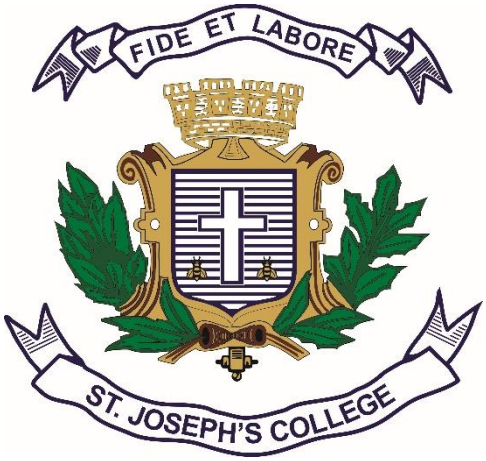
ST JOSEPH'S UNIVERSITY BANGALORE



Re-accredited with **'A++' GRADE with 3.79/4 CGPA** by
NAAC Recognized by UGC as College of Excellence

ST. JOSEPH'S INSTITUTE OF INFORMATION TECHNOLOGY

DEPARTMENT OF ADVANCED COMPUTING

SYLLABUS FOR UNDERGRADUATE PROGRAMME

SUMMARY OF CREDITS IN BCA (Data Analytics) - NEP
Revision Year - 2023

Department of Advanced Computing (UG)

Semester 1	**Code Number**	**Title**	**No. of Hours of Instructions per semester**	**Number of Hours of teaching per week**	**Number of credits**	**Max marks for SE - duration of examination**
Theory	BCADA1121	Fundamentals of computing and programming skills using C	45	03	03	60-2Hrs
Theory	BCADA1221	Exploratory Data Analysis	45	03	03	60-2Hrs
Theory	BCADA1324	Mathematics I	45	03	03	60-2Hrs
Theory		Language I	45	03	03	60-2Hrs
Theory		Language II	45	03	03	60-2Hrs
Theory	OE 01	Basics of Data Science	45	03	03	60-2Hrs
Theory	OE 02	Python Programming	45	03	03	60-2Hrs
Practical	BCADA1P1	C Programming Lab	40	04	02	25- 4 Hrs
Practical	BCADA1P2	Data Analysis Lab (Excel)	40	04	02	25- 4 Hrs
Total Number of credits:			**19**			

Semester 2	**Code Number**	**Title**	**No. of Hours of Instructions**	**Number of teaching hrs /week**	**Number of credits**	**Max marks for SE - duration of examination**
Theory	BCADA2121	Principles and Practices of Data Science	45	03	03	60-2Hrs
Theory	BCADA2221	Advanced Statistical Computing	45	03	03	60-2Hrs
Theory	BCADA2324	Mathematics II	45	03	03	60-2Hrs
Theory		Language I	45	03	03	60-2Hrs
Theory		Language II	45	03	03	60-2Hrs
Theory	OE 3	Machine Learning using R programming	45	03	03	60-2Hrs
Theory	OE 4	Digital design using HDL	45	03	03	60-2Hrs
Practical	BCADA2P1	Data Science Lab (Using R)	40	04	02	25- 4 Hrs
Practical	BCADA2P2	Statistical Computing Lab (Using R)	40	04	02	25- 4 Hrs
Total Number of credits:			**22**			

Semester 3	**Code Number**	**Title**	**No. of Hours of Instructions**	**Number of teaching hrs /week**	**Number of credits**	**Max marks for SE - duration of examination**
Theory	BCADA3122	Python for Data Analytics	45	03	03	60-2Hrs
Theory	BCADA3222	Econometrics	45	03	03	60-2Hrs

Theory	BCAD A3324	Mathematics III	45	03	03	60-2Hrs
Theory	OE 5	Business and Data Understanding	45	03	03	60-2Hrs
Theory	OE 6	Database Management System	45	03	03	60-2Hrs
Practical	BCAD A3P1	Python for Data Analytics Lab	40	04	02	25- 4 Hrs
Practical	BCAD A3P2	Econometrics Lab	40	04	02	25- 4 Hrs
Total Number of credits:					**16**	
Semester 4	**Code Number**	**Title**	**No. of Hours of Instructions**	**Number of teaching hrs /week**	**Number of credits**	**Max marks for SE - duration of examination**
Theory	BCAD A4122	Database Management System	45	03	03	60-2Hrs
Theory	BCAD A4222	OS using Linux	45	03	03	60-2Hrs
Theory	BCAD A4322	Multivariate Statistics	45	03	03	60-2Hrs
Practical	BCAD A4P1	Database Management System Lab	40	04	02	25- 4 Hrs
Practical	BCAD A4P2	Multivariate Statistics Lab	40	04	02	25- 4 Hrs
Theory	OE 7	Visualization techniques using Tableau	45	03	03	60-2Hrs
Total Number of credits:					**16**	
Semester 5	**Code Number**	**Title**	**No. of Hours of Instruction**	**Number of teaching**	**Number of credits**	**Max marks for SE - duration of**

			s	**hrs /week**		**examination**
Theory	BCAD A5123	Java Programming	45	03	03	60-2Hrs
Theory	BCAD A5223	Data Mining and Analytics	45	03	03	60-2Hrs
Theory	BCAD A5323	Cloud Computing	45	03	03	60-2Hrs
Theory	DSE-AC-01	Computer Networks	45	03	03	60-2Hrs
Practical	BCAD A5P1	Java Programming Lab	40	04	02	25- 4 Hrs
Practical	BCAD A5P2	Machine Learning Lab	40	04	02	25- 4 Hrs
Vocational Theory		AWS	45	03	03	60-2Hrs
Total Number of credits:					**19**	
Semester 6	**Code Number**	**Title**	**No. of Hours of Instructions**	**Number of teaching hrs /week**	**Number of credits**	**Max marks for SE - duration of examination**
Theory	BCAD A6123	Enabling Technologies for Data Science	45	03	03	60-2Hrs
Theory	BCAD A6223	Machine Learning	45	03	03	60-2Hrs
Theory	BCAD A6323	Optimization Techniques	45	03	03	60-2Hrs
Theory	DSE-AC- 02	AI & IOT	45	03	03	60-2Hrs
Practical	BCAD A6P1	Enabling Technologies for Data Science Lab	40	04	02	25- 4 Hrs

Practical	BCAD A6P2	Data Mining and Analytics Lab	40	04	02	25- 4 Hrs
Vocational Theory	VC-AC-01	Introduction to Power BI	45	03	03	60-2Hrs
Total Number of credits:					**19**	

Course Outcomes and Course Contents

SEMESTER I

Semester	I
Paper Code	BCADA1121
Paper Title	FUNDAMENTALS OF COMPUTING AND PROGRAMMING SKILLS USING C
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

The course is oriented to those who want to learn the fundamental concepts associated with the digital logic and circuit design and programming basics using C programming language as an implementation tool. It introduces the basic concepts and laws involved in the Boolean algebra and logic families and digital circuits and familiarize with the different number systems. Apart from learning digital basics it will provide students with understanding of programming essentials, including algorithms, data types, elementary control structures and functions used within the framework of imperative and structural programming paradigms.

COURSE OUTCOMES:

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

CO1: To understand different number systems and their conversions and to analyse and minimize Boolean expressions.
CO2: Understanding foundation concepts of information and information processing in computer systems: a matter of information, data representation, coding systems, Algorithm and Flowchart.
CO3: Understanding of programming language syntax and its definition by example of C language with the knowledge of basic principles of imperative and structural programming.
CO4: To gain insight knowledge of Functions, Arrays, Structures and Unions.
CO5: Learn the basics of pointers, File operations and Data Structures.

UNIT 1: INTRODUCTION TO DIGITAL ELECTRONICS **10 Hrs.**

INTRODUCTION TO DIGITAL ELECTRONICS, NUMBER SYSTEMS, OPERATIONS AND CODES
Introduction, Decimal numbers, Binary numbers, Decimal-to-Binary conversion, Binary Arithmetic, 1's and 2's Complements of Binary Numbers, signed numbers, Arithmetic operations with signed numbers, Hexadecimal Numbers, Octal numbers, Binary Coded Decimal(BCD), Digital Codes.

LOGIC GATES, BOOLEAN ALGEBRA AND LOGIC SIMPLIFICATION
The Inverter, the AND Gate, the OR gate, the NAND Gate, the NOR Gate, the Exclusive-OR and Exclusive-NOR Gates, Basics of Digital Integrated Circuits. Boolean Operations and Expressions, Laws and Rules of Boolean Algebra, DE Morgan's Theorems, Boolean Analysis of Logic Circuits, Simplification Using Boolean Algebra, Standard Forms of Boolean Expressions, Boolean Expressions and Truth Tables, The Karnaugh Map, Karnaugh Map SOP Minimization, POS Minimization.

UNIT 2: BASICS OF PROGRAMMING **7 Hrs.**

Introduction – The Problem-Solving aspect – Steps in Problem Solving – Types of Problems – Types of Programming Methodologies – Types of Computer Languages – Compiler – Interpreter – How to Write Algorithms – Implementation of Algorithms – Analysis of Algorithms – Flowchart – Pseudocode

UNIT 3: INTRODUCTION TO C LANGUAGE **9 Hrs.**

Overview of C – Constants, Variables and Data Types – Operators and Expressions – Managing Input / Output Operations – Formatted I/O – Decision Making - Branching –- IF, Nested IF – Switch – goto - Looping- While, do, for statements.

UNIT 4: ARRAYS, FUNCTIONS, STRUCTURES AND UNIONS **9 Hrs.**

Arrays – dynamic and multi-dimensional arrays - Character arrays and Strings – String handling Functions - User defined Functions – Categories of Functions – Recursion - Structures and Unions – Array of Structures – Structures and Functions

UNIT 5: POINTERS, FILES, DATA STRUCTURES **10 Hrs.**

Pointers – Declaration, Accessing a variable, character strings, pointers to functions and structures – Basic File Operations – Sorting – Searching – Stack – Queue - Trees

TEXTBOOKS:

1. M. Morris Mano “Digital Logic and Computer Design”, Pearson, 2013.
2. Harry H. Chaudhary, “C Programming The ultimate way to learn the fundamentals of the C language”, MIT- Createspace Inc. O-D-Publishing, LLC USA, 2016
3. Yashavant P. Kanetkar, “Let us C”, Fifth edition, BPB Publications, 2006
4. Ellis Horowitz and Sartaj Sahni, “Fundamentals of Data Structures”, Computer Science Press, 2012

SUGGESTED BOOKS:

1. Thomos L. Floyd, “Digital Fundamentals”, Tenth Edition, Pearson, 2015.
2. V. Anton Spraul, “Think Like a Programmer – An Introduction to Creative Problem Solving”, no starch press, 2014
3. Deitel and Deitel "C How to Program ", Sixth edition, Pearson, 2017

Semester	I
Paper Code	BCADA1221
Paper Title	EXPLORATORY DATA ANALYSIS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

- To apply principles of study design and data collection
- Produce and Interpret graphical summaries of data
- Graphically and numerically describe the relation between two or more variables

COURSE OUTCOME:

Upon successful completion of this course student should be able to

CO1: Understand the basic concepts of data and its measure
CO2: Develop the basic understanding of probability and random variable
CO3: Understand the usage of different probability distributions
CO4: Evaluate relationship among variables and using the same for prediction

CO5: Develop analytical ability to solve real-world problems using these methodologies.

UNIT I: STATISTICS –AN OVERVIEW **9 Hrs.**

Need for data, types of data, principles of measurement, sources of data, classification, tabulation and graphical presentation of data, Measures of central tendency: objectives of Averaging, requisites of a measure of Central Tendency, Mathematical Averages: mean, median and mode and quartiles, Measures of dispersion: significance of Measuring Dispersion, different measures of variation: range, variance, standard deviation, mean deviation, quartile deviation

UNIT II: FUNDAMENTALS OF PROBABILITY **9 Hrs.**

Concepts, the parallels between sets and events, Axioms of probability, Probability problems using permutations and combinations, The additive law, the idea of independence, Conditional probability, Bayes Theorem (simple problems. Problems involving conditional probability and dependence)**,** theory of random variables, expectation and variance of random variables, idea of dependent random variables.

UNIT III: PROBABILITY DISTRIBUTIONS **9 Hrs.**

Discrete Probability Distributions: Binomial, Poisson, Negative Binomial Distribution, Hypergeometric Distribution, Continuous Probability Distribution: Normal, Exponential Chi square, t and F distributions, Central Limit Theorem, Fitting distributions to data

UNIT IV: CORRELATION ANALYSIS **9 Hrs.**

Significance of measuring correlation, types of correlation, methods of correlation analysis, partial and multiple correlation

UNIT V: REGRESSION **9 Hrs.**

Some important information about straight lines, the method of least squares, assessing the goodness of fit, assessing each individual predictor - Case Studies related to the above discussed topics using Excel

TEXTBOOKS:

1. Statistics for Managers Using Microsoft Excel, Eighth Edition, David M. Levine, David F. Stephan Kathryn A. Szabat, Pearson Publications

SUGGESTED BOOKS:

1. A First Course in Statistics, Eighth Edition, Ronald Ross. Pearson

2. Probability and Statistics, Second Edition Murray R. Spiegel, John J. Schiller, R. Alu Srinivasan, Schuam Outline Series, Mac Graw Hill
3. Business Statistics, Second Edition, Pearson Education India.

Semester	I
Paper Code	BCADA1324
Paper Title	MATHEMATICS I
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

Introduce the concepts of set theory, relations and functions, concepts of calculus in one variable and higher dimensions and its applications.

COURSE OUTCOME:

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

CO1: Acquaint the students with mathematical/logical fundamentals including numerical techniques.

CO2: Solve problems involving recurrence relations and generating functions.

CO3: Explain the relationship between the derivative and linear approximation.

CO4: Solving real life maximum and minimum problems.

UNIT I: SET, RELATION AND FUNCTION **12 Hrs.**

Sets and Subsets, Set Operations and the Laws of Set Theory, Cartesian Products and Relations, Function: One-to-One, Onto Functions, Function Composition and Inverse Functions, Countable and Uncountable Sets.

UNIT II: DIFFERENTIAL CALCULUS OF ONE VARIABLE **12 Hrs.**

Basic properties of Functions: Functions and Their Graphs, Shifting and Scaling Graphs, Limits and Continuity: Limit of a functions and Limit laws, One sided Limit continuity, Limits involving infinity, Asymptotes of Graphs, Derivatives - Tangent Lines and derivative at a Point, The derivative as a function, Differentiation rules, The Chain Rule, Implicit Differentiation, Application of Derivatives: Extreme values of a functions on Closed Intervals, The Mean Value Theorem and related problems.

UNIT III: INTEGRAL CALCULUS **12 Hrs.**

Area and Estimating with Finite Sums, Limits of Finite, The Definite Integral, The Fundamental Theorem of calculus, Area sums, Indefinite Integral and Substitution Method, Area between curves, Arc Length, Area of Surface Revolution, Techniques of Integration , Geometric significance of integration.

UNIT IV: CALCULUS OF SEVERAL VARIABLE **9 Hrs.**

Functions of Several variables, Limit and Continuity in Higher Dimensions, Geometric significance of derivatives, Partial differentiations, The Chain Rule, Directional Derivatives and Gradient Vectors, Extremal Values and Saddle Point, Lagrange's Multipliers.

TEXT BOOKS:

1. "Thomas's Calculus" Pearson, 14th Edition by Joel Hass, Christopher Heil, Maurice D Weir
2. "Introduction to Real Analysis" by S K Mapa, Sarat Book Distributors, 2018,

SUGGESTED BOOKS:

1. Higher Engineering Mathematics by B.S.Grewal
2. "Principle of Real Analysis", 2nd Edition by S.C. Malik, New Academic Science, 2013

LABORATORY

Semester	I
Paper Code	BCADA1P1
Paper Title	C PROGRAMMING LAB
Number of teaching hrs per week	4 Hrs

Total number of teaching hrs per semester	40
Number of credits	2

1. Implementation of Basic Data types in C.
2. Implementation of various operators in C.
3. Implementation of Decision-making statement and Looping statement.
4. Execution of Break, Continue and Switch case statements.
5. Creation of functions in C.
6. Implementing array in C.
7. Creation of Unions and Structures in C.
8. Creation of Pointers in C.
9. Various File Operations in C.
10. Various sorting algorithms in C.
11. Various searching algorithms in C.
12. Stack and Queue implementation.
13. Implementation of Tree in C.

Semester	I
Paper Code	BCADA1P2
Paper Title	DATA ANALYSIS LAB(EXCEL)
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	40
Number of credits	2

1. Understanding Excel Interface for data Analysis : Getting Stated with Excel , Working with Data
2. Using the fill , series and column commands, Conditional Formatting in Excel
3. Using the visualization tools like charts and graphs in Excel for data comprehension
4. Using the summary statistics in Excel
5. Understanding the variations in data set using Excel commands
6. Creating random numbers in an Excel Spreadsheet
7. Generating probability distributions in Excel
8. Using Excel to understand relationship among Variables
9. Using regression as a prediction tool in Excel
10. Case Study 1
11. Case Study 2

SEMESTER II

Semester	II
Paper Code	BCADA 2121
Paper Title	PRINCIPLES AND PRACTICES OF DATA SCIENCE
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

To make the students learn the process of working with data in large scale. Make the student understand the existence of data with its wilderness and make use of it.

COURSE OUTCOME:

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

CO1: Understand the fundamental concepts of data.
CO2: Understand the fundamental concepts of data science process.
CO3: Understand the fundamental concepts of Machine Learning
CO4: Fundamental concepts of large data & Data Visualization

CO5: To gain knowledge about the recent trends of Data Science.

UNIT 1: PREPARING AND GATHERING DATA AND KNOWLEDGE **9 Hrs.**

Philosophies of data science - Data science in a big data world - Benefits and uses of data science and big data - facts of data- Overview of the data science process- Retrieving data –Data Preparation: Cleansing, integrating, and transforming data - Exploratory data analysis – Data Modeling: Model and variable selection, Model execution, Model diagnostic and model comparison - Presentation and automation: Presenting data, Automating data analysis

UNIT 2: BIG DATA **9 Hrs.**

Problems when handling large data – General techniques for handling large data – Case study – Steps in big data – Distributing data storage and processing with Frameworks – Case study.

UNIT 3: MACHINE LEARNING **9 Hrs.**

Application for machine learning in data science- Tools used in machine learning- Modeling Process – Training model – Validating model – Predicting new observations –Types of machine learning Algorithm : Supervised learning algorithms, Unsupervised learning algorithms.

UNIT 4: DATA VISUALIZATION **9 Hrs.**

Introduction to data visualization – Data visualization options – Filters – MapReduce – Dashboard development tools – Creating an interactive dashboard with dc.js-summary.

UNIT 5 : ETHICS AND RECENT TRENDS **9 Hrs.**

Data Science Ethics – Doing good data science – Owners of the data - Valuing different aspects of privacy - Getting informed consent - The Five Cs – Diversity – Inclusion – Future Trends.

TEXTBOOKS:

1. Introducing Data Science, Davy Cielen, Arno D. B. Meysman and Mohamed Ali, Manning Publications, 2016.
2. An Introduction to Statistical Learning: with Applications in R, Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Springer, 1st edition, 2013
3. Think Like a Data Scientist, Brian Godsey, Manning Publications, 2017.
4. Ethics and Data Science, D J Patil, Hilary Mason, Mike Loukides, O’ Reilly, 1st edition, 2018.

SUGGESTED BOOKS:

1. Doing Data Science, Straight Talk from the Frontline, Cathy O'Neil, Rachel Schutt, O’ Reilly, 1st edition, 2013.

2. Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman, Cambridge University Press, 2nd edition, 2014
3. An Introduction to Statistical Learning: with Applications in R, Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Springer, 1st edition, 2013

Semester	II
Paper Code	BCADA 2221
Paper Title	ADVANCED STATISTICAL COMPUTING
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

- To apply principles of sampling in data collection
- Learn the techniques of estimation
- Apply the regression techniques in solving real life problems

COURSE OUTCOME:

Upon successful completion of this course student should be able to -

CO1: Understand the basic concepts of sampling methods.
CO2: Understand the estimation procedures
CO3: Develop the skill to differentiate between parametric and non-parametric tests
CO4: Understand the use of multiple regression for prediction
CO5: Develop analytical ability to solve real-world problems using these methodologies

UNIT 1: SAMPLING AND SAMPLING DISTRIBUTIONS **9 Hrs.**
Principles of Sampling, Sampling methods, Sampling Distributions: mean, difference and proportions

UNIT 2: ESTIMATION AND CONFIDENCE INTERVALS **9 Hrs.**
Point Estimation, properties and drawback, Confidence Interval Estimation of population mean and proportions

UNIT 3: HYPOTHESIS TESTING **9 Hrs.**

General Procedure, Errors in Hypothesis Testing, testing related to parametric test like Z test, t –test, non-parametric statistics: advantages and limitations, the Chi-Square Distribution, applications of Chi-Square Test Statistic, Mann Whitney U-Test

UNIT 4: MULTIPLE REGRESSION ANALYSIS **9 Hrs.**

Assumptions, the basics, testing the accuracy of models, robust regression: bootstrapping, reporting the regression results, regression with categorical data, dummy coding

UNIT 5: ANALYSIS OF VARIANCE **9 Hrs.**

One Way and Two-Way Classification, assumptions, logic of F Ratio, post hoc procedures and violations of test assumptions - Case Study related to the above discussed topics using R

TEXTBOOK:

1. Statistical Inference: P. J. Bickel and K. A. Docksum, 2nd Edition, Prentice Hall.
2. Introduction to Linear Regression Analysis: Douglas C. Montgomery
3. An Introduction to Statistical Learning: with Applications in R, Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Springer, 1st edition, 2013

SUGGESTED BOOK:

1. Computer Age Statistical Inference by Bradley Efron and Trevor Hastie
2. Introduction to Statistical Learning by Gareth James

Semester	II
Paper Code	BCADA 2324
Paper Title	MATHEMATICS II
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

This course will enable students to acquire further skills in the techniques of linear algebra, as well as understanding of the principles underlying the subject. This course will prepare students for further courses in mathematics and/or related disciplines.

COURSE OUTCOME:

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

CO1: Acquaint the students with mathematical/logical fundamentals including numerical techniques

CO2: Understanding the numerical techniques of interpolation in various intervals in real life situations.

CO3: Understanding of numerical techniques of differentiation and integration which plays an important role in engineering and technology disciplines.

CO4: Able to define the various concepts of graphs and its implementations

CO5: Solve the real life problems using finite state machines

UNIT I: MATHEMATICAL LOGIC **10 Hrs.**

Basic Connectives and Truth Tables, Logic Equivalence – The Laws of Logic, Logical Implication – Rules of Inference, Fundamentals of Logic, the Use of Quantifiers, Quantifiers.

UNIT II: FORMAL LANGUAGES AND FINITE-STATE MACHINES **4 Hrs.**

Idea about formal languages, finite state machine.

Case studies

UNIT III: GRAPH THEORY **13 Hrs.**

Introduction to Graphs: Definition of Graph, Loop, Simple Graph, Graph as Models, Path and Cycle, Complete Graph, Bipartite Graph, Digraph, Tree, Isomorphism, Planar graphs, graph coloring, Hamilton circuits and Euler cycle, Shortest path algorithms.

Case studies

UNIT IV: INTERPOLATION AND APPROXIMATION **9 Hrs.**

Interpolation with unequal intervals – Lagrange's interpolation – Newton's divided difference interpolation – Cubic Splines – Difference operators and relations – Interpolation with equal intervals – Newton's forward and backward difference formulae.

UNIT V: NUMERICAL DIFFERENTIATION AND INTEGRATION 9 Hrs.

Approximation of derivatives using interpolation polynomials – Numerical integration using Trapezoidal, Simpson's 1/3 rule – Romberg's Method – Two point and three point Gaussian quadrature formulae – Evaluation of double integrals by Trapezoidal and Simpson's 1/3 rules.

TEXT BOOKS:

1. "Introduction to Graph Theory" PHI Learning Pvt Ltd 2012, Douglas B West
2. "Introductory Method to Numerical Analysis", Prentice Hall India Learning Private Limited; Fifth edition (1 January 2012), S. S. Sastry.
3. "Introduction to Automata Theory, Languages and Computation ", Pearson Education, 1979, J. Hopcroft, J. Ullman

SUGGESTED BOOKS:

1. Discrete Mathematics with Applications by Thomas Koshy
2. Discrete Mathematics with Applications by Susanna S. Epp

LABORATORY

Semester	II
Paper Code	BCADA2P1
Paper Title	DATA SCIENCE LAB (USING R)
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	40
Number of credits	2

1. Creating and manipulating vector in R
2. Creating matrix and manipulating matrix in R
3. Operations on Data Frames in R
4. Operations on Lists in R.
5. Programs on If – else statements in R
6. Programs on For Loops in R.
7. Customizing and Saving to Graphs in R.
8. PLOT Function in R to customize graphs
9. 3D PLOT in R to customize graphs
10. Implement in R Programming the concept to find Sum, Mean and Product of a Vector, ignore element like NA or NaN.
11. Implement in R Programming the concept to find missing values.
12. Implement the concept to create a list of data frames and access each of those data frames from the list using R.
13. Implement the concept of matrix multiplication and addition using R.
14. Implement linear regression model and compare predicted value with actual value using Visualization.
15. Implement logistic regression model and compare predicted value with actual value using Visualization.
16. Implement k-means clustering.

Semester	II
Paper Code	**BCADA2P2**
Paper Title	**STATISTICAL COMPUTING LAB (USING R)**
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	40
Number of credits	2

1. Getting Started with R environment : downloading , installing , using scripts , R workspace, installing packages in R
2. Getting data into R workspace : creating variables, creating data frames , organizing data
3. Manipulating Data : selecting parts of a data frame , data frames and matrices
4. Exploring data with graphs in R
5. Exploring the assumptions of normality in R
6. Understanding Interval Estimation in R

7. Parametric and Non-Parametric Tests in R
8. Testing the Regression models for accuracy
9. Comparing means Using ANOVA
10. Case Study 1
11. Case Study 2

OPEN ELECTIVES

Semester	I
Paper Code	OE 1
Paper Title	BASICS OF DATA SCIENCE
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

To make the students learn the process of working with data in large scale. Make the student understand the existence of data with its wilderness and make use of it.

COURSE OUTCOMES:

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

CO1: Understand the fundamental concepts of data.
CO2: Understand the fundamental concepts of data science process.
CO3: Understand the fundamental concepts of Machine Learning
CO4: Fundamental concepts of large data & Data Visualization
CO5: To implement the aspects of Data Science through case studies.

UNIT 1: PREPARING AND GATHERING DATA AND KNOWLEDGE 9 Hrs.

Philosophies of data science - Data science in a big data world - Benefits and uses of data science and big data - facts of data: Structured data , Unstructured data, Natural Language, Machine generated data, Audio, Image and video streaming data - The Big data Eco system: Distributed file system, Distributed Programming framework, Data Integration frame work, Machine learning Framework, NoSQL Databases, Scheduling tools, Benchmarking Tools, System Deployment, Service programming and Security.

UNIT 2: THE DATA SCIENCE PROCESS **9 Hrs.**

Overview of the data science process- Retrieving data –Data Preparation: Cleansing, integrating, and transforming data - Exploratory data analysis – Data Modeling: Model and variable selection, Model execution, Model diagnostic and model comparison - Presentation and automation: Presenting data, Automating data analysis

UNIT 3: MACHINE LEARNING **9 Hrs.**

Application for machine learning in data science- Tools used in machine learning- Modeling Process – Training model – Validating model – Predicting new observations –Types of machine learning Algorithm : Supervised learning algorithms, Unsupervised learning algorithms.

UNIT 4: VISUALIZATION **9 Hrs.**

Introduction to data visualization – Data visualization options – Filters – MapReduce – Dashboard development tools.

UNIT 5: CASE STUDIES **9 Hrs.**

Distributing data storage and processing with frameworks - Case study: e.g, Assessing risk when lending money.

TEXT BOOKS:

1. Introducing Data Science, Davy Cielen, Arno D. B. Meysman and Mohamed Ali, Manning Publications, 2016.
2. Think Like a Data Scientist, Brian Godsey, Manning Publications, 2017.

SUGGESTED BOOKS:

1. Doing Data Science, Straight Talk from the Frontline, Cathy O'Neil, Rachel Schutt, O' Reilly, 1st edition, 2013.
2. Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman, Cambridge University Press, 2nd edition, 2014
3. An Introduction to Statistical Learning: with Applications in R, Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Springer, 1st edition, 2013

Semester	II
Paper Code	OE 2
Paper Title	PYTHON PROGRAMMING

Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

The course is designed to provide Basic knowledge of Python. Python programming is intended for software engineers, system analysts, program managers and user support personnel who wish to learn the Python programming language.

COURSE OUTCOMES:

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

CO1: To understand the basic concepts in Python programming.
CO2: Learn how to write, debug and execute Python program.
CO3. Understand and demonstrate the use of Branching and Looping Structures.
CO4: To get insight knowledge related to advanced data types such as lists, tuples, dictionaries.
CO5: Acquire the basic knowledge of Object-Oriented Programming Concept and Exception Handling

UNIT I: INTRODUCTION **9 Hrs.**

Introduction to Python Programming, History of Python, its features, Scope of Python, Downloading and installing Python, Python code execution process, run a simple program on Python interpreter and IDLE.

UNIT II: DATA TYPES AND OPERATORS **9 Hrs.**

The concept of data types; variables, assignments; immutable variables; numerical types; arithmetic operators and expressions; comments in the program; understanding error messages; Illustrative programs.

UNIT III: BRANCHING AND LOOPING **10 Hrs.**

Conditions, Boolean logic, logical operators; ranges; Control statements: if-else, loops (for, while); short-circuit (lazy) evaluation. Illustrative programs

UNIT IV: LISTS, TUPLES AND DICTONARIES **10 Hrs.**

Lists, tuples, and dictionaries; basic list operators, replacing, inserting, removing an element; searching and sorting lists; dictionary literals, adding and removing keys, accessing and replacing values, Illustrative programs

UNIT V: OOPS AND EXCEPTION HANDLING **7 Hrs.**

Classes and OOP: classes, objects, attributes and methods; defining classes; design with classes, exception handling, Illustrative programs

TEXTBOOKS:

1. Python in easy steps - Mike McGrath, In Easy Steps Limited, Second Edition

2. "Hello World" - Computer Programming for Kids and other Beginners - Warren and Carter, Manning Publications, 2014

SUGGESSTED BOOKS:

1. Python3 Tutorial – Tutorialspoint
2. Mark Lutz, Programming Python, O`Reilly, 4th Edition, 2010

Semester	II
Paper Code	OE 3
Paper Title	MACHINE LEARNING
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

To make the students learn the statistics & mathematical concepts, Hypothesis & Dimension Reduction Technique, R Programming Concepts and Machine Learning.

COURSE OUTCOMES:

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

CO1: Understand the fundamental concepts of Statistics & Mathematics
CO2: Understand Hypothesis & Dimension Reduction Techniques

CO3: Hands on Experience in R Programming
CO4: Understand Machine Learning Concepts using R
CO5: To have basic knowledge of various predictive models.

UNIT I: STATISTICS & MATHEMATICAL ESSENTIALS **9 Hrs.**

Measure of Central Tendency - Mean, Median, Mode - Dispersion Technique - Range Inter Quartile Range - Variance, Standard Deviation - Mean Square Error & Root Mean Square - Probability Distribution.

UNIT II: HYPOTHESIS AND DIMENSION REDUCTION TECHNIQUE **9 Hrs.**

Types of Hypothesis - Sample testing - T-test - Z-test - Chi-square test - Anova test -. One Way Anova. Two Way Anova - Principle component analysis - Collinearity and multicollinearity

UNIT III: R PROGRAMMING CONCEPTS **9 Hrs.**

The Data types in R & its uses -Build in functions in R- Data Manipulation - Data import Techniques – Exploratory Data Analysis – Data Visualization.

UNIT IV: MACHINE LEARNING **9 Hrs.**

ML Fundamental & common use cases - Approach to Machine Learning - Understanding Supervised learning technique - Unsupervised learning technique

UNIT V: PREDICTIVE MODELLING IN R **9 Hrs.**

Introduction to predictive modeling - Regression Problem - Classification Problem - Linear Regression - Logistic Regression – Clustering - Distance measure types- K means clustering – Decision Tree Classifier – Random Forest Classifier – Support Vector Machine.

TEXTBOOK:

1. Introducing Data Science, Davy Cielen, Arno D. B. Meysman and Mohamed Ali, Manning Publications, 2016.
2. Think Like a Data Scientist, Brian Godsey, Manning Publications, 2017.

SUGGESTED BOOK:

1. Doing Data Science, Straight Talk from the Frontline, Cathy O'Neil, Rachel Schutt, O' Reilly, 1st edition, 2013.
2. Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman, Cambridge University Press, 2nd edition, 2014
3. An Introduction to Statistical Learning: with Applications in R, Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Springer, 1st edition, 2013

SEMESTER III

Semester	III
Paper Code	BCADA 3122
Paper Title	PYTHON FOR DATA ANALYTICS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

This course is designed to teach students how to analyse different types of data using Python. Students will learn how to prepare data for analysis, perform simple statistical analysis, create meaningful data visualizations and predict future trends from data.

COURSE OUTCOMES:

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

CO1: Understanding basics of python for performing data analysis
CO2: Use different python packages for mathematical, scientific applications and for web data analysis.
CO3: Able to get knowledge about Data Wrangling.
CO4: Develop the model for data analysis and evaluate the model performance.
CO5: Understanding the data, performing pre-processing, processing and data visualization to get insights from data.

UNIT I: DATA STRUCTURES AND OOP **9 Hrs**

Python Program Execution Procedure – Statements – Expressions – Flow of Controls – Functions – Numeric Data Types – Sequences – Strings – Tuples – Lists – Dictionaries. Class – Constructors – Object Creation – Inheritance – Overloading. Text Files and Binary Files – Reading and Writing.

UNIT II: NUMPY AND PANDAS PACKAGES **9 Hrs**

NumPy ndarray - Vectorization Operation - Array Indexing and Slicing - Transposing Array and Swapping Axes - Saving and Loading Array - Universal Functions - Mathematical and Statistical Functions in NumPy .

Series and DataFrame data structures in pandas - Creation of Data Frames – Accessing the columns in a DataFrame - Accessing the rows in a DataFrame - Panda's Index Objects - Reindexing Series and DataFrames - Dropping entries from Series and Data Frames - Indexing, Selection and Filtering in Series and Data Frames - Arithmetic Operations between Data Frames and Series - Function Application and Mapping.

UNIT III: DATA WRANGLING **9 Hrs**

Combining and Merging Data Sets – Reshaping and Pivoting – Data Transformation – String manipulations – Regular Expressions.

UNIT IV: DATA AGGREGATION AND GROUP OPERATIONS **9 Hrs**

Group By Mechanics – Data Aggregation – GroupWise Operations – Transformations – Pivot Tables – Cross Tabulations – Date and Time data types.

UNIT V: VISUALIZATION IN PYTHON **9 Hrs**

Matplotlib and Seaborn Packages – Plotting Graph - Controlling Graphs – Adding Text – More Graph Types – Getting and Setting Values – Patches.

REFERENCES:

1. Gowrishanker and Veena, "Introduction to Python Programming", CRC Press, 2019.

2. Python Crash Course, 2nd Edition, By Eric Matthes, May 2019

3. NumPy Essentials, By Leo Chin and Tanmay Dutta, April 2016

4. Joel Grus, "Data Science from scratch", O'Reilly, 2015.

5. Wes Mc Kinney, "Python for Data Analysis", O'Reilly Media, 2012.

6. Kenneth A. Lambert, (2011), "The Fundamentals of Python: First Programs", Cengage Learning

7. Jake Vanderplas. Python Data Science Handbook: Essential Tools for Working with Data 1st Edition.

Semester	III
Paper Code	BCADA 3222

Paper Title	ECONOMETRICS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

To equip the students with the knowledge of the techniques of modern econometrics, required for applied research in data analytic industry.

COURSE OUTCOMES:

Upon successful completion of this course student should be able to

CO1: To get familiar with the concept of econometrics
CO2: To get insights into classical linear regression
CO3: To familiarize the concept of relaxation of assumptions of classical regression model
CO4: To get insights of Non-linear relationship
CO5: To solve the Simultaneous equations using different methods

UNIT I : INTRODUCTION TO ECONOMETRICS 8 Hrs.

Nature and Scope of Econometrics - Meaning of Econometrics, Relationship Between Statistics, Mathematics and Economics, Economic and Econometric Models, the aims and methodology of Econometrics, statistical Vs deterministic relationship, regression Vs Causation, regression Vs correlation, terminology and notation, the nature and sources of data for Econometric analysis. Review of basic statistical concepts, Population and Sample, Random variables, Probability distribution function, Multivariate probability density functions, Expected value, Variance, Co variance, Correlation Co efficient.

UNIT II: CLASSICAL LINEAR REGRESSION 8 Hrs.

Two Variable Regression Analyses - the basic two Variable Regression model: OLS Estimation, Hypothesis testing, Extensions of two variable regression model – Functional forms of regression model. Interpretation of multiple regression coefficients , properties of regression coefficients, multicollinearity, goodness of fit.

UNIT III: RELAXING THE ASSUMPTIONS OF THE CLASSICAL REGRESSION MODEL 10 Hrs.

Multi-collinearity, Heteroscedasticity and Autocorrelation- Nature, Consequences, Detection and Remedial Measures

UNIT IV: UNDERSTANDING NON-LINEAR RELATIONSHIPS 6 Hrs.

Transformation of variables - Logarithmic transformations, Engel curve, semi- logarithmetic transformation, non linear regression, use of dummy variables in regression

UNIT V: SIMULTANEOUS EQUATIONS SYSTEM 8 Hrs.

Identification problem, Least Squares estimation , Bias Problem

SUGGESTED BOOKS:

1. Cristopher Dougherty Introduction to Econometrics
2. Johnson , Econometric Methods
3. Damodar Gujrati , Econometrics
4. The Econometrics of Financial Markets : J. Campbell, A.Lo and C. Mackinlay Econometric Analysis : William H. G

Semester	III
Paper Code	BCADA 3324
Paper Title	MATHEMATICS III
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

- To understand the basic concepts and techniques of solving algebraic equations.
- To introduce the numerical techniques of interpolation in various intervals in real life situations.
- To acquaint the student with understanding of numerical techniques of differentiation and integration which plays an important role in engineering and technology disciplines.

- To acquaint the knowledge of various techniques and methods of solving ordinary differential equations.
- To understand the knowledge of various techniques and methods of solving various types of partial differential equations.

COURSE OUTCOMES:

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

CO1: Understanding the basic concepts and techniques of solving algebraic equation and matrices techniques for solving systems of linear equations in the different areas of Linear Algebra.

CO2: Understanding the numerical techniques of interpolation in various intervals in real life situations.

CO3: Understanding of numerical techniques of differentiation and integration which plays an important role in engineering and technology disciplines.

CO4: Knowledge of various techniques and methods of solving ordinary differential equations

CO5: Understanding database the knowledge of various techniques and methods of solving various types of partial differential equations.

UNIT I: MATRICES **12 Hrs.**

System of linear equation, Row reduction and echelon forms, vector equation, Matrix equation, Solution sets of linear systems, Linear Independence, Introduction to Linear Transformation, The matrix of a Linear Transformation. Matrix Operations, The Inverse of a matrix, Characterization of Invertible Matrices, Subspaces of $R^n$and Dimension and rank.

UNIT II: SOLUTION OF EQUATIONS AND EIGENVALUE PROBLE **9 Hrs.**

Solution of algebraic and transcendental equations – Fixed point iteration method – Newton-Raphson method – Solution of linear system of equations – Gauss elimination method – Pivoting – Gauss Jordan method – Iterative methods of Gauss Jacobi and Gauss Seidel – Eigenvalues of a matrix by Power method and Jacobi's method for symmetric matrices.

UNIT III: VECTOR SPACES **12 Hrs.**

Vector spaces, subspaces, Null spaces, Column Spaces and Linear Transformation, linearly independent set and Bases, Coordinate system, The dimension of a vector space, Rank, Change of Basis, Eigenvectors and Eigenvalues, The characteristic equation, Diagonalization.

UNIT IV: DIFFERENTIAL EQUATIONS **9 Hrs.**

Introduction to differential equations, Formation of differential equations, Equations of First Order and First Degree: Solving equations of second and higher order homogeneous equations with constant coefficients, solving equations of second and higher order non-homogeneous equations with constant coefficients by finding the complementary function particular integral.

SELF-STUDY **3 Hrs.**

TEXT BOOKS:

1. "An introduction to Differential Equation", New Central Book Agency, 2011, R K Ghosh and K C Maity.
2. "Introductory Method to Numerical Analysis", Prentice Hall India Learning Private Limited; Fifth edition (1 January 2012), S. S. Sastry
3. "Linear Algebra and its Applications," 3rd edition, Pearson Education (Asia) Pte. Ltd, 2005. David C. Lay

SUGGESTED BOOKS:

1. "Linear Algebra with Applications, 8th Edition By Steve Leon
2. "Linear Algebra," 2nd edition, Pearson Education (Asia) Pte. Ltd/2004. Kenneth Hoffman and Ray Kunze
3. "Differential Equations", 3rd edition, Wiley Publisher, 2007, S.L Ross

LABORATORY

Semester	III
Paper Code	BCADA3P1
Paper Title	PYTHON FOR DATA ANALYTICS LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2

Semester	III

Paper Code	BCADA3P2
Paper Title	ECONOMETRICS LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2

Semester	III
Paper Code	OE 5
Paper Title	BUSINESS AND DATA UNDERSTANDING
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES :

The course aims to equip the students with basic understanding of data and business decision making

COURSE OUTCOMES:

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

CO1: To familiarize the students with business environment and its problems
CO2: To get an insight into the data generation in business and its analysis
CO3: To familiarize students with different methods for data analysis
CO4: To get an insight into the concepts of probability and distribution

CO5: To get an insight into the decision making framework

UNIT I: BUSINESS ENVIRONMENT **8 Hrs.**

An Overview of the business environment , how it operates -characteristics of business environment – types of environment – environment analysis- -socio culture environment -political and government environment

UNIT II: DATA UNDERSTANDING **8 Hrs.**

Data generation process in business environment - collecting and utilizing data for business solutions - variables selection and sampling process -role of softwares

UNIT III: EXPLORING DATA **12 Hrs.**

Describing the distribution of a single variable - descriptive measures for Categorical variable - descriptive measures for Numerical variable - Charts for Numerical Variables - Time series data - Outliers and missing values - Finding relationships among variables - Understanding Time series Data - components of time series data - measures of accuracy - testing for randomness- modeling time series data

UNIT IV: CRISP DECISION MAKING FRAMEWORK **8 Hrs.**

Probability and Distributions-Probability essentials - Distribution of single Random variable- summary measures of a Probability Distribution - Binomial -Poisson and Normal distributions and their applications

UNIT V: CRISP DECISION MAKING FRAMEWORK: **4 Hrs.**

Heart of Data Analysis: Modelling, model development and deployment

SELF STUDY **5 Hrs.**

SUGGESTED BOOKS:

1. Starling, Grower (1996) The changing Environment of Business Cincimmati, OH, South Western College Publishing

2. S. Christian Albright, Wayne L. Winston, Business Analytics: Data Analysis and Decision Making, Cengage Learning

3. S. Christian Albright & Wayne L. Winston, Business Analytics: Data Analysis and Decision making , Cengage Learning

4. Christian Heuman , Michael Schomaker and Shalabh : Introduction to Statistics and Data Analysis : With Exercises, Solutions and Applications in R,

5. Drew Bentley (2017) , Business Intelligence and Analytics , Library Press

SEMESTER IV

Semester	IV
Paper Code	BCADA 4122
Paper Title	DATABASE MANAGEMENT SYSTEM
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE DESCRIPTION:

To provide strong foundation for databases, tables, database management system and application area related to it and understand the underlying core concepts.

COURSE OBJECTIVES:

This course concentrates on introduction, principles, design and implementation of DBMS. It introduces about the distributed system and brief about data mining and data warehouse. To provide strong foundation of database concepts and develop skills for the design and implementation of a database application with a brief exposure to advanced database concepts.

COURSE OUTCOMES:

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

CO1: Understanding the fundamental concepts of Database Management systems
CO2: Understanding the concepts of Database models.
CO3: Understanding the core terms, concepts, and tools of relational database management systems.
CO4:Understanding database design and logic development for database programming.

.

UNIT I: DATABASE MANAGEMENT SYSTEM INTRODUCTION 10 Hrs.

Data- Database- Database management system- Characteristics of the database approach- Role of Database administrators- Role of Database Designers- End Users- Advantages of Using a DBMS-Data models, Schema and Instances –Database design - Database Engine – 1 tier architecture – 2 tier architecture- 3 tier architecture – History of Database Management systems- Types of Databases.

UNIT II: DATABASE MODELS AND IMPLEMENTATION **10 Hrs.**

Data Model and Types of Data Model- Relational Data Model- Hierarchical Model- Network Data Model- Object/Relational Model- Object-Oriented Model- Entity-Relationship Model- Modeling using E-R Diagrams- Notation used in E-R Model- Relationships and Relationship Types- Cardinalities. Subclasses, Super classes and Inheritance – Specialization and Generalization – Characteristics of Specialization and Generalization – Modeling of UNION types with categories- An example University EER Schema.

UNIT II: RELATIONAL DATABASES **10 Hrs.**

Structure of relational databases- Properties of relational databases and Tables –Structure of relational databases – Database Schema – Armstrong Axioms – Functional Dependency-Anomalies in a Database- Properties of Normalized Relations- First Normalization- Second Normal Form Relation- Third Normal Form- Boyce-Codd Normal Form (BCNF).

UNIT IV: SQLAND ADDITIONAL CONCEPTS **10 Hrs.**

Categories of SQL Commands; Data Definition; Data Manipulation Statements, SELECT - The Basic Form, Subqueries, Functions, GROUP BY Feature, Updating the Database, Data Definition Facilities. MongoDB Overview- MongoDB Data modeling.

SELF STUDY **5 Hrs.**

REFERENCE BOOKS:

1. Elmasri Ramez and Navathe Shamkant B, Fundamentals of Database Systems, Addison-Wesley, 6th Edition, 2010.
2. Silberschatz, Korth, Sudarshan, Database System Concepts, 5 Edition, McGraw Hill, 2006.
3. O`neil Patricand, O`neil Elizabeth, Database Principles, Programming and Performance, 2nd Edition, Margon Kaufmann Publishers Inc, 2008.

Semester	IV
Paper Code	BCADA 4222
Paper Title	OS USING LINUX
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45

Number of credits	3

COURSE OBJECTIVES :

1. To be able to read and understand sample open source programs and header files.
2. To understand how the processes are implemented in Linux along with File System.
3. To study Linux memory management data structures and algorithms.
4. To acquire the knowledge in the implementation of interprocess communication.
5. To understand various process communication and how program execution happens in Linux.

COURSE OUTCOMES:

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

CO1: Explain the components and functionality of Unix.
CO2: Learn the processes of Unix
CO3: Explain different types of files and its handling
CO4: Apply memory management of OS in Unix
CO5: Learn to execute programs inUnix

UNIT I: INTRODUCTION **9 Hrs.**

Basic Operating System Concepts - Process Management - Memory Management - Device Drivers.

Overview of Unix File System - Files - Links - Types - Inodes -Access Rights - System Calls - Overview of Unix Kernels -Model - Implementation - Reentrant Kernels - Address Space - Synchronization - Interprocess Communication.

UNIT II: PROCESSES **9 Hrs.**

Processes, Lightweight Processes, and Threads - Process Descriptor - State - Identifying a Process - Relationships among processes - Organization - Resource Limits - Creating Processes - System Calls - Kernel Threads - Destroying Processes -Termination - Removal.

UNIT III: FILE SYSTEM **9 Hrs.**

The Virtual File System (VFS) - Role - File Model -System Calls - Data Structures - Super Block, Inode, File, dentry Objects - dentry Cache - Files Associated with a Process - Filesystem Types - Special Filesystems - Filesystem Type Registration – File system Handling - Namespaces - Mounting - Unmounting - Implementation of VFS System Calls.

UNIT IV: MEMORY MANAGEMENT **9 Hrs.**

Page frame management -page descriptors - non-uniform memory access - memory zones - reserved page frames - zoned page frame allocator - kernel mappings - buddy system algorithm - page frame cache - zone allocator.

UNIT V: PROCESS COMMUNICATION AND PROGRAM EXECUTION 9 Hrs.

Process Communication - Pipes -Usage - Data Structures - Creating and Destroying a Pipe - Reading from and Writing into a Pipe. Program Execution - Executable Files - Process Credentials - Command-Line Arguments and Shell Environment - Libraries - Program Segments and Process Memory Regions - Execution tracing - Executable Formats - Execution Domains - The exec Functions

SUGGESTED BOOKS:

1. Daniel P. Bovet and Marco Cesati, "Understanding the Linux Kernel", 3rd Edition, O'Reilly Publications, 2005.
2. Harold Abelson, Gerald Jay Sussman and Julie Sussman, ―Structure and Interpretation of Computer Programs‖, Second Edition, Universities Press, 2013.
3. Maurice J. Bach, ―The Design of the Unix Operating System‖ 1 st Edition Pearson Education, 2003.
4. Michael Beck, Harald Bohme, Mirko Dziadzka, Ulrich Kunitz, Robert Magnus, Dirk Verworner, ―Linux Kernel Internals‖, 2nd Edition, Addison-Wesley, 1998.
5. Robert Love, ―Linux Kernel Development‖, 3 rd Edition, Addison-Wesley, 2010.

Semester	IV
Paper Code	BCADA 4322
Paper Title	MULTIVARIATE STATISTICS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE: The aim of this course is to appraise the students about the concepts and methods of multivariate analysis

COURSE OUTCOMES:

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

CO1: To get an insight into the application of multivariate analysis
CO2: To familiarize the students with the concept of multivariate normal distribution
CO3: To familiarize the students with multiple regression
CO4: To get insights into the concepts of data reduction using principal component analysis
CO5: To get insights into factor analysis

UNIT I: ASPECTS OF MULTIVARIATE ANALYSIS **5 Hrs.**

Application of Multivariate Techniques, the organization of data , data displays and pictorial representations , matrix algebra and random vectors

UNIT II: THE MULTIVARIATE NORMAL DISTRIBUTION **10 Hrs.**

The Multivariate Normal Density and its properties, mean vector , Covariance Matrix , Correlation Matrix , Relationship between Correlation and covariance matrix , Multivariate Normal Distribution – Geometric interpretation , Examining data for multivariate Normal Distribution , Multivariate Inferential Statistics : Hotelling's T^2, Confidence Region , Hypothesis Testing for Equality of two population mean Vectors

UNIT III: MULTIPLE LINEAR REGRESSION **10 Hrs.**

Assumptions and Estimation of Model Parameters , Sampling distribution of parameter estimates , Model Adequacy Test, Tests of assumptions , Remedy against violations of assumptions, multivariate Linear Regression ,

UNIT IV: PRINCIPAL COMPONENT ANALYSIS **10 Hrs.**

Conceptual Model , Extraction of Principal Components , Model Adequacy and Interpretation

UNIT V: FACTOR ANALYSIS **10 Hrs.**

Factor Analysis : Basic and Orthogonal Models , Types of Models , parameter estimation , Model Adequacy Tests and Factor Rotation

SUGGESTED BOOKS:

1. 1. Applied Multivariate Statistical Analysis by R A Johnson and D W Wichern, Sixth Edition, PHI, 2012.
2. 2. Multivariate data analysis by Joseph F. Hair Jr, Rolph E. Anderson, Ronald L Tatham, and William C. Black, Fifth Edition, Pearson Education, 1998.
3. 4. Analysing multivariate data by J Lattin, J D Carroll and P E Green, Cengage Learning, 2010.
4. 5. Applied multivariate analysis by N H Timm, Springer, 2002.
5. 6. An Introduction to multivariate Statistical Analysis , T.W Anderson , Third Edition , Wiley Student Edition

LABORATORY

Semester	IV
Paper Code	BCADA4P1
Paper Title	DATABASE MANAGEMENT SYSTEM LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2

Semester	IV
Paper Code	BCADA4P2
Paper Title	MULTIVARIATE STATISTICS LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2

SEMESTER V

Semester	V
Paper Code	BCADA 5123

Paper Title	JAVA PROGRAMMING
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE

This course introduces computer programming using the JAVA programming language with object-oriented programming principles. Emphasis is placed on event-driven programming methods, including creating and manipulating objects, classes, and using object-oriented tools such as the class debugger.

COURSE OUTCOMES:

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

CO1: Understand the basic Java programming concept using OOP principles.

CO2: Develop Java programs with the concepts of inheritance and interfaces.

CO3: Able to Build Java applications using exceptions and I/O streams

CO4: Gain in depth knowledge towards Java applications using threads, generic classes and Event Driven concepts.

UNIT I: INTRODUCTION TO OOP AND JAVA FUNDAMENTALS 10 Hrs.

Object Oriented Programming - Abstraction – objects and classes - Encapsulation- Inheritance - Polymorphism- OOP in Java – Characteristics of Java – The Java Environment - Java Source File - Structure – Compilation. Fundamental Programming Structures in Java – Defining classes in Java – constructors, methods -access specifiers - static members -Comments, Data Types, Variables, Operators, Control Flow, Arrays , Packages - Javadoc comments.

UNIT II: INHERITANCE AND INTERFACES 10 Hrs.

Inheritance – Super classes- sub classes –Protected members – constructors in sub classes- the Object class – abstract classes and methods- final methods and classes – Interfaces – defining an interface, implementing interface, differences between classes and interfaces and extending interfaces - Object cloning -inner classes, Array Lists - Strings

UNIT III: EXCEPTION HANDLING AND I/O 10 Hrs.

Exceptions - exception hierarchy - throwing and catching exceptions – built-in exceptions, creating own exceptions, Stack Trace Elements. Input / Output Basics – Streams – Byte streams and Character streams – Reading and Writing Console – Reading and Writing Files

UNIT IV: MULTI THREADING, GENERIC PROGRAMMING AND EVENT DRIVEN PROGRAMMING **10 Hrs.**

Differences between multi-threading and multitasking, thread life cycle, creating threads, synchronizing threads, Inter-thread communication, daemon threads, thread groups. Generic Programming – Generic classes – generic methods – Bounded Types – Restrictions and Limitations. Graphics programming - Frame – Components - working with 2D shapes - Using colour, fonts, and images - Basics of event handling - event handlers - adapter classes - actions - mouse events.

SELF STUDY **5 Hrs.**

TEXT BOOKS

1. Herbert Schildt, ―Java The complete reference‖, 8th Edition, McGraw Hill Education, 2011.

2. Cay S. Horstmann, Gary cornell, ―Core Java Volume –I Fundamentals‖, 9th Edition, Prentice Hall, 2013.

SUGGESTED BOOKS

1. Paul Deitel, Harvey Deitel, ―Java SE 8 for programmers‖, 3rd Edition, Pearson, 2015.

2. Steven Holzner, ―Java 2 Black book‖, Dreamtech press, 2011.

3. Timothy Budd, ―Understanding Object-oriented programming with Java‖, Updated Edition, Pearson Education, 2000.

Semester	VI
Paper Code	BCADA 5223
Paper Title	DATA MINING AND ANALYTICS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE DESCRIPTION:

To provide strong foundation for Data Mining and analytics concepts

COURSE OBJECTIVES:

- To Understand the basic concepts of Data Mining
- To gain knowledge about Data
- To understand the concepts of Data mining algorithms
- To apply the concepts of data mining in different business verticals

COURSE OUTCOMES:

On successful completion of this course, the student will able to

CO1: Understand basic concepts of data mining
CO2: Handle different types of data
CO3: Understand the concepts of data warehousing
CO4: Apply different data mining algorithms
CO5: Learn to apply data mining and analytics algorithm to solve problems in different business verticals

UNIT I: BASICS OF DATA MINING **9 Hrs**

Why Data Mining? Moving toward the Information Age Data Mining as the Evolution of information Technology, What kinds of Data Can Be Mined, Database Data, Data Warehouses,

Transactional Data, Other Kinds of Data, OLTP & Online Analytical, Processing (OLAP), Graphs Database

UNIT II: BASICS OF DATA **9 Hrs**

Getting to Know Your Data: Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Data Visualization, Measuring Data Similarity and Dissimilarity, Data Preprocessing: Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization

(ETL Operations)

UNIT III: DATA WAREHOUSING **9 Hrs**

Data Warehousing, Data Mart and difference Between Developer server to pro server for Data Wearhouse.

UNIT IV: ALGORITHMS **9 Hrs**

Building Actionable insights using, Market basket Analysis, Association rule Recommendation engine and Suggestive Dashboarding and Interactive Analysis

UNIT V: APPLICATIONS OF DATA ANALYTICS 9 Hrs

Data analytical applications in business verticals BSFI, Retail, Telecom & Healthcare

SUGGESTED BOOKS

1. Principles of Data Science Arabnia, H.R., Daimi, K., Stahlbock, R.,Soviany, C., Heilig, L., Brüssau, K. (Eds.)
2. Data Mining Concepts and Techniques: Jiawei Han, Micheline Kamber, Jian Pei

Semester	V
Paper Code	BCADA 5323
Paper Title	CLOUD COMPUTING
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

- To understand the concept of cloud computing.
- To appreciate the evolution of cloud from the existing technologies.
- To have knowledge on the various issues in cloud computing.
- To be familiar with the lead players in cloud.
- To appreciate the emergence of cloud as the next generation computing paradigm.

COURSE OUTCOMES:

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

CO1: Articulate the main concepts, key technologies, strengths and limitations of cloud computing.
CO2: Learn the key and enabling technologies that help in the development of cloud.

CO3: Develop the ability to understand and use the architecture of compute and storage cloud, service and delivery models.
CO4: Explain the core issues of cloud computing such as resource management and security and able to install and use current cloud technologies.

UNIT I: INTRODUCTION **10 Hrs**

Introduction to Cloud Computing – Definition of Cloud – Evolution of Cloud Computing – Underlying Principles of Parallel and Distributed Computing – Cloud Characteristics – Elasticity in Cloud – On-demand Provisioning

UNIT II: CLOUD ENABLING TECHNOLOGIES **10 Hrs**

Service Oriented Architecture – REST and Systems of Systems – Web Services – Publish-Subscribe Model – Basics of Virtualization – Types of Virtualization – Implementation Levels of Virtualization – Virtualization Structures – Tools and Mechanisms – Virtualization of CPU – Memory – I/O Devices –Virtualization Support and Disaster Recovery.

UNIT III: CLOUD ARCHITECTURE, SERVICES AND STORAGE **10 Hrs**

Layered Cloud Architecture Design – NIST Cloud Computing Reference Architecture – Public, Private and Hybrid Clouds - laaS – PaaS – SaaS – Architectural Design Challenges – Cloud Storage – Storage-as-a-Service – Advantages of Cloud Storage – Cloud Storage Providers – S3.

UNIT IV: RESOURCE MANAGEMENT AND SECURITY IN CLOUD **10 Hrs**

Inter Cloud Resource Management – Resource Provisioning and Resource Provisioning Methods – Global Exchange of Cloud Resources – Security Overview – Cloud Security Challenges – Software-as-a-Service Security – Security Governance – Virtual Machine Security – IAM – Security Standards.

SELF STUDY **5 Hrs**

TEXT BOOKS:

1. Kai Hwang, Geoffrey C. Fox, Jack G. Dongarra, "Distributed and Cloud Computing, From Parallel Processing to the Internet of Things", Morgan Kaufmann Publishers, 2012.

2. Rittinghouse, John W., and James F. Ransome, “Cloud Computing: Implementation, Management and Security”, CRC Press, 2017.

REFERENCE BOOKS:

1. Rajkumar Buyya, Christian Vecchiola, S. ThamaraiSelvi, “Mastering Cloud Computing”, Tata Mcgraw Hill, 2013.

2. Toby Velte, Anthony Velte, Robert Elsenpeter, "Cloud Computing - A Practical Approach”, Tata Mcgraw Hill, 2009.

3. George Reese, "Cloud Application Architectures: Building Applications and Infrastructure in the Cloud: Transactional Systems for EC2 and Beyond (Theory in Practice)”, O'Reilly, 2009.

Semester	V
Paper Code	BCADADE5423
Paper Title	COMPUTER NETWORKS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

- To understand the protocol layering and physical level communication
- To analyse the performance of a network
- To understand the various components required to build different networks
- To learn the functions of network layer and the various routing protocols
- To familiarize the functions and protocols of the Transport, Network and application layer

COURSE OUTCOMES:

On successful completion of this course, the student will able to

CO1: Comprehend the basic layers and its functions in computer networks.
CO2: Understand the basics of how data flows from one node to another.
CO3: Analyse and design routing algorithms.
CO4: Design protocols for various functions in the network.
CO5: Understand the working of various application layer protocols.

UNIT I: INTRODUCTION AND PHYSICAL LAYER 9 Hrs.

Networks – Network Types – Protocol Layering – TCP/IP Protocol suite – OSI Model – Physical Layer: Performance – Transmission media – Switching – Circuit-switched Networks – Packet Switching.

UNIT II: DATA-LINK LAYER & MEDIA ACCESS 9 Hrs.

Introduction – Link-Layer Addressing – DLC Services – Data-Link Layer Protocols – HDLC – PPP – Media Access Control – Wired LANs: Ethernet – Wireless LANs – Introduction – IEEE 802.11, Bluetooth – Connecting Devices.

UNIT III: NETWORK LAYER 9 Hrs

Network Layer Services – Packet switching – Performance – IPV4 Addresses – Forwarding of IP Packets – Network Layer Protocols: IP, ICMP v4 – Unicast Routing Algorithms – Protocols – Multicasting Basics – IPV6 Addressing – IPV6 Protocol.

UNIT IV: TRANSPORT LAYER 9 Hrs.

Introduction – Transport Layer Protocols – Services – Port Numbers – User Datagram Protocol – Transmission Control Protocol – SCTP.

UNIT V: APPLICATION LAYER 9 Hrs.

WWW and HTTP – FTP – Email –Telnet –SSH – DNS – SNMP.

TEXT BOOKS

1. Behrouz A. Forouzan, Data Communications and Networking, Fifth Edition TMH, 2013.

2. Larry L. Peterson, Bruce S. Davie, Computer Networks: A Systems Approach, Fifth Edition, Morgan Kaufmann Publishers Inc., 2012.

REFERENCE BOOKS

1. James F. Kurose, Keith W. Ross, “Computer Networking: A Top-Down Approach”, Seventh Edition, Pearson Education, 2017.

2. William Stallings, Data and Computer Communications, Tenth Edition, Pearson Education, 2013.

3. Nader F. Mir, Computer and Communication Networks, Second Edition, Prentice Hall, 2014.

4. Ying-Dar Lin, Ren-Hung Hwang and Fred Baker, Computer Networks: An Open Source Approach, McGraw Hill Publisher, 2011.

Semester	V
Paper Code	VC-CS-01
Paper Title	AWS
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

- To describe what a cloud service provider (CSP) is and the value they bring to computing.
- To identify basic security and compliance aspects of the AWS platform and the shared security model.
- To understand the billing, account management, and pricing models.
- Evaluate basic or core characteristics of deploying and operating in the AWS Cloud
- Analyse and differentiate between on-premises and cloud infrastructure

COURSE OUTCOMES:

On successful completion of this course, the student will able to

CO1: the concepts of AWS Cloud and explains the AWS pricing philosophy

CO2: explain the global infrastructure components of AWS and Describe the security and compliance measures of the AWS Cloud, including AWS Identity and Access Management (IAM)

CO3: Create a virtual private cloud (VPC) by using Amazon Virtual Private Cloud (Amazon VPC) and Demonstrate when to use Amazon Elastic Compute Cloud (Amazon EC2), AWS Lambda, and AWS Elastic Beanstalk

CO4: Explain the architectural principles of the AWS Cloud. Explore key concepts related to Elastic Load Balancing, Amazon CloudWatch, and Amazon EC2 Auto Scaling

UNIT I

Cloud Concepts Overview — 8 Hrs

Introduction to cloud computing, Advantages of the cloud, Introduction to AWS, Moving to the AWS Cloud.

Cloud Economics and Billing

Fundamentals of pricing, Total cost of ownership, Activity: Simple Monthly Calculator, Delaware North case study, AWS Organizations, AWS billing and cost management, Billing dashboards, Technical support models, Activity: Support Plan Scavenger Hunt.

UNIT II

AWS Global Infrastructure Overview — 10 Hrs

AWS global infrastructure, Demo: AWS global infrastructure, AWS services and service categories, Activity: AWS Management Console Clickthrough

Cloud Security

AWS shared responsibility model, Activity: AWS Shared Responsibility Model, AWS IAM, Demo: AWS IAM Console, Securing a new AWS account, Lab: Introduction to AWS IAM, Securing accounts, Securing data, Working to ensure compliance.

UNIT III

Networking and Content Delivery — 8 Hrs

Networking basics, Amazon VPC, VPC networking, Activity: Label This diagram, Demo: Amazon VPC Console, VPC security, Activity: Design a VPC, Lab: Build a VPC and Launch a Web Server, Route 53, CloudFront

Compute

Compute services overview, Amazon EC2 part 1, Amazon EC2 part 2, Amazon EC2 part 3, Demo: Amazon EC2, Lab: Introduction to Amazon EC2, Activity: Amazon EC2 Versus Managed Services, Demo: Amazon EC2 Part Console, Amazon EC2 cost optimization, Container services, Introduction to AWS Lambda, Activity: AWS Lambda, Introduction to AWS Elastic Beanstalk, Activity: AWS Elastic Beanstalk

UNIT IV

Storage **8 Hrs**

AWS EBS, Demo: Amazon Elastic Block Store Console, Lab: Working with EBS, AWS S3, Demo: AWS S3 Console, AWS EFS, Demo: AWS EFS Console, AWS S3 Glacier, Demo: AWS S3 Glacier Console, Activity: Storage Technology Selection

Databases

Amazon RDS, Demo: Amazon RDS Console, Lab: Build a Database Server, Amazon DynamoDB, Demo: Amazon DynamoDB, Amazon Redshift, Amazon Aurora, Activity: Database case study.

UNIT V

Cloud Architecture **8 Hrs**

AWS Well-Architected Framework design principles, Activity: AWS Well-Architected Framework Design Principles, Operational excellence, Security, Reliability, Performance efficiency, Cost optimization, Reliability & high availability, AWS Trusted Advisor, Activity: Interpret AWS Trusted Advisor Recommendations

Automatic Scaling and Monitoring

Elastic Load Balancing, Activity: Elastic Load Balancing, Amazon CloudWatch, Activity: Amazon CloudWatch, Amazon EC2 auto scaling, Lab: Scale & Load Balance your Architecture

SELF STUDY **3 Hrs**

REFERENCE BOOKS

1. AWS Certified Cloud Practitioner Study Guide: CLF-C01 Exam, Ben Piper, David Clinton, Sybex Publishers
2. AWS Basics: Beginner's Guides, Gordon Wong, CreateSpace Independent Publishing
3. AWS: The Beginners Guide to Amazon Web Services, Dennis Hutten, CreateSpace Independent Publishing
4. Amazon Web Services for Dummies, Bernard Golden, John Wiley & Sons
5. Learning Amazon Web Services (AWS): A Hands-On Guide to the Fundamentals of AWS Cloud, Mark Wilkins, Pearson Education
6. AWS Certified Cloud Practitioner (CLF-C01) Cert Guide, Anthony Sequeira, Pearson Education

LABORATORY

Semester	IV
Paper Code	BCADA1P2
Paper Title	JAVA PROGRAMMING LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2

Semester	IV
Paper Code	BCADA5P2
Paper Title	MACHINE LEARNING LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2

SEMESTER VI

Semester	VI
Paper Code	BCADA 6123
Paper Title	ENABLING TECHNOLOGIES FOR DATA SCIENCE

Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

This course explains the key issues in big data management and trains the students to have skills that will help them to solve complex real-world problems for prediction and decision making using different tools. We give them some idea about the cloud environment.

COURSE OUTCOMES:

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

CO1: Understand the key issues in big data management and its associated applications using Hadoop
CO2: Acquire fundamental enabling techniques and scalable algorithms like Map Reduce
CO3: Acquire fundamental enabling techniques of HIVE,SQOOP, PIG

UNIT 1: BIG DATA AND HADOOP **15 Hrs.**

Hadoop architecture, Hadoop Versioning and configuration, Single node & Multi-node Hadoop, Hadoop commands, Models in Hadoop, Hadoop daemon, Task instance, Illustrations.

UNIT 2: MAP-REDUCE **15 Hrs.**

Framework, Developing Map-Reduce program, Life cycle method, Serialization, Running Map-Reduce in local and pseudo-distributed mode, Illustrations

UNIT 3: Hive, SQOOP, PIG **15 Hrs.**

SUGGESTED BOOKS:

1. Hadoop in Action : Chuck Lam, 2010, ISBN : 9781935182191

2. Data-intensive Text Processing with Map Reduce : Jimmy Lin and Chris Dyer, Morgan& Claypool Publishers, 2010

Semester	IV
Paper Code	BCADA 6223

Paper Title	MACHINE LEARNING
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

This course will provide the students to understand the concepts of Machine Learning, supervised learning and their applications, the concepts and algorithms of unsupervised learning, the concepts and algorithms of advanced learning.

COURSE OUTCOMES:

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

CO1: Design a learning model appropriate to the application.
CO2: Design a supervised learning for an application of your choice.
CO3: Design an unsupervised learning for an application of your choice.
CO4: Identify applications dimensionality reduction suitable for different types of Machine Learning with suitable justification.

UNIT I MACHINE LEARNING BASICS **9 Hrs**

Introduction to Machine Learning (ML) - Essential concepts of ML – Types of learning – Machine learning methods based on Time – Dimensionality – Linearity and Non linearity – Early trends in Machine learning – Data Understanding Representation and visualization.

UNIT II MACHINE LEARNING METHODS **9 Hrs**

Linear methods – Regression -Classification –Perceptron and Neural networks – Decision trees – Support vector machines – Probabilistic models —Unsupervised learning – Featurization

UNIT III MACHINE LEARNING IN PRACTICE **9 Hrs**

Ranking – Recommendation System - Designing and Tuning model pipelines- Performance measurement – Azure Machine Learning – Open-source Machine Learning libraries – Amazon's Machine Learning Tool Kit: Sagemaker

UNIT IV MACHINE LEARNING AND DATA ANALYTICS **9 Hrs**

Machine Learning for Predictive Data Analytics – Data to Insights to Decisions – Data Exploration – Information based Learning – Similarity based learning – Probability based learning – Error based learning – Evaluation – The art of Machine learning to Predictive Data Analytics.

UNIT V APPLICATIONS OF MACHINE LEARNING 9 Hrs

Image Recognition – Speech Recognition – Email spam and Malware Filtering – Online fraud detection – Medical Diagnosis.

SUGGESTED BOOKS:

1. Tom Mitchell, "Machine Learning", McGraw-Hill, 1997.
2. Christopher Bishop, "Pattern Recognition and Machine Learning" Springer, 2007. \
3. Stephen Marsland, "Machine Learning – An Algorithmic Perspective", Chapman andHall, CRC Press, Second Edition, 2014.
4. Kevin P. Murphy, "Machine Learning: A Probabilistic Perspective", MIT Press, 2012.
5. Ethem Alpaydin, "Introduction to Machine Learning", MIT Press, Third Edition, 2014.

Semester	VI
Paper Code	BCADA 6323
Paper Title	OPTIMIZATION TECHNIQUES
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

The objective of this course is to enable the student to

- Formulate and solve linear programming problems (LPP)
- Evaluate Integer Programming Problems, Transportation and Assignment Problems.
- Obtain solution to network problems using CPM and PERT techniques.

- Able to optimize the function subject to the constraints.
- Identify and solve problems under Markovian queuing models.

COURSE OUTCOMES:

On successful completion of this course, the student will able to

CO1: Formulate and solve linear programming problems (LPP)
CO2: Evaluate Integer Programming Problems, Transportation and Assignment Problems.
CO3: Obtain solution to network problems using CPM and PERT techniques.
CO4: Able to optimize the function subject to the constraints.
CO5: Identify and solve problems under Markovian queuing models

UNIT I: LINEAR MODELS **9 Hrs.**

Introduction of Operations Research - mathematical formulation of LPP- Graphical Methods to solve LPP- Simplex Method- Big M method, Two-Phase method

UNIT II: INTEGER PROGRAMMING AND TRANSPORTATION PROBLEMS **9 Hrs.**

Integer programming: Branch and bound method- Transportation and Assignment problems -Travelling salesman problem.

UNIT III: PROJECT SCHEDULING **9 Hrs.**

Project network -Diagram representation – Floats - Critical path method (CPM) – PERT- Cost considerations in PERT and CPM

UNIT IV: CLASSICAL OPTIMISATION THEORY **9 Hrs.**

Unconstrained problems – necessary and sufficient conditions - Newton-Raphson method, Constrained problems – equality constraints – inequality constraints - Kuhn-Tucker conditions.

UNIT V: QUEUING MODELS **9 Hrs.**

Introduction, Queuing Theory, Operating characteristics of a Queuing system, Constituents of a Queuing system, Service facility, Queue discipline, Single channel models, multiple service channels.

TEXT BOOK:

1. Hamdy A Taha, Operations Research: An Introduction, Pearson, 10th Edition, 2017.

REFERENCE BOOKS:

4. ND Vohra, Quantitative Techniques in Management, Tata McGraw Hill, 4th Edition, 2011.
5. J. K. Sharma, Operations Research Theory and Applications, Macmillan, 5th Edition, 2012.
6. Hiller F.S, Liberman G.J, Introduction to Operations Research, 10th Edition McGraw Hill, 2017.
7. Jit. S. Chandran, Mahendran P. Kawatra, KiHoKim, Essentials of Linear Programming, Vikas Publishing House Pvt.Ltd. New Delhi, 1994.

8. Ravindran A., Philip D.T., and Solberg J.J., Operations Research, John Wiley, 2nd Edition, 2007.

Semester	VI
Paper Code	DSE-AC-02
Paper Title	AI & IoT
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVES:

- To Understand the basic concepts of intelligent agents
- To develop general-purpose problem solving agents and logical reasoning.
- To understand Smart Objects and IoT Architectures
- To learn about various IOT-related protocols.

COURSE OUTCOMES:

On successful completion of this course, the student will able to

CO1: Explain autonomous agents that make effective decisions in fully informed, partially observable, and adversarial settings
CO2: Choose appropriate algorithms for solving given AI problems Understand Smart Objects and IoT Architectures
CO3: Understand Fundamentals of IoT
CO4: Analyze various IoT protocols
CO5: Design a PoC of an IoT system using Rasperry Pi/Arduino

UNIT I: INTELLIGENT AGENTS **9 Hrs.**

Introduction to AI – Agents and Environments – Concept of rationality – Nature of environments – Structure of agents Problem solving agents – search algorithms – uninformed search strategies

UNIT II: PROBLEM SOLVING **9 Hrs.**

Heuristic search strategies – heuristic functions Local search and optimization problems – local search in continuous space – search with nondeterministic actions – search in partially observable environments – online search agents and unknown environments.

UNIT III: FUNDAMENTALS OF IoT **9 Hrs.**

Evolution of Internet of Things - Enabling Technologies – IoT Architectures: oneM2M, IoT World Forum (IoTWF) and Alternative IoT models – Simplified IoT Architecture and Core IoT Functional Stack – Fog, Edge and Cloud in IoT – Functional blocks of an IoT ecosystem – Sensors, Actuators, Smart Objects and Connecting Smart Objects.

UNIT IV: IoT PROTOCOLS **9 Hrs.**

IoT Access Technologies: Physical and MAC layers, topology and Security of IEEE 802.15.4, 802.15.4g, 802.15.4e, 1901.2a, 802.11ah and LoRaWAN – Network Layer: IP versions, Constrained Nodes and Constrained Networks – Optimizing IP for IoT: From 6LoWPAN to 6Lo, Routing over Low Power and Lossy Networks – Application Transport Methods: Supervisory Control and Data Acquisition – Application Layer Protocols: CoAP and MQTT.

UNIT V: DESIGN AND DEVELOPMENT **9 Hrs.**

Design Methodology - Embedded computing logic - Microcontroller, System on Chips - IoT system building blocks - Arduino - Board details, IDE programming - Raspberry Pi.

TEXT BOOKS:

1. Stuart Russel and Peter Norvig, "Artificial Intelligence: A Modern Approach", Fourth Edition, Pearson Education, 2020
2. David Hanes, Gonzalo Salgueiro, Patrick Grossetete, Rob Barton and Jerome Henry, "IoT Fundamentals: Networking Technologies, Protocols and Use Cases for Internet of Things, Cisco Press, 2017

REFERENCE BOOKS:

1. Arshdeep Bahga, Vijay Madisetti, "Internet of Things – A hands-on approach", Universities Press, 2015
2. Olivier Hersent, David Boswarthick, Omar Elloumi , "The Internet of Things – Key applications and Protocols", Wiley, 2012 (for Unit 2).
3. Dan W. Patterson, "Introduction to AI and ES", Pearson Education, 2007.
4. Kevin Night, Elaine Rich, and Nair B., "Artificial Intelligence", McGraw Hill, 2008

Semester	VI
Paper Code	VC-AC-01

Paper Title	INTRODUCTION TO POWER BI
Number of teaching hrs per week	3 Hrs
Total number of teaching hrs per semester	45
Number of credits	3

COURSE OBJECTIVE:

- Understand the basic concepts of data analysis, data visualization, and data modeling
- Navigate the Power BI Desktop interface.
- Import data from various sources into Power BI
- Create basic data models and relationships between tables.
- Use DAX (Data Analysis Expressions) to create calculated columns and measures.
- Create interactive reports and visualizations using various chart types, tables, and other visualization tools.
- Design and publish dashboards for stakeholders.
- Use Power BI service to share dashboards and reports with others.
- Understand the security and governance considerations in using Power BI.

COURSE OUTCOME:

On successful completion of this course, the student will able to

CO1: Understanding of basic concepts of Power BI
CO2: Creating and and applying data models in Power BI
CO3: Visualising datasets using Power BI
CO4: Understanding and applying data modelling and visualisation techniques using the advanced features of Power BI
CO5: Creating Dashboard using Power BI
CO6: Learning security, Collaboration and best practices of Power BI

UNIT I: INTRODUCTION TO POWER BI **7Hrs**

- Introduction to Power BI and its features
- Power BI Desktop and Power BI Service
- Data sources and connectors

UNIT II: DATA MODELING IN POWER BI **7 Hrs**

- Creating data models and relationships

- Introduction to DAX (Data Analysis Expressions)

UNIT III: DATA VISUALIZATION IN POWER BI **7 Hrs**

- Introduction to data visualization
- Creating charts, tables, and other visualization tools
- Using filters and slicers

UNIT IV: ADVANCED DATA MODELING AND VISUALIZATION **8 Hrs**

- Advanced data modeling techniques
- Advanced visualization techniques
- Custom visuals and importing from AppSource.

UNIT V: CREATING DASHBOARDS **8 Hrs**

- Introduction to dashboards
- Designing and creating dashboards in Power BI
- Sharing and collaborating on dashboards

UNIT VI: SECURITY AND GOVERNANCE IN POWER BI **8 Hrs**

- Security considerations in Power BI
- Sharing and collaboration in Power BI
- Best practices for governance and administration

SUGGESTED BOOKS:

1. Mastering Microsoft Power BI By Brett Powell
2. Beginning Power BI by Dan Clark

LABORATORY

Semester	VI
Paper Code	BCADA6P1
Paper Title	ENABLING TECHNOLOGIES IN DATA SCIENCE LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60

Number of credits	2

Semester	VI
Paper Code	BCADA6P2
Paper Title	DATA MINING AND ANALYTICS LAB
Number of teaching hrs per week	4 Hrs
Total number of teaching hrs per semester	60
Number of credits	2