S. S. JAIN SUBODH P.G. COLLEGE

(AUTOMOMOUS)

(Affiliated to University of Rajasthan, Jaipur)

DEPARTMENT OF COMPUTER SCIENCE

NATIONAL EDUCATION POLICY-2020

SYLABUS OF PG

Master of Science Information Technology (AI and Data Science)
(M.Sc. IT – AI and Data Science)

Syllabusfor1st to4th Semesters

M.Sc. IT – AI and Data Science Part I (I & II Sem)2025 – 26

M.Sc. IT – AI and Data Science Part II (III& IV Sem)2026 – 27

onwards

M.Sc.IT (AI and Data Science)

Two Year (Four Semesters) Course

The course of study of M.Sc. IT(AI and Data Science)shall extend over a period of **four semesters** spread over two years. On satisfactory completion of the course and after passing the examinations, a candidate will be awarded the degree of **Master of Science IT(AI and Data Science).**

Every academic year shall be divided into two semesters. First semester starts from July and ends in December. Second semester starts from January and ends in June and so on. There shall be an examinationattheendofeachsemester.Theexaminationsshallconsistoftheorypapers,practical papers. Mediumofinstructionsandexaminationwillbe**English**only.

EligibilityforAdmission:

The admission of the students for M.Sc. IT (AI and Data Science)I and III Semester will be taken for July to December session and II and IV Semester will be from January to June session.
A candidate who has qualified graduate degree with at least 50% marks (48% for SC/ST/OBC/SOBC category) of Rajasthan University or any other university recognized as equivalent shall be admitted to the first year of **M.Sc. IT(AI and Data Science)** Course.

Semester	CourseDuration	ExaminationTime
I,III	JulytoDecember	December
II,IV	JanuarytoJune	June

Examination Scheme:

Each theory paper shall be of 100 marks (70 marks for written examination of 3 hrs duration and 30 marks for internal assessment).
Each practical paper shall be of 100 marks (60 for practical exam and 40 for internal assessment). The basis for internal evaluation in theory shall be home assignment, internal test and regularities in the attendance.
The basis for internal assessment in the laboratory courses shall be timely submission of the lab. records,performance in the lab., internal tests etc.

Eachtheorypaperexaminationwillbeofthree-hourdurationandshallcarry70marks.Theory paper shall contain three parts.

a) Part-Awillcontain 12veryshort questionsof1markeach(studentwill attemptany 10).
b) PartBwillcontain4shortdescriptivetypesofquestionseachcarrying5marks,allare compulsory.
c) Part Cwill contain 4 long descriptive types of questions each carrying 10 marks, allquestions are compulsory with internal choice.

Each practical examination (Maximum marks 100) will be of two - hour duration on one day and carry 60 marks for exercise(s) assigned in the examination and Viva, and 40 marks for theInternal Assessment.

S.No.	QuestionPattern	Max.Marks (Theory)		Max.Marks (Practical)	
		ESE	CIA	ESE	CIA
1	PartA:12VeryShortQuestions (attempt any 10)	10 X1=10			
2	PartB:4ShortQuestionfromallUnits (Compulsory)	4 X5 =20			
3	PartC:4QuestionsfromeachUnit with Internal Choice	4 X10=40			
	SubTotal	**70**	**30**	**60**	**40**
		100		**100**	

Attendance:
Acandidateshallberequiredtoputinaminimumof75%attendanceatthelecturesand75% Attendanceatthepractical'sseparatelyineach paper.

Examination Scheme:

Sr. No.	Paper	ESE	CIA	Total
1	Theory	70%	30%	100
2	Practical	60%	40%	100

NumberofUnitsinSyllabus:

Sr. No.	Stream	No. of Units
1	**M.Sc. IT(AI and Data Science)**	4

MaximumMarksfortheCourseandNumberofUnitsofEachsubject2Years/4 Semesters:

Sr. No	Stream	Semester	NumberofPapers per semester		Total Marks (Credits)of1 Semester	SubTotal	Grand Total
			Theory (Credits)	Practical (Credits)			
1	**M. Sc. IT (AI and Data Science)**						
		I ,II,III Sem.	5 X100 (5 X4)	3 X100 (3 X4)	500 +300 = 800 (20 + 12 = 32)	800 X3 = 2400 (32 X3 = 96)	2400 (96)
		IVSem, Project	2 X100 (2 X4)	1 X200 (1 X8)	300+300 = 600 (8 +8 =16)	600 X1 =600 (16)	600 (16)
							3000 (112)

ProgrammeOutcomes :

PO1: The program equips students with essential domains of computer science and AI and Data Science. It empowers them to apply core concepts in the development of domain-specific applications.

PO2: The program fosters critical thinking, cultivates problem-solving skills, encourages evaluative learning of various techniques, and deepens comprehension of problem essence.

PO3: The program educates students about the latest industry technologies. The continuous review of syllabi adds value to graduates, preparing them to address dynamic industry demands.

PO4:Theprograminstructsstudentsinapplyingadvancedtoolstoaddressreal-worldchallenges.

PO5: The program trains students in designing and conceptualizing software architecture, managing complex product development processes, and making informed decisions for project management selection.

PO6: Real-world projects expose students to challenging industry environments, making thememployable and industry-ready through hands-on project development training.

PO7: The program hones students' teamwork skills and nurtures their ability to lead project management teams

SemesterStructure:

M.Sc. IT(AI and Data Science)SemesterI

S.No.	Subject Code	Course Title	Course Category	Credit	Contact Hours Per Week			EoSE Duration (Hrs.)	
					L	T	P	Thy	P
1.	MSAIDS101	Programming Fundamentals	DSC	4	4			3	
2.	MSAIDS102	DatabaseManagement System	DSC	4	4			3	
3.	MSAIDS103	Data Communication and Networking	DSC	4	4			3	
4.	MSAIDS104	Computer Architecture	DSC	4	4			3	
5.	MSAIDS105	OperatingSystem	DSC	4	4			3	
6.	MSAIDS151	‘C’ProgrammingLab	DSCP	4			6		2
7.	MSAIDS152	DBMSLab	DSCP	4			6		2
8.	MSAIDS153	OfficeManagementLab	DSCP	4			6		2

M.Sc. IT (AI and Data Science)SemesterII

S.No.	Subject Code	Course Title	Course Category	Credit	Contact Hours Per Week			EoSE Duration (Hrs.)	
					L	T	P	Thy	P
1.	MSAIDS201	ObjectOriented Programming Concepts using java Programming	DSC	4	4			3	
2.	MSAIDS202	DataStructureand Algorithms	DSC	4	4			3	
3.	MSAIDS203	WebDesigningand development	DSC	4	4			3	
4.	MSAIDS204	Fundamental of AI, Machine learning and Data Science	DSC	4	4			3	
5.	MSAIDS205	Big data Analysis	DSC	4	4			3	
6.	MSAIDS251	JavaProgrammingLab	DSCP	4			6		2
7.	MSAIDS252	DataStructureLab	DSCP	4			6		2
8.	MSAIDS253	WebDesigningLab	DSCP	4			6		2

M. Sc. IT (AI and Data Science)SemesterIII

S.No.	Subject Code	Course Title	Course Category	Credit	Contact Hours Per Week			EoSE Duration (Hrs.)	
					L	T	P	Thy	P
1.	MSAIDS301	Data Science with R	DSC	4	4			3	
2.	MSAIDS302	PythonProgramming With AI and Data Science	DSC	4	4			3	
3.	MSAIDS303	Data Visualization Tools and Techniques	DSC	4	4			3	
4.	MSAIDS304	Machine learning	DSC	4	4			3	
5.	MSAIDS305	Cloud computing	DSC	4	4			3	
6.	MSAIDS351	R Programming Lab	DSCP	4			6		2
7.	MSAIDS352	PythonLab	DSCP	4			6		2
8.	MSAIDS353	Machine learning Lab	DSCP	4			6		2

M. Sc. IT (AI and Data Science)SemesterIV

S.No.	Subject Code	Course Title	Course Category	Credit	Contact Hours Per Week			EoSE Duration (Hrs.)	
					L	T	P	Thy	P
1.	MSAIDS 401	**ElectivePaper-I** a) Data warehousing and data Mining b) Soft Computing	DSC	4	4			3	
2.	MSAIDS 402	**ElectivePaper-II** a) Natural language Processing b) Deep Learning	DSC	4	4			3	
3.	MSAIDS 451	**Project(Project,Report, Viva)** (12Weeks)	DSCP	8					2

M.Sc.IT (AI and Data Science) - FirstSemester

MSAIDS101:ProgrammingFundamentals

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks: 40*

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1:UnderstandtheconceptofAlgorithmsanddifferentsymbolsusedinflowcharts
CO2: Develop conditional and iterative statements to write C programs.
CO3:Utilizeuser-definedfunctionstosolvereal-timeproblems.
CO4: Create C programs that utilize pointers to access arrays, strings, and functions.
CO5:Applyuser-defineddatatypes,includingstructuresandunions,tosolveproblems.
CO6:Demonstratetheconceptoffilehandlingtoshowcaseinputandoutputoperationsin C programs.

Unit-I

ProblemSolvingwithComputers:Algorithms,and Flowcharts.Basic Structure of C Programs: Creating, Compiling, Linking and Execution of a C program Header Files, C Tokens (Keywords, Identifiers, Constants, and Special symbols), Data types, and Variable declaration, Operators, Expressions, Managing Data Input and Output Operations. Decision Making and Branching, Looping and Jumping Statements.

Unit-II

Arrays: Declaration, Definition, One and Two Dimensional Arrays, Character Arrays and Strings. **Functions:** Definition, Need for Functions, Standard and User-Defined Functions, Function Calls,Category of Functions, Recursion, Storage Class Specifiers

Unit-III

Structuresand Unions: Arrayof Structures,Nesting of Structures.**Pointers:**Declaration,DefinitionandUseof Pointers, Differencebetween Pointersand Structures.

Unit-IV

File Management in C:Creating, Opening and Closing a File, I/O Operations on Files, Command Line Arguments. Dynamic Memory Allocation, Pre-Processor.

Recommended Books:

1. Kerighan& RichietheC programminglanguage (PHI Publication)
2. ByronGottorfriedSchaum'soutlineofprogrammingwithC
3. E.BalaguruswamyProgramminginAnsi'C'(TataMcGrawHill)
4. Kanetkar"LetUsC",BPBPublications.(TataMcGrawHill)

MSAIDS102:DatabaseManagementSystem

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks: 40*

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1: Gain a solid understanding of the fundamental concepts of database management systems, including data models, data independence, database architecture, and components.
CO2: Learn the principles and techniques involved in relational database management systems.
CO3: Ability to transform user requirements into efficient and well-structured database schemas.
CO4: Develop proficiency in SQL

Unit-I

Introduction to Database Concepts: Database and Need for DBMS ,Characteristics of DBMS, Database Users, 3-tier architecture,(its advantages over 2- tier) Data Models, Views of data-schemes and instances, Independence, Data modeling using the Entity-Relationship approach, Entities, Relationships, Representation of entities, attributes, relationship set, Generalization , Aggregation.

Unit-II

Relational model: Overview of database models, Relational Model, Structure of relational database, different types of keys, Expressing M: N relation, relational algebra, Constraints.
Relational Database Design: Functional dependencies, Normalization Normal forms based on primary keys (1 NF, 2 NF, 3 NF, BCNF, 4 NF, 5 NF) Lossless joins and dependency preserving decomposition.

Unit-III

SQL: Introduction to SQL, SQL Data types and Literals, Types of SQL commands, SQL Operators and their procedures, Tables, Views and Indexes, queries and sub-queries, aggregate function, insert, delete and update operations, Joins, Unions, Intersections, Minus in SQL.
Query Processing: Query Processing Stages, Estimation of Query Processing Cost, Introduction to Transactions, States of Transactions, ACID Properties

Unit-IV

Concurrency Control: Concurrency Control, Lock Based Protocols, Two Phase Locking protocol, Timestamp-based Protocol, multiple granularity, Deadlock Handling, Deadlock Prevention,Deadlock Detection and Recovery.
Recovery: Failure classification, recovery concepts, database backup, recovery concepts based on deferred update and on immediate update. Shadow paging, check points, an overview of Distributed database and client-server Database

Recommended Books:

1. FundamentalofDatabaseSystemsby R.Elmasri,S.Navate;Benjamin Cummings
2. Introductiontodatabasesystems byC.J .Date
3. Databasesystemconcept by Korth
4. PrinciplesofDatabaseManagementbyJamesMartin
5. RelationaldatabasedesignforMicro computersApplicationbyPrenticeHall(Jackson)
6. DatabaseManagementSystemsbyBipin Desai

MSAIDS103:Data Communication and Networking

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks: 40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max. Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1:Abletounderstandnetworkcommunicationusingthelayeredconcept,OpenSystem Interconnect (OSI) and TCP/IP Model.
CO2:Understandvarioustypesoftransmissionmedia,networkdevices;andparametersof evaluation of performance for each media and device.
CO3:Understandtheconceptofflowcontrol,errorcontrolandLANprotocols
CO4:UnderstandtheworkingprinciplesofLANandtheconceptsbehindphysicalandlogical addressing, sub netting and supernetting.
CO5:UnderstandthefunctionsperformedbyaNetworkManagementSystemandtoanalyze connection establishment and congestion control with respect to TCP Protocol.

Unit-I

Data Communication Concepts: Introduction, Communication System, And Communicationmode, Data encoding: Analog and Digital data, digital and analog signal, Communication Channels, Synchronous and asynchronous transmission. Bandwidth concepts, channel capacity.
Introduction to Networking: Computer network, Characteristic & advantages of networking, types of network, LAN, MAN, WAN.

Unit-II

Transmission media & Network Topologies: Guided & Unguided media, Twisted pair, coaxial cable, Fiber optics, Radio. VHF and microwaves, Satellite link. Network topology, bus, star, ring, tree, mesh & hybrid topology. Advantages and disadvantages of these topology. Multiplexing Channels and Concept of multi channeling and modulation, pulse code modulation (PCM)Frequency Division multiplexing, Time Division multiplexing, CODECS.

Unit-III

Digital Data Communication Techniques: Asynchronous and Synchronous Transmission - Types of Errors, Error Detection and error Correction methodologies.
Network Standards: Introduction, Layered approach, OSI model functions & responsibilities of each layer.

Unit-IV

Internetworking: Principles of internetworking, Connectivity Devices, Switches, Bridges, Routers, Routing with bridges. Internet and e-mail protocols: SMTP, SLIP, POP, PPP, FTP, HTTP, Wi-Fi Network, Bluetooth, Broadband.

Recommended Books:

1. Introductionto Digitaland Data Communications,MichalAMiller, JAICO pub.
2. DataandComputerCommunication–WillamStaling,PHIpub.
3. DataCommunication&Network –Forouzan (TMH)
4. ComputerNetworks – A.Tanenbaum,(PHI pub.)
5. InternetworkingwithTCP/IPVol-I–Comer(PHIpub.)
6. DataCommunications anddistributed Networks-V.B,Black, (PrenticeHall pub.)

MSAIDS104:Computer Architecture

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1: Gain a comprehensive understanding of the components of a digital computer system CO2: Proficient in converting data between different digital number systems (binary, octal, decimal, hexadecimal)
CO3: Design and analyze combinational circuits, such as multiplexers, demultiplexers, encoders, and decoders
CO4: Familiar with the anatomy of a computer system, including components like motherboards, microprocessors (CISC and RISC), different types of memory modules (RAM, Cache), and expansion slots
CO5: Develop a strong grasp of input and output devices, including printers, plotters, and voice output devices.

Unit -I

Introduction to DigitalComputer:Overview ofthe DigitalComputer System(Processor, Memory, Input and Output Devices, Storage Devices, Representation of Data: Digital versus Analog, Digital number system (binary, octal, decimal and hexadecimal numbers,) conversion from one form to another, fractional numbers and signed numbers, 1'S, 2'S Complements, Fixed point and floating point representations, Boolean algebra (addition, subtraction, multiplication and division), Logic Gates(NOT,OR,AND,NAND,NOR,XOR,XNOR)typesCodes(ASCII,EBCDIC,Unicode) Combinational Circuits(Multiplexer, Demultiplexer, coder, Encoder)Sequential Circuits(Flip flops, registers, Counters).

Unit-II

Anatomy of a Computer: Mother Board (Special reference to Intel 810 Chipset motherboard),CISC Micro Processors (Special reference to Pentium, AMD, Cyrix), RISC, types of RAM, Flash, Cache, types of memory modules (SIMM, DIMM), System clock, Bus (Data, Address, Control), Expansion slots (ISA, MCA, EISA, PCI, AGP).

Unit-III

I\O and Storage Devices: Input devices, Output devices, Printers (Dot-Matrix, Line, Label, Ink-Jet, Laser, Color Laser, thermal wax, dye sublimation, fiely, IRIS), Plotters (Pen, Ink-jet, electrostatic), Voice output], Storage devices I Storage types (Magnetic, Optical, Magneto-optical, Solid state), randomversussequentialaccess,formatting,tracksandsectors,speed,storagecapacity,FloppyDisk (5.25 inch, 3.5 inch; 2HD, zip, Superdisk, HiFD) Hard Disk (tracks, cylinders, sectors; Hard Drive Interfaces (IDE, EIDE, Fast SCSI, Fast/wide SCSI, Uitra SCSI; Hard Disk Cartridges, RAID), Optical Disks [pits and lands, CD-ROM, R, RW, DVD-ROM, R, RAM)], Magnetic tape (reels, streamers, DAT,DLT, stripe, Smart card), Modem (Fax/Data/Voice).

Unit-IV

Computer Memory Systems: Architecture of Digital Computer, Processor Design Principles, Control Unit Design: Conventional and Micro programmed, Input- Output System. Memory and I/O Organization: Interfacing with CPU; Main Memory, Auxiliary Memory, Cache Memories, Associative Memory and Virtual Memory. I/O Interfacing with CPU, Addressing Data Transfer Techniques.

Recommended Books:

1. ComputerArchitectureAndOrganization:McgrawHill,2ndEdition,JohnHyaes.
2. ComputerSystemArchitecture:PHI,3rdEdition,M.MorriesMano.
3. ComputerOrganizationAndDesign:PrenticeHallOfIndia,ChaudhariP.P.
4. PerspectiveInComputerArchitecture:PrenticeHallOfIndia,RaoP.V.S.
5. ComputerSystemArchitecture:PrenticeHall,TannenbaumA.
6. ParallelComputerArchitecture:AHardware/SoftwareApproachbyDavid Culler

MSAIDS105: Operating System

Max Marks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1:Analyzevariousschedulingalgorithms.
CO2: Understand deadlock, prevention and avoidance algorithms.
CO3:Compareandcontrastvariousmemorymanagementschemes.
CO4:Understandthefunctionalityoffilesystems.
CO5:UnderstandtheOpensource operating system and basic Linux commands

Unit-I

Introduction to Operating System: Types of software (System/Application), Translators (Assemblers/ Interpreters/ Compilers), Operating system as a resource manager, Operating system need and services, Classification and Evolution of OS, Hierarchical/Layered Organization of OS.

Unit-II

Process Management: Process concept, Process Control Block, Process Life cycle, Type of Scheduler, Scheduling criteria, multiple processor scheduling, scheduling Algorithm, FCFS, SJF, Priority and round robin scheduling, critical section, semaphores. Asynchronous parallel process, multithreading at system/user level, Inter process communication, Process Synchronization & Deadlock,Monitors, Deadlockprevention&avoidance, DeadlockDetectionanddeadlock Recovery.

Unit-III

Memory Management: Memory Management Techniques; Single partition allocation, multiple partition allocation, Swapping, paging and segmentation, segmented-paged memory management techniques; logical and physical address space; address mapping. Demand paging, Virtual memory, protection and address mapping hardware, page fault, Page replacement and page removal algorithms.

Unit-IV

Device Management and I/O Programming: Disk structure, disk scheduling, access method and storage capacity; sharable and non sharable devices and their management.
Information Management & File System: File organization and access methods, logical and physical filestructure; physical file systemrealized with device management function; file allocation methods, linked and index allocation, logical file implemented on physical file system. File protection and security, Directory structure, single level, two level, tree structure, Free Space Management, Allocation Methods.

Recommended Books:

1. JamesL.Peterson&A.Silberschatz:OperatingSystemConcepts;2ndEdn., Addison Wesley, World Student Edition
2. AndrewS.Tenenbaum:ModernOperatingSystems;PrenticeHall,India
3. DietelH.M.:AnIntroductionToOperatingSystems;AddisonWesley,World Student Edition
4. SystemsProgramming &OperatingSystems,2ndEdn., TataMcGrawHill

MSAIDS151:'C'ProgrammingLab

Examination:Practical

Exercisesto be framedso asto coverthe topicsandtools coveredintheory paper.

Lab Course Outcomes:

CO1: Develop C programs based on a given task or algorithm.

CO2:Read,comprehend,andtracetheexecutionofCprograms.

CO3:ImplementCprogramsusingarrays,pointers,decision-makingstatements,andlooping statements.

CO4:Writeprogramsthatperformoperationsutilizingderiveddatatypes.

CO5:DevelopandimplementmodularapplicationsinCbyeffectivelyutilizingfunctions.

CO6: Develop applications in C that leverage structures and pointers.

MSAIDS152:DBMS Lab

Course Outcomes:

CO1:WritebothsimpleandcomplexSQLqueriestoretrieveinformationfromdatabaseswith many tables to support business decision making.

CO2:WriteSQLDDLtocreate,modifyanddropobjectswithinarelationaldatabase.

CO3:RetrieveandstoreinformationinarelationaldatabaseusingSQLinamulti-user environment.

Examination:Practical

Exercisesto beframed soas to coverthe topics andtools coveredintheory paper.

MSAIDS153:Office ManagementLab

Course Outcomes:

CO1: Able to use MS Office (word processor, spreadsheet and power point) professionally.

CO2: Develop understanding about the writing of effective business letters in computer through word processing.

CO3:Abletousespreadsheet programforbusinessdataprocessing.

CO4:Acquireskillsfordevelopmentandpresentationofpowerpointreport**.**

Examination:Practical

Exercisesto beframed so astocoverthe topicsandtools coveredintheory paper.

M.Sc.IT (AI and Data Science)–Second Semester

MSAIDS201:ObjectOrientedProgramming ConceptsusingJava Programming

MaxMarks:100(ESE: 70 CIA: 30) Passing marks: 40

QuestionPaperpatternforEndSemesterExam (ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1:Understandtheconceptsof OOPs
CO2:KnowledgeofbasicprogrammingconstructsofJava
CO3:UnderstandandImplementtheconceptsofClassesandObjects
CO4: Implement Polymorphism, Inheritance, and Multithreading
CO5:EnlightentheuseofInterfaces,PackagesandExceptionHandling
CO6: Perform Database Connectivity through JDBC
CO7:UnderstandString&MathClassesandtheirfunctions

Unit-I

Introduction: ObjectOrientedConcepts,IntroductiontoJava,FeaturesofJava,Javavirtual machine.
ProgrammingFundamentalsofJava: JavaTokens,Operators,Datatypes,ControlStructures, Array

Unit-II

ClassesandObjects: ObjectReferences,MethodOverloading,Constructors,Constructor Overloading, Inheritance. Static, Final, Abstract Keywords
PackagesandInterfaces: AccessSpecifiers,PackageCreation,Useofpackages,Basicsof Interfaces, Use of Interfaces
StringHandling: StringClass,Stringmethods.

Unit-III

ExceptionhandlinginJava: Exceptionclasses,Exceptionraising&handling,useofTry,Catch and finally, Throwing Exceptions, User Defined Exceptions.
Applets: Introduction toApplets, Appletlife cycle,CreatingApplets,AppletViewer
Graphics: Graphics Components, Color,Font,Drawing Objects

Unit-IV

ThreadinginJava: ConceptsofMulti-threading,ThreadLifeCycle,Threadclass,RunnableInterface, Thread Management in Java
JDBC: JDBC Drivers, Two Tier and Three Tier client server Architecture, Setting up a connectionto database, Creating and executing SQL statements

Reference Books:

1. C.Thomas Wu, AnIntroductiontoOOPwithJava, McGraw Hill.
2. Deiteland Deitel, Java,Howto Program,Pearson EducationAsia.
3. E.Balaguruswmy,Programming withJava,TataMcGrawHill.
4. Zukowski:Mastering Java2, BPB Publications.
5. HerbertScliildt, Java 2,Tata Mc GrawHill.

MSAIDS202:Data Structure andAlgorithms

MaxMarks:100(ESE: 70 CIA: 30) Passing marks:40

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions (1fromeachunitwith internalchoice)of10 marks.

Course Outcomes:

CO1:Analyzetheconceptsofalgorithmevaluationandfindtimeandspacecomplexities for searching and sorting algorithms.
CO2:Implementlineardatastructuresuchasstacks,queues,linkedlistsandtheirapplications.
CO3:Implement basic operations on binary trees
CO4:Demonstratetherepresentationandtraversaltechniquesofgraphsandtheir applications

Unit-I

Introduction, Types of Data Structures, Algorithm, Pseudocode, Characteristics of Algorithms, Algorithm Analysis, Algorithm Complexity – Space Complexity, Time Complexity, Abstract Data Types.

StringProcessing–BasicTerminology,StoringStrings,StringOperations,PatternMatching Algorithms.

Arrays:RepresentationinMemory,Operations-Insertion,Deletion,Searching–LinearSearch, Binary Search, Sorting – Bubble Sort, Multidimensional Arrays, Pointers, Pointer Arrays.

Unit-II

Sorting: Internal and External Sorting Techniques, Bubble Sort, Insertion Sort, Selection Sort,Merge Sort, Quick Sort, Radix Sort.

Linked lists: Linked List Vs. Arrays, Representation in Memory, Types of linked List - Singly Linked list, Doubly Linked list, Circular Linked list, Doubly Circular Linked list., Operations on Singly Link List – Insertion, Deletion, Traverse, Searching.

Unit-III

Stacks: Array and Linked List representation of Stacks, Operations – Insertion, Deletion, Traverse, Application of Stack – Recursion, Polish Notation.

Queues: Array and linked list representation in Memory, Type of Queues – Simple Queue, Circular Queue, Priority Queue, Double Ended Queue. Operations on Simple and Circular Queue – Insertion, Deletion, Traverse. Applications of Queues.

Unit-IV

Trees:BasicTreeConcepts,RepresentationofBinaryTreeinmemory,BinaryTreeTraversals, Binary Search Trees, Heapsort.

Graphs: Representations,sequentialrepresentation, Warshall'sAlgorithm,Linked Representationof graphs, Operations on Graph, Traversing Graph.

Reference Books:

1. S.Lioschutz: DataStructures,McGraw HillInternational Edition.
2. A.V.Aho.,J.E.Hopcroft,andJ.D.Ullman,DataStructuresandAlgorithms,PearsonEducation Asia.
3. A.MichaelBerman: DataStructuresvia C++,OxfordUniversity Press.
4. SaraBaaseandAllenVanGelder:Computer Algorithms,PearsonEducation Asia.
5. Jean-PaulTremblayandPaulG.Sorenson,AnIntroductiontoDatastructureswithapplications, TMH Publishing Co. Ltd.

MSAIDS203:Web Designing andDevelopment

MaxMarks:100(ESE: 70 CIA: 30) Passing marks:40

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

Course Outcomes:

CO1:UnderstandthefundamentalsofInternet,andtheprinciplesofwebdesign.
CO2: Able to construct websites using HTML and Cascading Style Sheets.
CO3: Able to build dynamic web pages with validation usingJava Script objects
CO4: Implement the event handling mechanisms.

Unit-I

The Internet: History of the World Wide Web, Hardware and Software Trend, Web Server, Web Client.

Creating Web Pages: Introduction to HTML, Types of tags, Formatting Tags, Common Tags, Hyperlinking, Images, Image Maps, Marquee Tag, Horizontal Rules, Lists.

Unit-II

AdvanceConcepts ofHTML:HTMLTables,HTMLForms,Frames

Dynamic HTML: CSS : introduction, inline styles, creating style sheets with the style element, conflicting styles, linking external style sheets, positioning elements, backgrounds, element dimensions, text flow and the box model, user style sheets, Filters and Transitions.

Unit-III

JavaScript: Introductiontoscripting language,ClientsideversusServersidescripting ,Advantages of Java Script, Features of Java Script, Keywords, Variables, Data Types, Constants, Comments,Java script control structures, Arrays, Array Library Methods, Java script Methods

Unit-IV

DOM:Introduction, Window,History,Navigator,Form,Frames,LocationObjects.

Event Handling in Java Script: Different types of events, key events, mouse events, loadingevents, Event Handlers,

Recommended Books:

1. M.L.Young: CompleteReferenceb:Internet;2nd Edition;Tata McGrawHill,2006.
2. ThomasA. Powel; WebDesign :C.R.; SecondEdition;TMH, 2009.
3. ThomasA. Powel; HTML& XHTML:C.R.; FourthEdition; TMH, 2008.
4. HarelyHahn:The Internet,TataMcGraw Hill.
5. G.Robertson: HandsonHTML, BPB Publications.
6. D.A.Tauber, B.Kienan: MicrosoftFromPage; BPB Publications.
7. PHPThe CompleteReference, StevenHolzer,TataMcGrawHill

MSAIDS204:Fundamental of AI, Machine Learning and Data Science

MaxMarks:100 (ESE: 70 CIA: 30) Passing marks:40

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwill contain 4 questions(1 from eachunitwith internalchoice) of10 marks.

COURSE OUTCOMES:

CO1: Understand the concept and significance of Artificial Intelligence (AI) and its various areas of application.

CO2: Learn different search techniques, such as depth-first search and breadth-first search, and their role in problem-solving.

CO3: Understand heuristic search methods and their applications, including hill climbing, bestfirst method, and graph search.

CO4: Able to gain insight view of knowledge representation schemes, including semantic nets, frames, conceptual dependency, and scripts.

CO5: Understand different types of reasoning, such as logical reasoning, statistical reasoning, and fuzzy logic.

CO6: Learn the concept of learning in AI, knowledge acquisition methods, and the stages involved in developing expert systems.

Unit – I

Concept of intelligence, Artificial intelligence, definition turning test, areas of application. Search techniques, state space, Production rules, problem characteristics. Production systemcharacteristic, depth first, breadth first search methods.

Unit – II

Heuristic search method, generate and test, hill climbing, best first method, graph search, AND OR search methods, constraint satisfaction, backtracking. Introduction to list and string processing, concept of knowledge, Logic, prepositional and predicate calculus, resolution.

Unit –III

Fundamentals of Machine Learning : Introduction to Machine Learning and its importance, Types of Machine Learning: Supervised, Unsupervised, and Reinforcement Learning, Basic concepts: features, labels, training data, etc., Popular Machine Learning algorithms: Linear Regression, Logistic Regression, Decision Trees, k-Nearest Neighbors, etc., Evaluation metrics for Machine Learning models: accuracy, precision, recall, F1-score, etc.

Unit –IV

Introduction to Data Science : What is Data Science and why it is important?, Role of Data Scientist and skills required, Data acquisition: sources of data, data formats, data cleaning, etc., Exploratory Data Analysis (EDA): statistical analysis, data visualization techniques, Introduction to libraries/tools: NumPy, Pandas, Matplotlib, Seaborn, etc

Reference Books:

1. "Artificial Intelligence: A Modern Approach" by Stuart Russell and Peter Norvig
2. "Introduction to Machine Learning" by EthemAlpaydin
3. "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking" by Foster Provost and Tom Fawcett

MSAIDS205:Big Data Analysis

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEnd SemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

COURSE OUTCOMES:

The students will be able to:

- Identify Big Data and its Business Implications.
- List the components of Hadoop and Hadoop Eco-System
- Access and Process Data on Distributed File System
- Manage Job Execution in Hadoop Environment
- Develop Big Data Solutions using Hadoop Eco System
- Analyze Infosphere BigInsights Big Data Recommendations.
- Apply Machine Learning Techniques using R.

Unit-I

Understanding Big Data :Introduction, Need, convergence of key trends, structured data Vs. unstructured data, application, web analytics- big data and marketing, fraud and big data risk, big data and algorithmic trading, big data and its applications in healthcare, advertising etc. Mining Data Streams: - Introduction To Streams Concepts - Stream Data Model and Architecture - Stream Computing - Sampling Data in a Real Time Sentiment Analysis- Stock Market Predictions.

Unit-II

Big Data Technologies: Hadoop: Open source technologies, cloud and big data, Crowd Sourcing Analytics, inter and trans firewall analytics. Introduction to Hadoop: Introduction, Data format, analyzing data with Hadoop, scaling out, Hadoop streaming, Hadoop pipes. Design of Hadoop distributed file system (HDFS), HDFS concepts - Java interface, data flow, Data Ingest with Flume and Sqoop. Hadoop I/O - data integrity, compression, serialization, Avro - file-based data structures.

Unit-III

Hadoop Related Tools: Introduction to Hbase: The Dawn of Big Data, the Problem with Relational Database Systems. Introduction to Cassandra: Introduction to Pig, Hive - data types and file formats - HiveQL data definition - HiveQL data manipulation - HiveQL queries. NOSQL Data Management: Introduction to NoSQL, aggregate data models, key-value and document data models, relationships, graph databases, schemaless databases,

Unit-IV

Predictive Analytics: Simple linear regression- Multiple linear regression- Interpretation 5 of regression coefficients. Visualizations - Visual data analysis techniques- interaction techniques - Systems and applications. Map Reduce Applications: MapReduce workflows, unit tests with MRUnit, test data and local tests, anatomy of MapReduce job run, classic Map-reduce - YARN, failures in classic Map-reduce and YARN, MapReduce Calculations.

Recommended Text / Reference Books:

1. Big Data, Black Book, DT Editorial Services, Dreamtech Press 2015
2. Professional NOSQL, Shashank Tiwari, Wrox, September 2011
3. Hadoop in Practice, Alex Homes, Dreamtech Press, 2015
4. HBase: The Definitive Guide, Lars George, O'Reilley, 2011.

5. Cassandra: The Definitive Guide, Eben Hewitt, O'Reilley, 2010.
6. Programming Pig, Alan Gates, O'Reilley, 2011.
7. NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, P. J. Sadalage and M. Fowler, Pearson Education, Inc. 2012.
8. Programming Hive, E. Capriolo, D. Wampler, and J. Rutherglen, O'Reilley, 2012
9. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
10. Tom White "Hadoop: The Definitive Guide" Third Edition, O'reilly Media, 2012

MSAIDS251:Java Programming Lab

CO1:AbletoimplementthebasicconceptssuchasFunctionOverloading,arrayandstringmanipulation in Java
CO2:Useutilityclassesintherealtimeapplications
CO3: Understand the types of inheritance
CO4:Implementpackages,manipulatethreadsandexceptionhandlingtechniques
CO5: Connect databases with Java programs

Examination: Practical
Exercisesto beframed soas to coverthe topics andtools coveredintheory paper

MSAIDS252:DataStructure Lab

CO1:Understandtheconceptofdatastructures,andapplyalgorithmforsolvingproblemslikeSorting, searching, insertion and deletion of data through C language.
CO2:Understandlineardatastructuresforprocessingoforderedorunordereddata.
CO3:ImplementvariousoperationsinCprogramondynamicdatastructureslikesinglelinkedlist, circular linked list and doubly linked list.
CO4:Exploretheconceptofnonlineardatastructuressuchastreesandgraphsthrough Cprogramming.

Examination:Practical
Exercisesto beframed soas to coverthe topics andtools coveredintheory paper

MSAIDS**253:WebDesigningLab**

CO1:CreatingwebpagesusingbasicHTMLtags
CO2: Styling Webpages using CSS
CO3:CreatingdynamicwebpagesusingJavascript
CO4: Implementation of DOM objects
CO5:Creatingwebpagesimplementingeventhandling,formValidationetc.

Examination:Practical
Exercisesto beframed soas to coverthe topics andtools coveredintheory paper

M.Sc. IT (AI and Data Science)–Third Semester

M.Sc.IT (AI and Data Science) - ThirdSemester

MSAIDS301: Data Science with R

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of 10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Demonstrate proficiency in the R programming environment by using data types, operators, control structures, functions, vectors, matrices, lists, and data frames for solving data analysis problems.

CO2: Import, clean, transform, and manage structured datasets using R packages and perform exploratory data analysis (EDA) to identify patterns, trends, and anomalies.

CO3: Apply descriptive and inferential statistical techniques in R to summarize data, test hypotheses, and interpret statistical results for decision-making.

CO4: Create effective data visualizations using R (such as **ggplot2**) and communicate analytical insights through charts, graphs, and dashboards.

CO5: Build, evaluate, and interpret basic predictive models using machine learning techniques in R, and apply the complete data science workflow to solve real-world problems.

Unit-I

Introduction to Data Science: What is Data Science, Need for Data Science, Components of Data Science, Big data, Facets of data: Structured data, Unstructured data, Natural Language, Machinegenerated data, Graph-based or network data, Audio, image and video, Streaming data, The need for Business Analytics, Data Science Life Cycle, Applications of data science.

Unit-II

Data Science Process: Overview of data science process, setting the research goal, Retrieving data, Cleansing, integrating and transforming data, Exploratory data analysis, Data Modeling, Presentation and automation, Types of Analytics: Descriptive analytics, Diagnostic analytics, Predictive analytics, Prescriptive analytics.

Unit-III

Statistics: Basic terminologies, Population, Sample, Parameter, Estimate, Estimator, Sampling distribution, Standard Error, Properties of Good Estimator, Measures of Centers, Measures of Spread, Probability, Normal Distribution, Binary Distribution, Hypothesis Testing,Chi-Square Test, ANOVA.

Unit-IV

Data Science Tools and Algorithms: Basic Data Science languages- R, Python, Knowledge of Excel, SQL Database, Introduction to Weka, Regression Algorithms: How Regression Algorithm Work, Linear Regression, Logistic Regression, K-

Nearest Neighbors Algorithm, K-means algorithm.

Recommended Books:

1. Samuel Burns, "Fundamentals of Data Science: Take the first Step to Become a Data Scientist", Amazon KDP Printing and Publishing, First Edition, 2019

2. Davy Cielen, Arno D.B. Meysman, Mohamed Ali, "Introducing Data Science", Manning Publications, 2016

3. Cathy O'Neil and Rachel Schutt, "Doing Data Science, Straight Talk From The Frontline", O'Reilly. 2014

MSAIDS302:Python with AI and Data Science

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam (ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Demonstrate proficiency in Python programming by using variables, data types, operators, control structures, functions, modules, file handling, and object-oriented programming concepts to solve computational problems.

CO2: Perform data collection, preprocessing, manipulation, and exploratory data analysis (EDA) using Python libraries such as NumPy and Pandas.

CO3: Develop effective data visualizations using Python libraries such as Matplotlib and Seaborn to analyze and communicate data-driven insights.

CO4: Apply fundamental Artificial Intelligence and Machine Learning techniques using Python libraries to build, train, and evaluate predictive models for real-world datasets.

CO5: Design and implement end-to-end AI and Data Science solutions using Python by integrating data preprocessing, model development, evaluation, and result interpretation while considering ethical and responsible AI practices.

Unit I

Introduction to Python Programming: Introduction to programming concepts, features and applications of Python, installation of Python and Integrated Development Environment (IDLE/Jupyter Notebook/Visual Studio Code), writing and executing Python programs, variables, identifiers, keywords, data types, input and output functions, operators and expressions, type conversion, comments and indentation. Decision-making statements including if, if-else, nested if, and match-case statement, looping constructs using for and while loops, use of break, continue and pass statements, and development of simple Python programs.

Unit II

Python Data Structures and Functions: Strings and string manipulation, lists, tuples, dictionaries and sets, basic operations on Python collections, user-defined functions, function arguments and return values, lambda functions (introduction), modules and packages, importing modules, exception handling using try, except and finally blocks, file handling for reading and writing text files, and development of menu-driven and real-life problem-solving programs.

Unit III

Python for Data Science: Introduction to Data Science, Data Science life cycle and its applications, introduction to NumPy library, creation and manipulation of arrays, introduction to Pandas, Series and DataFrames, importing data from CSV files, data cleaning, filtering, sorting and basic statistical analysis, introduction to data visualization using Matplotlib including line charts, bar charts, pie charts and histograms, and simple data analysis using real-world datasets.

Unit IV

Introduction to Artificial Intelligence using Python: Introduction to Artificial Intelligence, Artificial Intelligence, Machine Learning and Deep Learning concepts, applications of AI, ethical use of Artificial Intelligence, types of Machine Learning including supervised and unsupervised learning, introduction to model training, features and labels, introduction to Scikit-learn library, loading datasets, data splitting, basic concepts of linear regression and classification, model evaluation using accuracy, introduction to Generative AI tools, responsible use of AI applications, and development of simple Artificial Intelligence and Data Science applications using Python.

Reference Books:

1) "Python for Data Analysis" by Wes McKinney
2) "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by AurélienGéron
3) "Pattern Recognition and Machine Learning" by Christopher M. Bishop
4) "Deep Learning" by Ian Goodfellow, YoshuaBengio, and Aaron Courville

MSAIDS303:Data Visualization Tools and Techniques

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks: 40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Understand the principles, concepts, and importance of data visualization and identify appropriate visualization techniques for different types of data.

CO2: Collect, preprocess, and transform data for visualization using suitable tools and techniques to ensure data quality and accuracy.

CO3: Create effective charts, graphs, dashboards, and interactive visualizations using modern data visualization tools to represent complex datasets clearly and efficiently.

CO4: Analyze and interpret visualized data to identify patterns, trends, correlations, and insights that support data-driven decision-making.

CO5: Design and develop comprehensive visualization solutions and dashboards by applying best practices in visual analytics, storytelling, and user-centric design.

UNIT-I

Introduction to Data Visualization : Importance of Data Visualization Principles of Data Visualization Types of Data Visualization Data Representation Techniques Bar Charts and Histograms Line Charts and Area Charts Scatter Plots and Bubble Charts Pie Charts and Donut Charts.

UNIT-II

Data Visualization Tools : Introduction to Tableau, Introduction to Power BI and Python libraries (Matplotlib, Seaborn, **Plotly**) Introduction to R libraries (ggplot2, plotly (R))

UNIT-III

Interactive Data Visualization : Adding Interactivity to Visualizations, Tool-specific Interactivity Features, Best Practices for Interactive Visualizations

UNIT-IV

Design Principles : Color Theory in Data Visualization, Typography and Layout in Visualizations, Gestalt Principles in Visualization Design.

ReferenceBooks:

1. Data Points: Author: Nathan Yau, Publisher: Wiley, Year: 2013.
2. Information Dashboard Design: Author: Stephen Few, Publisher: O'Reilly Media, Year: 2013
3. The Visual Display of Quantitative Information: Author: Edward Tufte, Publisher: Graphics Press, Year: 2001.
4. Storytelling with Data: Author: Cole NussbaumerKnaflic, Publisher: Wiley, Year: 2015

MSAIDS304:Machine Learning

MaxMarks:100(ESE:70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks :

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Explain the fundamental concepts, types, and applications of Machine Learning, including data preprocessing, feature engineering, and model selection for solving real-world problems.

CO2: Apply supervised learning algorithms such as linear regression, logistic regression, decision trees, k-nearest neighbors, and support vector machines to develop predictive models.

CO3: Implement unsupervised learning techniques such as clustering and dimensionality reduction to discover patterns and extract meaningful insights from datasets.

CO4: Evaluate and optimize Machine Learning models using appropriate performance metrics, validation techniques, hyperparameter tuning, and methods to address overfitting and underfitting.

CO5: Design and implement end-to-end Machine Learning solutions using suitable tools and libraries, interpret model results, and demonstrate awareness of ethical, privacy, and responsible AI considerations.

Unit-I

Introduction :Machine Learning, Machine Learning Foundations-Overview, Applications, Types of Machine Learning, Basic Concepts in Machine Learning - Examples of Machine Learning. Perspectives/Issues in Machine Learning, Designing a learning system. Concept Learning: Introduction, a concept learning task, concept learning as search, Find-S algorithm, Version space and Candidate-Elimination algorithm,, Inductive bias

Unit-II

Supervised Learning :Introduction, Linear Models of Classification - Decision Trees, Naïve Bayes Classification, Linear Regression - Logistic Regression, Bayesian Logistic Regression, Probabilistic Models, Artificial Neural Network- perceptron, multilayer networks and back propagation algorithm, Ensemble Methods - Random Forest - Bagging - Boosting. Evaluating Hypothesis: estimating hypothesis accuracy, basics of sampling theory, comparing learning algorithms.

Unit-III

Unsupervised Learning :Clustering, K-Means Clustering, EM (Expectation Maximization). Mixtures of Gaussians, EM algorithm in General, The Curse of Dimensionality, Dimensionality Reduction, Factor Analysis, Principal Component Analysis, Probabilistic PCA, Independent Component Analysis. Challenges for Big Data Analytics.

Unit-IV

Instance based Learning-Nearest neighbor classification, k-nearest neighbor, locally weighted regression, lazy and eager learner Reinforcement Learning-Introduction, Elements of Reinforcement Learning, Difference between Reinforcement Learning and Supervised Learning, Applications of Reinforcement Learning, The Learning Task, Q learning, Nondeterministic rewards and actions, Temporal difference learning. Model based learning, Semi-Supervised Learning, Computational Learning Theory.

Recommended Text/ Reference Books:

1. Christopher Bishop, "Pattern Recognition and Machine Learning", Springer 2006

2. EthemAlpaydin, "Introduction to Machine Learning", Prentice Hall of India, 2005

3. Joel Grus, "Data Science from Scratch- First Principles with Python", O'Reilly, 2015

4. Tom Mitchell, " Machine Learning", McGraw-Hill, 1997

5. Stephen MarsLand, "Machine Learning-An Algorithmic Perspective", CRC Press, 2009

6. Kevin P. Murphy, "Machine Learning: A Probabilistic Perspective", MIT Press, 2012

7. M. Gopal, "Applied MACHINE LEARNING", McGraw-Hill, 2018

8. Mark Summerfield, "Programming in Python 3: A Complete Introduction to the Python Language", Addison Wesley, 2010

MSAIDS305:Cloud Computing

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions(1fromeachunitwithinternalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Explain the fundamental concepts, architecture, characteristics, deployment models, and service models (IaaS, PaaS, and SaaS) of Cloud Computing and their applications.

CO2: Demonstrate the use of cloud platforms and virtualization technologies to provision, configure, and manage cloud resources for computing, storage, and networking.

CO3: Design and deploy cloud-based applications and services using appropriate cloud technologies, virtualization, and containerization tools to meet organizational requirements.

CO4: Analyze cloud security, privacy, governance, disaster recovery, and performance issues, and apply suitable techniques to ensure secure and reliable cloud environments.

CO5: Evaluate and implement cloud computing solutions for real-world applications by considering scalability, cost optimization, resource management, and industry best practices.

Unit-I

Introduction of Cloud Computing: Nutshell of cloud computing, Enabling Technology, Vision, feature Characteristics and components of Cloud Computing. Challenges, Risks and Approaches of Migration into Cloud., Layer and Types of Clouds, Services models, Cloud Reference Model.

Unit-II

Cloud Computing Architecture: Data center Design and interconnection Network, Architectural design of Compute and Storage Clouds. Cloud Programming and Software: Features of cloud programming, Parallel and distributed programming paradigms-MapReduce, Hadoop, High level Language for Cloud. Service Oriented Architecture - REST and Systems of Systems - Web Services - Publish Subscribe Model

Unit-III

Virtualization Technology: Definition, Understanding and Benefits of Virtualization. Implementation Level of Virtualization, Virtualization Structure/Tools and Mechanisms, Hypervisor VMware, KVM, Xen. Virtualization: of CPU, Memory, I/O Devices, Virtual Cluster and Resources Management, Virtualization of Server, Desktop, Network, and Virtualization of data-center.

Unit-IV

Securing the Cloud: Cloud Information security fundamentals, Cloud security services, Design principles, Policy Implementation, Cloud Computing Security Challenges, Cloud Computing Security Architecture. Legal issues in cloud Computing. Data Security in Cloud: Risk Mitigation, Understanding and Identification of Threats in Cloud, SLA-Service Level Agreements, Trust Management Cloud Platforms in Industry: Amazon web services, Google AppEngine, Microsoft Azure Design, Aneka: Cloud Application Platform -Integration of Private and Public Clouds Cloud applications:
Protein structure prediction, Data Analysis, Satellite Image Processing, CRM and ERP, Social networking. Cloud Application- Scientific Application, Business Application.

Recommended Text / Reference Books:

1. Cloud Computing,Principle and Paradigms, Edited By RajkumarBuyya, JemesBroberg, Goscinski, Pub.-Wiley-2016
2. Kumar Saurabh, "Cloud Computing", Wiley Pub 2016 A.
3. Distributed and Cloud Computing, Kai Hawang, GeofreyC.Fox, Jack J. Dongarra Pub: Elservier, 2013
4. Krutz, Vines, "Cloud Security ", Wiley Pub,2010
5. Velte, "Cloud Computing- A Practical Approach",TMH Pub,2009
6. Katarina Stanoevska-Slabeva, Thomas Wozniak, SantiRistol, "Grid and Cloud Computing-A Business Perspective on Technology and Applications", Springer,2010

MSAIDS351:R Programming Lab

CO1: Develop R programs using fundamental programming constructs, including data types, operators, control statements, functions, vectors, matrices, lists, and data frames to solve computational problems.

CO2: Import, clean, transform, and manage datasets using R, and perform exploratory data analysis (EDA) with appropriate statistical techniques.

CO3: Create effective data visualizations using R libraries such as **ggplot2** and generate meaningful graphical representations for data interpretation.

CO4: Apply statistical methods and machine learning techniques using R packages to analyze data, build predictive models, and evaluate their performance.

CO5: Design and implement complete data analysis projects in R by integrating data preprocessing, visualization, statistical analysis, model development, and result interpretation for real-world applications.

Examination:Practical

Exercisesto beframed soas to coverthe topics andtools coveredintheory paper.

MSAIDS352: PythonLab

CO1: Develop Python programs using fundamental programming concepts, including data types, operators, control structures, functions, modules, file handling, exception handling, and object-oriented programming.

CO2: Apply Python's built-in data structures (lists, tuples, dictionaries, and sets) and standard libraries to solve computational and real-world problems efficiently.

CO3: Design and implement Python applications for data processing, file management, and database connectivity using appropriate libraries and tools.

CO4: Develop programs for data analysis and visualization using Python libraries such as NumPy, Pandas, Matplotlib, and Seaborn.

CO5: Build and test Python-based solutions for real-world applications by integrating programming concepts, debugging techniques, and problem-solving skills.

Examination:Practical

Exercisesto beframed soas to coverthe topics andtools coveredintheory paper.

MSAIDS353: Machine LearningLab

CO1: Implement data preprocessing techniques, including data cleaning, feature selection, feature scaling, and dataset splitting, using Machine Learning tools and libraries.

CO2: Develop and evaluate supervised Machine Learning models such as linear regression, logistic regression, decision trees, k-nearest neighbors, support vector machines, and Naïve Bayes for predictive analytics.

CO3: Apply unsupervised Machine Learning techniques, including clustering and dimensionality reduction, to discover hidden patterns and relationships in datasets.

CO4: Evaluate the performance of Machine Learning models using appropriate metrics, cross-validation techniques, and hyperparameter tuning to improve model accuracy and reliability.

CO5: Design, implement, and present end-to-end Machine Learning solutions for real-world problems by integrating data preprocessing, model development, evaluation, visualization, and result interpretation using Python-based ML libraries.

Examination:Practical

Exercisesto beframed soas to coverthe topics andtools coveredintheory paper.

M.Sc. IT (AI and Data Science)–Fourth Semester

M.Sc. IT(AI and Data Science)–Fourth Semester

MSAIDS401(Elective 1):(a)DataWarehousing andData Mining

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

MaxMarks:70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10). Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain 4questions(1 fromeachunitwithinternalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Explain the concepts, architecture, and components of Data Warehousing, including data warehouse design, ETL processes, multidimensional data models, and OLAP operations.

CO2: Apply data preprocessing techniques such as data cleaning, integration, transformation, reduction, and discretization to prepare datasets for effective data mining.

CO3: Implement data mining techniques, including classification, clustering, association rule mining, and outlier detection, to extract meaningful patterns and knowledge from large datasets.

CO4: Evaluate and compare data mining models using appropriate performance measures, validation techniques, and visualization methods to support data-driven decision-making.

CO5: Design and apply data warehousing and data mining solutions using modern tools and techniques to solve real-world business, healthcare, finance, and research problems while considering ethical and privacy issues

Unit –I

Introduction to Data Warehouse: Data warehouse uses, Data Warehouse Planning stages and Designingapproaches.DeliveryProcess-DataWarehouseDeliveryMethods. System Processes; data in Flow Process, Extract and load process, Clean and transformProcess, Backup and Archive process and Query Management Process. Process Architecture - Load manager, Warehouse manager, Query manager.

Unit –II

Database Schema: Star flake schema, Designing fact tables and dimension tables, Multi-dimension schemas. Horizontal and vertical partitioning, Hardware partitioning. Data Marts, Designing Data Marts. Metadata transformation and load Managers.

Hardware Architecture: Process, Server, Network and Client hardware. Contents of data warehousedatabase,Databasestructuresandlayoutandfilesystems.OLAPfunctionandtools, OLAP Servers, ROLAP, MOLAP.

Unit –III

Security: Security requirements, Backup strategies and disaster recovery, operations of Warehouse. Capacity Planning (Process Estimate load), Tuning the data warehouse (Aggregate performance, data load and queries). Testing data warehouse-Develop test plan Testing backup recovery, Testing operational environment, testing database, testing of the application. Data warehouse futures.

Unit –IV

Data Mining: Data mining concepts, business, technical and social context for data mining, data mining interface, data mining approaches, data mining methodologies, data preprocessing, data cleaning, data reduction, data transformation, technologies used for data mining, cluster

analysis, portioning method.

RecommendedBooks :

1. Datamining& warehousing[conceptsand techniques]:SaumyaBajpai.
2. Dataminingconcepts&techniques:jiaweihan,michelinekamber,jianpei.
3. SamAnahory,DennisMurray,"DataWarehousing",PearsonEducation pub.
4. MichelA.Berry,Gordon S.Linoff,"Mastering DataMining",Wiley Publishing.
5. MallachG, FrednE, "DecisionSupport Systemand DataWarehouse Sustems", TMH
6. John Poole, Dan Chang, DauglasTalbert,"CommonWarehourse Metadata Developer's Guide", Wiley pub

MSAIDS401Elective PaperI: (b)Soft Computing

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions (1fromeachunitwith internalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Explain the fundamental concepts, characteristics, and applications of Soft Computing, including fuzzy logic, artificial neural networks, evolutionary computing, and genetic algorithms.

CO2: Apply fuzzy logic principles to model uncertainty, develop fuzzy inference systems, and solve decision-making and control problems.

CO3: Design and implement artificial neural network models for pattern recognition, classification, prediction, and optimization problems.

CO4: Analyze and apply evolutionary computation techniques, including genetic algorithms and swarm intelligence, to solve complex optimization and search problems.

CO5: Develop intelligent Soft Computing solutions by integrating fuzzy systems, neural networks, and evolutionary algorithms to address real-world problems while evaluating their performance and limitations.

UNIT-I

Introduction to Soft Computing: Introduction of Hard and Soft Computing, Unique features of Soft computing, Components of Soft computing, Fuzzy Computing, Evolutionary Computation, Genetic Algorithm, Swarm Intelligence, Ant Colony Optimizations, Neural Network, Machine Learning Associative Memory, Adaptive Resonance Theory, Introduction to Deep Learning.

UNIT-II

Fuzzy Logic: Basic concepts of fuzzy logic, Fuzzy sets and Crisp sets, Fuzzy set theory and operations, Properties of fuzzy sets, Fuzzy and Crisp relations, Fuzzy to Crisp conversion, Membership functions, interference in fuzzy logic, fuzzy if-then rules, Fuzzy implications and Fuzzy algorithms, Fuzzyfications&Defuzzificataions, Fuzzy Inference Systems, Mamdani Fuzzy Model. Sugeno Fuzzy Model, Fuzzy Controller, applications.

UNIT-III

Neural Networks: Introduction and Architecture: Neuron, Nerve structure and synapse, Artificial Neuron and its model, Neural network architecture: single layer and multilayer feed forward networks, recurrent networks. Back propagation networks architecture: perceptron model, solution, single layer artificial neural network, multilayer perception model; back propagation learning methods, back propagation algorithm, applications.

UNIT-IV

Genetic Algorithms: Basic concepts of GA, working principle, procedures of GA, flow chart of GA, Genetic representations, (encoding) Initialization and selection, Genetic operators, Mutation, Generational Cycle, applications. Hybrid Systems: Integration of neural networks, fuzzy logic and genetic algorithms. GA Based Back Propagation Networks, Fuzzy Back Propagation Networks, Fuzzy Associative Memories, Simplified Fuzzy ARTMAP.

Recommended Text / Reference Books:

1. S. Rajasekaran and G.A.VijaylakshmiPai,"Neural Networks Fuzzy Logic, and Genetic Algorithms", Prentice Hall of India 2004.

2. K.H.Lee. First Course on Fuzzy Theory and Applications, Springer-Verlag.

MSAIDS402Elective Paper II:(a)Natural Language Processing

MaxMarks:100(ESE: 70 CIA: 30) *Passing marks:40*

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions (1fromeachunitwith internalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Explain the fundamental concepts, techniques, and applications of Natural Language Processing, including text preprocessing, linguistic analysis, tokenization, stemming, lemmatization, and part-of-speech tagging.

CO2: Apply text representation methods such as Bag of Words (BoW), TF-IDF, word embeddings, and language models to process and analyze textual data.

CO3: Develop NLP applications using machine learning and deep learning techniques for tasks such as text classification, sentiment analysis, named entity recognition, machine translation, and question answering.

CO4: Evaluate the performance of NLP models using appropriate evaluation metrics and optimize them through feature engineering, model selection, and hyperparameter tuning.

CO5: Design and implement end-to-end NLP solutions using modern frameworks and libraries (such as NLTK, spaCy, Hugging Face Transformers, and TensorFlow/PyTorch) while considering ethical issues, bias, privacy, and responsible AI practices.

UNIT-I

INTRODUCTION: Origins and challenges of NLP – Language Modeling: Grammar-based LM, Statistical LM – Regular Expressions, Finite-State Automata – English Morphology, Transducers for lexicon and rules, Tokenization, Detecting and Correcting Spelling Errors, Minimum Edit Distance WORD LEVEL ANALYSIS : Unsmoothed N-grams, Evaluating N-grams, Smoothing, Interpolation and Backoff – Word Classes, Part-of-Speech Tagging, Rule-based, Stochastic and Transformation-based tagging, Issues in PoS tagging – Hidden Markov and Maximum Entropy models.

UNIT-II

SYNTACTIC ANALYSIS: Context Free Grammars, Grammar rules for English, Treebanks, Normal Forms for grammar – Dependency Grammar – Syntactic Parsing, Ambiguity, Dynamic Programming parsing – Shallow parsing – Probabilistic CFG, Probabilistic CYK, Probabilistic Lexicalized CFGs – Feature structures, Unification of feature structures.

UNIT-III

SEMANTICS AND PRAGMATICS: Requirements for representation, First-Order Logic, Description Logics – Syntax-Driven Semantic analysis, Semantic attachments – Word Senses, Relations between Senses, Thematic Roles, selectional restrictions – Word Sense Disambiguation, WSD using Supervised, Dictionary & Thesaurus, Bootstrapping methods – Word Similarity using Thesaurus and Distributional methods.

UNIT-IV

BASIC CONCEPTS of Speech Processing : Speech Fundamentals: Articulatory Phonetics – Production And Classification Of Speech Sounds; Acoustic Phonetics – Acoustics Of Speech Production; Review Of Digital Signal Processing Concepts; Short-Time Fourier Transform, FilterBank And LPC Methods.

Text books:

1. Daniel Jurafsky, James H. Martin―Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, Pearson Publication, 2014.

2. Steven Bird, Ewan Klein and Edward Loper, ―Natural Language Processing with Python, First Edition, OReilly Media, 2009.

3. Lawrence Rabiner And Biing-Hwang Juang, "Fundamentals Of Speech Recognition", Pearson Education, 2003.

4. Daniel Jurafsky And James H Martin, "Speech And Language Processing – An Introduction To Natural Language Processing, Computational Linguistics, And Speech Recognition", Pearson Education, 2002.

5. Frederick Jelinek, "Statistical Methods Of Speech Recognition", MIT Press, 1997.

6. 1. Breck Baldwin, ―Language Processing with Java and LingPipe Cookbook, Atlantic Publisher, 2015.

7. Richard M Reese, ―Natural Language Processing with Java, OReilly Media, 2015.

MSAIDS**402ElectivePaperII:(b)Deep Learning**

MaxMarks:100(ESE: 70 CIA: 30) Passing marks:40

QuestionPaperpatternforEndSemesterExam(ESE)

Max Marks : 70

Part-Awillcontain12veryshortquestionsof1markeach(attemptany10).
Part-B will contain 4 questions (1 from each unit) of 5 marks.
Part-Cwillcontain4questions (1fromeachunitwith internalchoice)of10 marks.

COURSE OUTCOMES :

The student will be able to :

CO1: Explain the fundamental concepts, architectures, and applications of Deep Learning, including artificial neural networks, activation functions, optimization techniques, and learning paradigms.

CO2: Design and implement deep neural network models using modern Deep Learning frameworks to solve classification, regression, and pattern recognition problems.

CO3: Apply advanced Deep Learning architectures such as **Convolutional Neural Networks (CNNs)**, **Recurrent Neural Networks (RNNs)**, **Long Short-Term Memory (LSTM)** networks, and **Transformers** for image, text, and sequential data analysis.

CO4: Evaluate and optimize Deep Learning models using appropriate loss functions, performance metrics, regularization techniques, hyperparameter tuning, and model validation strategies.

CO5: Develop end-to-end Deep Learning solutions for real-world applications using frameworks such as TensorFlow and PyTorch while addressing ethical, privacy, bias, and responsible AI considerations.

UNIT-1

Introduction: What is a Neural Network?, The Human Brain, Models of a Neuron, Neural Networks Viewed As Directed Graphs, Feedback, Network Architectures, Rosenblatt‟s Perceptron: Introduction, Perceptron, The Perceptron Convergence Theorem, Relation Between the Perceptron and Bayes Classifier for a Gaussian Environment.

UNIT-II

Multilayer Perceptrons: Introduction, Batch Learning and On-Line Learning, The Back-Propagation Algorithm, XOR Problem, Heuristics for Making the Back- Propagation Algorithm Perform Better, Back Propagation and Differentiation.

UNIT-III

Regularization for Deep Learning: Parameter Norm Penalties - L2 Parameter Regularization, Dataset Augmentation, Semi-Supervised Learning. Optimization for Training Deep Models: Challenges in Neural Network Optimization – Ill Conditioning, Local Minima, Plateaus, Saddle Points and Other Flat Regions.

UNIT-IV

Convolution neural networks: The Convolution Operation, Motivation, Pooling, Convolution and Pooling as an Infinitely Strong Prior, Variants of the Basic Convolution Function, Structured Outputs, Data Types, Efficient Convolution Algorithms, Convolutional

Networks and the History of Deep Learning.

Text Book:

1. Simon Haykin, Neural networks and Learning Machines, Third Edition, Pearson, 2016
2. Ian Goodfellow, YoshuaBengio and Aaron Courville, Deep Learning, MIT Press, 2016. https://www.deeplearningbook.org/lecture_slides.html

MSAIDS451:Project(Project,Report,Viva)

Two typed and duly bound copies of project report shall be submitted at least 3 weeks beforecommencement of the Theory/Practical examination which ever commences earlier.

Generalinstructionsaboutpreparationoftrainingreport

1. ThePower PointPresentation hasto beprepared forthe projectreport.
2. ThePowerPointPresentationwillbearound10-15mins,andthenquestionanswers.So prepare the number of slides accordingly.
3. Thebinding forreports willbe spiral binding.
4. Theformat for thereportsshouldbe adheredwith exactly.
5. TheCoding of the Projectshouldnot be included in thereport.
6. ContentsofIndexpage shouldinclude thefollowing parts:
    a. Project Requirements.
    b. Feasibility Study.
    c. Detailed Designing:
    d. ListofFigures.
    e. List of DFD.
    f. Listof ER-Diagram.
    g. List of Tables.
    i. Testing.
    j. Future Scope.

PageFormatofProjectReportshouldbeasfollows.
Paper: A4
Font: Times New Roman, Bookman Old Style
ChapterHeading:16pt,Subheading:14pt.
RunningMatter: 12 pt
Alltopicswillbenumberedaccordingly. Paragraph
Gap: 6 Pt Maximum
Line Gap: 1.5
Margins:Left 1.5, Right, Topand Bottom1 inch

Please Note: Project report of live project in the given format has to be prepared in 3 sets. ThesereportsshouldhaveCDcontainingthesoftcopyandPowerPointPresentationofProject report.

Formatof FrontPage ofProject Report.

Project Report

SubmittedtotheS.S.JainSubodhP.G.(Autonomous)College,

University of Rajasthan, Jaipur

Logoof college **UORlogo**

inPartialfulfillmentoftherequirementforthedegreeof

MASTEROFSCIENCEIT (AI and Data Science)

Submittedby
(yourname)

Name of Internal Guide

Name of Training Incharge (fromthecompany,whereyou have undergone training)

MScAI and Data Science)

(AcademicSession)

(monthofdepositionandyear)

Formatof LiveTrainingCertificateon companyletter head

CERTIFICATE

Thisistocertifythat"**yourname,S/D/O---------- "**is/wasundertrainingfrom……(startdate) to….(enddate)inmysupervisionforpartialfulfillmentoftherequirementfortheawardofthe Degree of Master of Science (Information Technology).

Duringthisperiodhe/shehasworkedon…….(descriptionoftraining)projectas ----------. Iwishforhis/hergood future.

Date: **NameofTraining Incharge**

Designation