Rajiv Gandhi National Institute of Youth Development
(Institution of National Importance by the Act of Parliament No. 35/2012)
Ministry of Youth Affairs & Sports, Government of India
Sriperumbudur, Tamil Nadu - 602105, India

REGULATIONS, CURRICULUM & SYLLABUS FOR

MASTER OF SCIENCE (M. Sc.)
COMPUTER SCIENCE (DATA SCIENCE)

(For CBCS System in Rajiv Gandhi National Institute of Youth Development)

Effective from academic year 2022-23

REGULATIONS AND SYLLABUS FOR MASTER OF SCIENCE

COMPUTER SCIENCE (DATA SCIENCE)

(For CBCS System in RGNIYD)

(Effective from the Academic Year 2022-2023)

Eligibility for Admission

Candidates who have Any Computer Science Graduate / Any Information Technology Graduate/ Any UG degree with atleast 4 mathematics Papers or any other equivalent UG Computer Science Degree.

Duration of the Course

The course duration shall be for two years spread over four semesters. The maximum duration to complete the course shall be 3 years.

Medium

The medium of instruction shall be English.

Passing & Classification

Passing Eligibility & Classification for the award of the Degree are as per the RGNIYD Norms

REGULATIONS AND SYLLABUS FOR MASTER OF SCIENCE

COMPUTER SCIENCE (DATA SCIENCE)

(For CBCS System in RGNIYD)

(Effective from the Academic Year 2022-2023)

CURRICULUM

(Effective from the Academic Year 2022 - 2023)

COURSESTRUCTURE

Course Category	Number of Credits
Minimum credits for Core and departmental electives	**73**
Soft core course (English for Communication / Buddhist Philosophy / Ethics & IPR)	**02**
Course from Online Platform (SWAYAM, MOOC, NPTEL) (Courses are suggested by the Department)	**03**
Total	**78**

Important Note: If the candidates are undergoing a six-month dedicated project during their fourth semester course of study in any of the Industries / MNCs / Organizations, then they should take an additional course from SWAYAM / MOOC / NPTEL etc., instead of their regular course offered by RGNIYD during the fourth semester

M.SC. – DATA SCIENCE

Vision

Department of Computer science aims to become the an internationally recognized center of excellence in computer science with special focus on emerging areas like Artificial Intelligence, Machine Learning, Internet of Things, Data science and Computer Vision.

Mission

The mission of the department is

- To mentor students by providing them an environment that is supportive in fostering intellectual skills
- To produce employable graduates who are trained in cutting edge technologies of Computer Science
- To train the graduates with the skillsets required for collaborative multidisciplinary research
- To develop core competencies in the field of Artificial Intelligence, Machine Learning, Internet of Things, Data science and computer vision

Course description

M.Sc., in Data Science is a two-years masters' programme which imparts the technical skills and knowledge in the field of data science and its application to the graduates. This programme will equip the students with skills in data analytics, data mining, and business intelligence. The graduands will experience an industry-like environment wherein the learnt techniques will be employed to design and develop innovative & productive solutions. The course has a range of electives that kindle the thought process of the students and helps in selecting a focus area in the applications of Data Science.

Programme Objective

1. To impart technical knowledge, and skills in the field of computer science with specific focus on Data analytics
2. Transform the students in to technically competent industry ready graduands who can adapt to the pressing demands of the industry.
3. Encourage the students to pursue higher studies in Data Analytics

Course learning Objective

1. To impart fundamental knowledge, thinking skills and technical skills for superior mastery in the core areas of data science and industry 4.0
2. Promote the students in placement in leading business organizations globally
3. Demonstrate global standards to the students in the field of Data Science by providing hands-on training in the hot topics such as decision-making, statistical learning and so on.
4. Enable the students to go to higher studies in Data science or allied subjects like Business Analytics, Big Data and so on.

Course Duration

1. The course duration is of 24 months (2 years) spread over four semesters with credit hours as per the RGNIYD norms
2. The course has sufficient emphasis on mathematical skills as well along with its science and management parts

Course Curriculum Plan

The Course Curriculum is based on comparative analysis of existing MSc. Data Science and Information System curriculums of other Universities, ISI, IITs and NITs. The curriculum has sufficient exposure to hands-on skills and is much more directed towards higher employability. It is also well suited for upward accommodation of science graduates and Mathematics graduates.

Eligibility

Candidates who have Any Computer Science Graduate / Any Information Technology Graduate/ Any UG degree with atleast 4 mathematics Papers or any other equivalent UG Computer Science Degree.

First Semester

Course Code	Course Title	L	T	P	Credits	Core / Specialization
MSCS 101	Data Structures & Algorithms	3	0	0	3	Core
MSCS 102	Python Programming	2	0	0	2	Core
MSCS 103	Mathematical Foundations of Computer Science	2	1	0	3	Core
MSCS 104	Database Management System (DBMS)	3	0	0	3	Core
MSDS 101	Foundation to Data Science	3	0	0	3	Specialization
MSCS 105	Data structure & Algorithms Lab	0	0	4	2	Core
MSCS 106	Database Management System (DBMS) Lab	0	0	4	2	Core
MSCS 107	Python Programming Lab	0	0	4	2	Core
Total Credits		**13**	**1**	**12**	**20**	

Second Semester

Course Code	Course Title	L	T	P	Credits	Core / Specialization
MSCS 201	Operating System	3	0	0	3	Core
MSCS 202	Web Technology	3	0	0	3	Core
MSCS 203	Object Oriented Programming using JAVA	3	0	0	3	Core
MSCS 204	Machine Learning	3	0	0	3	Core
MSAI 101	Artificial Intelligence	3	0	0	3	Specialization
	Elective -1	3	0	0	3	Specialization
MSCS 205	Web Technology Lab	0	0	4	2	Core
MSCS 206	Object Oriented Programming using JAVA Lab	0	0	4	2	Core
Total Credits		**18**	**0**	**8**	**22**	

Third Semester

Course Code	Course Title	L	T	P	Credits	Core / Specialization
MSCS 301	Advanced Computer Networks	3	0	0	3	Core
MSAI 303	Deep Learning	3	0	0	3	Specialization
MSDS 301	Big Data Analytics	3	0	0	3	Specialization
MSDS 302	Data Mining & Warehousing	3	0	0	3	Specialization
	Elective-2	3	0	0	3	Specialization
MSDS 303	Big Data Analytics Lab	0	0	4	2	Specialization
Total Credits		**15**	**0**	**4**	**17**	

FOURTH SEMESTER

CourseCode	Course Title	L	T	P	Credits
	(MOOC/ SWAYAM/NPTEL/Elective as decided by the department)	2	0	0	2
MSCS 401	Project Work and Viva-Voce	0	0	24	12
Total Credits		**2**	**0**	**24**	**14**

SOFTCORE & MOOC'S

CourseCode	Course Title	L	T	P	Credits
	Softcore (English Communication / Buddhist Philosophy / Ethics & IPR)	2	0	0	2
	MOOC / SWAYAM / NPTEL online courses (Minimum 12 weeks, select any one course out of a pool of subjects which are related to the branch study and the list is provided by the Head of the Department)	3	0	0	3
Total Credits		**5**	**0**	**0**	**5**

Total Credits: 73+5 = 78 Credits

List of Electives

Course Code	Course Title	L	T	P	Credits
MSDS EL01	Speech Technology and Analytics	3	0	0	3
MSDS EL02	Health Care Analytics	3	0	0	3
MSDS EL03	Statistics in Data Science	3	0	0	3
MSDS EL04	Full Stack of Data Science	3	0	0	3
MSCS EL01	Information Retrieval Systems	3	0	0	3
MSCS EL02	Operational Systems	3	0	0	3
MSCS EL03	Theory of Computation	3	0	0	3
MSCS EL04	Video Processing and Analytics	3	0	0	3
MSCS EL05	Data Visualization Techniques	3	0	0	3
MSCS EL06	Cyber Security	3	0	0	3
MSCS EL07	Soft Computing	3	0	0	3
MSAI 302	Natural Language Processing	3	0	0	3
MSAI 201	Knowledge Engineering and Expert Systems	3	0	0	3

BRIDGE COURSES

MSBC0001	INTRODUCTION TO COMPUTERS	0	0	0	0
	Topics covered in this course should include but not limited to: i. Parts of a computer ii. Hardware – Software iii. System Configuration iv. Office Tools v. Internet Concepts				

MSBC0002	PROBLEM SOLVING TECHNIQUES USING C	0	0	0	0
	Topics covered in this course should include but not limited to: i. Summation of a set of numbers ii. Reversing the digits of an integer iii. Character to Number conversion iv. Smallest Divisor of an integer v. Linear Pattern Search vi. Computing nth Fibonacci number				

MSCS101	DATA STRUCTURES AND ALGORITHMS	3	0	0	3

Course Outcomes:

CO1: Understand the concepts of algorithm analysis and solving recurrence relations
CO2: Learn linear and non-linear data structures and their usage in applications
CO3: Understand and Analyze various searching and sorting algorithms
CO4: Applying greedy and dynamic approaches to solve challenging problems
CO5: Understand computational complexity classes and their importance in designing algorithms

Unit – 1	INTRODUCTION TO ALGORITHMS AND ANALYSIS	9 hrs

Overview and importance of algorithms and data structures – Fundamentals of algorithm analysis – Space and time complexity of an algorithm – Asymptotic Notations – Order of growth – Algorithm Efficiency – Best case, Worst Case, Average Case – Recurrence Relations – Solving recurrence relations using substitution method, recurrence tree method and Master Method.

Unit – 2	LINEAR AND NON-LINEAR DATA STRUCTURES	9 hrs

Linear Data Structures: Stacks – Queues – Lists – Applications. Non-linear Data Structures: Graphs – Trees – Binary Trees – Traversal Techniques – Binary Search Tree and its operations – AVL Trees.

Unit – 3	SEARCHING AND SORTING ALGORITHMS, DIVIDE AND CONQUER APPROACH	9 hrs

Search Problem – Linear Search – Binary Search – Sorting Problem – Bubble Sort – Insertion Sort – Heap Sort – Divide and Conquer Paradigm – Merge Sort – Quick Sort – Complexity analysis of searching and sorting algorithms.

Unit – 4	GREEDY METHOD AND DYNAMIC PROGRAMMING APPROACH	9 hrs

Greedy Method: Activity Selection Problem – Graph Traversal Algorithms. Dynamic Programming Paradigm: Knapsack problem – Matrix Chain Multiplication – All Pair Shortest Path – Single Source Shortest Path – Travelling Salesman Problem.

Unit –5	OTHER ALGORITHM PARADIGMS AND COMPUTATIONAL COMPLEXITY CLASSES	9 hrs

Backtracking: 8-Queens problem – Graph coloring. Brach and Bound: Least Cost 0/1 Knapsack problem. Tractable and Intractable problems – Decidable and Undecidable problems – P, NP and NP-Complete classes – Cook's Theorem (without proof) – NP-Hard problems.

Text Books:

1. Thomas H. Cormen, C.E. Leiserson, R L.Rivest and C. Stein, Introduction to Algorithms, Third edition, MIT Press February 2010

Reference Books:

1. Sanjoy Dasgupta, C.Papadimitriou and U.Vazirani , Algorithms , Tata McGraw-Hill, 2008.
2. A. V. Aho, J.E. Hopcroft and J. D. Ullman, The Design and Analysis of Computer Algorithms, Pearson, 1st edition, 2006.
3. Fundamentals of Data Structures in C -- by Horowitz, Sahni and Anderson-Freed (Silicon Press 2007).

MSCS102	PYTHON PROGRAMMING	2	0	0	2

Course Outcomes:

CO1: Understand the fundamental concepts of python and its main components.

CO2: Develop (Read and Write) python programs using variables, assignments, and conditional statements using functions.

CO3: Illustrate and implement different data structures.

CO4: Demonstrate Object-oriented concepts and file handling.

CO5: Analyze and plot data using python visualization libraries.

Unit – 1	Introduction to Python	6 hrs

Introduction to a programming language – History of Python- Python environment setup – Python 2 vs. Python 3 - Comments and documentation in Python- Keywords and Identifiers -Programming Errors - Writing and Running python programs.

Unit – 2	Variables and Conditionals in Python	6 hrs

Variables – Constants- Strings - Assignment statements – Expressions-Operators – Type Conversions- Control Flow statements and Loops- Functions.

Unit – 3	Data Structures in Python	6 hrs

List Basics- List Indexing and Slicing-Appending-Sorting and Ranging-Tuples-Creation-Deletion-Converting tuple to list- Assignment- Dictionaries-Adding-Modifying and Retrieving Values-Traversing all keys in the dictionary-Operations and methods-Sets- manipulating and accessing sets.

Unit – 4	Exception handling and File I/O	6 hrs

Exception handling - Catching and Handling Exceptions-Object Orientated Concepts (Basics) – Creating python class and Objects - Object properties and methods- Inheritance- Operator overloading- Polymorphism- File handling - Opening, Reading, Writing and Deleting files

Unit –5	Graph Plotting	6 hrs

Introduction to plotting python libraries -Plots and Graphs- Applied Visualizations – Seaborn - Matplotlib

Text Books:

1. Python: The Complete Reference , 2018
2. Python in easy steps, McGraw Hill, 2nd Reprint , 2014

References:

1. Python 3 Documentation, https://docs.python.org/
2. https://www.coursera.org/learn/python

MSCS103	MATHEMATICAL FOUNDATIONS OF COMPUTER SCIENCE	2	1	0	3

Course Outcomes:

CO1: Apply Mathematical thinking, Mathematical proofs, and algorithmic thinking, and be able to apply them in problem-solving.

CO2: Explain the concept of Sets, Relations and Functions and their properties.

CO3: Describe mathematical induction and probability.

CO4: Describe basic properties of graphs and related discrete structures, and be able to relate these to practical examples.

CO5: Describe the use of group theory and its applications

Unit – 1	SETS – RELATIONS – FUNCTIONS – BOOLEAN ALGEBRA	9 hrs

Set Operations – Representation and Properties of Relations – Equivalence Relations – Partially Ordering**.** Functions – one-one – onto bijective – composition of relations and functions – inverse functions. Boolean Algebra and Boolean Functions and its representations – Simplifications of Boolean Functions.

Unit – 2	MATHEMATICAL LOGIC	9 hrs

Propositional and Predicate Logic – Propositional Equivalences – Normal Forms – Predicates and Quantifiers – Nested Quantifiers – Rules of Inference.

Unit – 3	COUNTING – MATHEMATICAL INDUCTION AND DISCRETE PROBABILITY	9 hrs

Basics of Counting – Pigeonhole Principle – Permutations and Combinations – Inclusion-Exclusion Principle – Mathematical Induction – Probability – Bayes Theorem.

Unit – 4	GROUP THEORY	9 hrs

Groups – Subgroups – Semi Groups – Product and Quotients of Algebraic Structures – Isomorphism – Homomorphism – Automorphism – Rings - Integral Domains – Fields – Applications of Group Theory – Polya's theory of counting – Introduction to Error-Correcting Codes. Discrete Geometry: Some basic definitions – Ham-Sandwich theorem.

Unit – 5	GRAPH THEORY	9 hrs

Simple Graph – Multigraph – Weighted Graph – Paths and Circuits – Shortest Paths in Weighted Graphs – Eulerian Paths and Circuits – Hamiltonian Paths and Circuits – Planner graph – Graph Coloring – Bipartite Graphs.

Text Books:

1. Rosen, K.H. and Kamala Krithivasan, "Discrete Mathematics and its Applications", 8th Edition – July 2021

References:

1. Rosen, K.H., "Discrete Mathematics and Its Applications with Combinatorics and Graph Theory", 7th Edition – July 2017
2. Tremblay, J.P. and Manohar.R, " Discrete Mathematical Structures with Applications to Computer Science", McGraw Hill Education; 1st edition, July 2017
3. Grimaldi, R.P. "Discrete and Combinatorial Mathematics: An Applied Introduction", 4th Edition, Pearson Education Asia, Delhi, 2007.
4. Lipschutz, S. and Mark Lipson., "Discrete Mathematics", Schaum's Outlines, Tata McGraw Hill Pub. Co. Ltd., New Delhi, 3rd Edition, 2010.

MSCS104	DATABASE MANAGEMENT SYSTEMS	3	0	0	3

Course Outcomes:

CO1: Identify the basic concepts and apply relational database theory and recognize and identify the use of normalization and functional dependency, indexing technique used in database design

CO2: Describe relational algebra expression, tuple and domain relation expression for queries and query processing concepts.

CO3: Apply and relate the concept of transaction, concurrency control and recovery in database.

CO4: Familiar with basic database storage structures and access techniques: file and page organizations, indexing methods including B tree, and hashing.

CO5: Understand various concepts like Parallel, distributed and object-oriented databases.

Unit – 1	FUNDAMENTALS OF DBMS	9 hrs

File System versus DBMS – Advantages -Views – Data models – Database languages – Architecture – ER Model – Entities, Attributes and Entity Sets, Relationships and Relationship Sets – Features of ER Model – Conceptual design with ER – Extended E-R. Relational Model - Codd's rule – Integrity Constraints over Relations – Relational database design – Anomalies - Functional dependencies – Normalization – Normal Forms – Decomposition – Denormalization.

Unit – 2	RELATIONAL ALGEBRA AND RELATIONAL CALCULUS	9 hrs

Relational Query Languages – Relational Algebra – Tuple and Domain Relational Calculus – SQL – Query processing and optimization – Transformation of relational expressions – Evaluation Plans

Unit – 3	TRANSACTION CONCURRENCY CONTROL AND SECURITY	9 hrs

Transaction – Properties – Concurrent execution – Serializability – Concurrency control – Protocols – Recovery System – Database Security

Unit – 4	FILE ORGANIZATION	9 hrs

File organization – Organization of records in files – Indexing – B tree and B+ tree index files – Static hashing – Dynamic hashing

Unit –5	ADVANCED CONCEPTS	9 hrs

Parallel and distributed databases – Object-based databases - Mobile databases - XML and Web databases – Intelligent databases – Mongo DB – NOSQL – PostgreSQL

Text Books:-

1. A. Silberchatz, F. Korth, and S. Sudarshan, "Database System Concepts", 7th Edition, McGraw Hill, 2021.
2. R. Elmasri and S. B. Navathe, "Fundamentals of Database Systems", 7th Edition, Pearson 2016

References:-

1. Coronel, Carlos, and Steven Morris. Database systems: design, implementation, & management. Cengage Learning, 2016.
2. Ramakrishnan, R., Gehrke, J., & Gehrke, J. Database management systems , 3rd Edition. McGraw Hill 2006.

MSDS101	FOUNDATION TO DATA SCIENCE	3	0	0	3

Course Outcomes:

CO1: To understand the data science fundamentals and process.
CO2: learn to describe the data for the data science process.
CO3: To learn to describe the relationship between data.
CO4: To utilize the Python libraries for Data Wrangling.
CO5: To present and interpret data using visualization libraries in Python

Unit – 1	INTRODUCTION	9 hrs

Data Science: Benefits and uses – facets of data – Data Science Process: Overview – Defining research goals – Retrieving data – Data preparation – Exploratory Data analysis – build the model– presenting findings and building applications – Data Mining – Data Warehousing – Basic Statistical descriptions of Data

Unit – 2	DESCRIBING DATA	9 hrs

Types of Data – Types of Variables -Describing Data with Tables and Graphs –Describing Data with Averages – Describing Variability – Normal Distributions and Standard (z) Scores

Unit – 3	DESCRIBING RELATIONSHIPS	9 hrs

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 towards the mean

Unit – 4	PYTHON LIBRARIES FOR DATA WRANGLING	9 hrs

Basics of Numpy arrays –aggregations –computations on arrays –comparisons, masks, boolean logic – fancy indexing – structured arrays – Data manipulation with Pandas – data indexing and selection – operating on data – missing data – Hierarchical indexing – combining datasets – aggregation and grouping – pivot tables

Unit –5	DATA VISUALIZATION	9 hrs

Importing Matplotlib – Line plots – Scatter plots – visualizing errors – density and contour plots – Histograms – legends – colors – subplots – text and annotation – customization – three dimensional plotting – Geographic Data with Basemap – Visualization with Seaborn.

Text Books:-

1. David Cielen, Arno D. B. Meysman, and Mohamed Ali, “Introducing Data Science”, Manning Publications, 2016. (Unit I)
2. Robert S. Witte and John S. Witte, “Statistics”, Eleventh Edition, Wiley Publications, 2017. (Units II and III)
3. Jake VanderPlas, “Python Data Science Handbook”, O’Reilly, 2016. (Units IV and V)

References:-

1. Allen B. Downey, “Think Stats: Exploratory Data Analysis in Python”, Green Tea Press,2014.

MSCS105	DATA STRUCTURES AND ALGORITHMS LAB	0	0	4	2

Course Outcomes:

CO1: Understand the fundamental concepts of python and its main components.

CO2: Develop (Read and Write) python programs using variables, assignments, and conditional statements using functions.

CO3: Illustrate and implement different data structures.

CO4: Demonstrate Object-oriented concepts and file handling.

CO5: Analyze and plot data using python visualization libraries.

Programs should include but not limited to:

Implementation of

i. Stacks and Queues
ii. Lists
iii. Linear Search and Binary Search
iv. Sorting Algorithms
v. Graph Traversal Algorithms
vi. Tree Traversal Algorithms
vii. Shortest Path Algorithms
viii. Knapsack Problem
ix. Travelling Salesman Problem
x. N-Queens's Problem

MSCS106	DATABASE MANAGEMENT SYSTEMS LAB	0	0	4	2

Course Outcomes:

CO1: Identify the basic concepts and apply relational database theory and recognize and identify the use of normalization and functional dependency, indexing technique used in database design

CO2: Describe relational algebra expression, tuple and domain relation expression for queries and query processing concepts.

CO3: Apply and relate the concept of transaction, concurrency control and recovery in database.

CO4: Familiar with basic database storage structures and access techniques: file and page organizations, indexing methods including B tree, and hashing.

CO5: Understand various concepts like Parallel, distributed and object-oriented databases.

Programs should include but not be limited to:

Exercises on

i. DDL commands
ii. DML commands- Basic Queries
iii. Aggregate Commands
iv. Nested Queries
v. Joins
vi. Views
vii. Index
viii. Functions
ix. Procedures
x. Triggers

MSCS107	PYTHON PROGRAMMING LAB	0	0	4	2

Course Outcomes:

CO1: Understand the fundamental concepts of python and its main components.

CO2: Develop (Read and Write) python programs using variables, assignments, and conditional statements using functions.

CO3: Illustrate and implement different data structures.

CO4: Demonstrate Object-oriented concepts and file handling.

CO5: Analyze and plot data using python visualization libraries.

Programs should include but not limited to:

I. Test and Debug simple Python programs
II. Different datatypes in python (variables constants and strings)
III. Programs on different operators
IV. Control statements and Loops
V. Working on Functions
VI. Data structures in python (List, Tuple, Dictionary and Set)
VII. Objects and Classes manipulation using python
VIII. Open, Read and write data from/to files in Python
IX. Different plots using Matplotlib
X. Visualization of data using seaborn

MSCS201	OPERATING SYSTEMS	3	0	0	3

Course Outcomes:

CO1: Understand the functions, features, and concepts of operating systems.
CO2: Learn the concept of process and Analyze the Inter Process Communication methods
CO3: Analyze the scheduling algorithms and deadlocks
CO4: Understand the concept of memory management
CO5: Evaluate security mechanisms in operating computing systems

Unit – 1	OVERVIEW OF OPERATING SYSTEMS	9 hrs

Introduction to operating systems – Computer system organization, architecture – Operating system structure, operations – Process, memory, storage management – Protection and security – Distributed systems – Computing Environments – Open-source operating systems – OS services – User operating-system interface – System calls – Types – System programs – OS structure – OS generation – System Boot – Process concept, scheduling – Operations on processes – Cooperating processes – Inter-process communication – Examples – Multithreading models – Thread Libraries – Threading issues – OS examples

Unit – 2	PROCESS MANAGEMENT	9 hrs

Process Concept – Process Scheduling – Operations on Processes – Inter-Proces Communication (IPC) – IPC Examples – Client-Server Communication – Threads – Multi-core models – Multi-threading models – Threading issues – Process Synchronization – Critical Section Problem – Peterson's Solution – Synchronization Hardware – Mutex Locks – Semaphores - Monitors

Unit – 3	CPU SCHEDULING AND DEADLOCKS	9 hrs

Basic Concepts – Scheduling Criteria – Scheduling Algorithms – Thread Scheduling – Multi-processor Scheduling – Real-time CPU Scheduling. Deadlocks: System Model – Deadlock Characterization – Methods for Handling Deadlocks – Deadlock Prevention – Deadlock Avoidance – Deadlock Detection – Recovery from Deadlock.

Unit – 4	MEMORY MANAGEMENT	9 hrs

Main Memory Concept – Swapping – Contiguous Memory Allocation – Segmentation – Paging – Example Architectures – Virtual Memory concept – Demand Paging – Page Replacement - Frame Allocation – Thrashing – Memory Mapped Files – Kernel Memory Allocation. Secondary Storage: Disk Structure – Disk Scheduling – Disk Management – RAID structure

Unit – 5	I/O SYSTEMS AND CASE STUDIES	9 hrs

File concept – Access methods – Directory structure – File-system mounting –Protection –
Directory implementation – Allocation methods – Free-space management – Protection File systems. Case Studies: Linux System – Windows 7

Text Books:

1. Abraham Silberschatz, Peter B. Galvin, Greg Gagne, "Operating System Concepts", 10th edition John Wiley & Sons Inc., 2021

References:-

2. Andrew S. Tanenbaum, "Modern Operating Systems", Fourth Edition, Addison Wesley, 2016.
3. Charles Crowley, "Operating Systems: A Design-Oriented Approach", Tata McGraw Hill Education", 1996.
4. D M Dhamdhere, "Operating Systems: A Concept-based Approach", Second Edition, Tata McGraw-Hill Education, 2007.
5. William Stallings, "Operating Systems: Internals and Design Principles", Seventh Edition, PH, 2011

MSCS 202	WEB TECHNOLOGY	3	0	0	3

Course Outcomes:

CO1: Understand the principles of www and concepts of web clients and web servers
CO2: Interpret the key responsibilities and functionalities of different internet technologies
CO3: Demonstrate Markup languages and illustrate the working of it
CO4: Analyze the different client side scripting/programming languages
CO5: Explain and demonstrate server side , presentation and database technologies

Unit – 1	WEB ESSENTIALS	9 hrs

Internet Principles – basic web concepts – Client/ server model – Retrieving data from Internet –Internet Protocols and applications

Unit – 2	INTERNET TECHNOLOGIES	9 hrs

Streaming – Networking Principles – Sockets for Clients – Sockets for Servers – Protocol Handlers – Content Handlers – Multicast sockets – Remote method Invocation.

Unit – 3	MARKUP LANGUAGES AND STYLE SHEETS	9 hrs

HTML Elements – HTML Lists – HTML Tables – HTML Forms – Links and addressing – HTML Frames and Images – DHTML – Document Object model (DOM) – XML Schemas – CSS – Text Properties – Advanced CSS

Unit – 4	CLIENT SIDE TECHNOLOGY	9 hrs

Scripts – Javascript – VB Script – JQuery and Ajax

Unit – 5	SERVER SIDE TECHNOLOGY AND DATABASE	9 hrs

Server Scripts – Servlets – Sessions – Cookies – PHP– Node.js – Java Server Pages (JSP) – Active Server pages (ASP) – Simple Applications – On-line databases – Monitoring user events – Plug-ins – Database connectivity

Text Books:

1. Web Technologies: A Computer Science Perspective , Jeffrey Jackson.
2. Harvey M. Deitel and Paul J. Deitel, “Internet & World Wide Web How to Program”, 4th edition, 2008.

Reference:-

3. https://nptel.ac.in/courses/106/105/106105084/

MSCS203	OBJECT ORIENTED PROGRAMMING USING JAVA	3	0	0	3

Course Outcomes:

CO1: Understanding Object-Oriented programming concepts using basic syntaxes of control Structures, strings for developing skills of logic building activity using Java

CO2: Identification of classes, objects, members of a class and the relationships among them needed for a finding the solution to specific problem

CO3: Illustration to achieve reusability using inheritance, interfaces, and packages and describes faster application development can be achieved with exception handling mechanisms

CO4: Understanding concept of multithreading for robust faster and efficient application development and applications of collection interfaces in Java

CO5: Learning of various I/O operations, connecting Java with databases using JDBC and implementation of networking with Java

Unit – 1	OVERVIEW OF OOP AND INTRODUCTION TO JAVA	9 hrs

Structured Programming and its limitation – Object-Oriented Paradigm: Basic concepts of Object-Oriented Programming (OOP) – Structured Programming vs OOP – Benefits of OOP – Object Modeling – Association – Aggregation and Generalization. Introduction to Java – Evolution of Java – Features of Java – Java Environment – JVM – Data types – variables – operators – Decision statements – Control Structures – Loops – Arrays in Java – Strings in Java – String Buffer Class – Wrapper Classes.

Unit – 2	OBJECTS AND CLASSES	9 hrs

Introducing Classes - Class Fundamentals - Declaring Objects - Object Reference Variables - Introducing Methods - Constructors - this Keyword - Garbage Collection - A Closer Look at Methods and Classes - Overloading Methods - Objects as Parameters - Returning Objects - Access control - static - final - nested and inner class - Command line arguments

Unit – 3	INHERITANCE, PACKAGES, INTERFACES AND EXCEPTION HANDLING	9 hrs

Inheritance in Java – Constructors in Inheritance – super – Multilevel Inheritance – Overriding – Dynamic Method Dispatch – final Keyword – Interfaces – Packages – JAR files – Exception Handling

Unit – 4	MULTI THREADING AND JAVA COLLECTIONS	9 hrs

Threads – Multithreading in Java – Thread Priorities – Creating Multiple Threads – Inter Thread Communication – Synchronization – Suspending, Resuming and Stopping Threads – Collections Overview – Collection Interfaces – Collection Classes – Accessing collection via Iterator.

Unit –5	JAVA I/O, JDBC AND JAVA NETWORKING	9 hrs

I/O Basics - Reading Console Input - Java I/O Classes and Interfaces - Serialization. Networking Classes and Interfaces - InetAddress - TCP/IP Client Sockets - URL - URL Connection - JDBC Driver, Database Connection Steps, DriverManager Class, Statement Interface, ResultSet Interface.

Text Books:

1. Herbert Schildt, Java: The Complete Reference, Eleventh Edition, 11th Edition December 2018, McGraw-Hill, ISBN: 9781260440249

Reference Books:-

2. Javin Paul, Grokking the Java Interview: Prepare for Java interviews by learning essential Core Java concepts and APIs, 2020
3. Kishori Sharan, Adam L. Davis, Beginning Java 17 Fundamentals, Springer Link, 2022
4. Herbert Schildt and Dale Skrien, Java Programming: A Comprehensive Introduction, 2013, National Edition

MSCS 204	MACHINE LEARNING	3	0	0	3

Course Outcomes:

CO1: Understand the fundamentals of machine learning, model preparation and evaluation

CO2: Learn ability theory and its application in data understanding

CO3: Apply supervised learning algorithms and unsupervised learning algorithms on real data

CO4: Apply ensemble classifiers on data and compare the model performance

CO5: Explore data repositories and apply Machine Learning Algorithms on case studies

Unit – 1	INTRODUCTION – MODEL PREPARATION – EVALUATION	10 hrs

Introduction to Machine Learning: Human Learning and its Types; Machine Learning and its types; Well-Posed Learning Problem; Applications of Machine Learning; Issues in Machine Learning. **Preparing To Model:** Basic Data Types; Exploring Numerical Data; Exploring Categorical Data; Exploring Relationship Between Variables; Data Issues and Remediation; Data Pre-Processing. **Modelling And Evaluation:** Selecting A Model; Training Model – Holdout, *K*-Fold Cross-Validation, Bootstrap Sampling; Model Representation and Interpretability – Under-Fitting, Over-Fitting, Bias-Variance Tradeoff – Model Performance Evaluation – Classification, Regression, Clustering – Performance Improvement.

Unit – 2	FEATURE ENGINEERING – REVIEW OF PROBABILITY	6 hrs

Feature Engineering: Feature Construction – Feature Extraction – Feature Selection. **Brief Review of Probability:** Basic Concept Of Probability – Random Variables – Discrete Distributions – Binomial, Poisson – Bernoulli - Continuous Distribution – Uniform, Normal, Laplace; Central Theorem; Monte Carlo Approximation.

Unit – 3	CONCEPT LEARNING – SUPERVISED AND UNSUPERVISED LEARNING	12 hrs

Bayesian Concept Learning: Bayes Theorem – Prior and Posterior Probability, Likelihood; Concept Learning; Bayesian Belief Network. **Supervised Learning – Regression:** Simple Linear Regression; Other Regression Techniques. **Supervised Learning – Classification:** Basics of Supervised Learning – Classification; Logistic Regression – k-Nearestneighbour – Decision Tree – Support Vector Machine. **Unsupervised Learning:** Basics of Unsupervised Learning – Clustering Techniques; Association Rules.

Unit – 4	ENSEMBLE LEARNING	10 hrs

Concept of Ensemble Learning – Bagging and Boosting and its impact on bias and variance – Random Forest – Adaboost Classifier – Gradient Boosting Machines – XG Boost

Unit –5	PUBLIC DATASETS AND CASE STUDIES	7 hrs

Exploring UCI Machine Learning Repository – Kaggle Data Sets – Analysing Data Sets – Case Studies

Text Books:

1. Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das, “Machine Learning”, Pearson Publications, 2019

Reference Books:-

2. Tom Mitchell, —Machine Learning‖, McGraw Hill, 3rd Edition,1997
3. Christopher M. Bishop, —Pattern Recognition and Machine Learning‖, Springer 2011 Edition.
4. Ethem Alpaydin,"Introduction to Machine Learning‖, MIT Press, Prentice Hall of India, Third Edition 2014.

MSAI 101	ARTIFICIAL INTELLIGENCE	3	0	0	3

Course Outcomes:

CO1: Define basic terminologies of artificial intelligence

CO2: Solve simple search and decision related problems

CO3: Build intelligent agents for search and games

CO4: Learning optimization and inference algorithms for model learning

CO5: Design and develop programs for an agent to learn and act in a structured environment.

Unit – 1	Introduction	9 hrs

Introduction – Definition - Future of Artificial Intelligence – Characteristics of Intelligent Agents – Typical Intelligent Agents – Problem Solving Approach to Typical AI problems

Unit – 2	Problem solving Methods	9 hrs

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 – 3	Knowledge representation	9 hrs

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 – 4	Machine learning	9 hrs

Machine learning-Probability basics - Bayes Rule and its Applications - Bayesian Networks – Exact and Approximate Inference in Bayesian Networks - Hidden Markov Models - Forms of Learning - Supervised Learning - Learning Decision Trees - Regression and Classification with Linear Models - Artificial Neural Networks - Nonparametric Models - Support Vector Machines

Unit – 5	Statistical Learning	9 hrs

Statistical Learning - Learning with Complete Data - Learning with Hidden Variables- The EM Algorithm – Reinforcement Learning -AI applications – Language Models - Information Retrieval - Information Extraction – Natural Language Processing - Machine Translation – Speech recognition – Robot – Hardware – Perception – Planning – Moving

Text Books:

1. S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 3rd Edition, 2009
2. Bratko, I., Prolog Programming for Artificial Intelligence (International Computer Science Series), Addison-Wesley Educational Publishers Inc; 4th edition, 2011.

Reference Books:

1. David L. Poole, Alan K. Mackworth, Artificial Intelligence: Foundations of Computational Agents, Cambridge University Press, 2010.
2. 4.. M. Tim Jones, Artificial Intelligence: A Systems Approach (Computer Science), Jones and Bartlett Publishers, Inc; 1 edition, 2008
3. Ethem Alpaydin, Introduction to Machine Learning (Adaptive Computation and Machine Learning series), The MIT Press; second edition, 2009
4. Nils J. Nilsson, the Quest for Artificial Intelligence, Cambridge University Press, 2009.

MSCS205	WEB TECHNOLOGY LAB	0	0	4	2

Course Outcomes:

CO1: Understand the principles of www and concepts of web clients and web servers
CO2: Interpret the key responsibilities and functionalities of different internet technologies
CO3: Demonstrate Markup languages and illustrate the working of it
CO4: Analyze the different client side scripting/programming languages
CO5: Explain and demonstrate server side , presentation and database technologies

Programs should include but not be limited to:
Implementations of

i. Demonstration of client server programming
ii. Static and Dynamic web pages creation using HTML
iii. Inline, Internal and external style sheet using CSS
iv. Creation and manipulation of XML schema
v. Scripts – Simple programs
vi. Front end development using JavaScript, VBScript
vii. Programs based on sessions and cookies
viii. Web applications (Backend) using PHP, Servlets
ix. Exercises based on database connectivity
x. Web application for event tracking and monitoring

MSCS206	OBJECT ORIENTED PROGRAMMING USING JAVA LAB	0	0	4	2

Course Outcomes:

CO1: Understanding Object-Oriented programming concepts using basic syntaxes of control Structures, strings for developing skills of logic building activity using Java
CO2: Identification of classes, objects, members of a class and the relationships among them needed for a finding the solution to specific problem
CO3: Illustration to achieve reusability using inheritance, interfaces, and packages and describes faster application development can be achieved with exception handling mechanisms
CO4: Understanding concept of multithreading for robust faster and efficient application development and applications of collection interfaces in Java
CO5: Learning of various I/O operations, connecting Java with databases using JDBC and implementation of networking with Java

Programs should include but not limited to:
Programming Assignment on

i. Basic Problems
ii. Arrays
iii. Panagrams
iv. Duplicate Elements
v. Class – Objects – Constructors
vi. Access Specifiers
vii. Static and Non-static variables
viii. Packages and Interfaces
ix. Exception Handling
x. Multithreading and JDBC and Networking

MSCS301	ADVANCED COMPUTER NETWORKS	3	0	0	3

Course Outcomes:

CO1: Describe the fundamental concepts of Networking and Physical layer

CO2: Understand details and functionality of Data link layer.

CO3: Analyze switching protocols and routing algorithms

CO4: Analyze features, services and operations of various protocols of TCP/IP suite

CO5: Identify various application layer protocols and its functions

Unit – 1	FUNDAMENTAL CONCEPTS AND PHYSICAL LAYER	9 hrs

Basic Definitions - Basic Communication Models – Network Types – Protocol Layers and Service Models – OSI Model – TCP/IP protocol suite. Physical Layer: Data and Signals – Digital Transmission – Bandwidth Utilization – Transmission Media – Switching – Packet Switching – Circuit Switched Networks

Unit – 2	DATA LINK LAYER	9 hrs

Error Detection and Correction – Data Link Control – Multiple Access – Wired LANs – Wireless LAN – IEEE 802.11 – Bluetooth – Connecting Devices

Unit – 3	NETWORK LAYER	9 hrs

Circuit Switching – Packet Switching – Virtual Circuit Switching – IP – ARP – DHCP – ICMP – Routing – RIP – OSPF – Subnetting – CIDR – Interdomain Routing – BGP – IPV6 Basic Features – Multicast –Congestion Avoidance in Network Layer.

Unit – 4	TRANSPORT LAYER	9 hrs

Transport Layer Services – Port Numbers – Protocols – User Datagram Protocol (UDP) – Transmission Control Protocol (TCP) – Details - Flow Control – Congestion Control – Queuing Discipline - Introduction to Quality of services (QOS).

Unit – 5	APPLICATION LAYER AND INTERNET APPLICATIONS	9 hrs

Network Architecture – Layers - HTTP – DNS – E-Mail (SMTP, MIME, POP3, IMAP, Web Mail) – FTP – Telnet – SNMP.

Text Books:

1. Behrouz A. Forouzan, “Computer Networks - A top-down approach”, Tata McGraw-Hill, 2012.
2. Computer Networks -- Andrew S Tanenbaum, 6th Edition, Pearson Education

References:-

1. James F. Kurose, Keith W. Ross, “Computer Networking, A Top-Down Approach Featuring the Internet”, Sixth Edition, Pearson Education, 2012.
2. Larry L. Peterson, Bruce S. Davie, “Computer Networks: A Systems Approach”, Fifth
3. Edition, Morgan Kaufmann Publishers Inc., 2012.
4. Ying-Dar Lin, Ren-Hung Hwang and Fred Baker, Computer Networks: An Open Source
5. Approach “, McGraw Hill Publisher, 2011.

MSAI 303	DEEP LEARNING	3	0	0	3

Course Outcomes:

CO1: Recognize characteristics of deep learning models to solve real world problems

CO2: Understand the concepts of convolutional neural networks and its usage in classification problem

CO3: Learn deep learning supporting environments

CO4: Understand the concept of recurrent neural networks and apply them in solving real world problems

CO5: Generate auto encoders and generative models for real time applications

Unit – 1	INTRODUCTION TO DEEP LEARNING & ARCHITECTURES	9 hrs

Machine Learning Vs. Deep Learning – Representation Learning – Width Vs. Depth of Neural Networks - Activation Functions: RELU – LRELU – ERELU –Unsupervised Training of Neural Networks – Regularization – Dropout - Drop connect – Optimization methods for neural networks – Adagrad – Adadelta – Rmsprop – Adam – NAG.

Unit – 2	CONVOLUTIONAL NEURAL NETWORKS & TRANSFER LEARNING	9 hrs

Architectural Overview – Motivation - Layers – Filters – Parameter sharing – Regularization - Popular CNN Architectures: LeNet – ResNet – Vggnet – AlexNet. Transfer learning Techniques –DenseNet – PixelNet.

Unit – 3	TRAINING NEURAL NETWORKS	9 hrs

Deep Learning Hardware and Software – CPUs – GPUs – TPUs – PyTorch – TensorFlow – Dynamic vs Static computation graphs – Data Preprocessing – Data Augmentation – Batch Normalization – Transfer Learning – Deep Transfer Learning Strategies – Update Rules - Hyperparameter Tuning – Learning Rate Scheduling – Variants of CNN – ResNet – GoogleNet – Xception etc

Unit – 4	RECURRENT NEURAL NETWORKS	9 hrs

Recurrent Neural Networks – Bidirectional RNNs – Encoder-decoder sequence to sequence Architectures – Backpropagation Through Time for training RNN – Long Short Term Memory Networks.

Unit – 5	AUTO ENCODERS AND DEEP GENERATIVE MODELS	9 hrs

Under complete Autoencoders – Regulraized Autoencoders – Sparse Autoencoders – Denoising Autoencoders –Representational Power – Layer – Size and Depth of Autoencoders – Stochastic Encoders and Decoders – Contractive Encoders – Deep Belief networks – Boltzmann Machines – Deep Boltzmann Machine – Generative Adversarial Networks

Text Books:

1. Ian Goodfellow, Yoshua Bengio and Aaron Courville, " Deep Learning", MIT Press, 2017.
2. Josh Patterson, Adam Gibson "Deep Learning: A Practitioner's Approach", O'Reilly Media, 2017

Reference Books:

1. Kevin P. Murphy "Machine Learning: A Probabilistic Perspective", The MIT Press, 2012.
2. Ethem Alpaydin,"Introduction to Machine Learning", MIT Press, Prentice Hall of India, Third Edition 2014.
3. Giancarlo Zaccone, Md. Rezaul Karim, Ahmed Menshawy "Deep Learning with TensorFlow: Explore neural networks with Python", Packt Publisher, 2017.

MSDS 301	BIG DATA ANALYTICS	3	0	0	3

Course Outcome:

CO1: To understand the competitive advantages of big data analytics

CO2: To understand the big data frameworks

CO3: To learn data analysis methods

CO4: To learn stream computing

CO5: To gain knowledge on Hadoop related tools such as HBase, Cassandra, Pig, and Hive for big data analytics

UNIT I	INTRODUCTION TO BIG DATA	9 Periods

Big Data – Definition, Characteristic Features – Big Data Applications – Big Data vs Traditional Data – Risks of Big Data- Structure of Big Data – Challenges of conventional Systems – Web Data – Evolution of Analytics Scalability – Evolution of Analytic Processes, Tools and methods – Analysis vs Reporting – Modern Data Analytics Tools.

UNIT II	HADOOP FRAMEWORK	9 Periods

Distributed File Systems – Large-Scale File System Organization – HDFS concepts- MapReduce Execution, Algorithms using MapReduce, Matrix-Vector Multiplication – Hadoop YARN

UNIT III	DATA ANALYSIS	9 Periods

Statistical Methods: Regression modeling, Multivariate Analysis – Classification: SVM & Kernal Methods – Rule Mining – Cluster Analysis, Types of Data in Cluster Analysis, Partitioning Methods, Hierarchical Methods, Density Based Methods, Grid Based Methods, Model Based Clustering Methods, Clustering High Dimensional Data – Predictive Analytics – Data analysis using R

UNIT IV	MINING DATA STREAMS	9 Periods

Streams: Concepts – Stream Data Model and Architecture – Sampling data in a stream – Mining Data Streams and Mining Time – series data- Real Time Analytics Platform (RTAP) Applications – Case Studies – Real Time Sentiment Analysis, Stock Market Predictions

UNIT V	BIG DATA FRAMEWORKS	9 Periods

Introduction to NoSQL – Aggregate Data Models – Hbase: Data Model and Implementations – Hbase Clients – Examples – Cassandra: Data Model – Examples – Cassandra Clients – Hadoop Integration. Pig-Grunt – Pig Data Model – Pig Latin – Developing and testing Pig Latin scripts. Hive- Data Types and File Formats – HiveQL Data Definition – HiveQL Data Manipulation – HiveQL Queries.

Text Books:-

1. Michael Berthold, David J. Hand, “Intelligent Data Analysis”, Springer, Second Edition, 2007.
2. Michael Minelli, Michelle Chambers and Ambiga Dhiraj, “Big Data, Big Analytics:Emerging

References:

1. Business Intelligence and Analytics Trends for Today’s Businesses”, Wiley, 2013.
2. P.J Sadalage and M.Fowler, “NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence”, Addison – Wesley Professional, 2012.
3. Richard Cotton, “Learning R – A Step-by-Step Function Guide to Data Analysis”, O’Reilly Media, 2013.

MSDS 302	DATA MINING & WAREHOUSING	3	0	0	3

Course Outcome:

CO1: This course will allow learners to discuss diverse modalities of data by outlining the core difference in the nature of data and to demonstrate computational methods for outlier detection, classification, and clustering that explores interesting patterns

UNIT I	Introduction	9 Periods

Introduction: What is data; How to Mine Data; Principles of Data Mining; Web content mining; Features and Weighted Features; Class Labels – Binary and Multistage; Data Visualization; Data Discretization and Concept Hierarchy Generation; Feature Interaction; Pattern & Analysis; Noise & Outliers

UNIT II	Outlier Analysis	9 Periods

Outlier Analysis: What is an Outlier; Outlier in Text Data; Outlier in Higher Multimedia Context; Principal Component Analysis; Singular Value Decomposition; Z-score method; Extreme Value Method; Clustering method; Bayesian Outlier Detection.

UNIT III	Data Pre-processing	9 Periods

Data Pre-processing: Data Transformation; Missing Value Detection & Imputation – Types of Imputation; Binning Methods; Regression Models; Geo-Imputation; Rare Pattern Approach.

UNIT IV	Classification	9 Periods

Classification: Introduction to classifiers, Naïve Bayes algorithm, Decision tree, Random Forest, KNN algorithm, Support Vector Machine, Kernel Functions, Back propagation neural network, Symbolic dynamic filter, Extreme learning machines, Metrics for classification

UNIT V	Clustering	9 Periods

Clustering: Introduction to clustering, Categorization of Major Clustering Methods, Partion clustering, Hierarchial clustering, Density-based clustering, Grid-based clustering, Metric for clustering

Text Book:-

1. J. Han and M. Kamber, Data Mining – Concepts & Techniques, Morgan Kaufmann Publisher, 2011.

References:-

2. C. W. Hsu, C. C. Chang & C. J. Liu, Practical Guide to SVM , Taiwan Publishers, 2021
3. F. Sun, K. A. Toh, M. G. Romay & K. Mao, Extreme Learning Machines Algorithms & Applications, Springer Publisher, 2013

MSDS 303	BIG DATA ANALYTICS LABORATORY	0	0	4	2

Course Outcome:

CO1: To implement Map Reduce programs for processing big data

CO2: To realize storage of big data using H base, Mongo DB

CO3: To analyse big data using linear models

CO4: To analyse big data using machine learning techniques such as SVM / Decision tree classification and clustering

List of Experiments

Hadoop:

1. Install, configure and run Hadoop and HDFS
2. Implement word count / frequency programs using MapReduce
3. Implement an MR program that processes a weather dataset

R :

4. Implement Linear and logistic Regression
5. Implement SVM / Decision tree classification techniques
6. Implement clustering techniques
7. Visualize data using any plotting framework 8. Implement an application that stores
8. big data in Hbase / MongoDB / Pig using Hadoop / R.

List of Software:

Hadoop
YARN
R Package
Hbase
MongoDB

References:

1. Alan Gates and Daniel Dai, "Programming Pig – Dataflow scripting with Hadoop", O'Reilley, 2nd Edition, 2016.
2. Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, —An Introduction to Statistical Learning with Applications in R‖, Springer Publications, 2015(Corrected 6 th Printing)
3. Hadley Wickham,‖ggplot2 – Elegant Graphics for Data Analysis‖, Springer Publications,2nd Edition, 2016
4. Kristina Chodorow, "MongoDB: The Definitive Guide – Powerful and Scalable Data Storage", O'Reilley, 2nd Edition, 2013.
5. Lars George, "HBase: The Definitive Guide", O'Reilley, 2015.
6. Tom White, —Hadoop: The Definitive Guide – Storage and Analysis at Internet Scale‖, O'Reilley, 4th Edition, 2015.

	Intellectual Property Rights	2	0	0	2

Course Outcomes:

CO1: To acquaint the learners with the basic concepts of Intellectual Property Rights.
CO2: To develop expertise in the learners in IPR related issues and sensitize the learners with the emerging issues in IPR and the rationale for the protection of IPR.

UNIT I	Introduction	6 Hrs

Introduction to IPRs, Basic concepts and need for Intellectual Property – Meaning and practical aspects of Patents, Copyrights, Geographical Indications, IPR in India and Abroad. **Nature** of Intellectual Property, Industrial Property, technological Research, Inventions and Innovations – Important examples of IPR.

UNIT II	Intellectual Property Rights	6 Hrs

The IPR tool kit, Patents, the patenting process, Patent cooperation treaties: International Treaties and conventions on IPRs: Trade Related Aspects of Intellectual Property Rights Agreement, Patent Cooperation Treaty, Patent Act of India, Patent Amendment Act, Design Act, Trademark Act, Geographical Indication Act.

UNIT III	Intellectual Property Protections	6 Hrs

IPR of Living Species, protecting inventions in biotechnology, protections of traditional knowledge, biopiracy and documenting traditional knowledge, Digital Innovations and Developments as Knowledge Assets – IP Laws, Cyber Law and Digital Content Protection. Case studies: The basmati rice issue, revocations of turmeric patent, revocation of neem patent.

UNIT IV	Exercising and Enforcing of Intellectual Property Rights	6 Hrs

Rights of an IPR owner, licensing agreements, criteria for patent infringement. Case studies of patent infringement, IPR – contract, unfair competitions and control, provisions in TRIPS,

UNIT V	Role of Patents in Product Development & Commercialization	9 Hrs

Recent changes in IPR laws impacting patents and copy rights, intellectual cooperation in the science and allied industry. Patentable and non-patentable research. Case studies

Text Books:

1. P.B. Ganguli, *Intellectual Property Rights: Unleashing the Knowledge Economy.* Tata Mc Graw Hill, 2001.
2. Kompal Bansal and Praishit Bansal. *Fundamentals of IPR for Engineers*, 1st Edition, BS Publications, 2012.

References:

3. Prabhuddha Ganguli. *Intellectual Property Rights.* 1st Edition, TMH, 2012.
4. R Radha Krishnan & S Balasubramanian. *Intellectual Property Rights.* 1st Edition, Excel Books, 2012.
5. M Ashok Kumar & Mohd. Iqbal Ali. *Intellectual Property Rights.* 2nd Edition, Serial Publications, 2011.
6. VinodV. Scople, *Managing Intellectual Property.* Prentice Hall of India PvtLtd, 2012.
7. Deborah E. Bouchoux. *Intellectual Property: The Law of Trademarks, Copyrights, Patents and Trade Secrets.* Cengage Learning, 3rd ed. Edition, 2012.
8. Prabuddha Ganguli. *Intellectual Property Rights: Unleashing the Knowledge Economy.* McGraw Hill Education, 2011.
9. Edited by Derek Bosworth and Elizabeth Webster.*The Management of Intellectual Property.* Edward Elgar Publishing Ltd., 2013.
10. B.S. Patil, Legal Aspects of Building and Engineering Contracts, 1974.
11. Wadhera (2004), Intellectual Property Rights, Universal Law Publishing Co.
12. Ramappa (2010), Intellectual Property Rights Law in India, Asia Law House

SOFTCORE (ENGLISH)	2	0	0	2
This is a Spoken English Training course offered by the department of English. There are many people who know reading and writing of English but find it difficult to speak while communicating with others. As we know, without speaking practice it's impossible to learn and speak any language.				
English is the most used language in this entire world. Speaking English immediately opens doors to opportunities regardless of your race, color, and caste anywhere in this world. Not only has that, speaking fluent English had its major advantages in terms of employability.				
Softcore has carefully taken into account the general mistakes and errors most commonly used by non-English speaking speakers and also aim to reduce the heavy mother tongue influence of non-native speakers.				

MSDS EL01	SPEECH TECHNOLOGY AND ANALYTICS	3	0	0	3

Course Outcome:

CO 1: To understand the need for morphological processing and their representation
CO 2. To know about the various techniques used for speech synthesis and recognition
CO 3. To appreciate the syntax analysis and parsing that is essential for natural language processing
CO 4. To learn about the various representations of semantics and discourse
CO 5. To have knowledge about the applications of natural language processing

UNIT I	SPEECH PROCESSING	9 hrs

Phonetics –Articulatory Phonetics -Phonological Categories -Acoustic Phonetics and Signals - Speech Synthesis –Text Normalization –Phonetic and Acoustic Analysis -Diphone Waveform synthesis – Evaluation-Automatic Speech Recognition –Architecture -Hidden Markov Model to Speech -MFCC vectors -Acoustic Likelihood Computation -Evaluation. Triphones – Discriminative Training -Modeling Variation. Computational Phonology- Finite-State Phonology –Computational Optimality Theory - Syllabification -Learning Phonology and Morphology

UNIT II	SPEECH ANALYSIS	9 hrs

Features, Feature Extraction and Pattern Comparison Techniques: Speech distortion measures – mathematical and perceptual – Log Spectral Distance, Cepstral Distances, Weighted Cepstral Distances and Filtering, Likelihood Distortions, Spectral Distortion using a Warped Frequency Scale, LPC, PLP and MFCC Coefficients, Time Alignment and Normalization – Dynamic Time Warping, Multiple Time – Alignment Paths

UNIT III	SPEECH MODELING	9 hrs

Hidden Markov Models: Markov Processes, HMMs – Evaluation, Optimal State Sequence – Viterbi Search, Baum-Welch Parameter Re-estimation, Implementation issues.

UNIT IV	SPEECH RECOGNITION	9 hrs

Large Vocabulary Continuous Speech Recognition: Architecture of a large vocabulary continuous speech recognition system – acoustics and language models – n-grams, context dependent sub-word units; Applications and present status.

UNIT V	SPEECH SYNTHESIS	9 hrs

Text-to-Speech Synthesis: Concatenative and waveform synthesis methods, sub-word units for TTS, intelligibility and naturalness – role of prosody, Applications and present status.

Text Book:

1. Jurafsky and Martin, “Speech and Language Processing”, Pearson Prentice Hall, Second Edition, 2008.
2. Lawrence Rabinerand Biing-Hwang Juang, “Fundamentals of Speech Recognition”, Pearson Education, 2003

References:

1. Steven W. Smith, “The Scientist and Engineer’s Guide to Digital Signal Processing”, California Technical Publishing.
2. Thomas F Quatieri, “Discrete-Time Speech Signal Processing – Principles and Practice”, Pearson Education.
3. Claudio Becchetti and Lucio Prina Ricotti, “Speech Recognition”, John Wiley and Sons, 1999.
4. Ben gold and Nelson Morgan, “Speech and audio signal processing”, processing and perception of speech and music, Wiley- India Edition, 2006 Edition.
5. Frederick Jelinek, “Statistical Methods of Speech Recognition”, MIT Press

MSDS EL02	Health Care Analytics	3	0	0	3

Course Outcomes:

CO 1: Understand the health data formats, health care policy and standards
CO 2: Learn the significance and need of data analysis and data visualization
CO 3: Understand the health data management frameworks
CO 4: Learn the use of machine learning and deep learning algorithms in healthcare
CO 5: Apply healthcare analytics for critical care applications

Unit – 1	INTRODUCTION TO HEALTHCARE ANALYSIS	9 hrs

Overview - History of Healthcare Analysis Parameters on medical care systems- Health care policy- Standardized code sets – Data Formats – Machine Learning Foundations: Tree Like reasoning , Probabilistic reasoning and Bayes Theorem, Weighted sum approach.

Unit – 2	ANALYTICS ON MACHINE LEARNING	9 hrs

Machine Learning Pipeline – Pre-processing –Visualization – Feature Selection – Training model parameter – Evaluation model : Sensitivity , Specificity , PPV ,NPV, FPR ,Accuracy , ROC , Precision Recall Curves , Valued target variables –Python: Variables and types, Data Structures and containers , Pandas Data Frame :Operations – Scikit –Learn : Pre-processing , Feature Selection.

Unit – 3	HEALTH CARE MANAGEMENT	9 hrs

IOT- Smart Sensors – Migration of Healthcare Relational database to NoSQL Cloud Database – Decision Support System – Matrix block Cipher System – Semantic Framework Analysis – Histogram bin Shifting and Rc6 Encryption – Clinical Prediction Models – Visual Analytics for Healthcare.

Unit – 4	HEALTHCARE AND DEEP LEARNING	9 hrs

Introduction on Deep Learning – DFF network CNN- RNN for Sequences – Biomedical Image and Signal Analysis – Natural Language Processing and Data Mining for Clinical Data – Mobile Imaging and Analytics – Clinical Decision Support System.

Unit –5	CASE STUDIES	9 hrs

Predicting Mortality for cardiology Practice –Smart Ambulance System using IOT –Hospital Acquired Conditions (HAC) program- Healthcare and Emerging Technologies – ECG Data Analysis.

Text Book:

1. Chandan K.Reddy, Charu C. Aggarwal, “Health Care data Analysis”, First edition, CRC, 2015.
2. Vikas Kumar, “Health Care Analysis Made Simple”, Packt Publishing, 2018.
3. Nilanjan Dey, Amira Ashour , Simon James Fong, Chintan Bhatl, “Health Care Data Analysis and Management, First Edition, Academic Press, 2018.
4. Hui Jang, Eva K.Lee, “HealthCare Analysis : From Data to Knowledge to Healthcare Improvement”, First Edition, Wiley, 2016.
5. Kulkarni , Siarry, Singh ,Abraham, Zhang, Zomaya , Baki, “Big Data Analytics in HealthCare”, Springer, 2020.

MSDS EL03	STATISTICS IN DATA SCIENCE	3	0	0	3

Course Outcomes:
CO 1: To study the basic inferential statistics and sampling distribution.
CO 2: To understand the concept of estimation of parameters using fundamental tests and testing of hypotheses.
CO3: To understand the techniques of analysis of variance.
CO4: To gain knowledge in predictive analytics techniques.
CO5: To perform a case study with any available sample data sets.

Unit – 1	INFERENTIAL STATISTICS I	9 hrs

Populations – samples – random sampling – probability and statistics Sampling distribution – creating a sampling distribution – mean of all sample means – standard error of the mean – other sampling distributions Hypothesis testing – z-test – z-test procedure – statement of the problem – null hypothesis – alternate hypotheses – decision rule – calculations – decisions - interpretations

Unit – 2	INFERENTIAL STATISTICS II	9 hrs

Why hypothesis tests? – Strong or weak decisions – one-tailed and two-tailed tests – case studies Influence of sample size – power and sample size32 Estimation – point estimate – confidence interval – level of confidence – effect of sample size

Unit – 3	T-TEST	9 hrs

t-test for one sample – sampling distribution of t – t-test procedure – degrees of freedom – estimating the standard error – case studies t-test for two independent samples – statistical hypotheses – sampling distribution – test procedure – p-value – statistical significance – estimating effect size – meta analysis t-test for two related samples

Unit – 4	ANALYSIS OF VARIANCE	9 hrs

F-test – ANOVA – estimating effect size – multiple comparisons – case studies Analysis of variance with repeated measures Two-factor experiments – three f-tests – two-factor ANOVA – other types of ANOVA Introduction to chi-square tests

Unit – 5	PREDICTIVE ANALYTICS	9 hrs

Linear least squares – implementation – goodness of fit – testing a linear model – weighted resampling Regression using StatsModels – multiple regression – nonlinear relationships – logistic regression – estimating parameters – accuracy Time series analysis – moving averages – missing values – serial correlation – autocorrelation Introduction to survival analysis

TEXT BOOKS

1. Robert S. Witte and John S. Witte, “Statistics”, Eleventh Edition, Wiley Publications, 2017.
2. Allen B. Downey, “Think Stats: Exploratory Data Analysis in Python”, Green Tea Press, 2014. [Unit V]

REFERENCES

1. David Spiegelhalter, “The Art of Statistics: Learning from Data”, Pelican Books, 2020.
2. Peter Bruce, Andrew Bruce, and Peter Gedek, “Practical Statistics for Data Scientists”, Second Edition, O’Reilly Publishers, 2020.
3. Charles R. Severance, “Python for Everybody: Exploring Data in Python 3”, Shroff Publishers, 2017.
4. Bradley Efron and Trevor Hastie, “Computer Age Statistical Inference”, Cambridge University Press, 2016.

MSDS EL04	Full Stack of Data Science	3	0	0	3

Course Outcomes:

CO1: To introduce core programming basics required for data science using Python language

CO2: To read and write simple Python programs

CO3: To develop Python programs with conditionals and loops

CO4: To use Python data structures – lists, tuples, dictionaries

CO5: To introduce the important data science modules NumPy, SciPy and Matplotlib

CO6: To introduce the input/output with files in Python and statistical processing of a data using Pandas

Unit – 1	Algorithmic Problem Solving	9 hrs

Algorithms, building blocks of algorithms (statements, state, control flow, functions); algorithmic problem solving; iteration, recursion. Illustrative problems: finding minimum in a list, guess an integer number in a range, factorial of a number

Unit – 2	Data, Expressions, Statements in Python	9 hrs

Python Strengths and Weakness; Installing Python; IDLE - Spyder – Jupyter; Mutable and Immutable Data Types, Naming Conventions; String Values; String Operations; String Slices; String Operators; String functions – split, join, chr, ord; Numeric Data Types; Arithmetic Operators and Expressions; Comments in the Program; Understanding Error Messages

Unit – 3	Data Collection and Language Component of Python	9 hrs

List; Tuples; Sets; Dictionaries; Sorting Dictionaries; Control Flow and Syntax; Indenting; The if statement; Relational Operators; Logical Operators; Bit-wise Operators; The while Loop – break and continue statements; The for Loop; List Comprehension

Unit – 4	Functions and Modules in Python	9 hrs

Functions - Introduction; Defining your own functions; parameters; local and global scope; passing collections to a function; variable number of arguments; passing functions to a function; Lambda function; map; filter; Modules: Introduction; Standard Modules – sys, math, time

Unit – 5	Python Modules for Data Science	9 hrs

NumPy arrays – 1-d, multidimensional arrays and matrices; Mathematical operations with arrays; Slicing and addressing arrays; Boolean masks; Difference between lists and arrays SciPy – Scientific Computing library of Python – Introduction, Basic functions, Special functions, scipy.integrate, scipy.optimize, scipy.interpolate

TEXT BOOKS

1. David J. Pine, Introduction to Python for Science and Engineering, CRC Press, 2019.
2. Jake vander Plas, Python Data Science Handbook – Essential Tools for Working with Data, O'Really Media, 2017

REFERENCES

1. Robert Johansson, Numerical Python – Scientific Computing and Data Science Applications with NumPy, SciPy and Matplotlib, Apress, 2019
2. Robert Sedgewick, Kevin Wayne, Robert Dondero, Introduction to Programming in Python: An Inter-disciplinary Approach, Pearson India Education Services Pvt. Ltd., 2016
3. Nelli, F., Python Data Analytics: with Pandas, NumPy and Matplotlib, Apress, 2018.

MSCS EL01	INFORMATION RETRIEVAL TECHNIQUES	3	0	0	3

Course Outcomes:

CO1: To understand the basics of information retrieval with pertinence to modeling, query operations and indexing

CO2: To get an understanding of machine learning techniques for text classification and clustering.

CO3: To understand the various applications of information retrieval giving emphasis to multimedia IR, web search

CO4: To understand the concepts of digital libraries

UNIT I	Introduction	9 Periods

Motivation: Basic Concepts – Practical Issues - Retrieval Process – Architecture - Boolean Retrieval – Retrieval Evaluation – Open Source IR Systems–History of Web Search – Web Characteristics– The impact of the web on IR —IR Versus Web Search–Components of a Search engine

UNIT II	Modeling	9 Periods

Taxonomy and Characterization of IR Models – Boolean Model – Vector Model – Term Weighting – Scoring and Ranking –Language Models – Set Theoretic Models - Probabilistic Models – Algebraic Models – Structured Text Retrieval Models – Models for Browsing, Zipf's law, Porter stemmer

UNIT III	Indexing	9 Periods

Static and Dynamic Inverted Indices – Index Construction and Index Compression.
Searching - Sequential Searching and Pattern Matching. Query Operations -Query Languages – Query Processing - Relevance Feedback and Query Expansion - Automatic Local and Global Analysis – Measuring Effectiveness and Efficiency

UNIT IV	Classification and Clustering	9 Periods

Text Classification and Naïve Bayes, Improved smoothing for document retrieval – Vector Space Classification – Support vector machines and Machine learning on documents. Flat Clustering – Hierarchical Clustering –Matrix decompositions and latent semantic indexing – Fusion and Meta learning 41

UNIT V	Searching the Web	9 Periods

Applications to web search and information organization, Searching the Web –Structure of the Web –IR and web search – Static and Dynamic Ranking – Web Crawling and Indexing – Link Analysis - XML Retrieval Multimedia IR: Models and Languages – Indexing and Searching Parallel and Distributed IR – Digital Libraries, Privacy, Fairness, Fake news and disinformation, Deep Learning for IR

Text Books:

1. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schutze, —Introduction to Information Retrieval, Cambridge University Press, First South Asian Edition, 2008.
2. Implementing and Evaluating Search Engines, The MIT Press, Cambridge, Massachusetts London, England, 2010

Reference Books:

1. Ricardo Baeza – Yates, Berthier Ribeiro – Neto, —Modern Information Retrieval: The concepts and Technology behind Search‖ (ACM Press Books), Second Edition, 2011.
2. Stefan Buttcher, Charles L. A. Clarke, Gordon V. Cormack, —Information Retrieval

MSCS EL02	OPERATIONS RESEARCH	3	0	0	3

Course Outcomes:

CO1: Formulate various real-life problems as Operations Research models and to study methodologies to solve the problems.

CO2: Introduce Linear Programming, Transportation and Assignment problems and to discuss methods to find optimum solutions.

CO3: Study the theory of duality and sensitivity analysis in linear programming.

CO4: Learn network flow problems and their solution techniques.

CO5: Explore dynamic programming problem and its applications.

UNIT I	Introduction	9 Periods

Introduction – Models in Operations Research – Linear Programming Problems – Simplex Method – Big-M Method – Two-Phase Method – Special Cases – Degeneracy and Cycling – Unbounded Solutions – Alternative Optima.

UNIT II	Dual Linear Programs	9 Periods

Dual Linear Programs – Duality Theorems – Dual Simplex Method - Transportation Problems - Finding an Initial Basic Feasible Solution - Optimality Condition – MODI method – Degeneracy – Assignment Problems – Hungarian Method.

UNIT III	Revised Simplex Method	9 Periods

Revised Simplex Method – Sensitivity Analysis – Parametric Programming.

UNIT IV	Network Analysis	9 Periods

Network Analysis – Shortest Route Problems – Maximal Flow Problems – Critical Path Method (CPM) – Program Evaluation and Review Techniques (PERT).

UNIT V	Dynamic Programming	9 Periods

Dynamic Programming – Introduction – Principle of Optimality – Forward and Backward recursions – Discrete Dynamic Programming – Continuous Dynamic Programming – Applications.

Text Books:

1. Ravindran, Don T. Phillips and James J. Solberg, *Operations Research- Principles and Practice*, John Wiley, 2014.
2. Hamdy A. Taha, *Operations Research-An Introduction*, Prentice Hall of India, 2000.

Reference Books:

1. Frederick S. Hillier and Gerald J. Lieberman, *Introduction to Operations Research*, McGraw Hill, 2010.
2. KantiSwarup, P.K. Gupta and Man Mohan, *Operations Research*, Sultan Chand,2014.

MSCS EL03	THEORY OF COMPUTATION	3	0	0	3

Course Outcomes:

CO1: Use regular language and grammar for designing real world problems

CO2: Perform computational complexity analysis

CO3: Define Turing machines performing simple tasks.

CO4: Understand formal languages

Unit – 1	Introduction to theory of computation	9 hrs

Introduction to theory of computation and Finite Automata: Mathematical preliminaries – Basic concepts – Applications – DFA – NFA – Equivalence – Reduction of states.

Unit – 2	REGULAR EXPRESSIONS AND LANGUAGES	9 hrs

Regular Language (RL) , Regular Grammar, Properties of RL: Regular Expressions (RE) – Relation between RE and RL – Regular Grammars – Properties – Context Free Grammars (CFG)

Unit – 3	CONTEXT FREE GRAMMAR AND LANGUAGES	9 hrs

Simplification of Context Free Grammars & Normal Forms: Methods for transforming Grammars – Chomsky and Greibach Normal Forms Push Down Automata (PDA)Non-deterministic PDA – PDA and Context Free Languages (CFL) – Deterministic PDA and CFL

Unit – 4	PROPERTIES OF CONTEXT FREE LANGUAGES	9 hrs

Properties of CFL and Turing Machines: Pumping lemma –closure properties Turing machines ™ – the standard TM – Turings' thesis – Linear Bounded Automata

Unit –5	UNDECIDABILITY	9 hrs

Non Recursive Enumerable (RE) Language — Undecidable Problem with RE —Undecidable Problems about TM — Post Correspondence Problem, The Class P and NP.

Text Books:

1.Peter Linz, An introduction to Formal Languages and Automata, 2012, Fifth Edition,
Jones & Bartlett Learning

Reference Books

1. Automata, Computability and Complexity: Theory and Applications, Pearson Education India; 1 edition 2012.

2. Moore, Cristopher, and Stephan Mertens. The Nature of Computation. Oxford University Press, 2011.

3. Arora, Sanjeev, and Boaz Barak. Computational Complexity: A Modern Approach. Cambridge University Press, 2009

MSAI EL04	VIDEO PROCESSING AND ANALYTICS	3	0	0	3

Course Outcomes:

CO1: To have a better knowledge about videos.

CO2: To enrich students with data analytics.

CO3: To understand the video content analysis.

Unit – 1	Video Fundamentals	9 hrs

Video Fundamentals : Basic concepts and Terminology-Monochrome Analog video – Colour in Video – Analog video standards – Digital video basics – Analog to Digital conversion – Colour representation and chromasub sampling – Digital video formats and standards Video sampling rate and standards conversion.

Unit – 2	Video Segmentation and Video Features	9 hrs

Video Segmentation and Video Features : Fundamentals of Motion Estimation – Optical flow - Pixel Video Features - colour, shape features, Textural features -Feature selection and Dimensionality Reduction.

Unit – 3	Introduction to Analytics	9 hrs

Introduction to Analytics : Big-Data - Descriptive data analysis - Analytic Processes and Tools - Regression – Classification - Clustering algorithms - Validation - Multimodal approach to Image and Video data mining - Probabilistic semantic mode - Model based annotation and video mining.

Unit – 4	Video Content Analysis and Analytics :	9 hrs

Video Content Analysis and Analytics : Introduction- Detecting Shot Boundaries in Video – Parsing a Video into Semantic Segments – Video Indexing and Abstraction for Retrievals – Affective Video Content Analysis – Automatic Video Trailer Generation - Video database - Video categorization - Video query categorization.

Unit –5	Emerging Trends	9 hrs

Emerging Trends : Object Segmentation and Tracking in the Presence of Complex Background – Video In painting –Video Summarization - Forensic video analysis.

Text Book:

1. Michael Berthold, David J.Hand, Intelligent Data Analysis, Springer, 2007.

Reference Books

1. Oges Marques, Practical Image and Video Processing Using MATLAB, Wiley-IEEEPress, 2011.
2. Anand Rajaraman and Jeffrey David Ullman, Mining of Massive Datasets,CambridgeUniversity Press, 2012.

MSAI EL05	DATA VISUALIZATION TECHNIQUES	3	0	0	3

Course Outcome:

CO1: To develop skills to both design and critique visualizations.

CO2: To introduce visual perception and core skills for visual analysis.

CO3: To understand visualization for time-series, ranking, deviation, distribution, correlation, multivariate analysis.

CO4: To understand issues and best practices in information dashboard design.

UNIT I	INTRODUCTION	9 hrs

Core Skills for Visual Analysis: Information visualization – effective data analysis – traits of meaningful data – visual perception –making abstract data visible – building blocks of information visualization – analytical interaction – analytical navigation – optimal quantitative scales – reference lines and regions – trellises and crosstabs – multiple concurrent views – focus and context – details on demand – over-plotting reduction – analytical patterns – pattern examples.

UNIT II	KNOWLEDGE REPRESENTATION AND REASONING	9 hrs

Time-Series, Ranking, and Deviation Analysis: Time-series analysis – time-series patterns – time-series displays – time-series best practices – part-to-whole and ranking patterns – part-to-whole and ranking displays – best practices – deviation analysis – deviation analysis displays – deviation analysis best practices.

UNIT III	UNCERTAINTY	9 hrs

Distribution, Correlation, and Multivariate Analysis: Distribution analysis – describing distributions – distribution patterns – distribution displays – distribution analysis best practices – correlation analysis – describing correlations – correlation patterns – correlation displays – correlation analysis techniques and best practices – multivariate analysis – multivariate patterns – multivariate displays – multivariate analysis techniques and best practices.

UNIT IV	LEARNING	9 hrs

Information Dashboard Design: Information dashboard – Introduction– dashboard design issues and assessment of needs – Considerations for designing dashboard-visual perception – Achieving eloquence.

UNIT V	INTELLIGENCE AND APPLICATIONS	9 hrs

Information Dashboard Design: Advantages of Graphics _Library of Graphs – Designing Bullet Graphs – Designing Sparklines – Dashboard Display Media –Critical Design Practices – Putting it all togetherUnveiling the dashboard.

Text Book:

1. Ben Fry, "Visualizing data: Exploring and explaining data with the processingenvironment", O'Reilly, 2008.

References:

1. Edward R. Tufte, "The visual display of quantitative information", Second Edition,Graphics Press, 2001.
2. Evan Stubbs, "The value of business analytics: Identifying the path to profitability",Wiley, 2011.
3. Gert H. N. Laursen and Jesper Thorlund, "Business Analytics for Managers: Takingbusiness intelligence beyond reporting", Wiley, 2010.

MSCS EL06	CYBER SECURITY	3	0	0	3

Course Outcomes:

CO1: To understand the principles of Data warehousing and Data Mining
CO2: To be familiar with the Data warehouse architecture and its Implementation
CO3: To know the Architecture of a Data Mining system
CO4: To understand the various Data preprocessing Methods
CO5: To perform classification and prediction of data

Unit – 1	COMPUTER SECURITY CONCEPTS	9 hrs

Definitions – Threats – Harm – Vulnerabilities – Unintentional (Non-Malicious) Programming – Malicious Code – Malware Countermeasures

Unit – 2	ELECTRONIC PAYMENTS AND SAFEGUARDS	9 hrs

Concept of E-payments – ATM and Tele Banking – Immediate Payment Systems – Mobile Money Transfer and E-Wallets – Unified Payment Interface – Cybercrimes in Electronic Payments – Precautions in Electronics Money Transfer – Precautions in Electronic Money Transfer – RBI Guidelines of Customer Protection in Unauthorized Banking Transactions – KYC: Concept, cases and safeguards

Unit – 3	CYBER CRIMES AND SAFETY	9 hrs

Introdution to cybercrimes – Kinds of Cybercrimes: Phishing – Identity Theft – Cyber Stalking – Cyber Terrorism – Cyber Obscenity – Computer Vandalism – Ransomware – Identity Theft Forgery – fraud from Mobile Devices

Unit – 4	INTRODUCTION TO SOCIAL NETWORKS	9 hrs

Social Network and its contents – Blogs – Safe and Proper use of Social Networks – Inappropriate Content on Social Networks – Flagging and Reporting of Inappropriate Content – Laws regarding posting of Inappropriate Content

Unit – 5	INTRODUCTION TO INFORMATION AND TECHNOLGOY ACT, 2000 (IT ACT) AND ITS USE IN CYBER SPACE	9 hrs

Concepts as defined in IT Act – Communication Device – Computer, Cyber Security, Data Security – Secure System – Online Gaming and its Risks – Basic Concepts of Blockchain and Cryptocurrency

Text Books:

1. W.A.Coklin, G.White, Principles of Computer Security: Fourth Edition, McGrawHill, 2016
2. William Stallings, Cryptography and Network Security Principles and Practices, Seventh Edition,Pearson

Reference Books:

1. Pfleeger, C.P., Security in Computing 5th Edition, Prentice Hall, Copyright 2010,
2. ISBN 0-13-239077-9
3. Schneier, Bruce. Applied Cryptography, Second Edition, John Wiley & Sons, 1996..
4. J Michael Stewart, "Network Security, Firewalls And Vpns", Jones and Bartlett Publishers, Inc; 2nd edition, 2013
5. Charlie Kaufman and Radia Perlman, "Network Security: Private Communication in a Public World", Prentice Hall, 2002

MSCS EL07	SOFT COMPUTING	3	0	0	3

Course Outcomes:

CO1: To provide a strong foundation of fundamental concepts in Artificial Intelligence.
CO2: To provide a basic exposition to the goals and methods of Artificial Intelligence.

UNIT I	Introduction to Soft Computing	9 Periods

Introduction: Introduction to soft computing - brief description of separate theories, Introduction to biological and artificial neural network, Classification algorithms- Decision Trees, Bayesian classifier - Neural Networks and Probabilistic Reasoning

UNIT II	Neural Networks	9 Periods

Neural Networks: Basic concepts of neural networks, Neural network architectures, Learning methods, Supervised and un- supervised learning, Architecture of a back- propagation network, Applications

UNIT III	Fuzzy Sets	9 Periods

Fuzzy Sets: Fundamentals of fuzzy sets and fuzzy logic theory, fuzzy inference principle, Examples of use of fuzzy logic in control of real-world systems

UNIT IV	Optimization	9 Periods

Optimization: Derivative-based Optimization – Descent Methods – The Method of Steepest Descent – Classical Newton's Method – Step Size Determination – Derivative-free Optimization – Genetic Algorithms –Simulated Annealing – Random Search Downhill Simplex Search

UNIT V	Applications of Computational Intelligence	9 Periods

Applications of Computational Intelligence: AI Search Algorithm-Predicate calculus- rules of interface - Semantic networks – frames – objects - Hybrid models, Applications -Printed Character Recognition – Inverse Kinematics Problems – Automobile Fuel EfficiencyPrediction – Soft Computing for Colour Recipe Prediction

Text Books:

1. Jang J.S.R.,Sun C.T and Mizutami E - Neuro Fuzzy and Soft ComputingPrentice hall,New Jersey,1998.
2. Munakata, T.: Fundamentals of the New Artificial Intelligence, Springer-Verlag NewYork, Inc.,1998.

References:

1. Goldberg, Introduction to Genetic Algorithms.
2. Jang, Nero-Fuzzy & Soft Computing, Pearsons.
3. Cordón, O., Herrera, F., Hoffman, F., Magdalena, L.: Genetic Fuzzysystems,World Scientific Publishing Co. Pte. Ltd., 2001.
4. Kecman, V.: Learning and Soft Computing, The MIT Press, 2001

MSAI 302	NATURAL LANGUAGE PROCESSING	3	0	0	3

Course Outcomes:

- To learn about the concepts and principles of natural language processing.
- To explore both theoretical and practical issues of natural language processing.
- To develop skills of finding solutions and building software using natural languageprocessing techniques.

Unit – 1	Introduction to Natural Language Processing	9 hrs

Introduction: Natural Language Processing tasks in syntax, semantics, and pragmatics – Issues - Applications - The role of machine learning - Probability Basics –Information theory – Collocations -N-gram Language Models - Estimating parameters and smoothing - Evaluating language models.

Unit – 2	Morphology and part of speech tagging	9 hrs

Morphology and part of speech tagging: Linguistic essentials - Lexical syntax- Morphology and Finite State Transducers - Part of speech Tagging - Rule-Based Part of Speech Tagging - Markov Models - Hidden Markov Models – Transformation based Models - Maximum Entropy Models. Conditional Random Fields

Unit – 3	Syntax Parsing	9 hrs

Syntax Parsing - Grammar formalisms and treebanks - Parsing with Context Free Grammars - Features and Unification -Statistical parsing and probabilistic CFGs (PCFGs)-Lexicalized PCFGs.

Unit – 4	Semantic analysis	9 hrs

Semantic analysis- Representing Meaning – Semantic Analysis - Lexical semantics –Word-sense disambiguation - Supervised – Dictionary based and Unsupervised Approaches - Compositional semantics, Semantic Role Labeling and Semantic Parsing – Discourse Analysis.

Unit – 5	Applications	9 hrs

Applications- Named entity recognition and relation extraction- IE using sequence labeling- Machine Translation (MT) - Basic issues in MT-Statistical translation-word alignment- phrase- based translation – Question Answering

Text / Reference Books:

1. Daniel Jurafsky and James H. Martin Speech and Language Processing (2nd Edition), Prentice Hall; 2 edition, 2008
2. Foundations of Statistical Natural Language Processing by Christopher D. Manning and Hinrich Schuetze, MIT Press, 1999
3. Steven Bird, Ewan Klein and Edward Loper Natural Language Processing with Python, O'Reilly Media; 1 edition, 2009
4. Roland R. Hausser, Foundations of Computational Linguistics: Human- C o m p u t e r Communication in Natural Language, Paperback, MIT Press, 2011
5. Pierre M. Nugues, An Introduction to Language Processing with Perl and Prolog: An Outline of Theories, Implementation, and Application with Special Consideration of English, French, and German (Cognitive Technologies) Softcover reprint, 2010
6. James Allen, Natural Language Understanding, Addison Wesley; 2 edition 1994

MSAI 201	KNOWLEDGE ENGINEERING AND EXPERT SYSTEMS	3	0	0	3

Course Outcomes:

CO1: Demonstrate the knowledge of fundamental elements and concepts related to Intelligent Systems

CO2: Demonstrate the fundamental and advanced modules of KEespecially with Searching methods, Representation of knowledge and different reasoningtechniques.

CO3: Ability to work with Predicate logic, back propagation with respect to the CNNs model parameters and implementing the models successfully.

CO4: Apply the higher order logics for handling uncertainty

CO5: Implement an expert system to solve critical problems of medical domain, application of business intelligence and robotics in real life problems.

Unit – 1	Knowledge Engineering Concepts	**9 hrs**

Definition of Knowledge Engineering – Knowledge base Systems – Knowledge base systems Vs Database systems – Rules Vs Triggers – Domain Expert – Expert Systems –Heuristic Search – A*, AO* and Mini-max algorithms - Knowledge representation – Semantic Networks – Frames- Conceptual Dependency – Scripts – Ontology – Semantic Web– Reasoning Methods

Unit – 2	First Order Logic	**9 hrs**

Role of Logic –Propositional logic – Predicate logic – Syntax – Semantics – Interpretations – Denotation – Satisfaction and models – Pragmatics – Explicit and Implicit Beliefs - Logical Consequence – Expressing Knowledge - Basic and Complex Facts – Terminological facts – Entailment –Abstract Individuals - Other Sorts of Facts –Resolution – The Propositional Case – Predicate Logic – Handling Variables and Quantifiers –First Order Resolution- Answer Extraction – Skolemization – Clause Form – Equality - Dealing with Computational Intractability - The First-Order Case - Herbrand Theorem - The Propositional Case - The Implications - SAT Solvers - Most General Unifiers - Other Refinements

Unit – 3	Knowledge Representation – Using Rules	**9 hrs**

Procedural Versus Declarative Knowledge - Logic Programming - Forward versus Backward Reasoning – Rule Matching – Rules in Production Systems- Working Memory- Conflict Resolution- Rete's Algorithm – Discriminant Networks - Control Knowledge –Reasoning with Horn Clauses – Computing Selective Linear Definite clause resolution Derivatives – Rule Formation and Search Strategy – Algorithm Design – Specifying Goal order – Committing to Proof methods – Controlling Back Tracking – Negation as Failure – Dynamic Databases

Unit – 4	Expert Systems and Learning	**9 hrs**

Expert Systems – Shells for Expert Systems – Inference Engine – Forward and Backward Chaining Inference – MYCIN - DENDRAL –Knowledge Acquisition - Rote Learning – Learning from Examples – Machine Learning- Neural Networks – Regression Analysis- Predictive Models - Deep Learning

Unit –5	Applications of Knowledge Base Systems	**9 hrs**

Factory Automation -Field and Service Robotics–AssistiveRobotics -Military Applications -Medicare–Education – Business Intelligence – Recommendation Systems – Social Network Analysis – Natural Language Processing – Information Retrieval Systems

Text Books:

Ronald Brachman, Hector Levesque, Knowledge Representation and Reasoning, 1st Edition, Morgan Kaufmann, 2004

2. Richard A Frost, "Introduction to Knowledge Based Systems", Macmillan Publishing Co, 1986.

Reference Books:

Elaine Rich, Kevin Knight, Shivashankar B. Nair, "Artificial Intelligence", Third Edition, Tata McGraw-Hill Education Pvt. Ltd., 2010.

2. Donald A Waterman,"A Guide to Expert Systems", Addison Wesley, 1986.

3. Schall, Daniel,"Social Network-Based Recommender Systems", Springer, 2015