School of Computer Science, Engineering & Applications

Bharathidasan University

Tiruchirappalli - 620023

M.Sc., Data Science - Course Structure and Syllabus

(Applicable to the candidates admitted from the academic year 2022-2023 onwards)

BHARATHIDASAN UNIVERSITY

SCHOOL OF COMPUTER SCIENCE, ENGINEERING & APPLICATIONS

MASTER OF DATA SCIENCE

(CHOICE-BASED CREDIT SYSTEM)

REGULATIONS

(w.e.f 2022-2023)

1. **Name of the Programme**: Bharathidasan University is offering a two year M.Sc., (Data Science) Programme to be conducted in the Department of Computer Science, Bharathidasan University.
2. **Aim of the Course**: The course strives to inculcate job-oriented and value based quality education in Information Technology and Commercial Application Development. At the end of the course, the students will be well-versed, particularly in core subjects with quality in inter-personal and professional skills.
3. **Eligibility for Admission to the Programme:** A Candidates who has passed B.Sc.(Computer Science, Computer Technology, Information Technology, Information Science, Information Systems, Software Science, Software Development, Statistics, Mathematics), BCA of this University or from a recognized University.
4. **Duration of the Course:** The Course duration shall be for two years consisting of four semesters. In order to be eligible for the award of the degree of the candidate shall successfully complete the course in a maximum period of four years from the date of enrolment for the first semester of the course.
5. **Choice Based Credit System:** The University follows the 'Choice Based Credit System (CBCS)' for all its programmes. Each credit is worth 12 hours of student study time, comprising all learning activities. Thus a four-credit course involves 48 study hours. This helps the student to understand the academic effort and to successfully complete a course.
6. **Structure of the Course and Evaluation Pattern:**

Internal Marks: 25 External Marks: 75

The duration of University examination for both theory and practical subjects shall be 3 hours. The maximum marks for each theory and practical course is 100. Continues Internal

Assessment (CIA) will be 25. The University theory examination will be conducted for 75 marks, which will be added with Continuous Internal Assessment to make 100 marks for the course. For the conduct of University examinations in theory, the question paper for the theory examination will be set by the external examiner and for practical examination; the question paper will be set by both internal and external examiners appointed by the Department.

7. Attendance:

- Students should gain at least 75% attendance in each course.
- In each semester every candidate must compulsorily register for the examination in all the courses pertaining to that semester.
- Candidate who has less than 75% attendance of the working days in a semester will be permitted to take the ESE pertaining to that semester unless he/she gets a condonation certificate.
- On the day on which a course is concluded, the course teacher of the course shall intimate the Head of the Department, of the particulars of all the students who have a shortage of attendance in the course offered by him/her.
- The Head of the Department should announce the names of all the students who will not be eligible to take the end-semester examination in the various courses due to the shortage of attendance.
- Condonation of shortage of attendance shall be given as per the provisions given below:
- The Head of the Department may condone lack of attendance for a student in a course when the student had put in at least 65% attendance for the course concerned and pays a condonation fee of the suitable amount to be prescribed by the University from time to time. A candidate who has put in less than 65% attendance must repeat the course.

8. Procedures for Awarding Marks for Internal Assessment

THEORY COURSES

- For regularity and discipline - 5 Marks
- For two assignments (Equal weightage) - 5 Marks

- For two tests to be conducted (Equal weightage) - 5 Marks
- Model examination - 10 Marks

Total –25 Marks

PRACTICAL COURSES

- For regularity and discipline - 5 Marks
- Completion of all Experiments prescribed for the course - 5 Marks
- Observation Note - 5 Marks
- For model Examination - 10 Marks

Total - 25 Marks

In the case of CIA, a candidate who secures not less than 40% of total marks prescribed for any course shall be declared to have passed for that course, failing which the candidate has to redo the academic activities prescribed for the Continuous Internal Assessment (CIA).

9. Pattern of Question Paper (Theory)

Time 3 hours Max Marks 75

Section - A: (10 X 2 = 20 Marks)

All questions are to be answered. Two questions from each Unit

Section – B: (5X 5 = 25 Marks)

Five Questions are to be answered, Two questions from each unit in the either or Pattern

Section - C :(3 x 10 = 30 Marks)

Three Questions are to be answered out of five questions – One question from each unit

10. Passing Minimum for a course:

A candidate shall be declared to have passed a certain course if he/she secures not less than 40% marks in the ESE and not less than 50% in the aggregate taking CIA and ESE together. In

terms of marks, the student should secure not less than 38 marks and not less than 50 in the aggregate taking CIA and ESE together.

A candidate shall be declared to have passed in the project work if he/she gets not less than 40% in the valuation of dissertation and not less than 50% in the aggregate of both the marks for valuation of dissertation and viva-voce examination to pass in project work.

11. Supplementary Examinations:

Any candidate, for whom, when the results of the exam conducted in April/May of the ESE are published there is just one course for which credit is to be earned and it is from the final semester, will be eligible to appear for the supplementary exam which may be conducted in the supplementary July/August provided that he/she appears for the ESE for that courses in April/May.

M.Sc., DATA SCIENCE PROGRAMME COURSE STRUCTURE UNDER CBCS

Programme Outcomes (POs):

S.No.	Programme Outcomes
PO1	**Technical Knowledge:** Apply knowledge of mathematics, statistics, science and computing appropriately to model the software applications, configure software platforms and analyze real-time data in heterogeneous domains.
PO2	**Problem Analysis:** Design a system, component or process, or tools to meet desired needs within realistic constraints such as economic, environmental, social, ethical and safety contexts.
PO3	**Design/development of Solutions**: Have the ability to design, implement, evaluate, analyze, and interpret complex problems and data, provide sustainable computational solutions and synthesis the information to provide a valid conclusion for domains of business, healthcare, and the environment.
PO4	**Modern tool usage:** Create, Select and apply appropriate technologies, tools, and techniques for data modelling, processing complex problems, data analysis, and prediction analysis.
PO5	**Communication:** Communicate effectively with the computing community, and with society, about complex computing activities by being able to comprehend and write effective reports, design documentation, demographics and make effective presentations.
PO6	**Project management:** Manage projects and function effectively as an individual, member or leader in diverse teams, and multidisciplinary settings.
PO7	**Sustainability:** Understand the impact of professional analytical solutions in societal and environmental contexts and apply the knowledge to benefit individuals for sustainable development.
PO8	**Life-long learning:** Recognize the need for, and prepare them to engage in independent lifelong learning in the context of techno advancements for the betterment of individuals, organizations, research community and society.
PO9	**Ethics:** Apply ethical principles, and commit to professional ethics and responsibilities and human values.
PO10	**Societal Contribution:** Utilize the knowledge education on an understanding of data, management principles, and computing solutions to apply to one's own work, as a member and leader of the projects to manage projects in multidisciplinary environments and societal contexts.

M.Sc., Data Science - Program Specific Outcomes (PSOs)

After the successful completion of the M.Sc., Data Science Programme, the students are expected to

PSO1	Evolve AI-based efficient domain-specific processes for effective decision-making in several domains such as business and governance domains.
PSO2	Acquire knowledge of Data Science Principles and Components Data Acquisition, Data Transformations, and Big Data Platforms for analysis and Interpretation.
PSO3	Arrive at actionable Foresight, Insight, and hindsight from data for solving business and engineering problems.
PSO4	Sound Knowledge of constructing data into meaningful structures by data curation and reporting to predict and gather valuable Data Insight.
PSO5	Knowledge of using Statistics, Mathematics in designing Models and Algorithms for achieving Business Objectives.
PSO6	Sound Knowledge of Data Science, Big Data Technology Tools, Natural Language Processing, Visualization, Database Management, Machine Learning and Programming for Analytics of Large scale Data to support business processes and functions.
PSO7	Apply data science methods in assessing data requirements and integrating data analytic problem framework for domain-specific applications.
PSO8	Communicate data assumptions, analysis and insights in written and visual dashboards and articulate data stories.
PSO9	Carry knowledge of professionals for ethical responsibility on data ownership and data privacy.
PSO10	Carry out fundamental research to cater for the critical needs of society through cutting-edge technologies of AI.

Master of Data Science – Course Structure and Syllabus under CBCS

(Applicable to the candidates admitted from the academic year 2022 – 2023 onwards)

Course Number	Course Code	Course Name	L	P	C	Marks		Total
						Int.	Ext.	
Semester I								
MDS22011	CC-I	Principles of Data Science	5	0	4	25	75	100
MDS22012	CC-II	Mathematical Foundation for Data Science	5	0	4	25	75	100
MDS22013	CC-III	Problem solving using Python	5	0	4	25	75	100
MDS22014	CC-IV	Database Systems	4	0	4	25	75	100
MDS22015	CC-V	Statistical Foundation for Data Analytics	5	0	4	25	75	100
MDS22016P	CC-VI	Python Programming Lab	0	3	2	25	75	100
MDS22017P	CC-VII	Database Systems Lab	0	3	2	25	75	100
22VAC01	VAC-I	Value Added Course - I	0	0	0	25	75	100
		Total	24	6	24			**800**
Semester II								
MDS22021	CC-VIII	Data Visualization	5	0	4	25	75	100
MDS22022	CC-IX	Artificial Intelligence & Machine Learning	5	0	4	25	75	100
MDS22023	CC-X	Big Data Analytics Framework	4	0	4	25	75	100
MDS22024	EC-I	Elective-I	4	0	4	25	75	100
MDS22025	EC-II	Elective-II	4	0	4	25	75	100
MDS22026P	CC-XI	Data Visualization Lab	0	3	2	25	75	100
MDS22027P	CC-XII	Big Data Analytics Lab	0	3	2	25	75	100

EDC-I		Extra Disciplinary Course - I	2	0	2	25	75	100
MDS22028	Online Course - I	SWAYAM / MOOCs / NPTEL	-	-	2	25	75	100
Total			24	6	28			**900**
Summer Internship								
Semester III								
MDS22031	CC-XIII	Principles of Deep Learning	5	0	4	25	75	100
MDS22032	CC-XIV	Predictive Analytics	5	0	4	25	75	100
MDS22033	CC-XV	Social Media Analytics	4	0	4	25	75	100
MDS22034	EC-III	Elective – III	4	0	4	25	75	100
MDS22035	EC-IV	Elective-IV	4	0	4	25	75	100
MDS22036P	CC-XVI	Deep Learning Lab	0	3	2	25	75	100
MDS22037P	CC-XVII	Predictive Analytics Lab	0	3	2	25	75	100
EDC-II		Extra Disciplinary Lab	2	0	2	25	75	100
MDS22038	Online Course-II	SWAYAM / MOOCs / NPTEL			2	30	70	100
22VAC02	VAC-II	Value Added Course - II	0	0	0	25	75	100
		Total	24	6	28	**1000**		
Semester IV								
MDS22041	CC-XVIII	Major Project	0	0	14			100
TOTAL					**94**	**2800**		

Elective - I			**Elective - II**
a)	Machine Learning Operations (MLOps)	a)	Image and Video Analytics
b)	Internet of Things	b)	Digital Forensics
c)	Cloud Computing	c)	Regression Analytics
Elective - III			**Elective - IV**
a)	Project Management	a)	Business Analytics
b)	IBM Watson Tools	b)	Evolutionary Computing
c)	Time Series Analysis and Forecasting	c)	Sensor Analytics

Value Added Courses		
22VAC01	VAC - I	Augmented Reality & Virtual Reality
22VAC02	VAC - II	Robotics

M.Sc., DATA SCIENCE PROGRAMME COURSE STRUCTURE UNDER CBCS

SEMESTER I

Course Number: MDS22011 **CC-I**
L-P:5-0 **Credits: 4**

PRINCIPLES OF DATA SCIENCE

Objectives:

- To gain knowledge in the basic concepts of Data Analysis
- To acquire skills in data preparatory and preprocessing steps.
- To understand the mathematical skills in statistics.
- To understand the concepts of Artificial Intelligence Roles and Skills in Data Science.
- To understand the role of Data Science in Real-time applications

UNIT I INTRODUCTION

Need for data science – benefits and uses – facets of data – data science process – setting the research goal – retrieving data – cleansing, integrating, and transforming data – exploratory data analysis – building the models – presenting and building applications.

UNIT II DESCRIBING DATA - I

Frequency distributions – Outliers – relative frequency distributions – cumulative frequ distributions – frequency distributions for nominal data – interpreting distributions – grap averages – mode – median – mean – averages for qualitative and ranked data – descri variability – range – variance – standard deviation – degrees of freedom – interquartile ran variability for qualitative and ranked data.

UNIT III DESCRIBING DATA - II

Normal distributions – z scores – normal curve problems – finding proportions – finding scores – more about z scores – correlation – scatter plots – correlation coefficient for quantitative data – computational formula for correlation coefficient – regression – regression line – least squares regression line – standard error of estimate – interpretation of r2 – multiple regression equations – regression toward the mean.

UNIT IV AI: ROLES AND SKILLS

AI: Cognitive Computing: Learning Perceptions – Terminologies - Machine Learning – Neural Networks – Deep Learning - NLP – Speech Processing – Big Data and AI – Ethics in

AI Research - Advanced Applications – AI Myths – Data Science Roles Data Scientist, Data Architect, Data Analyst – Machine Learning Engineer – Skills.

UNIT V DATA SCIENCE USE CASES

Data Science Use cases Specifications and Discussion – Data Sources Identification – Data Types –Data Classification – Data Characteristics of Big V‟s – Data Science P‟s – Applications of AI: Domains: Customer Insights – Behavioral Analysis –- Marketing – Retails – Insurance – Risk and Security –Health care – Supply Chain Logistics.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. David Cielen, Arno D. B. Meysman, and Mohamed Ali, “Introducing Data Science”, Manning Publications, 2016. (first two chapters for Unit I)

2. Robert S. Witte and John S. Witte, “Statistics”, Eleventh Edition, Wiley Publications, 2017. (Chapters 1–7 for Units II and III)

3. Joel Grus, “Data Science from Scratch”, 2nd Edition, O′Reilly Publisher, ISBN: 9781492041139, May 2019 (for Unit IV and V)

Reference Books:

1. Lillian Pierson, Jake Porway, “Data Science for Dummies”, Second Edition, John Wiley & Sons, Publishers, ISBN: 9781119327639, 2017 **(**EBook**)**

2. Sinan Ozdemir, Sunil Kakade, “Principles of Data Science”, Second Edition **(**EBook**)**

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

- On successful completion of this course, students would be able to
- Understand the foundational concepts of Data Science.
- Understand the nature of Data.
- Determine the relationship between data dependencies using statistics.
- Understand the concepts of Artificial Intelligence Roles and Skills in Data Science.

Course Number: MDS22012 **CC-II**
L-P:5-0 **Credits: 4**

MATHEMATICAL FOUNDATION FOR DATA SCIENCE

Objectives:

- To understand the properties of Linear Programming
- To understand the properties of dynamic Programming
- To use the properties of Linear Equations and Vector Spaces.
- To demonstrate proficiency on the topics of Eigenvalues, Eigenvectors and Inner Product Spaces.
- To understand the properties of multivariable calculus and graphs.

UNIT I LINEAR PROGRAMMING PROBLEM

Introduction to Operations Research: Basics definition - scope – objectives - phases - models - limitations of Operations Research - Linear Programming Problem - Formulation of LPP - Graphical solution of LPP - Simplex Method - Artificial variables - Big-M method - Two-phase method - Degeneracy - Unbound solutions – Duality in Linear Programming Problems – Dual Simplex - Introduction to optimization - gradient descent method - convex optimization.

UNIT II DYNAMIC PROGRAMMING

Introduction - Characteristics of dynamic programming – Dynamic programming approach for Priority Management employment smoothening – capital budgeting – Stage Coach/Shortest Path – cargo loading and Reliability problems.

UNIT III GEOMETRY LINEAR EQUATIONS AND VECTOR SPACES

The Geometry of Linear Equations - An Example of Gaussian Elimination- Matrix Notation and Matrix Multiplication - Triangular Factors and Row Exchanges- Inverses and Transposes. Introduction to Vector Spaces, Definition of Vector spaces, Subspaces, sums of Subspaces, Direct Sums, Span and Linear Independence, bases, and dimension.

UNIT IV EIGEN VALUES, EIGEN VECTORS AND INNER PRODUCT SPACES

Eigenvalues and Eigenvectors - Eigenvectors and Upper Triangular matrices – Eigenspaces and Diagonal Matrices - Inequalities on Linear Spaces - Norms on Linear Spaces - Inner products - Orthogonality - Unitary and Orthogonal Matrices - Norms for matrices.

UNIT V CALCULUS OF SEVERAL VARIABLES AND BASIC GRAPH THEORY

Functions of Several Variables - Limits and continuity in Higher Dimensions – Partial

Derivatives - The Chain Rule - Directional Derivative and Gradient vectors - Tangent Planes and Differentials - Extreme Values and Saddle Points - Lagrange Multipliers. Graphs - subgraphs - factors - Paths - cycles - connectedness - trees - Euler tours -Hamiltonian cycles - Planar Graphs - Digraphs.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. K Sharma, "Operations Research Theory &Applications", 6th Edition, Laxmi Publications, 2017.
2. Gilbert Strang, "Linear Algebra and Its Application", 5th Edition, Wellesley Cambridge Press, ISBN: 9780980232776, 2017.
3. S. Axler, "Linear algebra done right", Springer, 2017.
4. Eldén Lars, "Matrix methods in data mining and pattern recognition, Society for Industrial and Applied Mathematics", 2007.
5. M. D. Weir, J. Hass, and G. B. Thomas, "Thomas' calculus", Pearson, 2016.
6. D. Jungnickel, "Graphs, networks and algorithms". Springer, 2014.

Reference Books:

1. P. K. Gupta and D. S. Hira, "Operations Research", S. Chand & co., 2017
2. David C. Lay, Steven R. Lay, Judi J. McDonald, "Linear Algebra and Its Applications", 5th Edition, Pearson Education, 2016.
3. J. V. Kepner and J. R. Gilbert, "Graph algorithms in the language of linear algebra", Society for Industrial and Applied Mathematics, 2011.
4. D. A. Simovici, "Linear algebra tools for data mining", World Scientific Publishing, 2012.
5. P. N. Klein, "Coding the Matrix: linear algebra through applications to computer science", Newtonian Press, 2015.
6. J. Patterson and A. Gibson, "Deep learning: a practitioner's approach", O'Reilly Media, 2017.
7. S. Sra, S. Nowozin, and S. J. Wright, "Optimization for machine learning", MIT Press, 2012.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

- On successful completion of this course, students would be able to
- Understand the properties of Linear Programming
- Understand the properties of dynamic Programming
- Use the properties of Linear Equations and Vector Spaces
- Demonstrate proficiency on the topics Eigenvalues, Eigenvectors and Inner Product Spaces.
- Understand the properties of multivariable calculus and graphs.

Course Number: MDS22013 **CC-III**
L-P:5-0 **Credits: 4**

PROBLEM SOLVING USING PYTHON

Objectives:

- To acquire programming skills in core Python
- To study data structures in Python.
- To get familiar in modules and packages.
- To develop the skill of designing web applications in Python.
- To develop the ability to write database applications in Python.

UNIT I INTRODUCTION TO PYTHON PROGRAMMING

Python interpreter and interactive mode; values and types variables, expressions, statements, Order of operations, comments, debugging; modules and functions: function Calls, adding new functions, Definitions and Uses, flow of execution, parameters and arguments, Fruitful functions. Conditionals: Boolean values and operators, conditional (if), alternative (if-else), chained conditional (if-elif-else); Iteration: state, while, for, range, break, continue, pass; recursion; Strings: string slices, immutability, string functions and methods, string module; Lists as arrays.

UNIT II LISTS, TUPLES, DICTIONARIES

Lists: Traversing a List, list operations, list slices, list methods, Map, Filter and Reduce, list loop, mutability, aliasing, cloning lists, list parameters; Dictionaries: operations and methods; advanced list processing - list comprehension; Tuples: tuple assignment, tuple as a return value

UNIT III FILES, MODULES, PACKAGES

Files and Input/output; text files, reading and writing files, format operator; command line arguments, Errors and Exception; detecting and handling an exception, raising exceptions and Assertions, Modules; importing modules, Features, Packages: PANDAS and NUMPY

UNIT IV NETWORK AND WEB PROGRAMMING

Network programming; Socket communication, socket server module, Internet client programming, Transferring files, Electronic mail and related modules, Web Programming: Creating simple web client, Process client data and building CGI applications

UNIT V DATABASE AND GUI PROGRAMMING

Introduction to tkinter, Top Level Windows, Dialogs, Message and Entry, Event Handling, Menus, List boxes and Scrollbars, Text, SQL Database interfaces with sqlite: Basic operations and table load scripts

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Allen B. Downey, "Think Python: How to Think Like a Computer Scientist", 2nd edition, Updated for Python 3, Shroff/O Reilly Publishers, 2016.
2. Wesley. Chun, "Core Python Programming", 2ndEdition, Pearson Education, 2007.

References:

1. Guido van Rossum and Fred L. Drake Jr, "An Introduction to Python – Revised and updated for Python 3.2", Network Theory Ltd., 2011.
2. Mark Lutz, "Learning Python", O Reily, 4th Edition,2009,
3. Magnus Lie Hetland, "Beginning Python: From Novice To professional", Second Edition, 2005.
4. Mark Summerfield, "Programming in Python 3", Pearson Education, 2010.

E-Resources:

- http://www.network-theory.co.uk/docs/pytut"
- http://www.network-theory.co.uk/docs/pytut
- http://docs.python.org/tutorial/
- http://docs.python.org/tutorial/
- http://index-f.es/Python/Core.Python.Programming.2nd.Edition.Wesley.Chun.2006.pdf

Course Outcomes:

- After completing the course students are able to understand the basics of python programming.
- Explain the use of the built-in data structures list, sets, tuples and dictionary.
- Make use of modules and create python modules.
- Build real-world applications using network and web programming in python.
- Create database applications and GUI-based applications in python.

Course Number: MDS22014 **CC-IV**
L-P:4-0 **Credits: 4**

DATABASE SYSTEMS

Objectives:

- To understand the fundamental concepts of Database systems.
- To understand the SQL Statements.
- To understand competency in designing NoSQL database management systems
- To understand the competency in selecting a particular NoSQL database for specific use cases.
- To understand the concepts of MongoDB.

UNIT I INTRODUCTION

Database Systems, View of Data Models, Database Languages, DBMS Architecture, Database Users and Data Independence. ER Modeling, relation types, role and Structural Constraints, Extended ER Modeling Features, Design of an ER Database Schema, Reduction of ER Schema to Tables.

UNIT II STRUCTURED QUERY LANGUAGE (SQL)

Introduction to SQL: SQL data types and literals, Types of SQL commands, SQL operators, Tables, views and indexes, Queries and subqueries, Aggregate functions, and Cursors in SQL.

UNIT III NOSQL INTRODUCTION

Why NoSQL – Value of Relational Database – Emergence of NoSQL – Aggregate data models – More details on data models: Relationships, Graphs DB, Schema less DB, Materialized views – Distribution models: Single server, sharding, replication – Consistency: Update, read, relax consistency.

UNIT IV NOSQL DATABASES

Key value databases: What is Key-Value store, Features of Key value DB, Suitable use cases, When not to use it – Document databases: Definition, features, Suitable use cases, when not to use – Column family stores: Definition, features, suitable use cases, when not to use – Graph databases: Definition, features, use case, when not to use – Schema migration – Polyglot persistence - Beyond NoSQL – Choosing your database.

UNIT V MONGODB

Document – Collection – Database – Data types – Creating, deleting, updating documents – Querying – Indexing – Aggregation: Pipeline, Aggregation commands – Application design

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Thomas M. Connolly and Carolyn E. Begg, "Database Systems: "A Practical Approach to Design, Implementation, and Management", 6th Edition, Pearson, 2015.

2. Pramod J. Sadalage, Martin Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence". Addison-Wesley, ISBN: 0321826620, 2012.

3. Kristina Chodorow, MongoDB: The Definitive Guide, 2ed, Oreilly Publishers, 2013.

Reference Books:

1. Eric Redmond, Jim R. Wilson, "Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement", Pragmatic Bookshelf, ISBN: 1934356921, 2012.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

- On successful completion of this course, students would be able to
- Understand the fundamental concepts of Database systems
- Understand the SQL Statements.
- understand competency in designing NoSQL database management systems
- Understand the competency in selecting a particular NoSQL database for specific use cases.
- Understand the concepts of MongoDB.

Course Number: MDS22015 **CC-V**
L-P:5-0 **Credits:4**

STATISTICAL FOUNDATION FOR DATA ANALYTICS

Objectives:

- To study the various facts and problems.
- To gain knowledge from data
- To develop in-depth understanding of data analysis and data measurement
- To practice problems on Correlation and Regression
- To understand the concept of distribution

UNIT I

Analysis of Data: Measure of central tendency –Arithmetic Mean – Weight mean – median – mode – geometric mean – harmonic mean – choice of an average – characteristics of a good average. Measure of dispersion: Range – quartile deviation – Mean deviation – standard deviation – relative measure of dispersion – coefficient of variance.

UNIT II

Data Measurement: collection of data – classification of data – Introduction of skewness, Moments and kurtosis – Difference between variation and skewness – Measure of skewness Karl Pearson's coefficient of skewness – Moments and skewness – kurtosis – the concept of kurtosis Types of the curve.

UNIT III

Correlation: Definition of correlation – scatter diagram – karl Pearson's coefficient of linear correlation – coefficient of correlation and probable error of r – coefficient of determination – merits and limitation of coefficient of correlation – spearman's rank correlation.

UNIT IV

Regression Analysis: Regression and correlation [Intro] – Difference between correlation and Regression Analysis – Linear Regression equation – least square method – Regression lines – properties of Regression coefficient – standard error of the estimate.

UNIT:V

Distribution: Descriptive statistics and sampling distribution – population – sampling – standard normal distribution – chi-square distribution - t-distribution - F- distribution. Test of Hypothesis: Testing for Attributes – mean of a normal population – one-tailed and two-tailed test, F - test and chi-square test.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Book:

1. Gupta, S.C & Kapoor. V. K, ”Fundamentals of Mathematical Statistics” sultan & Chand & Sons, New Delhi, 2011

Reference Book:

1. T. Veerarajan, “Probability, Statistics and Random processes with Queuing theory and queuing Network, Third edition.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students will be able to,

- Recall the basis of statistics and correlation techniques.
- Demonstrate the application of various regression methods.
- Determine and test various Hypotheses and sampling Distribution.

Course Number: MDS22016P **CC-VI**
L-P:0-3 **Credits: 2**

PYTHON PROGRAMMING LAB

Objectives:

- To acquire programming skills in core Python.
- To understand the data structures in Python.
- To develop the skill on files and modules.
- To develop the skill of designing Graphical user Interfaces in Python.
- To develop the ability to write database applications in Python.

List of Exercises

1. Exercise programs on basic control structures &loops.
2. Exercise programs on Python Script.
3. Exercise programs on Lists.
4. Exercise programs on Strings.
5. Exercise programs on functions.
6. Exercise programs on recursion & parameter passing techniques.
7. Exercise programs on searching & sorting Techniques.
8. Exercise programs on Files
9. Exercise Program on Modules
10. Exercise Program on Exception handling
11. Exercise programs on Networking
12. Exercise programs on GUI, Graphics
13. Exercise Programs on Database Connection

Course Outcomes:

At the end of the course, students will be able to

- Understand the basics of python programming.
- Explain the use of the built-in data structures list, sets, tuples and dictionary.
- Make use of functions, strings and their applications.
- Demonstrate the use of modules, packages and their applications.
- Build real-world applications using databases, GUI and Networking.

Course Number: MDS22017P **CC-VII**
L-P:0-3 **Credits: 2**

DATABASE SYSTEMS LAB

Objectives:

- To Model data using ER diagrams
- To create tables from SQL and NoSQL
- To create queries of SQL and NoSQL
- To implement databases using any RDBMS and MongoDB
- To solve real-world problems using SQL and NoSQL.

Exercises:

1. Develop applications that will demonstrate the following concepts using SQL and NoSQL
2. Drawing E-R Diagrams for the given business problem.
3. Designing tables from E-R Diagrams
4. Creating tables using SQL and NoSQL features
5. Working with queries and subqueries
6. Performing Join operations
7. Creating Indexes
8. Solving real-world business use cases

Course Outcomes:

On successful completion of this course, students would be able to Model data using ER diagrams

- Creating tables using SQL and NoSQL
- Create queries of SQL and NoSQL
- Implement databases using any RDBMS and MongoDB
- Solve real-world problems using SQL and NoSQL.

SEMESTER II

Course Number: MDS22021 **CC-VIII**
L-P:5-0 **Credits: 4**

DATA VISUALIZATION

Objectives:

- To understand the basic concepts of Data visualization.
- To develop a maximum readability table.
- To design and implement chart animations.
- To know to learn advanced data visualization.
- To load and filter external data.

UNIT I

Introduction to Data Visualization: Acquiring and Visualizing Data - Simultaneous acquisition and visualization - Applications of Data Visualization - Keys factors of Data Visualization (Control of Presentation, Faster and Better JavaScript processing, Rise of HTML5, Lowering the implementation Bar) Exploring the Visual Data Spectrum: charting Primitives (Data Points, Line Charts, Bar Charts, Pie Charts, Area Charts), Exploring advanced Visualizations (Candlestick Charts, Bubble Charts, Surface Charts, Map Charts, Infographics). Making use of HTML5 CANVAS, Integrating SVG.

UNIT II

Basics of Data Visualization – Tables: Reading Data from Standard text files (.txt, .csv, XML), Displaying JSON content Outputting Basic Table Data (Building a table, Using Semantic Table, Configuring the columns), Assuring Maximum readability (Styling your table, Increasing readability, Adding dynamic Highlighting), Including computations - Using data tables library - relating data table to a chart.

UNIT III

Visualizing data Programmatically: Creating HTML5 CANVAS Charts (HTML5 Canvas basics, Linear interpolations, A Simple Column Chart, Animations), Starting with Google charts (Google Charts API Basics, A Basic bar chart, A basic Pie chart, Working with Chart Animations).

UNIT IV

Introduction to D3.js: Getting setup with D3 - Making selections - changing selection’s attribute -Loading and filtering External data: Building a graphic that uses all of the population distribution data - Data formats you can use with D3 - Creating a server to upload

your data - D3's function for loading data - Dealing with Asynchronous requests - Loading and formatting Large Data Sets.

UNIT V

Advanced Data Visualization: Making charts interactive and Animated: Data joins - updates and exits - interactive buttons - Updating charts - Adding transactions - using keys Adding a Play Button: wrapping the update phase in a function - Adding a Play button to the page - Making the Play button go - Allow the user to interrupt the play – sequence.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Jon Raasch, Graham Murray, Vadim Ogievetsky, Joseph Lowery, "JavaScript and jQuery for Data Analysis and Visualization", WROX, 2014.
2. Ritchie S. King, "Visual story telling with D3", Pearson, 2015.
3. Ben Fry, "Visualizing data: Exploring and explaining data with the processing environment", O'Reilly, 2008.

Reference Books:

1. A Julie Steele and Noah Iliinsky, "Designing Data Visualizations: Representing Informational Relationships", O'Relly (Ebook),
2. Andy Kirk, Data Visualization: A Successful Design Process, PAKT
3. Scott Murray, Interactive Data Visualization for Web, O'Relly
4. Nathan Yau, "Data Points: Visualization that means something", Wiley, 2013.
5. Tamara Munzner, Visualization Analysis and Design, AK Peters Visualization Series, CRC Press, Nov. 2014

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to

- Understand the basic concepts of Data visualization.
- Explore the visual data spectrum.
- Can visualize data programmatically.
- Can setup D3.js.
- Can make interactive and animated charts.

Course Number: MDS22022 **CC-IX**
L-P:5-0 **Credits: 4**

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

Objectives:

- To understand the various characteristics of Intelligent agents
- To learn the different search strategies in AI
- To learn to represent knowledge in solving AI problems
- To understand supervised and unsupervised learning.
- To know the various applications of AI.

UNIT I

Introduction –Foundation of Artificial Intelligence –Intelligent Agents– structure of agents – Definitions of a rational agent, reflex, model-based, goal-based, and utility-based agents, the environment in which a particular agent operates.

UNIT II

Problem solving Methods - Search Strategies- Uninformed - Informed - Heuristics - Local Search Algorithms and Optimization Problems - Searching with Partial Observations - Constraint Satisfaction Problems – Constraint Propagation - Backtracking Search - Game Playing - Optimal Decisions in Games – Alpha - Beta Pruning – Stochastic Games

UNIT III

Knowledge Representation: First Order Predicate Logic – Prolog Programming – Unification – Forward Chaining-Backward Chaining – Resolution – Knowledge Representation - Ontological Engineering-Categories and Objects – Events - Mental Events and Mental Objects - Reasoning Systems for Categories - Reasoning with Default Information

UNIT IV

Machine Learning: Supervised learning- learning decision trees- Linear Regression and Classification- Nonparametric Models - Unsupervised learning – Reinforcement learning- Passive Reinforcement Learning - Active Reinforcement Learning-

UNIT V

Applications of Artificial Intelligence- Deep Learning for Natural Language Processing, Computer vision-classifying images, Robotics.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", Prentice Hall, Third Edition, 2010.
2. Bratko, "Prolog: Programming for Artificial Intelligence", Fourth edition, Addison-Wesley Educational Publishers Inc.,2011.

References:

1. M. Tim Jones, "Artificial Intelligence: A Systems Approach (Computer Science)", Jones and Bartlett Publishers, Inc.; First Edition, 2008
2. Nils J. Nilsson, "The Quest for Artificial Intelligence", Cambridge University Press, 2009.
3. William F. Clocksin and Christopher S. Mellish, "Programming in Prolog: Using the ISO Standard", Fifth Edition, Springer, 2003.
4. Gerhard Weiss, "Multi Agent Systems", Second Edition, MIT Press, 2013.
5. David L. Poole and Alan K. Mackworth, "Artificial Intelligence: Foundations of Computational Agents", Cambridge University Press, 2010.

E-Resources :

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.

Course Outcome:

On successful completion of this course, students would be able to

- Gained knowledge of Intelligent Agents
- Learnt problem-solving techniques of Artificial Intelligence.
- Design applications for NLP that use Artificial Intelligence.
- To apply supervised and unsupervised learning techniques on real-time problems.
- Gained knowledge of applications of Artificial Intelligence.

Course Number: MDS22023 **CC-X**
L-P:4-0 **Credits: 4**

BIG DATA ANALYTICS FRAMEWORK

Objectives:

- To understand Big Data concepts and technologies.
- To develop applications using Hadoop.
- To store and manipulate data using HDFS.
- To explore very large datasets using Pig.
- To perform Data Warehousing operations using Hive.

UNIT I INTRODUCTION TO BIG DATA

What is Big data – Industrial examples of Big Data: Digital Marketing, fraud, risk, trading, healthcare, medicine and advertising – Big Data Technology: Hadoop, cloud, BI, crowdsourcing analytics – Business Analytics.

UNIT II MAPREDUCE-I AND HDFS

MapReduce model: Weather dataset, Analyzing data with Hadoop, Combiner functions, Hadoop streaming with Python. Hadoop Distributed File System: Block, Name node, Data node, Caching – File system operations in command line – Java Interface to Basic Hadoop - Reading data and writing data – Anatomy of File Write

UNIT III MAPREDUCE-II AND SPARK

Steps of developing MapReduce application - Working of MapReduce: Running Jobs, failure, Shuffle and sort, Task execution - MapReduce Types: Input formats - Output formats - MapReduce features: Counters, Sorting, Joins. **SPARK**- Introduction to Data Analysis with Spark, In-memory Computing with Spark, Spark Basics, Interactive Spark with PySpark, Writing Spark Applications.

UNIT IV EXPLORING LARGE DATASETS USING PIG

Structure, Statements, Expressions, Types, Schemas, Functions, Macros - User-Defined Functions: Filter UDF, Eval UDF, Load UDF - Data Processing Operators: Loading and Storing Data, Filtering Data, Grouping and Joining Data, Sorting Data, Combining and Splitting Data.

UNIT V DATA WAREHOUSING USING HIVE

Comparison with Traditional Databases - HiveQL: Data Types, Operators and Functions - Tables: Managed Tables and External Tables, Partitions and Buckets, Storage Formats, Importing Data, Altering Tables, Dropping Tables - Querying Data: Sorting and Aggregating,

MapReduce Scripts, Joins, Subqueries, Views - User-Defined Functions: Writing a UDF, Writing a UDAF.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Michael Minelli, Michele Chambers and Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", 1ed, Wiley CIO Series, ISBN 9781118147603, 2013.
2. Tom White, "Hadoop: The Definitive Guide", Fourth Edition, O'reilly Media, 2015.

Reference Books:

1. Nathan Marz and James Warren, "Big Data Principles and Practice of Scalable Real-Time Data Systems", Manning Publications. 2015.
2. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
3. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", John Wiley& sons, 2012.
4. Glenn J. Myatt, "Making Sense of Data", Volume I and II. John Wiley & Sons, 2007.
5. Mark Grover, Ted Malaska, Jonathan Seidman, Gwen Shapira, "Hadoop Application Architecture", Shroff Publishers, 2015.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to

- Understand Big Data concepts and technologies.
- Develop applications using Hadoop.
- Store and manipulate data using HDFS.
- Explore very large datasets using PIG
- Perform Data Warehousing operations using Hive.

Course Number: MDS22026P **CC-XI**

L-P:0-3 **Credits: 2**

DATA VISUALIZATION LAB

Objectives:

- To practice spreadsheet models.
- To apply oracle database connectivity.
- To visualize semi-structured data.
- To learn about tableau.
- To implement interactive plots in python.

Exercises:

1. Visualization of Spreadsheet Models.
2. Oracle Database Connectivity using Python.
3. Visualization of Semi-Structured Data.
4. Introduction to Tableau and Aggregation Methods in Tableau.
5. Visual Encodings and Basic Dashboards in Tableau.
6. Interactive Plots in Python.

Course Outcomes:

On successful completion of this course, students would be able to

- Practice Tableau.
- Apply spread sheets model.
- Visualize semi-structured data.
- Can implement visual encodings in tableau.
- Perform different types of data visualization techniques.

Course Number: MDS22027P **CC-XII**
L-P:0-3 **Credits: 2**

BIG DATA ANALYTICS LAB

Objectives:

- Know the concept of MapReduce architecture.
- Know the concept of Programming Methods.
- Understand the HDFS Commands
- Know the concept of data types in Hadoop
- Know the concept of MapReduce program.

Exercises:

1. To study big data analytics and Hadoop Architecture
2. To Understand the Overall Programming architecture of MapReducee API. Implement MapReduce Programming
3. To Study HDFS Commands
4. To Study serializes and deserializes data of integer type in Hadoop.
5. To run a basic Word Count MapReduce program to understand MapReduce
6. Paradigm.
7. Basic CRUD operations in MongoDB.
8. Store the basic information about students such as roll no and name using
9. various collection types Map
10. To run a Grep program on Hadoop to understand MapReduce Paradigm:
11. To count words in a given file, To view the output file, and calculate execution time.
12. To Study the Hive commands using HQL (DDL and DML).

Course Outcomes:

On successful completion of this course, students would be able to

- Understand Hadoop Architecture
- Implement the concept of MapReduce
- Understand the HDFS commands
- Implement basic operations in MongoDB

SEMESTER III

Course Number: MDS22031 **CC-XIII**
L-P:5-0 **Credits: 4**

PRINCIPLES OF DEEP LEARNING

Objectives:

- To learn TensorFlow programming.
- To design and implement CNN for classification.
- To perform sequence analysis.
- To design machine translation networks using RNN.
- To implement a reinforcement learning system.

UNIT I TENSORFLOW BASICS

TensorFlow: variables, operations, placeholder Tensors, sessions – Navigating variable scopes and shared variables – Managing models over CPU and GPU – Logistic Regression in TensorFlow–Training Logistic Regression model – Visualizing using TensorBoard – Building multilayer model in TensorFlow.

UNIT II CONVOLUTIONAL NEURAL NETWORK

Shortcomings of Feature Selection – Width, height and depth of layers – Filters and feature maps – Describing convolutional layer – Max pooling - Architectural Description of Convolution Networks – Recognizing handwritten digits using CNN for MNIST dataset – Image preprocessing pipelines - Training with Batch normalization.

UNIT III AUTOENCODERS AND SEQUENCE ANALYSIS

Autoencoders and Sequence Analysis Embedding – Principal Component Analysis - Architecture of Autoencoders – Implementing autoencoders in TensorFlow–Denoising - Word2Vec framework for language modelling. Sequence Analysis: seq2seq problem – Dependency parsing – Beam search.

UNIT IV RECURRENT NEURAL NETWORKS

Single neuron and fully connected recurrent layer – Challenges of vanishing gradients - LSTM architecture – TensorFlow primitives for RNN models – Implementing Sentiment analysis Model –Solving seq2seq tasks with RNN – Augmenting RNN with Attention – Designing Neural Translation Network.

UNIT V DEEP REINFORCEMENT LEARNING

Reinforcement Learning: Markov Decision Processes, Policy, Future return, Discounted future return, Balancing Explore-Exploit dilemma, Annealed e-Greedy – Policy learning and Value learning - Solving Pole Cart problem with Policy Gradients - QLearning -Deep QNetworks – Deep Q Recurrent Networks – UNREAL Learning.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Nikhil Buduma, Nicholas Locascio. Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms. O'Reilly Media. 2017.

2. Ian Goodfellow, YoshuaBengio, Aaron Courville. Deep Learning (Adaptive Computation and Machine Learning series). MIT Press, 2017.

Reference Books:

1. Francois Chollet. Deep Learning with Python. 1ed, Manning Publications, 2017. ISBN 978-1617294433.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to

- Learn TensorFlow programming.
- Design and implement CNN for classification.
- Perform sequence analysis.
- Design machine translation networks using RNN.
- Implement a reinforcement learning system

Course Number: MDS22032 **CC-XIV**
L-P:5-0 **Credits: 4**

PREDICTIVE ANALYTICS

Objectives:

- To learn predictive analytics with business intelligence.
- To design data visualization.
- To solve the problem using descriptive models.
- To understand the data using statistical models.
- To implement a reinforcement learning system.

UNIT I PREDICTIVE ANALYTICS OVERVIEW

Overview of Predictive Analytics: What Is Analytics? - What Is Predictive Analytics? - Business Intelligence - Predictive Analytics vs. Business Intelligence - Predictive Analytics vs. Statistics - Predictive Analytics vs. Data Mining - Who Uses Predictive Analytics? - Challenges in Using Predictive Analytics - What Educational Background Is Needed to Become a Predictive Modeler? Setting Up the Problem: Predictive Analytics Processing Steps: CRISP-DM - Business Understanding - Defining Data for Predictive Modeling - Defining the Target Variable - Defining Measures of Success for Predictive Models - Doing Predictive Modeling Out of Order - Case Study: Recovering Lapsed Donors - Case Study: Fraud Detection.

UNIT II UNDERSTANDING THE DATA

Data Understanding: What the Data Looks Like - Single Variable Summaries - Data Visualization in One Dimension - Histograms - Multiple Variable Summaries - Data Visualization, Two or Higher Dimensions - The Value of Statistical Significance - Pulling It All Together into a Data Audit. Data Preparation: Variable Cleaning - Feature Creation.

UNIT III RULES

Item sets and Association Rules: Terminology - Parameter Settings - How the Data Is Organized - Measures of Interesting Rules - Deploying Association Rules - Problems with Association Rules - Building Classification Rules from Association Rules. Descriptive Modeling: Data Preparation Issues with Descriptive Modeling - Principal Component Analysis - Clustering Algorithms.

UNIT IV DESCRIPTIVE MODELS

Interpreting Descriptive Models: Standard Cluster Model Interpretation. Predictive Modeling: Decision Trees - Logistic Regression - Neural Networks - K-Nearest Neighbor - Naïve Bayes -

Regression Models - Linear Regression - Other Regression Algorithms.

UNIT V PREDICTIVE MODELS

Assessing Predictive Models: Batch Approach to Model Assessment - Assessing Regression Models. Model Ensembles: Motivation for Ensembles – Bagging – Boosting - Improvements to Bagging and Boosting - Model Ensembles and Occam's Razor - Interpreting Model Ensembles.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Dean Abbott, "Applied Predictive Analytics –Principles and Techniques for the Professional Data Analyst", 2014.

Reference Books:

1. Eric Siegel, "Predictive Analytics: The Power to predict who will click, buy, lie or die", 2013.
2. John D. Kelleher, Brian Mac Namee, Aoife D'Arcy, "Fundamentals of Machine learning for predictive data analytics", 2015.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to

- Learn predictive analytics with business intelligence.
- Design data visualization.
- Solve the problem using descriptive models.
- Understand the data using statistical models.
- Implement a reinforcement learning system.

Course Number: MDS22033 **CC-XV**
L-P:4-0 **Credits: 4**

SOCIAL MEDIA ANALYTICS

Objectives:

- To understand how accurately analyze voluminous complex data set in social media and other sources.
- To understand the models and algorithms to process large data sets.
- To apply supervised and unsupervised learning algorithms on social media.
- To discover community and follow information diffusion in social networks.
- To understand social behavior and recommendation challenges and methodologies.

UNIT I SOCIAL MEDIA MINING

Social Media Mining - Introduction – Atoms – Molecules – Interactions – Social Media mining Challenges - Graphs - Basics – Nodes – Edges – Degree of Distribution- Types – Directed – Undirected – Weighted - Graph Connectivity - Tress and Forests – Bipartite graphs – Complete Graphs – Sub graphs – Planar Graphs - Graph Representation - Graph Traversal Algorithms – Shortest path algorithms Dijkstra‟s - Spanning tree algorithms – Prims - Bipartite matching - FordFulkerson algorithm.

UNIT II SOCIAL NETWORK MODELS

Network Measures: Centrality – Transitivity, reciprocity – Balance and status – Similarity. Network Models: Properties – Random graphs – Small world models – Preferential attachment model.

UNIT III SOCIAL MEDIA COMMUNITIES

Social media Communities – Social Communities – Member based Detection – Node degree – Node Similarity – Node reachability - Group Based detection methods - balanced – robust - modular – dense - hierarchical - Spectral Clustering : Balanced Community algorithm Community Evolution - Evaluation.

UNIT IV INFORMATION DIFFUSION AND INFLUENCE IN SOCIAL MEDIA

Information Diffusion: Herd behaviour – Information cascades – Diffusion of innovations – Epidemics. Influence and Homophily: Measuring Assortativity – Measuring and modelling influence – Measuring and modelling homophily – Distinguishing influence and homophily.

UNIT V RECOMMENDATION AND BEHAVIOUR ANALYSIS IN SOCIAL MEDIA

Recommendation in Social Media: Challenges – Classical recommendation algorithms – Recommendation using social context – Evaluating recommendations. Behaviour Analysis: Individual behaviour – Collective behaviour. Events Analytics in Social Media.

Text Books:

1. Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu, "Social Media Mining: An Introduction", Cambridge University Press, 2014.

Reference Books:

1. Matthew A. Russell, "Mining the Social Web". 2nd Edition. O'Reilly Media. 2013.
2. Jennifer Golbeck, Morgn Kaufmann, "Analyzing the Social Web", ISBN 978-0124055315, 2013.
3. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, "Modern Information Retrieval: The Concepts and Technology behind Search", 2ed, ACM Press Books, ISBN 978-0321416919, 2011.
4. Charu C. Aggarwal, "Social Network Data Analytics". Springer. 2011.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to

- Understand how accurately analyze voluminous complex data set in social media and other sources.
- Understand the models and algorithms to process large data sets.
- Apply supervised and unsupervised learning algorithms on social media.
- Discover community and follow information diffusion in social networks.
- Understand social behavior and recommendation challenges and methodologies.

Course Number: MDS22036P **CC-XVI**
L-P:0-3 **Credits: 2**

DEEP LEARNING LAB

Objectives:

- To learn deep neural networks and apply for simple problems
- To Learn and apply Convolution Neural Network for image processing.
- To Learn and apply Recurrent Neural Network and its variants for text analysis To augment data using generative models
- To explore real world applications with deep neural networks

Exercises:

1. Solving XOR problem using Multilayer perceptron
2. Implement character and Digit Recognition using ANN.
3. Implement the analysis of X-ray image using auto encoders
4. Develop a code to design object detection and classification for traffic analysis using CNN.
5. Implement online fraud detection of share market data using any one of the data analytics tools.
6. Implement Sentiment Analysis using LSTM.

Course Outcomes:

On successful completion of this course, students would be able to

- Apply deep neural network for simple problems.
- Apply Convolution Neural Network for image processing.
- Apply Recurrent Neural Network and its variants for text analysis.
- Apply generative models for data augmentation.
- Develop a real world application using suitable deep neural networks

Course Number: MDS22037P **CC-XVII**
L-P:0-3 **Credits: 2**

PREDICTIVE ANALYTICS LAB

Objectives:

- Understand modern notions in predictive data analysis.
- Select data, model selection, model complexity and identify the trends.
- Understand a range of machine learning algorithms along with their strengths and weaknesses.
- Build predictive models from data and analyze their performance.

Exercises:

1. Write a python program to compute
   - Central Tendency Measures: Mean, Median, Mode
   - Measure of Dispersion: Variance, Standard Deviation
2. Study of Python Basic Libraries such as Statistics, Math, Numpy and Scipy
3. Study of Python Libraries for ML applications such as Pandas and Matplotlib
4. Write a Python program to implement Simple Linear Regression
5. Implementation of Multiple Linear Regression for House Price Prediction using sklearn
6. Implementation of a Decision tree using sklearn and its parameter tuning
7. Implementation of KNN using sklearn
8. Implementation of Logistic Regression using sklearn
9. Implementation of K-Means Clustering

Course Outcomes:

- On successful completion of this course, students would be able to,
- Understand the basis python libraries (Math and Statistics) used for mathematical calculations
- Understand multidimensional array and matrix processing using Numpy library
- Understand data pre-processing and data visualization
- Understand the regression on actual dataset to do prediction

SEMESTER IV

Course Number: MDS22041 **CC-XVIII**
L-P:0-0 **Credits: 14**

MAJOR PROJECT

Objectives:

- To apply coding, debugging and testing tools and construct a new software system.
- To prepare documentation and user manuals by following the standard guidelines.
- To learn technical report writing and oral presentation skills

Methodology for Project Implementation:

For research projects, students will propose methodology based on the proposal report and validate their research with various experimental results. This way, they are expected to advance the current state of the art in their chosen area of research.

Course Outcomes:

On successful completion of this course, students would be able to

- Apply coding, debugging and testing tools and construct a new software system
- Prepare documentation and user manuals by following the standard guidelines.
- Learn technical report writing and oral presentation skills

M.Sc., DATA SCIENCE PROGRAMME COURSE STRUCTURE UNDER CBCS

ELECTIVES

Course Number: MDS22024 **EC-I**

L-P:4-0 **Credits: 4**

MACHINE LEARNING OPERATIONS (MLOPS)

Objectives:

- To learn MLOP with risk management.
- To form life cycle for monitoring MLOP.
- To create runtime environment.
- To understand the monitoring tools.
- To implement a reinforcement learning system.

UNIT I WHY MLOP?

Why Now and Challenges: Defining MLOps and Its Challenges - MLOps to Mitigate Risk - MLOps for Scale. People of MLOps: Subject Matter Experts - Data Scientists - Data Engineers - Software Engineers – DevOps - Model Risk Manager/Auditor - Machine Learning Architect.

UNIT II FEATURES OF MLOP

Key MLOps Features: A Primer on Machine Learning - Model Development - Productionalization and Deployment – Monitoring - Iteration and Life Cycle - Governance. Developing Models: What Is a Machine Learning Model? - Data Exploration - Feature Engineering and Selection – Experimentation - Evaluating and Comparing Models - Version Management and Reproducibility.

UNIT III PRODUCTION

Preparing for Production: Runtime Environments - Model Risk Evaluation - Quality Assurance for Machine Learning - Quality Assurance for Machine Learning - Key Testing Considerations - Reproducibility and Auditability - Machine Learning Security - Model Risk Mitigation. Deploying to Production: CI/CD Pipelines - Building ML Artifacts - Deployment Strategies – Containerization - Scaling Deployments - Requirements and Challenges.

UNIT IV MONITORING AND FEEDBACK

Monitoring and Feedback Loop: How Often Should Models Be Retrained? - Understanding

Model Degradation - Drift Detection in Practice - The Feedback Loop. Model Governance: Who Decides What Governance the Organization Needs? - Matching Governance with Risk Level - Current Regulations Driving MLOps Governance - The New Wave of AI-Specific Regulations - The Emergence of Responsible AI - Key Elements of Responsible AI - A Template for MLOps Governance.

UNIT V PRACTISING MLOP

MLOps in Practice: Consumer Credit Risk Management: Background: The Business Use Case - Model Development - Model Bias Considerations - Prepare for Production - Deploy to Production. MLOps in Practice: Marketing Recommendation Engines: The Rise of Recommendation Engines - Data Preparation - Design and Manage Experiments - Model Training and Deployment - Pipeline Structure and Deployment Strategy - Monitoring and Feedback.

UNIT VI

Current Contours (for Continuous Internal Assessment Only): Contemporary Developments Related to the Course during the Semester Concerned.

Text Book:

1. Mark Treveil & Dataiku Team, "Introducing MLOps – How to Scale Machine Learning in Enterprises", 2020.

Reference Books:

1. David Sweenor, Steven Hillion, Dan Rope, Dev Kannabiran, Thomas Hill, Michael O'Connell, "MLOps: Operationalizing Data Science", 2020.
2. Noah Gift & alfredo Deza, " Practical MLOps – Operationalizing Machine Learning Models", 2021.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to,

- Learn MLOP with risk management
- Form life cycle for monitoring MLOP.
- Create runtime environment.
- Understand the monitoring tools.
- Implement a reinforcement learning system.

Course Number: MDS22024 **EC-I**

L-P:4-0 **Credits: 4**

INTERNET OF THINGS

Objectives:

- To gain insight about the architecture and enabling technologies of Internet of Things
- To understand Arduino microcontroller and IDE
- To understand Smart Objects and IoT Architectures
- To learn about various IOT-related protocols
- To develop simple IoT Applications for different domains

UNIT I FUNDAMENTALS OF IoT

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 II BASIC ELECTRONICS FOR IOT&ARDUINO IDE

Understanding basic electronic components and power elements Electric Charge, Resistance, Current and Voltage – Resistors, Capacitors, Diodes, LED, Potentiometer, circuit boards - Analog and digital circuits – Microcontrollers – Electronic Signals – A/D and D/A Conversion – Pulse Width Modulation Arduino IDE: Installation and Set-up-Programming Fundamentals with Cusing Arduino IDE Program Structure in C-Basic Syntax-Data Types/Variables/Constants-Operators, Conditional Statements and Loops - Using Arduino C Library functions for Serial, delay and other invoking functions.

UNIT III ARDUINO MICROCONTROLLER AND SENSORS

Working with Arduino: LED and Switch - Data acquisition with IOT Devices - Understanding Sensors and Devices-Understanding the Inputs from Sensors-Working with Temperature Sensors-Working with Ultrasound Sensor-Working with humidity sensor-Working with Motion Sensor-Working with IR Sensor -Working with Proximity Sensor-Working with Accelerometer and vibration sensor.

UNIT IV MEDICAL SENSORS AND ACTUATORS

Understanding Medical Sensors: Flow Sensor-Optical Sensor-Body Temperature Sensor – Blood Pressure Sensor-Airflow sensor (breathing)-Patient position sensor(accelerometer)-Pulse and oxygen in blood sensor(SPO2)-Galvanic skin response (GSR-sweating) sensor. Understanding the Outputs through Actuators-Activating LED Lights-Activating Relays-Activating Buzzer -Running DC Motors-Running Stepper Motors and Servo Motors.

UNIT V DATA COMMUNICATION FROM IOT DEVICES

Building and Using Communication Devices to transfer data from IOT Devices - Understanding the Communication Principles to Transfer the data from IOT Devices; Using WIFI to Transfer the data from IOT Sensor; Programming Fundamentals with Web Applications for handling Data Communication from IOT Device; Remote Communication to cloud/external application.

Text Books:

1. 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.
2. Arshdeep Bahga, Vijay Madisetti, "Internet of Things: A Hands-On Approach", Universities Press, 2015.
3. Boris Adryan, Dominik Obermaier, Paul Fremantle, "The Technical Foundations of IoT", Artech Houser Publishers, 2017.
4. Michael Margolis, "Arduino Cook book", 2ndEdition, O'Reilly Media, 2012.
5. Marco Schwartz, "Internet of Things with ESP8266", Packt Publishing, 2016.

Reference Books

1. Charles Platt, "Make Electronics –Learning by discovery", O'Reilly Media, 2015.
2. Michael Miller, "The Internet of Things", Pearson India, 2015.

E-Learning Resources:

- https://nptel.ac.in/courses/106/105/106105166/
- https://www.arduino.cc/
- https://www.ibm.com/smarterplanet/us/en/?ca=v_smarterplanet

Course Outcomes:

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

- To learn the importance of smart objects and smart environment
- To understand and use the micro controller and various sensors

- To create programs using Arduino IDE and extract data
- To perform WiFi data communications, remote data storage in cloud, and handle the data using web applications
- To identify potential problems and develop solutions using IOT

Course Number: MDS22024 **EC-I**

L-P:4-0 **Credits:4**

CLOUD COMPUTING

Objectives:

- Gain knowledge on cloud computing, cloud services, architectures and applications.
- Enable the students to learn the basics of cloud computing with real time usage
- Know how to store and share, in and from cloud.

UNIT I INTRODUCTION

Cloud Computing Introduction, From, Collaboration to cloud, Working of cloud computing, pros and cons, benefits, developing cloud computing services, Cloud service development, discovering cloud services.

UNIT II CLOUD COMPUTING FOR EVERYONE

Centralizing email communications, cloud computing for community, collaborating on schedules, collaborating on group projects and events, cloud computing for corporation, mapping, schedules, managing projects, presenting on road.

UNIT III USING CLOUD SERVICES

Collaborating on calendars, Schedules and task management, exploring on line scheduling and planning, collaborating on event management, collaborating on contact management, collaborating on project management, collaborating on word processing, spreadsheets, and databases.

UNIT IV OUTSIDE THE CLOUD

Evaluating web mail services, Evaluating instant messaging, Evaluating web conference tools, creating groups on social networks, Evaluating on line groupware, collaborating via blogs and wikis.

UNIT V STORING AND SHARING

Understanding cloud storage, evaluating on line file storage, exploring on line book marking services, exploring on line photo editing applications, exploring photo sharing communities, controlling it with web based desktops.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books

1. Michael Miller, "Cloud Computing", Pearson Education, New Delhi, 2009.

Reference Books

1. Anthony T. Velte, "Cloud Computing: A Practical Approach", 1st Edition, Tata McGraw Hill Education Private Limited, 2009.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes::

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

- Understand the concepts of Cloud and its services
- Collaborate Cloud for Event & Project Management
- Analyze on cloud in – Word Processing, Spread Sheets, Mail, Calendar, Database
- Analyze cloud in social networks
- Explore cloud storage and sharing.

Course Number: MDS22025 **EC-II**

L-P:4-0 **Credits: 4**

IMAGE AND VIDEO ANALYTICS

Objectives:

- Will gain knowledge in the basic techniques for digital image and video processing.
- To acquire skills to build image representation, filtering, colors.
- To understand the concept of compression in image and video.
- To understand object detection & tracking.
- To understand the role of applications in video processing

UNIT II IMAGE REPRESENTATION AND PROCESSING

Digital image representation- Visual Perception- Sampling and Quantization- Basic Relations between Pixels- Mathematical Tools Used in Digital Image Processing: Fundamental Operations –Vector and Matric Operations- Image Transforms (DFT, DCT, DWT, Hadamard).

UNIT II IMAGE FILTERING

Fundamentals of spatial filtering: spatial correlation and convolution-smoothing, blurring-sharpening-edge detection - Basics of filtering in the frequency domain: smoothing-blurring-sharpening--Histograms and basic statistical models of image.

UNIT III COLORS AND COMPRESSION

Color models and Transformations – Image and Video segmentation-Image and video demonising- Image and Video enhancement- Image and Video compression.

UNIT IV OBJECT DETECTION AND TRACKING

Object detection and recognition in image and video-Texture models Image and Video classification models- Object tracking in Video.

UNIT V APPLICATIONS

Applications and Case studies- Industrial- Retail- Transportation & Travel- Remote sensing- Video Analytics in WSN: IoT Video Analytics Architectures.

Text Books

1. R.C. Gonzalezand R. E. Woods, “Digital Image Processing”, 3rd Edition, Addison Wesley, 2007.

2. Pratt, W. K, “Digital image processing: PIKS scientific inside”, 4ed. NewYork: John Wiley, 2007 .

Reference Books

1. W. Härdle, M. Müller, S. Sperlich, A. Werwatz, “Nonparametric and Semiparametric Models”. Springer, 2004.
2. Rick Szelisk, “Computer Vision: Algorithms and Applications”, Springer, 2011.
3. Jean-Yves Dufour. Intelligent Video Surveillance Systems.Wiley,2013.
4. CaifengShan, FatihPorikli, TaoXiang, ShaogangGong. “Video Analytics for Business Intelligence”, Springer, 2012.
5. Asier Perallos, Unai Hernandez-Jayo, Enrique Onieva, Ignacio Julio García Zuazola. Intelligent Transport Systems: Technologies and Applications. Wiley, 2015.
6. Basudeb Bhatta, “Analysis of Urban Growth and Sprawl from Remote Sensing Data, Springer, 2010

E-Learning Resources:

1. Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

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

- Learn the fundamentals of digital image processing, image and video analysis
- Understand and appreciate there all time use of image and video analytics.
- Demonstrate real time image and video analytics applications.
- Understand the importance of image and video analytics.
- Ability to use and understand generalizations of the image and video analytics.

Course Number: MDS22025 **EC-II**

L-P:4-0 **Credits:4**

DIGITAL FORENSICS

Objectives:

- To Introduce the Concepts of Digital Forensics.
- To be familiar with the types of Forensics Technology and of Computer Forensics Systems.
- To understand the concept of Computer Forensic Analysis
- To understand the Information Warfare.
- To know the Processing Evidence and Report Preparation and Future Issues.

UNIT I

Introduction: Computer Forensics Fundamentals – Types of Computer Forensics Technology – Types of Computer Forensics Systems – Vendor and Computer Forensics Services.

UNIT II

Computer Forensics Evidence and Capture: Data Recovery – Evidence Collection and Data Seizure - Duplication and Preservation of Digital Evidence - Computer Image Verification and Authentication.

UNIT III

Computer Forensic Analysis: Discover of Electronic Evidence - Identification of Data – Reconstructing Past Events – Fighting against Macro Threats – Information Warfare Arsenal – Tactics of the Military – Tactics of Terrorist and Rogues – Tactics of Private Companies

UNIT IV

Information Warfare: Arsenal – Surveillance Tools – Hackers and Theft of Components – Contemporary Computer Crime - Identity Theft and Identity Fraud – Organized Crime &Terrorism – Avenues Prosecution and Government Efforts – Applying the First Amendment to Computer Related Crime-The Fourth Amendment and other Legal Issues.

UNIT V

Computer Forensic Cases: Developing Forensic Capabilities – Searching and Seizing Computer Related Evidence –Processing Evidence and Report Preparation – Future Issues.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. John R. Vacca, –Computer Forensics: Computer Crime Scene Investigation‖, Laxmi Publications; First Edition, 2015. (CHAPTERS 1 – 18). (UNIT I –IV)

2. Marjie T Britz, – Computer Forensics and Cyber Crime: An Introduction‖, Pearson Education, 3rd Edition, 2013. (CHAPTERS 3 – 13). (UNIT IV –V)

Reference Books:

1. Mari E-Helen Maras, Computer Forensics: Cybercriminals, Laws, and Evidence‖, Jones & Bartlett Learning; 2nd Edition,2014.

2. Majid Yar, Cybercrime and Society, SAGE Publications Ltd, Hardcover, 2nd Edition, 2013.

E-Resources:

- https://www.cisa.gov/uscert/sites/default/files/publications/forensics.pdf

Course Outcomes:

At the end of the course the student should be able to

- Understand the Basic Concept of Computer Forensics
- Know about Computer Forensics Evidence and Capture
- Know about Tactics of the Military and of Terrorist and Rogues and of the Tactics of Private Companies
- Understanding the Surveillance Tools
- Know about Computer Forensic Cases

Course Number: MDS22025 **EC-I I**

L-P:4-0 **Credits: 4**

REGRESSION ANALYSIS

Objectives:

- Will gain knowledge in the basic concepts of Regression Analysis.
- To acquire skills to build simple and multiple regression models.
- To understand the concept of linear regression and its application.
- To understand the implementation of regression in data science.
- To understand the role of residual analysis in Real-time Applications

UNIT I SIMPLE LINEAR REGRESSION

Introduction to regression analysis: Modelling a response – overview and applications of regression analysis - major steps in regression analysis. Simple linear regression (Two variables): assumptions - estimation and properties of regression coefficients - significance and confidence intervals of regression coefficients - measuring the quality of the fit.

UNIT II MULTIPLE LINEAR REGRESSION

Multiple linear regression model: assumptions - ordinary least square estimation of regression coefficients - interpretation and properties of regression coefficient- significance and confidence intervals of regression coefficients.

UNIT III CRITERIA FOR MODEL SELECTION

Mean Square error criteria - R2 and $\hat{R}^2$criteria for model selection; Need of the transformation of variables - Box-Cox transformation – Forward - Backward and Stepwise procedures.

UNIT IV RESIDUAL ANALYSIS

Residual analysis – Departures from underlying assumptions, Effect of outliers - Collinearity - Non-constant variance and serial correlation - Departures from normality - Diagnostics and remedies.

UNIT V NON LINEAR REGRESSION

Introduction to nonlinear regression - Least squares in the nonlinear case and estimation of parameters - Models for binary response variable - estimation and diagnosis methods for logistic and Poisson regressions - Prediction and residual analysis.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. D.C Montgomery, E.A Peck and G.G Vining, "Introduction to Linear Regression Analysis", John Wiley and Sons, Inc.NY, 2003.
2. S. Chatterjee and A Hadi, "Regression Analysis", 4th Ed., John Wiley and Sons, Inc, 2006
3. Seber, A.F. and Lee, A. J, " Linear Regression Analysis", John Wiley, Relevant sections from chapters3, 4, 5, 6, 7, 9, 10, (2003).

Reference Books:

1. Iain Pardoe, " Applied Regression Modeling", John Wiley and Sons, Inc, 2012.
2. P. McCullagh, J. A. Nelder, "Generalized Linear Models", Chapman &Hall, 1989.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

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

- Develop a deeper understanding of the linear regression model.
- Learn about R-square criteria for model selection.
- Understand the forward, backward and stepwise methods for selecting the variables.
- Understand the importance of multicollinearity inregression modeling.
- Ability to use and understand generalizations of the linear model to binary and count data.

Course Number: MDS22034 **EC-III**

L-P:4-0 **Credits: 4**

PROJECT MANAGEMENT

Objectives:

- To outline the need for Project Management
- To highlight different techniques of activity planning
- To understand the Project Planning & Management
- To understand and apply appropriate approaches to plan a new project and develop project schedule.
- To understand and identify the important risks facing in a new project.

UNIT I INTRODUCTION TO PROJECT MANAGEMENT ANDPROJECT SELECTION

Objectives of Project Management- Importance of Project Management- Types of Projects Project Management Life Cycle- Project Selection – Feasibility study: Types of feasibility Steps in feasibility study.

UNIT II PROJECT PLANNING AND IMPLEMENTATION

Project Scope- Estimation of Project cost – Cost of Capital – Project Representation and Preliminary Manipulations - Basic Scheduling Concepts - Resource Levelling –Resource Allocation.

UNIT III PROJECT MONITORING AND CONTROL

Setting a base line- Project management Information System – Indices to monitor progress. Importance of Contracts in projects- Teamwork in Project Management -Attributes of a good project team – Formation of effective teams – stages of team formation.

UNIT IV PROJECT CLOSURE

Project evaluation - Project Auditing – Phases of project Audit- Project closure reports Guidelines for closeout reports.

UNIT V SPECIAL TOPICS IN PROJECT MANAGEMENT

Computers, e-markets and their role in Project management- Risk management- Environmental Impact Assessment. Case studies in Project management.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text books

1. Adrienne Watt, Management", "Project Management", BCcampus Publisher, 2014, ISBN 13: 9781774200131(Ebook)

Reference Book

1. James P. Lewis , "Fundamentals of Project Management", 1995 (Ebook)
2. Tom Kendrick, "Identifying and Managing Project Risk",2013(Ebook)

E-Learning Resources

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcome

On successful completion of this course, students would be able to

- Understand project characteristics and various stages of a project.
- Understand the conceptual clarity about project organization and feasibility analyses Market, Technical, Financial and Economic.
- Analyze the learning and understand techniques for Project planning, scheduling and Execution Control.
- Apply the risk management plan and analyse the role of stakeholders.
- Understand the contract management, Project Procurement, Service level Agreements and productivity.

Course Number: MDS22034 **EC-III**

L-P:4-0 **Credits: 4**

IBM Watson Tools

Objective:

- To understand the project data and jobs using Watson tool
- To understand project Streams Flow
- To understand to create predictive model with streaming data
- To understands the metrics and trouble shooting of the stream flow
- To understand the neural network and deep learning models

UNIT I

Overview-IBM Watson Studio, About Assets, Asset Types, Lineage, Profiles, Previews, IBM Watson APIs- Projects, Exporting Projects, Environment, Stopping Active Runtimes, Jobs

UNIT II

Preparing Data in a Project, Refining Data, Data Annotation, Python Model Operator, Buckets, File paths, and Partitions in Cloud Object Storage, Cloud Function Operator, Event Streams (Source) Operator, Installing Python Libraries, MQTT (Source and Target) operators, Buckets, file paths, and partitions in Cloud Object Storage, Cloud Function operator, SPSS model operator, Event Streams (Source) operator, Congestion, Clickstream

UNIT III

Data Science, The cells in a Jupyter notebook, Deep learning libraries for notebooks, Using project-lib for Python, Using Python functions to work with Cloud Object Storage, Scheduling a notebook, Classification and Regression, Clustering, Survival Analysis, SPSS model visualizations, Random trees, RStudio

UNIT IV

Machine Learning & AI Models, Training, Developing Apps, Machine Learning, Watson machine learning plans and compute usage, Watson Machine Learning authentication, AutoAI tutorial: Build a binary classification model, Machine Learning tutorials, Deep Learning sample apps, Python Flask tutorial: Build a web app that recognizes hand-drawn digits, Deploying an AutoAI model, Synthesized Neural Networks, Downloading and using NeuNetS models.

UNIT V

Modeling, SQL optimization, End-to-end example for running a deep learning training run, Deep Learning Experiment Builder, Importing models into Watson Machine Learning, Deploying a model, Overview: Deploying Python functions in Watson Machine Learning, Managing models

UNIT VI

Catalogs with IBM Watson Knowledge Catalog, Integrating with Information Governance Catalog, Governance Data governance with Watson Knowledge Catalog, Data protection

Text Books

1. Arindam Ganguly, "IBM Watson Solutions for Machine Learning:Achieving Successful Results Across Computer Vision, Natural Language Processing and AI Projects Using Watson Cognitive Tools", BPB Publications (1 July 2021), ISBN-13:978-9390684700

Reference Books

1. James D. Miller, "Hands-On Machine Learning with IBM Watson", Packt Publishing, March 2019, ISBN: 9781789611854

E-Learning Resources

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcome

On successful completion of this course, students would be able to

- Understand project data and jobs using Watson tool and build a model with the project data and jobs using Watson tool
- Understand project Streams Flow and the use of cloud storage
- Understand the predictive model with streaming data
- Understand the metrics and trouble shooting of the stream flow
- Understand the neural network and deep learning models and python functions

Course Number: MDS22034 **EC-III**

L-P:4-0 **Credits: 4**

TIME SERIES ANALYSIS AND FORECASTING

Objectives:

- To understand basic introduction to modern time series analysis.
- To analyze time series regression and exploratory data analysis, ARMA and ARIMA models,
- To understand model identification, estimation and linear operators
- To analyze Fourier analysis, spectral estimation, and state space models.
- To understand different concepts of Time series analysis and forecasting

UNIT I BASIS TIME SERIES MODELS

Examples of Nature of Time series data – Time series statistical models – Measures of dependence - Stationary. Time series regression – Detrending and differencing – Smoothing a time series

UNIT II AR MODELS, FORECASTING AND ESTIMATION

Auto Regressive models – Moving Average models - ARMA models – Auto Correlation Function - Partial Auto Correlation Function – Forecasting algorithms – Estimation: Yule-Walker, Method of moments, MLE and LSE

UNIT III ARIMA AND GARMA MODELS

Basics of ARIMA models: random models with drift, Steps to fitting ARMA model – Multiplicative Seasonal ARIMA models: Mixed, GARMA – Generalized Auto Regressive Conditionally Heteroscedastic (GARCH) models

UNIT IV SPECTRAL ANALYSIS

Cyclical Behaviour and Periodicity: concepts, Periodic Series, Star Magnitude - The Spectral Density: Periodic stationary process–Periodogram: Spectral analysis as ANOVA, Principal Component Analysis

UNIT V STATE SPACE MODELS

Dynamic Linear Models – Examples of DLMs – Filtering DLM – Smoothing DLM: Kalman, Lag One covariance – Forecasting DLM – Maximum Likelihood Estimator for DLMs

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned

Text Books:

1. Shumway and Stoffer, "Time Series Analysis and its applications, with examples in R". 4ed, Springer. 2016.

Reference Books:

1. Brockwell& Davis. Introduction to Time Series and Forecasting, 3rd edition, Springer. 2016
2. Cryer& Chan. Time Series Analysis with Applications in R, Springer. 2008
3. Prado & West. Time Series: Modeling, Computation, and Inference Chapman & Hall. 2010
4. Petris, Petrone, Campagnoli. Dynamic Linear Models with R, Springer. 2009
5. Ruppert& Matteson. Statistics and Data Analysis for Financial Engineering with Rexamples,2ed, Springer.2016.

E-Learning Resources:

1. Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

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

- Learn basic analysis of time series data
- Learn basic concepts in time series regression
- Learn auto-regressive and model averaging models
- Learn basic concepts of spectral analysis and space-time models
- Understand the aspect oriented Time series analysis.

Course Number: MDS22035 **EC-IV**

L-P:4-0 **Credits: 4**

BUSINESS ANALYTICS

Objectives:

- To understand the Analytics Life Cycle.
- To comprehend the process of acquiring Business Intelligence
- To understand various types of analytics for Business Forecasting
- To model the supply chain management for Analytics.
- To apply analytics for different functions of a business

UNIT I INTRODUCTION TO BUSINESS ANALYTICS

Analytics and Data Science – Analytics Life Cycle – Types of Analytics – Business Problem Definition – Data Collection – Data Preparation – Hypothesis Generation – Modeling – Validation and Evaluation – Interpretation – Deployment and Iteration

UNIT II BUSINESS INTELLIGENCE

Data Warehouses and Data Mart - Knowledge Management – Types of Decisions - Decision Making Process - Decision Support Systems – Business Intelligence – OLAP, Analytic functions

UNIT III BUSINESS FORECASTING

Introduction to Business Forecasting and Predictive analytics - Logic and Data Driven Models – Data Mining and Predictive Analysis Modeling – Machine Learning for Predictive analytics.

UNIT IV HR & SUPPLY CHAIN ANALYTICS

Human Resources – Planning and Recruitment – Training and Development - Supply chain network - Planning Demand, Inventory and Supply – Logistics – Analytics applications in HR & Supply Chain

UNIT V MARKETING & SALES ANALYTICS

Marketing Strategy, Marketing Mix, Customer Behavior – selling Process – Sales Planning – Analytics applications in Marketing and Sales

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned

Text Books:

1. R. Evans James, “Business Analytics”, 2017
2. RN Prasad, Seema Acharya, “Fundamentals of Business Analytics”, 2016

Reference Books:

1. Philip Kotler and Kevin Keller, “Marketing Management”, 15th edition, PHI, 2016
2. VSPRAO, “Human Resource Management”, 3rd Edition, Excel Books, 2010.
3. Mahadevan B, “Operations Management-Theory and Practice” , 3rd Edition, Pearson Education, 2018.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes::

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

- Explain the real world business problems and model with analytical solutions.
- Identify the business processes for extracting Business Intelligence
- Apply predictive analytics for business forecasting
- Apply analytics for supply chain and logistics management
- Use analytics for marketing and sales.

Course Number: MDS22035 **EC-IV**

L-P:4-0 **Credits: 4**

EVOLUTIONARY COMPUTING

Objectives:

- To understand the evolutionary and heuristic technique
- To understand the value representation.
- To understand the Optimization Algorithm, Genetic Algorithm
- To understand the Neural Networks.
- Tounderstandmulti-objectiveoptimizationandapplicationsofheuristic technique.

UNIT I INTRODUCTIONTOEVOLUTIONARYCOMPUTING

Introductiontoevolutionaryandheuristictechniques-PrinciplesandHistoricalPerspectives; Application potential in optimization- dimensionality reduction- data mining and analytics – Genetic Algorithms –Evolutionary Strategies – Evolutionary Programming

UNIT II OPTIMIZATIONALGORITHMS

Introduction to Representations - Binary Strings- Real-Valued Vectors - Various Selection Strategies Introduction to Search Operators- Crossover and Mutation- Ant Colony Optimization Pheromonemediated search and Exploration and Exploitation strategies – Particles warm optimization basic PSO strategies and variants –different neighborhood topologies

UNIT III ARTIFICIALNEURALNETWORKS

Fundamentals of Artificial neural networks – Architecture – Learning Paradigms – Activation Functions - Multi-Objective optimization problem- principles of Multi objective optimization–Dominanceandpareto-optimality-ParetoFrontandNon-dominatedSolutions– Classical Methods

UNIT IV FUZZYLOGIC

Fuzzylogic-FuzzySets–OperationsonFuzzySets–FuzzyRelations–MembershipFunctions-FuzzyRulesandFuzzyReasoning–FuzzyInferenceSystems–FuzzyExpertSystems–FuzzyDecisionMaking-AdaptiveNeuro-FuzzyInferenceSystems.

UNIT V OPTIMIZATIONINDATAANALYTICS

Applications of evolutionary & Heuristic techniques in large scale Optimization – Combinatorial & Function optimization–NSGA –Applications to large scale clustering classification – rule mining and Data driven Modeling –Variable Selection and Informative Data reduction and parameter optimization in predictive data analytics.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned

Text Books:

1. David E Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning, Pearson Education India, 2013.
2. S. Rajasekaran, G.A.Vijayalakshmi Pai, "Neural Networks, Fuzzy Logic and Evolutionary
3. Algorithms: Synthesis &Applications", Prentice-Hall of India Pvt. Ltd., II edition, 2017.
4. S.N.Sivanandam and S.N.Deepa,"PrinciplesofSoftComputing",3rdedition, Wiley India Pvt Ltd,2018.
5. Andries P. Engel brecht, "Fundamentals of Computational Swarm Intelligence", Wiley publications,2005.

Reference Books

1. Xin She Yang, "Nature-Inspired Computation and Swarm Intelligence-Algorithms, Theory and Applications", 1st Edition, Academic Press, 2020.
2. Marco Dorigo, Thomas Stutzle, "Ant Colony Optimization", MIT Press, 2010.
3. Oded Maimon, Lior Rokach (Eds), "Data Mining and Knowledge discovery handbook", Springer,2005.

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]
- Introduction to Soft Computing, ,https://nptel.ac.in/courses/106/105/106105173/

Course Outcomes:

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

- Develop knowledge of evolutionary computation methodologies in the Context of modern heuristic methods
- Gain experience in matching various evolutionary computation methods And algorithms for particular classes of problems
- Understand Single objective and Multi-objective optimization problems
- Solve optimization problems using suitable algorithms
- Develop evolutionary algorithms for real-world applications

Course Number: MDS22035 **EC-IV**

L-P:4-0 **Credits: 4**

SENSOR ANALYTICS

Objectives:

- To learn sensor analytics with model based sensor data.
- To form topologies of WSNS.
- To create time series sensor using PCA.
- To understand the real time data analytics using sensor networks.
- To implement a reinforcement learning system.

UNIT I SENSOR DATA ANALYTICS OVERVIEW

An Introduction to Sensor Data Analytics. A Survey of Model-based Sensor Data Acquisition and Management: Introduction - Model-Based Sensor Data Acquisition - Model-Based Sensor Data Cleaning - Model-Based Query Processing - Model-Based Sensor Data Compression.

UNIT II QUERY PROCESSING

Query Processing in Wireless Sensor Networks: Introduction - Limitations of Sensor Nodes - Topologies of WSNS - Data Storage - Data Acquisition and Aggregation. Event Processing in Sensor Streams: Events and Event Processing - Event Processing in Sensor Streams - Event Processing over RFID Streams - Advanced Topics on Complex Event Processing for Sensor Streams.

UNIT III TIME SERIES SENSOR

Dimensionality Reduction and Filtering on Time Series Sensor Streams: Introduction - Broader Overview - Principal Component Analysis (PCA) - Auto-Regressive Models and Recursive Least Squares – MUSCLES - Tracking Correlations and Hidden Variables: SPIRIT - An Application-driven View: Putting Correlations to Work. Mining Sensor Data Streams: Introduction - Sensor Stream Mining Issues - Stream Mining Algorithms - Sensor Applications of Stream Mining - Conclusions and Research Directions.

UNIT IV REAL-TIME DATA ANALYTICS IN SENSOR NETWORKS

Real-Time Data Analytics in Sensor Networks: Introduction - Data Collection - Data Processing – Discussion. Distributed Data Mining in Sensor Networks: Introduction - Clustering in Wireless Sensor Networks - Classification in Wireless Sensor Networks - Outlier Detection in WSN.

UNIT V SOCIAL SENSING

Social Sensing: Introduction - Technological Enablers of Social Sensing - Data Collection,

Architectural and System Design Challenges - Privacy Issues in Social Sensing - Trust in Social Sensing - Implied Social Networks: Inference and Dynamic Modeling - Trajectory Mining for Social Sensing - Social Sensing Applications - Future Challenges and Research Directions.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Text Books:

1. Charu. C. Aggarwal, "Managing and Mining Sensor Data"

Reference Books:

1. Sensor Data Analytics Challenges and Methods for Data-Intensive Applications by Felipe Ortega and Emilio López Cano(Eds.)

2. Sensor Data Analysis and Management - The Role of Deep Learning by A. Suresh

E-Learning Resources:

- Related Online Contents [MOOC, SWAYAM, NPTEL, Websites etc.]

Course Outcomes:

On successful completion of this course, students would be able to

- Learn sensor analytics with model based sensor data.
- Form topologies of WSNS.
- Create time series sensor using PCA.
- Understand the real time data analytics using sensor networks.
- Implement a reinforcement learning system.

VALUE ADDED COURSES

Course Code: 22VAC01 **VAC-I**

AUGMENTED REALITY AND VIRTUAL REALITY

Objectives:

- To introduce the importance of augmented realityand its need.
- To create awareness on augmented reality and its application for the society.
- Visual Perceptionis used to develop the future Business and Industry.
- To know about Virtual Realityand their Applications.
- To know about Interface to the Virtual World

UNIT I Introduction to Augmented Reality

History of augmented reality, The Relationship Between Augmented Reality and Other Technologies-Media, Technologies, Other Ideas Related to the Spectrum Between Real and Virtual Worlds, applications of augmented reality

UNIT II Augmented Reality Hardware

Audio Displays, Haptic Displays, Visual Displays, Other sensory displays, Visual Perception, Requirements and Characteristics, Spatial Display Model. Processors – Role of Processors, Processor System Architecture, Processor Specifications..

UNIT III Introduction to Virtual Reality

Defining Virtual Reality, History of VR, Human Physiology and Perception, Key Elements of Virtual Reality Experience, Virtual Reality System, Interface to the Virtual World-Input & output- Visual, Aural & Haptic Displays, Applications of Virtual Reality.

UNIT IV Representing the Virtual World

Representation of the Virtual World, Visual Representation in VR, Aural Representation in VR and Haptic Representation in VR

UNIT V Visual Perception

Visual Perception - Perception of Depth, Perception of Motion, Perception of Color, Combining Sources of Information

Text Books:

1. Designing for Mixed Reality, Kharis O'Connell Published by O'Reilly Media, Inc., 2016, ISBN: 9781491962381
2. Augmented Reality: Principles & Practice by Schmalstieg / Hollerer, Pearson Education India; First edition (12 October 2016),ISBN-10: 9332578494

Reference books:

1. Understanding Virtual Reality: Interface, Application and Design, William R Sherman and Alan B Craig, (The Morgan Kaufmann Series in Computer Graphics)". Morgan Kaufmann Publishers, San Francisco, CA, 2002
2. Developing Virtual Reality Applications: Foundations of Effective Design, Alan B Craig, William R Sherman and Jeffrey D Will, Morgan Kaufmann, 2009.

E-Resources:

- https://www.vttresearch.com/sites/default/files/pdf/science/2012/S3.pdf
- http://lavalle.pl/vr/book.html

Course Outcomes::

- Students can able to understand the importance of augmented reality and its need.
- Students can acquire in-depth knowledge on augmented reality and its application.
- Students can be familiarizing with Visual Perception is used to develop the future.
- Students can understand the basic concept Virtual Reality.
- Students can able get interpretation about the Virtual Reality and augment reality.

Course Number: 22VAC02 **VAC-II**

ROBOTICS

Objectives:

- To introduce the importance of Robotics for Industrial needs.
- To create awareness on Robotics and its application for the society.
- Robot programming Methods is used to develop the future Business and Industry.
- To know about Operational Capabilities level of Robot and their Applications.
- To know about Internal Sensors and External Sensors for Robot.

UNIT I FUNDAMENTAL OF ROBOTICS

Historical Development of Robotics – Industrial Robot – Classification – Degree of Freedom and Degree of Motion – Manipulation of Robot Components.

UNIT II ROBOT PROGRAMMING

Robot programming Methods-Advantages and Disadvantages of Robot-Requirements for a Robot In a Industry –Specification for a Robot

UNIT III MODULAR COMPONENTS

Operational Capabilities level of robot-Modular Robot Components –Wrist Mechanism-Numerical Examples.

UNIT IV ROBOT SENSORS

Internal Sensors - External Sensors – Force – Sensors – Thermocouples -Performance Characteristics Of Robot-Static Performance Characteristics- Dynamic Performance Characteristics.

UNIT V ADVANCED ROBOT SYSTEMS

Heuristics Decision for Robot-Fuzzy logic for robot control-Artificial Neural in Robotics-Robot Calibration.

UNIT VI

Current Contours (for Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

Reference books:

1. Appuu Kuttan, K.K, “Robotics”, I.K International Publishing House Pvt. Ltd (Ebook).
2. Danny Staple, “Learn Robotics Programming”, Packt Publisher (Ebook).
3. Bijoy K. Ghosh, Bhaskar Kumar Ghosh, Ning, “Control in Robotics and Automation: Sensor-based Integration”, Academic Press (Ebook).

E-Resources:

1. https://www.geeksforgeeks.org/robotics
2. https://www.britannica.com/technology/robot-technology

Course Outcomes::

- Students can able to understand the importance of the importance of Robotics.
- Students can acquire in-depth knowledge on Robotics and its application for the society.
- Students can be familiarizing with Robot programming Methods.
- Students can understand the basic concept and the operational capabilities of Robot.
- Students can able get interpretation about Internal Sensors and External Sensors for Robot.