicationCompany
2. Bernard Kolman, RobertC.Busby and Sharon Ross(2004).Discrete MathematicalStructures, New Delhi: Prentice Hall

**Pedagogy:**Lectures,DemonstrationandCaseStudies

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **LINEARALGEBRA** |  | **Category** | **L** | **T** | **P** | **Credit** |
|  | **Theory** | **56** | **4** | **-** | **4** |

**Preamble** The course introduces the principles underlying linear equations and vector spaces. It also provides the concepts of Eigen values, Eigen vectors and Positive Definite Matrices

# CourseOutcomes

Uponthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| CO1 | DemonstratecompetencewiththebasicideasoflinearAlgebraincluding the concepts of vector spaces, Determinants, Eigen values and Eigenvectors and positive definite matrices | K2 |
| CO2 | Theabilityto understandtheprinciplesofLinearAlgebra | K3 |
| CO3 | Applypropertiesoflinearspacestospecificmathematicalstructures | K4 |
| CO4 | Composeclearand accurateproofsusingtheconceptsoflinearAlgebra | K5 |
| CO5 | Appreciatethesignificanceofvectorspacesandpositivedefinitematrices | K6 |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | S | S | S |
| CO2 | S | S | S | M |
| CO3 | S | S | S | S |
| CO4 | S | S | M | S |
| CO5 | S | S | S | S |

S-Strong;M-Medium;L-Low

# Syllabus UNIT I (11 Hours)

The Geometry of Linear Equations- An Example of Gaussian Elimination- Matrix Notation and Matrix Multiplication - Triangular Factors and Row Exchanges- Inverses and Transposes

# UNITII (11Hours)

**Vector Spaces**: Vector Spaces and Subspaces – Solving Ax=0 and Ax=b - Linear Independence, Basis, and Dimension- The Four Fundamental Subspaces- Graphs and Networks- Linear Transformations

# UNITIII (11Hours)

**Determinants:** Introduction- Properties of the Determinant- Formulas for the Determinant- Applications of Determinants

# UNITIV (11Hours)

**Eigenvalues and Eigenvectors:** Introduction- Diagonalization of a Matrix .- Difference Equations and Powers A k- DifferentialEquations and e At - Complex Matrices- Similarity Transformations

# UNITV (12Hours)

**Positive Definite Matrices**: Minima, Maxima, and Saddle Points - Tests for Positive Definiteness- Singular Value Decomposition

# TextBook

GilbertStrang(2006).LinearAlgebraandItsApplication,FourthEdition,AcademicPress.

# Reference Books

1. DavidC.Lay,StevenR.Lay,JudiJ.McDonald(2014).LinearAlgebraandItsApplications, Pearson Education.
2. PeterD.Lax(2007).LinearAlgebraandItsApplications,SecondEdition,WileyPublication

**Pedagogy:**Lectures,DemonstrationandCaseStudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATASTRUCTURES** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

# Preamble

This course covers the various data structures, including arrays, structures, stacks and queues. It includes sorting and searching techniques and effective search methods in Binary trees. This course also deals with graph data structures.

# Prerequisite

* + Discretemathematics.

# CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| CO1. | Understandtheconceptsofarrays, stringsandalgorithmsforbasicoperations. | **K2** |
| CO2. | Applyconceptofstacks,queues, linkedlistandalgorithmsforbasicoperations. | **K3** |
| CO3. | Identifythefamiliaritywithmajoralgorithmsanddatastructures | **K4** |
| CO4. | Analyzeappropriatealgorithmsanddatastructuresforvariousapplications | **K5** |
| CO5. | Formulatethe computationalcomplexityofvariousalgorithms | **K6** |

#

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | M | M | M |
| CO3. | S | S | M | M |
| CO4. | S | S | S | M |
| CO5. | S | S | S | M |

S-Strong;M-Medium;L-Low

# Syllabus

## UNITI (11Hours)

**Basics:** Algorithm Specification – Data Abstraction – Performance Analysis – Arrays – Dynamic Allocated Arrays – Structures and Unions – Polynomials - Sparse Matrices- Representation of Multidimensional Arrays – Strings.

# UNITII (12Hours)

**Stacks and Queues:** Stacks – Stacks Using Dynamic Arrays - Queues - Circular Queues Using DynamicArrays-EvaluationofExpressions-MultipleStacksAndQueues **Linked Lists:** Singly Linked List And Chains – Representing Chains in C – Linked Stack And Queues – Polynomials - Additional List Operations - Sparse Matrices – Doubly Linked List.

# UNITIII (11Hours)

**Searching**:Introduction-SequentialSearch- BinarySearch.**Sorting:**Introduction- Insertion Sort

-SelectionSort-MergeSort-QuickSort-HeapsandHeapSort.

# UNITIV (11Hours)

**EfficientBinarySearchTrees:**OptimalBinarySearchTrees–AVLTrees-KDTrees.

# UNITV (11Hours)

**Graphs:**ThegraphAbstractDataType-Elementarygraphoperations-Minimumcostspanning trees- shortest paths and transitive closure- AOV networks –AOE networks.

**TextBook**

1. Ellis Horowitz,Sartaj Sahni and Anderson Freed (2009), Fundamentals of data structures in C,University Press
2. Ellis Horowitz, Sartaj Shani, SanguthevarRajasekaran (2013), Fundamentals of computer algorithms, Galgotias Publications private limited
3. RobertLKruse(2008).DataStructures&ProgramDesign,NewDelhi:PrenticeHall

# Reference Books

1. MarkAllenWeiss(2012).DataStructuresandAlgorithmAnalysisinC++;PearsonEducation
2. Sartaj Sahni (2010). Data Structures, Algorithms, and Applications in C++; McGraw-Hill International Edition
3. AlfredV.Aho,John E.Hopcroft, Jeffrey D.Ullman (2000). Data structures and algorithms; Pearson Education, Asia.
4. AdamDrozdek(2013).DataStructuresandAlgorithminC++,4thEdition.

**Pedagogy:**Lectures,GroupDiscussion,casestudy

#

**RDBM SANDSQL**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course presents the fundamental concepts of Database Management Systems, database design and Relational model. It provides concepts ofhowto applythese inpractice and learn howto usethe structured query language to work and analyse databases.

# Prerequisite

* + Database
	+ Programmingconcepts

# CourseOutcomes

Onthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1 | Understandtheconceptsofrelationaldatabases,databasedesignUsing ER diagram | K2 |
| CO2 | DemonstratevariousSQLqueriesbyapplyingRDBMSconcepts | K3 |
| CO3 | AnalyzevariousrealtimeapplicationsforapplyingRDBMSconcepts. | K4 |
| CO4 | Evaluateconstraintsondataandidentifysituationstoapplytheconstraints. | K5 |
| CO5 | Designdifferentdatabasesforvarioussituations | K6 |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | S | S | S |
| CO2 | S | S | S | S |
| CO3 | S | S | S | S |
| CO4 | S | S | S | S |
| CO5 | S | S | S | S |

S-Strong;M-Medium;L-Low

# Syllabus

## UNITI: (11Hours)

Overview of database systems: Managing Data - A Historical Perspective - File System versus DBMS – Advantages of a DBMS - Describing and Storing data in a DBMS - Queries in a DBMS - Transaction Management - Structure of a DBMS.

**UNITII: (11Hours)**

Database Design & ER diagrams - Entities, Attributes and Entity Sets - Relationships and Relationship Set - Additional features of the ER model- Conceptual Database design with ER Model.

**UNITIII: (12Hours)**

Relational Model: Introduction - Integrity Constraints Over Relations - Enforcing Integrity ConstraintsonRelationalData - LogicalDatabase Design: ERto Relational- Introductionto Views - Destroying / Altering Tables and Views - Relational Algebra and Calculus.

**UNITIV: (11Hours)**

SQL Queries, Constraints, Triggers: The form of a Basic SQL Query - UNION, INTERSECT and EXCEPT - Nested Queries - Aggregate Operators - Null Values - Complex integrity constraints in SQL - Triggers and Active Data bases – Query Evaluation.

**UNITV: (11Hours)**

PL/SQL Programming: Functions and Procedures, Triggers, Queries, Forms, Reports, Cursors, Exceptions. Introduction to NoSQL – Types.

**Text Book** Raghu Ramakrishnan and Johannes Gehrke (2003). Database Management System, Third edition, McGraw-Hill.

# ReferenceBooks

1. Abraham Silberschatz, Henry F. Korth and Sudarshan S(2005). Database System Concepts,5/e, McGraw-Hill.
2. DateCJ(2003).AnIntroductiontoDatabaseSystems,8/e, PearsonEducation.
3. MichaelMclaughlin,(2010).OracleDatabase11gPL/SQLProgramming,McGrawHill.
4. ShashankTiwari(2011).ProfessionalNoSQL,JohnWiley&Sons

**Pedagogy:**Lectures,GroupDiscussion,casestudy

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ORACLEANDSQL LAB** | **Category** | **L** | **T** | **P** | **Credit** |
| **Practical** | **-** | **-** | **75** | **3** |

**Preamble** This course provides sound introduction to implement the relational database management systems concepts in SQL. This course also provides various exercises to implement the integrity constraints on databases, functions, procedures, cursors, triggers, exception handling, forms and reports.

# Prerequisite

* + Databaseconcepts
	+ Programming concepts

# CourseOutcomes

Onthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO6. | ImplementthedatabasesconceptsandSQLqueriesasperimplementation. | **K2** |
| CO7. | ApplyspecificSQLcommandsonrelationaltablesfordifferentsituations | **K3** |
| CO8. | Analyseusecasesandcreateconstraintssuitableforthegivensituation. | **K4** |
| CO9. | CreateandanalyseadatabaseusingSQLDML/DDLcommands | **K6** |
| CO10. | DesignandbuildaGUI application | **K6** |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO6. | S | S | S | S |
| CO7. | S | S | S | S |
| CO8. | S | S | S | S |
| CO9. | S | S | S | S |
| CO10. | S | S | S | S |

S-Strong;M-Medium;L-Low

# Syllabus

* + - Exercisestoimplementtheconceptsofnullconstraint,uniqueconstraint,integrity constraints, check constraints.
		- Exercisestoimplement nestedqueries.
		- Exercisestoimplementtheconceptsofpartitioningqueries
		- Exercisestocreateaviewfromthetables
		- Exercisestocreatefunctionsandprocedures
		- Exercisetocreatetriggersandqueries
		- Exercisetocreateformsandreports.
		- Exercisestocreatecursorsand exceptions
		- ExercisestocreateSimpleapplications

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATAANALYTICSLABI** | **Category** | **L** | **T** | **P** | **Credit** |
| **Practical** | **-** | **-** | **75** | **3** |

**Preamble** The course deals with the implementation of linear algebra concepts in Scilab. This course also provides various exercises to implement the statistical functions using R and SPSS.

# Prerequisite

* + Statisticalconcepts
	+ Basicconceptsofalgebra

# CourseOutcomes

Onthesuccessfulcompletionofthecourse,studentswillbeable to

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **COStatement** | **Knowledge Level** |
| CO1 | UnderstandthebasicconceptsofLinear algebraandstatistics | **K2** |
| CO2 | ImplementthealgebraicandstatisticalproblemsusingScilabandR | **K3** |
| CO3 | ApplytheconceptsofLinearalgebraandstatisticsinrealtimeproblems | **K4** |
| CO4 | Analyserealtimedatausingvariousstatisticalmeasuresinscilab | **K5** |
| CO5 | ConstructmodelsusingvariousstatisticalmethodsinRandSPSS | **K6** |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO11. | S | S | S | S |
| CO12. | S | S | S | S |
| CO13. | S | S | S | S |
| CO14. | S | S | S | S |
| CO15. | S | S | S | S |

S-Strong;M-Medium;L-Low

# Syllabus

* ExercisestoimplementthebasicmatrixoperationsinScilab.
* ExercisestofindtheEigenvaluesandeigenvectorsinScilab.
* Exercises to solve equationsby Gauss elimination,Gauss Jordan Method and Gauss Siedel inScilab.
* Exercisestoimplementtheassociative,commutativeanddistributivepropertyinamatrixin Scilab.
* Exercisestofindthereduced rowechelon formofamatrixinScilab.
* Exercisestoplotthefunctionsandtofinditsfirstandsecondderivatives inScilab.
* ExercisestopresentthedataasafrequencytableinSPSS.
* ExercisestofindtheoutliersinadatasetinSPSS.
* Exercisestofindthemost riskyproject outoftwomutuallyexclusiveprojectsinSPSS
* Exercises to draw a scatter diagram, residual plots, outliers leverage and influential data pointsin R
* ExercisestocalculatecorrelationusingR
* Exercisestoimplement TimeseriesAnalysisusingR.
* ExercisestoimplementlinearregressionusingR.

# SEMESTERII

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PROBABILITYANDDISTRIBUTIONS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course introduces the fundamental concepts of probability and random variables .It also provides knowledge in discrete and continuous distributions. It deals with various sampling distributions like t,F, chi-square distributions etc.

# CourseOutcomes

Onthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1 | Understandtheconceptsofprobabilityanditsdistributions. | K2 |
| CO2 | Demonstrateskillsinhandlingproblemsinvolvingrandomvariablesandtheir functions | K2 |
| CO3 | Applysamplingdistributionstocontributetotheprocessofmakingrational decisions in analytical problems. | K3 |
| CO4 | Analysethevarioussampling testsandchoosetestsuitablefortheproblem | K4 |
| CO5 | Formulatethehypothesistestforvariouscomplex problems | K6 |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | M | M |
| CO2 | S | S | M | M |
| CO3 | S | S | S | M |
| CO4 | S | M | S | M |
| CO5 | S | S | S | M |

S-Strong;M-Medium;L-Low

# Syllabus UNIT I (11 Hours)

Random Experiment: Sample space, Different types of events. Approaches of measuring probability: Mathematical, Statistical and Axiomatic probability, Laws of probability-conditional probability – Baye’s theorem

# UnitII (11Hours)

Randomvariable- Discreteand ContinuousRandomvariables- Distributionfunctionanditsproperties

– Expectation and Moment Generating function: Raw and Central moments-relationship between central and raw moments-moments about an arbitrary value-moment generating function-properties of moment generating moment-characteristic function- simple problems.

# UnitIII (11Hours)

DiscreteDistributions:Binomial,Poisson,Geometricdistributionsandtheirproperties.Continuous

Distributions: Rectangular, Exponential, Normal, lognormal distributions and their properties-Simple problems

# UnitIV (11Hours)

Centrallimittheorem-Confidenceintervalforonemeananddifferenceoftwomeans. Testing of Hypotheses : Introduction –relation between confidence interval and testing of hypothesis- level of significance and p-value. Classification of hypothesis tests-Large sample tests: Single mean, Difference of two means, Single proportion, Difference of two proportions

# UnitV (12Hours)

Small sample tests:t-test for single mean,difference between two means-F-test for equality of two population variances-Chi-square test for single mean, Chi-square test for goodness of fit- Chi-square test for independence of attributes and homogeneity and equality of proportions. Applications to machine learning

**TextBooks**

Ravichandran.J,ProbabilityandStatisticsforEngineers,Wiley,2015. Unit I: Section: 1.1-1.4

UnitII:Section:2.1-2.5, 3.1-3.2,3.3.1-3.3.6

UnitIII:Section:4.1,4.2.1-4.2.4,5.1-5.3,5.5

Unit IV:Section:6.4, 10.3.1,10.3.2,11.1-11.4

UnitV:Section–11.5.1-11.5.7

# ReferenceBooks

.GoonA.M.,GuptaM.K.&DasguptaB,AnOutlineofStatisticalTheory(Vol-1),WorldPress, 1994.

1. RohatgiV.K.,AnIntroductiontoProbabilityTheoryandMathematicalStatistics, JohnWiley, 1984
2. ScymourLipschuts,Probability,Schaum’sOutline,1996.
3. S.C.GuptaandV.K.Kapoor, FundamentalsofMathematicalStatistics,SultanChandandsons, 2002

**Pedagogy:**Lectures,Casestudies, Demonstrations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **MACHINELEARNING** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course introduces the fundamentals of Machine Learning and algorithms. It also covers various supervised and unsupervised learning algorithms used for classification, prediction and clustering.

# Prerequisite

* Statisticsandprobability
* Linear Algebra

# CourseOutcomes

Onthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1 | Understandtheconceptsofmachinelearning. | K2 |
| CO2 | Applysupervisedandunsupervisedlearningalgorithmsforclassification, prediction and clustering. | K3 |
| CO3 | Analyzethelogicbehindtheexecutionofvarious classifiers. | K4 |
| CO4 | Evaluatetheperformanceofdifferentalgorithmsavailableforminingdata. | K5 |
| CO5 | Predictsolutionforrealworldproblems. | K6 |

# MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | M | M |
| CO2 | S | M | M | M |
| CO3 | S | S | S | M |
| CO4 | S | S | S | M |
| CO5 | S | S | S | M |

S-Strong;M-Medium;L-Low

# Syllabus UNIT I (11 Hours)

Introduction:Machinelearning–ExamplesandApplications-PerspectivesandIssuesinMachine

learning - Input: Concepts, Instances, and Attributes - Output: Knowledge Representation-Credibility: Evaluating What’s Been Learned: Training and Testing - Predicting Performance - Cross Validation - Other Estimates - Counting the cost.

# UNITII (11Hours)

Decision Tree Learning: Decision tree representation – Decision tree learning – Random forest -Issues in decision tree learning- Bayesian Learning: Naïve Bayes classifier - Instance Based Learning: Introduction – k- nearest neighbor Learning - RadialBasis Function, Case based reasoning.

# UNITIII (11Hours)

ArtificialNeuralNetwork-Introduction–NeuralNetworkRepresentation-Perceptrons-Multilayer

NetworksandBackpropagationAlgorithm-LinearmodelsforRegression-LinearDiscriminant Analysis - PCA – Kernel PCA

# UNITIV (12Hours)

SVM : Introduction – Kernel methods - formulation and computation- SVM Linear classifier – SVM with two variables –non-linear classifier-Polynomial kernels- Radial Basis Function Kernels - Clustering Methods.- Introduction – K- Means- Expectation-Maximization Algorithm- Hierarchical Clustering - Choosing the Number of Clusters

# UNITV (11Hours)

Deep Learning – Deep feed forward network – Convolutional neural network– Autoencoders –Deep Belief Networks -Recurrent Neural Network - Use Cases: Finding similar users in Twitter (Mahout), Email marketing system (Mahout)

# ReferenceBooks

1. Ian Witten, Data mining: PracticalMachine Learning Tools and Techniques, Fourth edition, Morgan Kaufmann Publishers
2. TomM.Mitchell(1997).MachineLearning,TataMcGraw-Hill,NewDelhi
3. K.P.Soman,MachineLearningwithSVMandOtherKernelMethods,2011,PHIPublishing

**Pedagogy:**Lectures,GroupDiscussions, Demonstrations.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **BIGDATAFRAMEWORK** | **Category** | **L** | **T** | **P** | **Credits** |
| **Theory** | **56** | **4** | **-** | **4** |

 **Preamble** This course introduces big data framework, technologies, storage and Hadoop ecosystem. It also deals with the concepts of MapReduce, Pig and Scala programming.

# Prerequisite

* + DatabaseManagementsystems
	+ Programmingconcepts

# CourseOutcomes

Onsuccessfulcompletionofthecourse,studentswillbeable to

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **COStatement** | **Knowledge****Level** |
| CO1 | Understandthecharacteristicsofbigdata,conceptsofHadoop ecosystem and scala programming. | K2 |
| CO2 | ApplyMapreduceprogrammingmodeltoprocessbig data. | K2 |
| CO3 | AnalyzeSparkapplicationsforitsuseinbigdataprocessing. | K3 |
| CO4 | EvaluatetechniquesusedforhandlingNoSQLdatabasesMongoDBandHbase. | K4 |
| CO5 | DesignprogramsforbigdataapplicationsusingHadoopcomponents. | K6 |

**MappingwithProgramOutcomes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | S | S | M |
| CO2 | S | S | S | M |
| CO3 | S | S | S | M |
| CO4 | S | S | S | M |
| CO5 | S | S | S | M |

#  Syllabus UNIT I (11 Hours)

**Introduction to Big data:** Introduction – Big Data- Characteristics of Big Data – Bigdatamanagement architecture- Examining Big Data Types – Big Data TechnologyComponents -– Big data analytics –Big data analytics examples - Web Data Overview – Web Data in Action.

# UNITII (11Hours)

**Hadoop :**Introduction – History of Hadoop - Hadoop Ecosystem- Analyzing data with Hadoop - Hadoop Distributed File System- Design- HDFS concepts - Hadoop filesystem–Data flow – Hadoop I

/ O - Data integrity – Serialization - Setting up a Hadoop cluster - Cluster specification - cluster setup and installation – YARN

# UNITIII (12Hours)

**MapReduce:** Introduction – Understanding Map, Reduce functions - Scaling out - Anatomy of a MapReduce Job Run - Failures – Shuffle and sort - Mapreduce types and formats - features – counters

-sorting-UnittestwithMRunit-localtest

# UNITIV (11Hours)

**Data Analytics using Pig:** Introduction – Downloading and installing pig -Grunt – Pig’s Data model–Types–Schemas.IntroductiontoPigLatin–Preliminarymatters-InputandOutput–Relational operators-DevelopingandTestingPigLatinScripts.**SCALA:**Introduction-Classesandobjects- Basic types and operators - built-in control structures - functions and closures - inheritance

# UNITV 11 Hours)

**NoSQL Databases:** Introduction to NoSQL- **MongoDB**: Introduction – Data types – Creating, Updating and deleing documents -Querying – Introduction to indexing – Capped collections. **Hbase**: Concepts - Hbase Vs RDBMS - Creating records- Accessing data – Updating and deleting data – Modifying data- exporting and importing data.

**USECASES:**Calldetailloganalysis,Creditfraudalert,Weatherforecast

# TextBook

TomWhite(2012).Hadoop:TheDefinitiveGuide,ThirdEdition,O’ReillyMedia

# ReferenceBooks

1. BillFranks(2012).TamingtheBigDataTidalwave, JohnWiley&Sons
2. Martin Odersky,LexSpoon, Bill Venners(2010),Programming in Scala, Second Edition, Artima Press, California.
3. ShashankTiwari(2011).ProfessionalNoSQL,JohnWiley&Sons
4. Borislublinsky,Kevint.Smith,Alexey,Yakubovich(2015).ProfessionalHadoopSolutions,Wiley
5. ChrisEaton,Dirkderoosetal.(2012).UnderstandingBigdata,McGrawHill
6. MinChen(2014). BigData:RelatedTechnologies,ChallengesandFutureProspects,Springer
7. JudithHurwitz(2013).BigDataforDummies,JohnWiley&Sons

**Pedagogy:**Lectures,GroupDiscussions,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATASCIENCEWITHPYTHON** | **Category** | **L** | **T** | **P** | **Credits** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course introduces the concepts of Python for data science as well as programmingin general.It covers the core programming concepts of python, advanced concepts like regularexpressions, exception handling, multithreading, web programming and data base programming. It also introduces the python libraries that are most commonly used for data analysis.

# Prerequisite

* BasicunderstandingofOpensourcesoftware
* Databaseconcepts

# CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1.** | UnderstandtheprogrammingconstructsofPython | **K2** |
| **CO2.** | DemonstratetheconceptsofPythoninsimpletasks | **K3** |
| **CO4.** | Analyse the utilities of Python with other object oriented programming languages | **K4** |
| **CO3.** | Identify features to be appliedforapplications developmentin various domains | **K5** |
| **CO5.** | Designsolutionsfordataanalyticsproblems | **K6** |

**MappingwithProgramOutcomes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | S | S | M |
| CO2 | S | S | S | M |
| CO3 | S | S | S | M |
| CO4 | S | S | S | M |
| CO5 | S | S | S | M |

**Syllabus**

# UnitI (10Hrs)

Introduction: Origins - Features - Getting started-Python Basics - Python Objects - Numbers – Sequences- Strings, ListsandTuples: - Strings- Stringsandoperators - Stringonlyoperators - Built-in functions - String - Built-in methods - Special features of strings

## UnitII (12Hrs)

Lists - Operators – Built in functions - List Type built-in methods - Special features of Lists, Tuples - Operators and Built-in functions - Special features of Tuples – Mapping and setting Dictionaries – Operators- Built-inand factoryfunctions - Mapping types-built-inmethods- Dictionarykeys. Set types

-Operators-Built-infunction-Settypebuilt-inmethods

**UnitIII (11Hrs)**

Conditionalsandloops.Functionsandfunctionalprogramming–Modules-Regularexpressions- Multithreaded programming

## UnitIV (11Hrs)

Files & I/O: File objects – Builtin Functions – Methods – Builtin Attributes – Standard files –Commandline arguments– File System – File Execution – Storage Modules.Exception Handling: Exception - Exception Handling - Except clause – Try- Finally clause - User Defined Exceptions.

## UnitV (12Hrs)

Python Libraries – NumPy, SciPy, MatplotLib, Pandas, NLTK Note:Examples,exercisesandcasestudiesrelatedto dataanalytics

# TextBook

WesleyJ.Chun(2010).CorePythonprogramming,2/e,Pearsoneducation.

# ReferenceBooks

1. MarkLutz(2010).ProgrammingPython,4/e,O’ReillyMedia.
2. MarkSummerfield(2009),ProgramminginPython3,PearsonEducation.
3. AlbertoBoschetti,LucaMassaron(2016),PythonDataScienceEssentials,PacktPublishing.

**Pedagogy:**Lectures**,**Demonstrations,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATAANALYTICSLABII** | **Category** | **L** | **T** | **P** | **Credit** |
| **Practical** | **-** | **-** | **75** | **3** |

 **Preamble** This course provides implementation of statistics concepts in R and SPSS. It also provides sound introduction to implement the Hadoop framework. This course also provides various exercises to implement in the distributed environment through map reduce programming.

# Prerequisite

* + Statistics
	+ Big data framework
	+ Java

# CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1.** | DemonstrateconceptsofprobabilityinR | **K2** |
| **CO2.** | Applydifferent probabilitydistributions inSPSS | **K3** |
| **CO3.** | ApplyspecificMapperandreducerfunctionsfordifferentsituations | **K3** |
| **CO4.** | Designandbuildahadoopcluster.createNoSQLdatabasesusing MongoDB. | **K6** |
| **CO5.** | DevelopDML,DDLcommandsinHbase. | **K6** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **COs** | **PO1** | **PO2** | **PO3** | **PO4** |
| **CO1.** | S | S | M | M |
| **CO2.** | S | S | M | M |
| **CO3.** | S | S | M | M |
| **CO4.** | S | S | S | M |
| **CO5.** | S | S | S | M |

S-Strong;M-Medium;L-Low

**MappingwithProgrammeOutcomes**

# Syllabus

* + Setting up ahadoopenvironment
	+ Exercisestoimplement filemanagementtasksusingHadoop
	+ ExercisestoimplementMapreduceprogramthatminesweatherdata
	+ ExercisestoimplementStockcountMapreduceprogram.
	+ ExercisestoimplementPigLatinscriptstosort,group,join,project,andfilterdata.
	+ ExercisestoimplementCRUD,AggregatingandindexingoperationsinMongoDB.
	+ ExercisestoapplytheconceptofMapReduceinMongoDB.
	+ ExercisestoimplementDDL,DMLcommandsusingHBase.
	+ ExercisestoimplementconceptsofprobabilityanddistributionsinR
	+ ExercisestoimplementconceptsofprobabilityanddistributionsinSPSS

**Pedagogy:**Demonstrations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **MACHINELEARNINGAND****PYTHONLAB** | **Category** | **L** | **T** | **P** | **Credits** |
| **Practical** | **-** | **-** | **75** | **3** |

**Preamble** This course introduces sci-kitlearn, the popularmachine learning library in Python. It also provides exercises to implement sci-kit learn syntax and tools to apply machine learning algorithms.

# Prerequisite

* + - * KnowledgeinOpensource-Python
			* MachineLearning

# CourseOutcomes

Onsuccessfulcompletionofthecourse,studentswillbeable to

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **Co Statement** | **Knowledge Level** |
| **CO1.** | AnalyzebasicconceptsofPython | **K3** |
| **CO2.** | Build modelsusingclassificationalgorithmforrealworldproblems | **K6** |
| **CO3** | Buildmodelsusingclusteringalgorithmforrealworld problems | **K6** |
| **CO4.** | Createclassificationandclustering models | **K6** |
| **CO5** | Testandevaluatethe models | **K6** |

|  |
| --- |
| **MappingwithProgrammeOutcomes** |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | S | S | M |
| CO2. | S | S | S | M |
| CO3. | S | S | S | M |
| CO4. | S | S | S | M |
| CO5 | S | S | S | M |

S-Strong;M-Medium;L-Low

# Syllabus

* + - * + ExercisetodevelopsimplewebapplicationsinPython
				+ Exercisetomanipulatedatausingdifferentqueries
				+ ExercisestohandleExceptions, Multithreading
				+ Exercisestoloaddatasetintosci-kitlearn
				+ ExerciseforBuildingmodelsinsci-kitlearn
				+ Exercisetoextractfeaturesfromdatasets
				+ ExercisetoimplementRegression
				+ ExercisetoimplementClassification
				+ ExercisetoimplementClustering
				+ ExercisesforModelselectionandevaluation
				+ ExercisestoBuildadatapipeline

**Pedagogy:**Demonstrations

## SEMESTERIII

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **BIGDATAWITHSPARKANDHIVE** | **Category** | **L** | **T** | **P** | **Credits** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course introduces the fundamental concepts of Spark and Hive. It also provides knowledge on SQL in Spark and understanding Hive data model.The course also covers data definition and data manipulation operations using HiveQL and applying these to perform analytics.

## Prerequisite

* + BigDataFramework
	+ SQL

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1** | UnderstandtheprogrammingconstructsofSparkand databasemanagementusingHiveQL | **K2** |
| **CO2** | Applythe conceptsofSpark QLand Hiveinsimpletasks.UnderstandwritingscriptsusingSQLinSparkandperformvariousHiveQLqueries by applying RDBMS concepts | **K3** |
| **CO3** | AnalyzedifferenttypesofdataanditssourcesforuseinSparkApplications. | **K4** |
| **CO4** | AppraisetechniquesfordatamanipulationinHive. | **K5** |
| **CO5** | DesignanddevelopscriptsinSparkandHiveforvarioussituations | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| **CO1** | S | S | M | M |
| **CO2** | S | S | M | M |
| **CO3** | S | S | S | M |
| **CO4** | S | S | S | M |
| **CO5** | S | S | M | M |

S-Strong;M-Medium;L-Low

## Syllabus UNIT I: (11 Hrs)

OverviewofBigdataandspark-SparkPhilosophy-HistoryofSpark-RunningSpark-Spark

Architecture - Data Frames - Transformations - End to end example –Spark Toolset - Spark run on cluster - Developing spark Applications - Deploying Spark

## UNITII: (11Hrs)

Structure API -Structured Spark Types -API Execution - Basic structured operations -Working with Different typesofData-Booleans,strings,dates,complextypes -WorkingwithJson-Aggregations – Grouping – Joins - Data sources.

## UNITIII: (12Hrs)

SparkSQL-RunningSparkSQLQueries–Tables–Views-Databases-Advancedsubqueries–

Dataset - Transformations\_RDD - Creating RDD – Transformations – Actions - saving files - Advanced RDD - Key-value RDD - Distributed shared variables - Accumulators. Streaming: Streaming Fundamentals - Processing Design Points - Structure streaming Basics - core concepts - Input and output.

## UNITIV: (11Hrs)

Developing Hive – Services in Hive – Understanding Hive Data model-HiveQL: Data definition – Databases – Alter database – Create Tables – Alter Tables – Manage Tables – Drop Tables.HiveQL**: Data manipulation** – Load Data into managed tables – Insert into tables from Queries - Creating Tables and Loading Them in One Query - Exporting Data.

## UNITV: (11Hrs)

HiveQL: Queries – SELECT … FROM Clauses –Where Clauses – Group by clauses - Join Statements–ORDERBYandSORTBY-DISTRIBUTEBYwithSORTBY-CLUSTERBY–

casting-UNION ALL-HiveQL:Views-HiveQL:Indexes-SchemaDesign–Tuning–Functions

* CustomizingHiveFileandRecordformats-CaseStudies.

## TextBooks

* 1. BillChambersandMateiZaharia(2018).“Spark:TheDefinitiveGuide”,O’ReillyMedia.
	2. JasonRutherglen,DeanWampler,EdwardCapriolo(2012).ProgrammingHive,First

edition,O’ReillyMedia.

## ReferenceBooks

1. HanishBansal,SaurabhChauhan(2016).ApacheHiveCookbook,Packtpublishing.
2. MikeFrampton,MasteringApachespark,PackPublishing.
3. Jules.S.,BrookeWenig,TathagataDas&DennyLee(2020).”LearningSpark”,O’ReillyMedia.
4. HoldenKarau&RachelWarren,(2017),”HighPerformanceSpark”, O’ReillyMedia.

**Pedagogy:**Lectures**,**Demonstrations,Casestudies

|  |  |  |  |
| --- | --- | --- | --- |
| **Unit** | **Topic** | **Activity** | **WebReferences** |
| Unit I | ApacheSpark Architecture- | Sequenceof Questions | https://intellipaat.com/blog/tutorial/spark-tutorial/spark- architecture/https:/[/www](http://www.edureka.co/blog/spark-architecture/).[edureka.co/blog/spark-architecture/](http://www.edureka.co/blog/spark-architecture/) |
| Unit II | Aggregations- Grouping-Joins- | Packetof Problems | <http://sqlandhadoop.com/spark-dataframe-aggregate-functions/>[https://medium.com/@achilleus/https-medium-com-joins-in-](https://medium.com/%40achilleus/https-medium-com-joins-in-)apache-spark-part-1-dabbf3475690 https://mungingdata.com/apache-spark/aggregations/ |
| Unit III | Spark SQL Queries-tables-Views | Student- generatedcontent | <https://docs.databricks.com/spark/latest/spark-sql/language-manual/create-view.html> |
| Unit IV | Data manipulation | Group Discussion | https:/[/www](http://www.oracle.com/technetwork/topics/bigdata/articles/hive-).[oracle.com/technetwork/topics/bigdata/articles/hive-](http://www.oracle.com/technetwork/topics/bigdata/articles/hive-) and-pig-hol-1937050.pdfhttps:/[/www](http://www.rcvacademy.com/big-data/hiveql-data-manipulation/).[rcvacademy.com/big-data/hiveql-data-manipulation/](http://www.rcvacademy.com/big-data/hiveql-data-manipulation/) |
| Unit V | Queries–SELECTFROMClausesWhereClauses–Group by clauses | Discussion activities | https:/[/www](http://www.guru99.com/hive-queries-implementation.html).[guru99.com/hive-queries-implementation.html](http://www.guru99.com/hive-queries-implementation.html) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **APPLIEDSTATISTICS** | **Category** | **L** | **T** | **P** | **Credits** |
| **Theory** | **56** | **4** | **-** | **4** |

 **Preamble** Thiscoursepresents thedifferent aspectsofstatistics likeANOVA, designofanexperiments, control charts,multivariateanalysis,Correlationandregression,whichenablesthestudenttoanalyze,

organize,presentandinterpret dataeffectively.

## CourseOutcomes

Uponthesuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1 | Understandthebasicconceptsofanalysisofvariance | **K2** |
| CO2 | Applytheconceptofpartial,multiplecorrelationandregressionforsolvingproblems | **K3** |
| CO3 | Applystatisticalinferenceandcontrolchartstoapplications | **K4** |
| CO4 | Analyseandinterpretdatausingtechniqueslikemultivariatestatistics | **K5** |
| CO5 | Designtheexperimentsusing methodslikeRandomizedblockandLatinsquares | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | S | S |
| CO2 | S | S | S | M |
| CO3 | S | S | M | M |
| CO4 | M | S | S | M |
| CO5 | S | S | S | M |

S-Strong;M-Medium;L-Low

## Syllabus UNIT I (11 Hrs)

Analysisofvariance:OnewayandTwowayanalysis:Introduction–Singlefactorexperimentand

linear statistical model – Fixed effects mode and ANOVA – Random effects model and ANOVA – Computation for sum of squares – Multiple comparison test : Grouping of means – Single factor experimental and linear statistical model – Fixed effective model for two way ANOVA – Random effective model for two way ANOVA – Computation for sum of squares.

## UNITII (11Hrs)

Introduction – Randomized block design- Advantages of a completely randomized experimental design– Latin squares– SignificanceofLatinsquares – Assumption inthe AnalysisofLatinsquares – Randomized block Vs. Latin squares – Latin Cubes – Factorial experiment.

## UNITIII (11Hrs)

Introduction – Statistical quality control: Relation Between Confidence Limit and control limit- Types ofControlchart’s – Controlcharts for variables*X* -chart, R-chart, S-chart, X-chart – Control chartforattributes:p-chart,C-chart,–Outofcontrolsituationsincontrolchartandprocess

monitoring – Process capability and process capability index – Six sigma: Six sigma metrics- Sigma Levels and Process Capabilities.

Usecases:ProcessControlandProductControlinIndustry

## UNITIV (11Hrs)

Multivariate Analysis – Basic concepts – Measurement Scales – Measurement Error – Statistical significance – Types of multivariate techniques – guidelines for multivariate analyses and interpretation – structured approach to multivariate model building – preparing data for amultivariate analysis – graphical examination of the data – missing data and its approaches – methods of detection of outliers – testing the assumptions of multivariate analysis

## UNITV (12Hrs)

Correlation ratio: -Multiple and partialcorrelation:Yule’s Notation-Plane of regression-Properties of residuals :Variance of residual-Coefficient of multiple correlation:Properties of multiple correlation coefficient-Coefficient of partial correlation-Multiple correlation in terms of total and partial correlation-Multiple correlation in terms of total and partial correlation- Multiple Regression Analysis-Multiple Regression-Expression for regression coefficient in terms of regressioncoefficient coefficients of lower order.

## ReferenceBooks

1. Johnson R.A. and Wichern,W (2001): Applied Multivariate Statistical Analysis, Fifth edition, Prentice Hall.
2. S.P.Gupta:StatisticalMethods,ThirdRevisedEdition2004.
3. Dr.J.Ravichandran,ProbabilityandStatisticsforEngineers,Wiley2015.

**Pedagogy:**Lectures**,**Demonstrations,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATAMINING** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course presents the basic concepts of data mining and data warehousing, various data mining techniques like classification, clustering, association rule mining. The course also introducesvarious applications of data mining, data visualization using Tableau

## Prerequisite

* + DatabaseManagement Systems
	+ ProbabilityandStatistics

## CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1.** | Understandbasicconceptsofdatamininganddatawarehousing | **K2** |
| **CO2.** | Applydataminingtechniquestocarryoutsimpledataminingtasks | **K3** |
| **CO3.** | Examinedataminingtechniqueslikeclassifications,clustering, | **K4** |
| **CO4.** | ComparedataminingalgorithmsusingTools | **K5** |
| **CO5.** | Designsolutionswithdatavisualizationusing Tableau | **K6** |

**MappingwithProgramOutcomes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| **CO1** | S | S | S | S |
| **CO2** | S | S | S | S |
| **CO3** | S | S | S | S |
| **CO4** | S | S | S | S |
| **CO5** | S | S | S | S |

S-Strong;M-Medium;L-Low

**Syllabus**

## UNITI (11Hrs)

Introduction to data mining: Mining fromdatabase - Data mining functionalities - Mining patterns - Classification of data mining systems - Major issues in Data mining. Data warehouse - Multi dimensional Data Model - Data warehouse Architecture -Data warehouse Implementation - Data warehousing to Data Mining.

## UNITII (11Hrs)

Data Preprocessing: Need for preprocessing – Data summarization – Data cleaning – Dataintegration - Data transformation – Data reduction – Data discretization.

## UNITIII (12Hrs)

AssociationRuleMining:Basicconcepts-Frequent Itemset MiningMethods-PatternEvaluation Methods –DensityBased Clustering- Grid-Based Methods – Clustering High-DimensionalData – Evaluation of clustering

## UNITIV (11Hrs)

Outlier Detection:Outliers and Outlier Analysis - Outlier Detection Methods -Proximity-Based Approaches - Outlier Detection in High-Dimensional Data – Data visualization : Foundations for building visualizations-visualizing data – working with data in Tableau – Moving from foundational to Advanced visualizations.

## UNITV (12Hrs)

MiningComplexData-Time-Series,Symbolic Sequences,and BiologicalSequences–Graph mining

* Applications of Data Mining: Financial Data Analysis - Retail and Telecommunication Industries - Intrusion Detection and Prevention-Recommender Systems- Data mining applications in functional areas like Banking, Marketing, Stock Market, Retail Marketing.

## TextBook

* 1. JaiweiHan,MichelineKamber(2006).DataMining-conceptsandtechniques,2/e, Morgan Kaufmann Publishers, San Francisco
	2. JoshuaN.Milligan(2015),LearningTableau,PacktPublishing

## ReferenceBooks

1. MarkA.Hall,IanH.Witten,EibeFrank(2011).DataMining:PracticalMachineLearning
2. ToolsandTechniques,3/e,MorganKaufmannPublishers,SanFrancisco
3. ArunK.Pujari(2001).DataMiningTechniques;UniversitiesPress, Hyderabad
4. SomanKP(2005).Dataminingfromtheorytopractice,2/e,PHILearningPvt. Ltd., NewDelhi

**Pedagogy:**Lectures**,**Demonstrations,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **RESEARCHMETHODOLOGY** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **30** | **-** | **-** | **2** |

**Preamble** Thiscoursepresentstheconceptsofresearch,typesofresearch,researchdesign,literature review, writing reports and adhering to research ethics.

## Prerequisite

This course is most appropriate for post graduate students who are interested in research but do not have prior research experience.

## CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1.** | Understandtheconceptsofresearch,researchdesign,researchprocessformeasurement,scalinganddatacollection | **K2** |
| **CO2.** | Applystatisticalmethodsforprocessingtheresearchdata. | **K3** |
| **CO3.** | Examinedifferentmethodsofdesignandresearch approaches. | **K4** |
| **CO4.** | Analyzetheprocessofhypothesistestingandreportwriting | **K5** |
| **CO5.** | Designsolutionsforresearchproblemsinaresponsibleandethicalmanner | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | M | M | M |
| CO3. | S | S | S | M |
| CO4. | S | S | S | M |
| CO5. | S | S | S | S |

S-Strong;M-Medium;L-Low

##  Syllabus UNIT I (6 Hrs)

AnIntroduction:–Meaningofresearch–Objectiveofresearch–Typesofresearch–Research

approaches – Significance of research – Research methods versus Methodology – Research and scientific method – Research process.

Defining the research problem: What is research problem? – Selecting the problem – Necessity of defining the problem. Research Design: Meaning of research design – Need for research design – Features of a good design – Different research designs.

## UNITII (6Hrs)

Measurement and scaling techniques: Measurement in research – Measurement scales – Sources of errorinmeasurement –Testsofsound measurement –Techniqueofdeveloping measurementtools – scaling – Meaning of scaling – Scale classification bases.

MethodsofData Collection: Collectionofprimarydata – Observation method –Interview method – Collection of data through schedules - Collection of secondary data – Selection of appropriate method for data collection.

## UNITIII (6Hrs)

Processing and Analysis of Data: Processing operations – Some problems in processing – Elements/Types of Analysis - Statistics in research – Measures of central tendency – Measures of dispersion – Measures of Asymmetry – Measures of relationships

## UNITIV (6Hrs)

Testing of Hypotheses: Introduction to Hypothesis – Basic concepts concerning testing of Hypotheses - Procedure forHypothesis testing –Measuring the powerofa Hypothesis testing -Tests of Hypothesis. Interpretation and report writing: Meaning of interpretation – Technique of interpretation – Significance of report writing – Different steps in writing report – Layout of the research report – Types of reports.

## UNITV (6Hrs)

Research Ethics **–**Brief history and analytical basis of research ethics, responsible conduct in research (Honesty in Science: Integrity, Authorship, Conflicts of Interest, Privacy and Confidentiality,InformedConsent,Risk/Benefit Assessment),The legalregulationofresearchethics in India (From UGC, MHRD and other governing agencies), Regulatory requirements relevant to international research.

## TextBook

Kothari, C.R (2013), Research Methodology – Methodsand Techniques, 2/e.Wiley Eastern Limited

## ReferenceBooks

1. R. Panneerselvam**(**2014), Research Methodology**,** 4/e. Prentice Hall India Learning Private Limited**.**
2. Ranjit Kumar(2019),ResearchMethodology–Astep-by-stepguide forbeginners, 5/e. Pearson Education.
3. DeepakChawlaandNeenaSondh(2011),ResearchMethodology, ConceptsandCases, Vikas Publishing House Pvt. Ltd.
4. OnBeingaScientist,AGuidetoResponsibleConductinResearch:ThirdEdition(2009)
5. RoleoftheEthicsCommittee:HelpingToAddressValueConflictsorUncertainties Author links open overlay panel Mark P.Aulisio, Robert M.Arnold.
6. https:/[/www](http://www.glos.ac.uk/docs/download/Research/handbook-of-principles-and-).[glos.ac.uk/docs/download/Research/handbook-of-principles-and-](http://www.glos.ac.uk/docs/download/Research/handbook-of-principles-and-) procedures.pdf.
7. ResearchRegulatoryCompliance1st Edition(Mark Suckow,BillYateseBookISBN: 9780124200654)
8. RecentresearchethicspolicyfromGovernmentofIndia.

**Pedagogy:**Lectures,Demonstrations,CaseStudies,GroupDiscussions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **BUSINESSANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **30** | **-** | **-** | **2** |

**Preamble** This course provides the fundamentals of domain knowledge in the functional areas of Data Analytics. Various important functional areas such as Health care analytics, Banking and Finance, Telecommunication and Retail Analytics are elaborated in this course. Few Use cases are also covered.

## Prerequisite

* FoundationsofData Science
* Bigdata framework

## CourseOutcomes

Uponsuccessfulcompletionofthiscourse,studentsshould beable to:

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| **CO1.** | Compare variousdomainareasandtheirchallenges | **K2** |
| **CO2.** | Applythe conceptsofanalytics tomakebetterdecisions | **K3** |
| **CO3.** | Examineusecasesfordifferent domains. | **K4** |
| **CO4.** | Evaluatethechallenges facedinvariousdomainsandchooseappropriateanalyticssolutions inalldomains | **K5** |
| **CO5.** | Proposesuitableanalyticssolutionsasrequiredbytheusecases. | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | M | M | M |
| CO3. | S | S | M | M |
| CO4. | S | M | S | M |
| CO5. | S | S | S | M |

S-Strong;M-Medium;L-Low

## Syllabus UNIT I (6 Hrs)

Healthcareanalytics–Introduction-Potentialcontributions-Challengesofhealthcareindustry-

currentandfuturestateofhealthcareanalytics–tophealthcareanalyticsadaptations

## UNITII (6Hrs)

Banking and Finance: Systems of Banking – Commercial Banking – New Financial Services:Overview of Analytics in Insurance: Key Insurance Analytics – Challenges – Health Insurance Analytics, Life Insurance Analytics- Types of Insurance – Housing Finance.

## UNITIII (6Hrs)

Telecommunication:Introduction-End-UserNeedsandDemands-TelecomBusiness

## UNITIV (6Hrs)

Retailanalytics–Understandingthenewconsumer–Marketinginaconsumer-drivenera - Managing the brand to drive loyalty

## UNITV (6Hrs)

Casestudies:Walmart,Netflix,Facebook,Uber,Amazon,Kaggle

## ReferenceBooks

1. DwightMcNeill(2013).AFrameworkforApplyingAnalyticsinHealthcare:WhatCanBe Learned from Best Practices in Banking , Retail, Politics and Sports, Pearson Education
2. GomezClifford(2011).Bankingand FinanceTheoryLawandpractice,PHI Learning
3. PatriciaL.Saporito(2014).AppliedInsuranceAnalytics: AFrameworkforDrivingMoreValue from Data Assets, Technologies and Tools, Pearson Education LTD.
4. AndersOlsson(2004).UnderstandingChangingTelecommunications,WileyPublications
5. Jennifer LeClaire, Danielle Dahlstrom, Vivian Braun. Business analytics in Retail for dummies,2 nd IBM Limited edition
6. AlistairCroll(2013)Leananalytics:UseDatatoBuildaBetterStartupfaster,OReilly Publishers
7. BernardMarr(2016).BigDatainPractice–How45successfulcompaniesusedbigdataanalytics to deliver extraordinary results, Wiley Publications
8. PurbaHaladyRao(2013).BusinessAnalytics.AnapplicationFocus,PHILearningprivateltd.

**Pedagogy:**Lectures,GroupDiscussions,Demonstrations,Casestudies

|  |  |
| --- | --- |
| **EVALUATION** |  |
| CIAI | - | 5 |
| CIAII | - | 5 |
| Case studyDiscussion | - | 10 |
| Mini-Project | - | 20 |
| ExternalVivaVoce | - | 10 |
| **Total** | - | **50** |

**QuestionpaperpatternforCA**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bloom’s****Category** | **Section** | **Marks** |  | **Total** |
| K3,K4 | A–5 outof8 x 8 marks | 40 | 300 words | 50 |
| K4,K5 | B–1 outof2x 10marks | 10 | 500 words |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATAANALYTICSLABIII** | **Category** | **L** | **T** | **P** | **Credit** |
| **Practical** | **-** | **-** | **75** | **3** |

**Preamble** This course provides implementation of the Hadoop components like Hive and Spark. This coursealso provides various exercises to implement the components in the distributed environment through mapreduce programming.

## Prerequisite

* + - Bigdata framework
		- NoSQLconcepts

## CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| **CO1.** | UnderstandHadoopcomponentsforbigdataprocessing. | **K2** |
| **CO2.** | ApplysimpleprocessingoperationsinSpark. | **K3** |
| **CO3.** | Applyconceptstotransfervariousfile formatsintoHive forProcessing. | **K3** |
| **CO4.** | DevelopSparkSQLfordataprocessing | **K6** |
| **CO5.** | CreateHivecommandsforbigdataapplicationsusingHadoop. | **K6** |

**MappingwithProgrammeOutcomes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **COs** | **PO1** | **PO2** | **PO3** | **PO4** |
| **CO1.** | S | S | S | M |
| **CO2.** | S | S | S | M |
| **CO3.** | S | S | S | M |
| **CO4.** | S | S | S | M |
| **CO5.** | S | S | S | M |

S-Strong;M-Medium;L-Low

## Syllabus

* + - ExercisestoimplementHiveQLtosort,order,group,distributeandcluster.
		- ExercisestoimplementpartitioningandbucketinginHive.
		- Exercisestocreatejoins,viewsand indexesinHive.
		- ExercisestotransferthecontentsofXML,JSONandORCfilesinto Hive for processing.
		- ExercisestoimplementsimpleprocessingtasksinSpark
		- ExercisestoimplementbasicoperationsinSparkSQL.

**Pedagogy:**Demonstrations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **DATAMININGANDVISUALIZATIONLAB** | **Category** | **L** | **T** | **P** | **Credits** |
| **Practical** | **-** | **-** | **75** | **3** |

**Preamble** This course provides exercises to implement data mining techniques such as classification, clustering, associationrule mining, text mining using R and also introducesdata visualizationusing Tableau.

## Prerequisite

* + Datamining
	+ Datavisualizationconcepts

## CourseOutcomes

Onsuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1.** | Constructprogramsforimplementedtheassociationrulemining,classification,clustering,PredictionalgorithmusingR | **K3** |
| **CO2.** | Applydataminingtechniquestorealworldproblem | **K3** |
| **CO3.** | Analyzedatausingdatavisualizationtoolandprovideinterpretation. | **K4** |
| **CO4.** | EvaluatethefeaturesofvariousdataminingandvisualizationtoolsusingTableau. | **K5** |
| **CO5** | Developalgorithmbasedsolutionsandvisualizationsforthegivenproblems. | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| **CO1.** | S | S | M | M |
| **CO2.** | S | S | M | M |
| **CO3.** | S | S | M | M |
| **CO4.** | S | S | S | M |

S-Strong;M-Medium;L-Low

## Syllabus

* + - ExercisetoimplementtheconceptofclassificationinR
		- Exercisetoimplementtheconceptofclustering inR
		- Exercisetofindassociateditemsindatasetusing R.
		- ExercisetoperformtextclassificationusingthemoviereviewdatasetinR.
		- ExercisetocreateacorpusofdocumentsandpreprocesstheminRusingstemming,stop word removal, whitespace removal, convert them to lowercase and remove punctuations.
		- ExercisetocreateatermdocumentmatrixforacorpusinR.
		- ExercisetofindthefrequenttermsinadocumentandremovesparsetermsinR
		- ExercisestovisualizedatausingBarchart,Linechart,Piechart,ScatterplotandHistogram.
		- ExercisestocreateDashboard, analyticsreportforadataset.
		- Exercisesto createStorybycombining worksheets/dashboards.

**Pedagogy:**Lectures**,** Demonstrations

## SEMESTERIV

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **PROJECT WORKANDVIVAVOCE** | **Category** | **L** | **T** | **P** | **Credits** |
| **Project** | **-** | **-** | **-** | **8** |

**Preamble**

Tobuildproblemsolvingabilityandtechnicalskillsthroughtheapplicationoftheoreticalconceptsfor modeling the real world problems using latest technologies in data analytics.

## PreRequisited

* + Domainknowledge
	+ DataSet Description
	+ Machinelearningtoolsfor algorithms.
	+ Technologies

## Outcomes

Onsuccessfulcompletionoftheproject,thestudentswillbeableto

* + Designsolutionsusingdata setsforrealworldproblems
	+ Developsolutionsinvariousdomainsforcontemporaryscenarios

## Methodology

* + Identifythedomainandproblem
	+ Decidethesoftwaretooltocarryoutthework
	+ Designanddevelopthesolution
	+ Analyzeandinterpretthefindings
	+ Documentthework

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ADVANCEDLEARNERCOURSE-1 CLOUD COMPUTING** | **Category** | **L** | **T** | **P** | **Credits** |
| **Theory** | **-** | **-** | **-** | **2** |

**Preamble** Thiscoursecoversvariousservicesofcloudcomputing.Italsopresentscloudcomputing collaborations and applications. It covers new concept of virtualization and applications

## Prerequisite

* + WebTechnology

## CourseOutcomes

Onsuccessfulcompletionofthecourse,studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge****Level** |
| **CO1.** | UnderstandtheconceptsofcloudArchitectureanditsservices | **K2** |
| **CO2.** | Learntoidentifydifferentservicesproviders,itsservicesandtools. | **K3** |
| **CO3.** | Analyzevariouswebbasedapplicationsforcollaboratingeveryone inthecloudcomputing. | **K4** |
| **CO4.** | Evaluatethebestserviceproviderforcloudcomputingintermsofstorage,services | **K5** |
| **CO5.** | Choosefromvariousindustrialplatformsfor development | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| **CO1.** | S | M | M | M |
| **CO2.** | S | M | M | M |
| **CO3.** | S | S | M | M |
| **CO4.** | S | M | S | M |
| **CO5.** | S | S | S | M |

S-Strong;M-Medium;L-Low

##  Syllabus UNIT I

Introduction:BenefitsandLimitations-CloudArchitecture–Storage–Services–ServiceProviders

-TypesofCloudServiceDevelopment–ServicesandTools

## UNITII

Collaborating on Contact Management - Collaborating on Project Management- Collaborating on Word Processing, Spreadsheet, Presentations, Databases- Sharing Files and Photographs

## UNITIII

Cloud Virtualization Technology –Virtualization Defined–Virtualization Benefits–Server Virtualization – Virtualization for x86 Architecture – Hypervisor Management Software – Logical Partitioning – VIO Server – Virtual Infrastructure Requirements

## UNITIV

Deep Dive: Cloud Virtualization –Introduction - Storage Virtualization–Storage Area Networks– Network Attached Storage – Cloud Server Virtualization – Virtualized Data Center - Classification of Data center - Overview of Datacenter environment

## UNITV

Industrial platforms and new developments - Amazon web services: Compute services -Storage services - Communicationservices - Additionalservices - Google AppEngine: Architecture and core concepts - Application life cycle - Cost modelMicrosoft Azure: Azure coreconcepts - SQLAzure - Windows Azure platform appliance - Cloud Applications: Scientific Applications – Business and Consumer applications

## ReferenceBooks

1. MichaelMiller(2011).CloudComputing: Web-Based ApplicationsThat Changethe WayYou Work and Collaborate Online, Pearson publication.
2. Dr.KumarSaurabh(2011).CloudComputing:InsightsintoNewEraInfrastructure,WileyIndia
3. Rajkumar Buyya,Christian Vecchiola,S. ThamaraiSelvi (2013).Mastering Cloud Computing Foundations and Applications Programming, Morgan Kaufmann is an imprint of Elsevier
4. RishabhSharma(2014).CloudComputing:Fundamentals,IndustryApproachandTrends,Wiley India edition.
5. PaulMehner(2013).CloudComputingwiththewindowsAzurePlatform,MicrosoftPress US.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ADVANCEDLEARNERCOURSE–2 CYBERSECURITY** | **Category** | **L** | **T** | **P** | **Credits** |
| **Theory** | **-** | **-** | **-** | **2** |

**Preamble** Thiscourseprovidestheclassificationofcybercrime, Botnets, attacksonthemobiledevices, tools and methods used in cybercrime, laws of cybercrime and cyber forensics.

## Prerequisite

* + BasicfundamentalknowledgeofNetworking
	+ WebApplication
	+ MobileApplication
	+ RelationalDatabaseManagement System

## CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| **CO1** | Understandthebasicconceptsofcybercrimeandcyberforensics | **K2** |
| **CO2** | Applymethodsofthecybercrime | **K3** |
| **CO3** | Analyzethetechniquesused forcybercrimeand forensics. | **K4** |
| **CO4** | Evaluatetoolsandmethodsappliedincybercrime | **K5** |
| **CO5** | Createmethodologiesto securedataintherealworld | **K6** |

**MappingofCourseOutcomewithProgrammeOutcome**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | S | M | M |
| CO2. | S | S | M | M |
| CO3. | S | S | S | S |
| CO4. | S | S | S | S |
| CO5. | S | S | S | S |

**S-Strong;M-Medium;L-Low**

**Syllabus Unit I (11 hrs)**

Introduction to Cybercrime:Introduction, Classifications of Cybercrimes: E-Mail Spoofing, Spamming, Cyber defamation, Internet Time Theft, Newsgroup Spam/Crimes from Usenet Newsgroup, Industrial Spying/Industrial Espionage, Hacking, Online Frauds, PornographicOffenses, Software Piracy, Password Sniffing, Credit Card Frauds and Identity Theft. Cyber offenses: How Criminals Plan that attack, Categories of Cybercrime, How Criminals Plan the Attacks: Passive Attack, Active Attacks, Scanning/Scrutinizing gathered Information, Attack (Gaining and Maintaining the System Access), Social Engineering, Cyber stalking, Cyber cafe and Cybercrimes, Botnets: The Fuel for Cybercrime, Attack Vector and Cloud Computing

## UnitII (12hrs)

Cybercrime- Mobile and Wireless Devices: Introduction, Proliferation of Mobile and Wireless Devices,TrendsinMobility,CreditCardFraudsinMobileandWirelessComputingEra,Security

Challenges Posed by Mobile Devices, RegistrySettings for Mobile Devices, Authentication Service Security, Attacks on Mobile/Cell Phones, Mobile Devices: Security Implications for Organizations, Organizational Measures for Handling Mobile, Organizational Security Policies and Measures in Mobile Computing Era and Laptops.

## UnitIII (11hrs)

Tools and Methods Used in Cybercrime: Introduction, Proxy Servers and Anonymizers, Phishing, Password Cracking, Keyloggers and Spywares, Virus and Worms, Trojan Horses and Backdoors, Steganography, DoS and DDoS Attacks, SQL Injection, Buffer Overflow, Attacks on Wireless Networks. Phishing and Identity Theft: Introduction, Phishing, Identity Theft (ID Theft): Types of Identity Theft, Techniques of ID Theft, Identity Theft Countermeasures, How to Protect your Online Identity.

## UnitIV (11hrs)

Cybercrimes and Cybersecurity: The Legal Perspectives Introduction, Why Do We Need Cyberlaws: The Indian Context, The Indian IT Act, Challenges to Indian Law and Cybercrime Scenario in India, Consequences of Not Addressing the Weakness in Information Technology Act,, Amendments to the Indian IT Act, Cybercrime and Punishment, Cyberlaw, Technology and Students: Indian Scenario.

## UnitV (11hrs)

sIntroduction, HistoricalBackgroundof Cyber forensics, DigitalForensics Science, TheNeedfor Computer Forensics, Cyber forensics and Digital Evidence, Forensics Analysis of E-Mail : RFC282, Digital Forensics LifeCycle, Chainof CustodyConcept, NetworkForensics, Approachinga Computer Forensics Investigation, Setting up a Computer Forensics Laboratory: Understanding the Requirements, Computer Forensics and Steganography, Relevance of the OSI 7 Layer Model to Computer Forensics, Forensics and Social Networking Sites: The Security/Privacy Threats, Challenges in Computer Forensics, Special Tools and Techniques, Forensics Auditing and Anti forensics.

## Pedagogy

Lectures,GroupDiscussions,Demonstrations

NinaGodbole,SunitBelapur,“CyberSecurityUnderstandingCyberCrimes,Computer

**TextBooks**

ForensicsandLegalPerspectives”,WileyIndiaPublications,April, 2013

## ReferenceBooks

1. JamesGraham,RicharHoward,RyanOlson,“CyberSecurityEssentials”,CRCPress,

TailorandFrancisGroup,2013

1. RobertJones,“InternetForensics:UsingDigitalEvidencetoSolveComputerCrime”, O’Reilly Media, October, 2005
2. ChadSteel,“WindowsForensics:Thefieldguideforconductingcorporatecomputer investigations”, Wiley India Publications, December, 2006

.

# ELECTIVES

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **INFORMATIONRETRIEVAL** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** Thiscoursepresentstheconceptsofdocumentrepresentation,documentindexing,digital information storage,retrieval anddistribution.It alsointroduceseffectivesearch strategiesforIR

systems,vectorspacemodel,textclassificationandevaluationmethodsofIRsystems.

## Prerequisite

* + DatabaseManagement systems
	+ Datamining

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1 | Demonstratetheconceptsofdocumentrepresentation,documentindexing,digitalinformationstorage, retrievalanddistribution | K2 |
| CO2 | Applytheconceptsofvectorspacesandclassifierstoperformdocumentclassification. | K3 |
| CO3 | Examinetheadvantagesanddisadvantagesofdifferentinformation-retrievalmodels | K4 |
| CO4 | DeterminetheeffectivesearchstrategiesforIRsystems | K5 |
| CO5 | Abletodevelopinformationretrievalsystemforspecificusecases. | K6 |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | M | M | M |
| CO3. | S | S | M | M |
| CO4. | S | M | S | M |
| CO5. | S | S | S | M |

S-Strong;M-Medium;L-Low

**Syllabus UNIT I (11Hrs)**

Booleanretrieval:Informationretrievalproblem-ProcessingBooleanqueries-Booleanmodel

versus ranked retrieval. The term vocabulary and postings lists: Document delineation andcharacter sequence decoding - Determining the vocabulary of terms - Faster postings list intersection via skip pointers - Positional postings and phrase queries

**UNITII (12Hrs)**

Dictionaries and tolerant retrieval: Search structures for dictionaries - Wildcard queries - Spelling correction-Phoneticcorrection.Index construction:Hardwarebasics-Blockedsort-basedindexing

- Single-pass in-memory indexing - Distributed indexing - Dynamic indexing - Other types of indexes

**UNITIII (11Hrs)**

Scoring, termweighting and the vector space model: Parametric and zone indexes - Term frequency and weighting - The vector space model for scoring. Evaluation in information retrieval:

InformationRetrievalsystem–StandardtestCollection–Evaluationofunrankedretrievalsets–

Evaluationofrankedretrievalresults–Assessing relevance

**UNITIII (11 Hrs)**

XML retrieval : Basic XML concepts – A vector space model for XML retrieval – Evaluation of XML retrieval - Text-centric vs.data-centric XML retrieval.Textclassification and Naive Bayes: Thetextclassificationproblem-NaiveBayestextclassification-PropertiesofNaiveBayes- Feature selection - Evaluation of text classification

**UNITV (11Hrs)**

Vector space classification: Document representations and measures of relatedness in vector spaces– Rocchio classification -Flat clustering: Clustering in information retrieval -Evaluation of clustering- K-means–Websearchbasics- Webcharacteristics- Advertisingastheeconomic model – Search user experience – Basic Page Rank

**Text Book** ChristopherD.Manning,PrabhakarRaghavan,HenrichSchutze (2008).Introductionto Information Retrieval, 1/e; New York: Cambridge University Press

## ReferenceBooks

1. StefanButtcheret.al(2012).InformationRetrieval-ImplementingandEvaluating, MITPress
2. DrRicardoBaeza-Yateset.al(2011).ModernInformationRetrieval:TheConceptsand Technology, Addison Wesley
3. DavidA.GrossmanandOphirFrieder(2010).InformationRetrieval,2/e,UniversitiesPress

**Pedagogy:**Lectures,Demonstrations,GuestLecture,VideoLectures

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **INFORMATIONSECURITY** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course covers the concepts of information security, ethical hacking, policies standards and security practices,riskmanagement,implementation and maintenance processes.It alsodeals with

managerial, technical aspects, physical security and cryptographic techniques of informationsecurity.

## Prerequisite

* + Cryptography
	+ Internet,firewalls,attacks,andthreats

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1. | UnderstandtheneedofInformationSecurity, polices,standards and security blue print of an organization | K2 |
| CO2. | Applycryptographicalgorithms inrealtime applications | K3 |
| CO3. | Analysedifferenttypesofphysicalsecuritytechniques | K4 |
| CO4. | Assessthebehaviorofdifferentthreatsand attacks | K5 |
| CO5. | Proposesolutionsforcybersecurityissuesusingvariousmodels. | K6 |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | M | M | M |
| CO3. | S | M | M | M |
| CO4. | S | S | M | M |

S-Strong;M-Medium;L-Low

## Syllabus

**UNITI (11Hrs)**

Introduction to Information Security: Introduction- - Security - Critical Characteristics ofInformation - NSTISSC Security Model - Components ofan Information System - Security Components- Approaches to Information Security Implementation - The Systems Development LifeCycle - The Security Systems Development Life Cycle

## UNITII (12Hrs)

The Need for Security: Business Needs First - Threats - Attacks - Secure Software Development. Risk management: Overview - Risk Identification - Risk Assessment- Risk Control Strategies - Selecting a Risk Control Strategy - Quantitative Versus Qualitative Risk Control Practices -PlanningforSecurity:InformationSecurityPolicy,StandardsandPractices-TheInformation

SecurityBlue print

## UNITIII (11Hrs)

SecurityTechnology:FirewallsandVPNS -IntrusionDetection,AccessControl,andotherSecurity Tools - Intrusion Detection and Prevention Systems - HoneyPots - HoneyNets, and Padded Cell Systems - Scanning and Analysis Tools - Access Control Devices

## UNITIV (11Hrs)

Cryptography: Foundations of Cryptology- cipher methods- cryptographic algorithms – cryptographic tools. PhysicalSecurity: Physicalaccess controls – fire securityand safety– failure of supporting utilities and structural collapse – Interception of data – mobile and portable systems – special considerations for physical security threats

## UNITV (11Hrs)

ImplementationofInformationSecurity:InformationSecurityProjectManagement-Technical Topics of Implementation - Non Technical aspects of Implementation

InformationSecurityMaintenance:SecurityManagement Models - TheMaintenanceModel- Digital forensics

**TextBook** Michael E. Whitman and Herbert J. Ma ttord (2014). Principles of Information Security, 5/e,Cengage Learning,Indian edition

## ReferenceBooks

1. CharlesA.Sennewald(2020).EffectiveSecurityManagement,7**/**e,Elsevier
2. DhirenR.Patel(2008).InformationSecurity:TheoryandPractice,PrenticeHallofIndiaPvtLtd
3. S.M.Bhaskar,S.I.Ahson(2008).InformationSecurity:APracticalApproach,AlphaScience
4. GeraldL.Kovacich(2016).InformationSystemSecurityOfficer’sGuide(3/e),Butterworth Hinemann

## Pedagogy:

Lectures,CaseAnalysis,GroupDiscussions,Demonstrations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **WEBDATAANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** Thiscoursegives insightsabout leveraging webdatato achievestrategicbusinessobjectives. It deals with the various techniques for analysing web data like click stream analysis. The course also provides ways to execute competitive intelligence analysis and to analyze emerging social, mobile and video data.

## Prerequisite

* + FoundationsofData Science
	+ DataMining

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1. | Understand thetechniquesofwebdataanalytics | K2 |
| CO2. | Applywebdataanalyticsonsocial,mobileandvideodata | K3 |
| CO3. | Analyzetechniquesformeasuringthesuccessofawebsite | K4 |
| CO4. | Assessthevariouscasesto applywebdata analytics | K5 |
| CO5. | Proposenewmetricsbasedsolutionsforuserwebsites. | K6 |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | S | M | M |
| CO3. | S | M | S | M |
| CO4. | S | S | S | M |
| CO5. | S | S | M | M |

S-Strong;M-Medium;L-Low

## Syllabus

**UNITI (11Hrs)**

Introduction : Web Analytics 2.0 - Clickstream- multiple outcome analysis-experimentation and testing- voice of customer – competitive intelligence- the tactical shift -Optimal strategy forchoosing web analytics

## UNITII (11Hrs)

Clickstream analysis: Metrics-Eight critical web metrics-web metrics demystified –strategically aligned tactics for impactful web –Web analytics report-Foundational analytical strategies- clickstream analysis made actionable-challenges

## UNITIII (12Hrs)

Measuring Success-Actionable Outcome KPIs- Moving beyond conversion rates- Micro and macro conversion-Measuringsuccess foranon–ecommercewebsite- Leveragingqualitativedata:Surveys- Web enabled emerging user research options

## UNITIV (11Hrs)

A/B Testing - Multivariate testing-Actionable testing ideas-Controlled experiments-Competitive intelligence analysis-CI data source, types, secrets- website traffic analysis-Search and keyword analysis- audience identification and segmentation analysis

## UNITV (11Hrs)

Emerging analytics: Social. mobile, video: Measuring social web - the data challenge- analyzing mobile customer experiences-measuring the success of blogs- quantifying the impact of Twitter – Analyzing the performance of videos.

## TextBook

1. Avinash Kaushik (2010) , Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity, 1st Edition, Wiley Publishing.
2. DietmarJannach,MarkusZanker(2011),Recommendersystem-Anintroduction,Cambridge University Press
3. BingLiu(2012), SentimentAnalysisandopinionmining,MorganandclaypoolPublishing

## ReferenceBooks:

1. EricEnge,StephanSpencer,JessieStricchiola,TheArtofSEO:MasteringSearchEngine Optimization, 3rdEdition.
2. KristinaHalvors,ContentStrategyfortheWeb,1stEdition.

**Pedagogy:**Lectures, Demonstrations,GroupDiscussions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **SOCIALMEDIAANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** The course covers concepts and techniques for retrieving, exploring, visualizing, and analyzingsocial network and social media data. Students learn the key metrics to assess social media goals, perform social network analysis to apply social media analytics process and formulate effective strategies based on the analytics.

## Prerequisite

* + - FoundationsofData Science

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1 | Understandsourcesand limitationsofsocialmediadata. | **K2** |
| CO2 | Applysocialmediaanalyticsprocessandevaluatemetrics. | **K3** |
| CO3 | Examinedifferentsocialmediaplatformsandtheir associatedtools | **K4** |
| CO4 | Applysocialmediainformationtocreatedashboardsandreportsforvisualization. | **K5** |
| CO5 | Designeffectivestrategybasedonthesocialmedia analyticsdata | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | M | M |
| CO2 | S | S | M | M |
| CO3 | S | M | S | M |
| CO4 | S | S | S | M |
| CO5 | S | S | M | M |

S-Strong;M-Medium;L-Low

## Syllabus Unit I : (11 Hrs)

**FoundationforAnalytics**:–DigitalGap–SocialMediaDataSources–DefiningSocialMedia

Data – Data Sources – Estimated vs. Factual Data Sources – Data Gathering in Social Media Analytics. **From Data to Insights :**Actionable Analytics – Focus on objective – Plan to shape data to insights – Choosing a good analytics tool – Data Aggregation calculations and display – Data display – Social media and Big data – Potential Challenges. **Data Identification :**Professional networking sites- social sites – information sharing sites – microblogging sites – blogs /wikis.

## UnitII: (12Hrs)

**AnalyticsinsocialMedia:**Typesofanalytics.DedicatedVs.HybridTools–Dedicatedtools–

Hybridtools–DataIntegrationTools–BestSetup.**SocialNetworkLandscape:**ConceptandUX

on social networks – Interactivity of social network – Content flow on social network – Interaction Pattern between users – Social Media as a two way channel.

## UnitIII: (11Hrs)

**Analytics Process:** Analysis – Insight – Investigation beyond socialanalytics – Shaping a method – analysis cycle – Community Activity – Resources – Attention span – Dynamic cycles – ShortPeriods – Long Periods – Analyst Mindset – Instinctive Analyst. **Metrics:** Introduction – Defaultand custom metrics – Metrics Categories – Graph Types – Metric Capabilities – Metrics andStrategy – Estimated Metrics – Metrics and Tactics.

## UnitIV: (11Hrs)

**Dashboards:** Purpose and Objectives – Default Vs. Custom Dashboards – Linearity and order of metrics – Metrics Positioning and Correlation – Metric and dashboard layout – Graphic design – Data Integration dashboards. **Reports:** Elements of reporting – Reporting approaches and formats – Animation and effects in reporting – Stake holders and feedback – Reporting with teams.

## UnitV: (11Hrs)

**Strategy:** Strategyinsocialmediaanalytics –Strategicplanning –Dataavailabilityanddatasources – Knowledge beyond social media – Tools and technology preparation – Team Preparation – Goals and objectives – Contingency plans – application of social media analytics strategy – Strategy and tactics – Evaluation of a strategic analytics cycle.

**Case Studies :**Targeting the audience using Facebook Analytics, Tracking profile analytics in LinkedIn, Analysis of Political Tweets, ROI Analytics using Facebook, Marketing Strategy in Pinterest.

**TextBook** Alex Goncalves(2017). Social Media Analytics Strategy-Using Data to Optimize Business Performance. Alex Goncalves. APress

## ReferenceBooks

1. Ganis,Kohirkar(2016).SocialmediaAnalytics,IBMPressPTG,1stEdition
2. NancyFlynn(2012).TheSocialMediaHandbookPolicies,andBest Practices,Wiley

**Pedagogy:** Lectures,Demonstrations,GroupDiscussions,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **CUSTOMERANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course introduces the fundamental concepts of Customer data analytics. The course alsoprovides the ways to recognize customer life time value, enhance customer loyalty, develop and execute analytics plan on customer data.

## Prerequisite

* + FoundationsofData Science
	+ Bigdata framework

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| **CO1** | Understandthecustomeranalyticsandvisualizetheresults | **K2** |
| **CO2** | Applythe conceptsofanalytics tomakebetterdecisions | **K3** |
| **CO3** | Analysetheissuesincustomeranalytics | **K4** |
| **CO4** | Appraisevariousdatatypesincustomeranalyticsbeforeformulationofstrategies. | **K5** |
| **CO5** | Proposesolutionsforthevariouscasesincustomer analytics | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | M | M |
| CO2 | S | S | M | M |
| CO3 | S | M | S | M |
| CO4 | S | S | S | M |
| CO5 | S | S | M | M |

S-Strong;M-Medium;L-Low

## Syllabus

**UNITI: (11 Hrs)**

Understanding Customer Analytics - Mining Data for Insights - Visualizing Your Results -Acquiring and Segmenting - Valuable Customers - Acquiring the Ideal Customer - Using Segmentation Techniques - Identifying Customer Preferences through Micro Segmentation

## UNITII: (11Hrs)

Recognizing Customer Lifetime Value - Defining CLV - Looking at Best Practices for CLV -Getting Started with Predictive Modeling - Personalizing Cross-Sells and Upsells

## UNITIII: (11Hrs)

Enhancing Customer Loyaltyand Retention - Coping withCustomer Churn - Increasing Customer Retention- Operationalizing Analyticsto MakeBetterDecisions - Growing Customer Loyaltyand Advocacy- ExtractingValueOutofSocialMedia - SeeingSocialAnalytics inAction- Employing Sentiment Analysis - Incorporating All Data Types

## UNITIV: (12Hrs)

TenKeyWaysto GetStartedwithCustomerAnalytics - StartwiththeStrategicEnd inMind- Ask the Right Questions - ClarifyYour Objectives - Choose YourSuccessMetrics - Secure Leadership and IT Buy-inBefore You Start - Get the Right People for the Job - Evaluate Your Data - Pick the Right Tools to Complete the Project - Develop Your Analytics Plan - Execute

## UNITV: (11Hrs)

CustomerAnalytics:4PhaseProcess–Casestudy:CardPioneerMicroinsuranceDataAnalysis. Use Cases:Call Center Audio Analysis, Precision Medicine, Churn Prediction using Telecom database

## TextBook

1. Stephanie Diamond and Marygrace Bateman, Customer Analytics For Dummies, John Wiley & Sons, 2013
2. CGAPteam:YaninaSeltzer,LisaStahl,andGerhardCoetzee,“CustomerAnalyticsToolKit,

CustomerCentricityThroughAnalytics“,CGPA2017

## ReferenceBooks

1. CustomerAnalyticswithR,Gitbook.
2. CustomerAnalyticsforDummies,JeffSauro,JohnWiley&sons2015

**Pedagogy:**Lectures,Demonstrations,GroupDiscussions,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **MARKETINGANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** This course contains the strategies of marketing analytics, forecasting techniques, Customer Needs and retailing. This will also offer the market research tools involved in retailing.

## Prerequisite

FoundationsofData science

## CourseOutcomes

Onsuccessfulcompletionofthecourse,thestudentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| **CO1** | UnderstandtheconceptsofmarketingdataandeffectiveuseofMicrosoftExcel | **K2** |
| **CO2** | ApplyForecastingTechniquestoimproveresponseratesfor marketingcampaigns | **K3** |
| **CO3** | AnalyseMarketsegmentationbasedonCluster Analysis | **K4** |
| **CO4** | Estimatemethodsforprocessing marketingdatainpredictionand marketsegmentation. | **K5** |
| **CO5** | rmulatesolutionsinexcelforvarioususecasesofmarketinganalyticsmethods, | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | L | M |
| CO2 | S | L | L | M |
| CO3 | S | S | L | M |
| CO4 | M | M | S | M |
| CO5 | S | M | M | M |

S-Strong;M-Medium;L-Low

## Syllabus UNIT I (11Hrs)

MarketingDataSummarization-SlicingandDicingMarketingDatawithPivotTables-UsingExcel

ChartstoSummarizeMarketingData -UsingExcelFunctionstoSummarizeMarketingData.

## UNITII (12Hrs)

Forecasting Techniques: Simple Linear Regression and Correlation - Using Multiple Regression to Forecast Sales - Forecasting inthe Presence ofSpecialEvents -Modeling Trend and Seasonality- Ratio to Moving Average Forecasting Method - Winter’s Method - Using NeuralNetworksto Forecast Sales.

## UNITIII (11Hrs)

CustomerNeeds:ConjointAnalysis- LogisticRegression-DiscreteChoiceAnalysis–Customer Value

-IntroductiontoCustomer value,Benefits

## UNITIV (11Hrs)

Market segmentation: Cluster Analysis - User-Based Collaborative Filtering - Collaborative Filtering - Using Classification Trees for Segmentation.

## UNITV (11Hrs)

Retailing and market research tools : Retailing - Introduction to retailing, Market Basket Analysis andLift - Marketing Research Tools - Principal Components Analysis

## TextBook

1. Wayne.L.Winston(2014).MarketingAnalytics:DatadriventechniqueswithMS-Excel,WileyPublications

## ReferenceBooks

1. StephanSorger(2013). Marketing Analytics: Strategic models and metrics, CreateSpaceIndependent Publishing Platform
2. ChuckHemann,KenBurbary(2013).DigitalMarketing Analytics,PearsonPublication

**Pedagogy:**Lectures,Demonstrations,GroupDiscussions,Casestudies

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **IOTANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble** The course covers basic concepts for IOT Analytics, collection of data for IOT, Integration of IOT with Cloud, Big Data Environments. Students can learn about applying geospatial analytics and applying machine learning to the IOT data. The course also covers the organization ofthe IOT data, cost benefits of using IOT and review of IOT in various sectors.

## Prerequisite

* + - FoundationsofData Science
		- BasicsofIOT

## CourseOutcomes

Onthesuccessfulcompletionofthecourse, studentswillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| **CO1** | UnderstandIOTDataAnalytics, lifecycle,IOTbasedgeospatialanalyticsand machine learning application in IOT | **K2** |
| **CO2** | ApplyIOTconcepts inGeospatialanalyticsandMachinelearning | **K3** |
| **CO3** | ExamineconceptsofcloudbasedIOT, BigdataandIOTinvariousdomains. | **K4** |
| **CO4** | Appraisetechniquesandstrategiesfordatacollectionwithreferencetobigdata. | **K5** |
| **CO5** | ProposestrategiesfororganizationofIOTdataandoptimizecostbenefitsinusingIOTdata. | **K6** |

## MappingwithProgrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1 | S | M | M | M |
| CO2 | S | S | M | M |
| CO3 | S | S | S | M |
| CO4 | S | S | S | M |
| CO5 | S | S | S | M |

S-Strong;M-Medium;L-Low

## Syllabus Unit I (11Hrs)

IntroducingIOTAnalytics:Introduction–IOTDataandBigData–ChallengesofIOTanalytics

Applications – IOT analytics Lifecycle and Techniques.IOT Cloud and Big Data Integration – Cloud based IOT platform – Data Analytics for IOT – Data Collection – WAZIUP software Platform – Ikaas Software Platform. Searching the Internet of things: Introduction – Search architectures for social and physical sensors – local Event Retrieval – Use of Sensor meta data streams.

## UnitII: (12Hrs)

IOTDevicesand NetworkingProtocols:IOTdevices indifferent domains.IOTAnalytics fortheCloud – Building elastic analytics – elastic analytics concepts – designing for scale – Cloud security and analytics – AWS overview - AWS key services for IOT analytics. Thingworxoverview . Creating an AWS Cloud Analytics environment.

## UnitIII (11Hrs)

Strategies and Techniques in Data collection :Designing data processing for analytics – Applying big datatostorage–ApacheSpark forIOTdataprocessing.ExploringIOTData:Exploreandvisualize data – Tableau – Attribute identification – Solving industry specific problems

## UnitIV (11Hrs)

Geospatial Analytics to IOT Data : Basics – Vector and Raster based methods – Processing geospatial data. Data Science for IOT analytics – Machine learning basic – Forecasting IOT data using ARIMA – Deep learning with IOT data.

## UnitV (11Hrs)

Organize IOT data – Linked analytics datasets – Managing data lakes – data retention strategy for IOT data. Economics of IOT data – Cloud computing and open source – cost considerations – Revenue – Predictive maintenance.IOT review : IOT data flow – IOT exploratory analytics – IOT data science – Building revenue – Sample project.

**Use Cases:** Real time data analysis for manufacturing sector, IOT analytics for healthcare gamechanger.

## TextBook

AndrewMinteer(2017).AnalyticsfortheInternetofthings.PacktPublishing.

## ReferenceBooks

1. PrasantKumarPattnaik,RaghvendraKumar,SouvikPal,S.N.Panda(2020).IOTandAnalytics in Agriculture.Springer.
2. JohnSoldatos(Editor).BuildingblocksforIOTAnalytics:Internet-of-ThingsAnalytics.River Publisher Series in Signal Image and Speech Processing.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ARTIFICIALINTELLIGENCEAND ANALYTICS** | **Category** | **L** | **T** | **P** | **Credit** |
| **Theory** | **56** | **4** | **-** | **4** |

**Preamble:** Thecoursecoversanoverviewof implementationofAnalyticsandAIstrategiesinBusinessand provide details of disruption and transformation broughtin various domainslike Banking,Healthcare and Life sciences, Retail and Exponential technologies

## Prerequisite

* + BasicKnowledgeofAnalyticsandArtificialIntelligence
	+ Knowledgeindifferent domains

## Courseoutcomes:

Uponthesuccessfulcompletionofthecourse,studentwillbeableto

|  |  |  |
| --- | --- | --- |
| **CO****Number** | **CO Statement** | **Knowledge Level** |
| CO1. | Understand the need of Analytics and AI strategy in BusinessTransformation | K2 |
| CO2. | ApplyAnalyticsandAIstrategyindifferentdomains | K3 |
| CO3. | AnalysemitigationofFraudandcustomerretentionusingAIindifferentdomains | K4 |
| CO4. | Assessthebehaviorofdifferentcyber threatsandattacks | K5 |

## MappingwithprogrammeOutcomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cos** | **PO1** | **PO2** | **PO3** | **PO4** |
| CO1. | S | M | M | M |
| CO2. | S | M | M | M |
| CO3. | S | M | M | M |
| CO4. | S | S | M | M |

S-Strong;M-Medium;L-Low

##  Syllabus UNIT I (11 Hrs)

Analyticsand AIStrategyforBusinessTransfer:Re-engineeringBusinesstothinkAIandAnalytics – Robust Data Monetization Strategy – Accelerated Decision-making with Real-Time Analytics – Analytics as a Service Model – Analytics-Led Enterprise Transformation.

## UNITII (12Hrs)

Banking Industry Transformed by Analytics and AI :Redefining Banking Industry – AI powered Financial services – Fraud Mitigation through AI – Reorienting Customer Retention and Risk Management –AdvantageofAI inFintechCompanies–AI-DrivenTransformationsinInsurance – Adopting Digital Based Insurance Model.

## UNITIII (11Hrs)

Redefining Healthcare and Life Sciences :AI adoption in Healthcare – Real-world Evidence Based AnalyticsimprovingTreatment outcomes–LeveragingPatient andDrugsimilarityAnalytics –AI:A Boon to the Life Science Industry – Analytics and Genomics.

## UNITIV (11Hrs)

Analytics and AI in Retail :AI-powered shopping experience – Emergence of Smart Consumers – RecommendationEnginesforPersonalizingExperiences –EvolutionofSmart Retailers–Omnichannel Experiences – Fluid Supply Chain Transformation.

## UNITV (11Hrs)

Exponential Technologies underpinned by Analytics and AI :Beating Cyberattacks with Analytics – ConnectedCarTechnologyreshapingAutomotiveIndustry–IoTAnalytics–CryptocurrencyAnalytics – Chatbots – Redefining the Talent Landscape.

## TextBook

SameerDhanrajani, “AIandAnalytics-AcceleratingBusinessDecisions”,Wiley,2018.

## ReferenceBooks

1. StuartRusselandPeter Norvig,“ArtificialIntelligence– AModernApproach”,Pearson

EducationPress,2011.

1. KevinKnight,ElaineRich,B.Nair,“ArtificialIntelligence”,McGrawHill, 2008.
2. GeorgeF.Luger,“ArtificialIntelligence”,PearsonEducation,2001.
3. Nils J. Nilsson, “Artificial Intelligence: A new Synthesis”, MorganKaffman, 2002. **Pedagogy:** Lectures, Case Analysis, Group Discussions, Demonstrations