UG CURRICULUM (for 2022 Batch)

B.TECH. COMPUTER SCIENCE & ENGINEERING (DATA SCIENCE)

A. COURSES / CREDITS DISTRIBUTION

I. Institute Core Courses (ICC) – 37 Credits
   a. Basic Science Courses (BSC) – 16 Credits
   b. Engineering Science Courses (ESC) – 17 Credits
   c. General Science Courses (GSC) – 4 Credits

II. Humanities, Communication and Management Elective Courses (HSSMEC) – 12 Credits

III. Departmental Core Courses (DCC) – 39 Credits
(including 3 credits of Minor Project)

IV. Departmental Elective Courses (DEC) – 20 Credits

V. Institute Open Elective Courses – 30 Credits
   a. Open Elective Courses – 24 Credits
   b. Project (Compulsory Major Project) – Interdisciplinary – 6 Credits

VI. Internship (Optional) / Course Work – 12 Credits
(Students Opting for course work will do department elective (4 credits), open elective (4 credits) and Project Work (4 Credits)

VII. Non-Academic Courses (NAC) – 10 Credits

Total Credits (without Honours): 160
Honours: 16 Credits
Total Credits (with Honours): 160 + 16 = 176

- **Minor Specialization:**
  Minor Specialization in CSE (DS) will be given to a student of outside the CSE department who earns 16 credits (within 160 credits) from the basket of Open Elective Courses offered by the CSE department as Minor Specialization Courses (MSC).

B. DEPARTMENT ELECTIVE COURSES (DEC)

1. Introduction to Data Analytics and Visualisation
2. Statistical Thinking for Data Science
3. Discrete Structures
4. Software Engineering
5. Agile Software Development and Modern Practices
6. Web Information Retrieval and Crawling
7. Recommendation Systems
8. Image Processing and Computer Vision
9. Natural Language Processing
10. Deep Learning
11. Advanced Artificial Intelligence
12. Knowledge Representation and Reasoning
13. Data and Information Security
14. Database Systems
15. Cloud Data Management
16. Mining Massive Datasets
17. Mathematical Techniques for Data Science
18. Architectures for Management of Large Data Sets
19. Graph Theory and Social Network Analysis
20. Optimization Strategies in Big Data Analytics

C. TEACHING SCHEME

First Year

Semester I

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	OR1101	Orientation Course (2 Credits) (including Introduction to Discipline Engg -1 credit 14 hours) (ii) Orientation (iii) Introduction to Data Science	 1	 0	 0	 1 1	NAC
2.	MA1101	Calculus and Ordinary Differential Equations	3	1	0	4	BSC-I
3.	CH1101	Applied Chemistry-I	3	0	2	4	BSC-IV
4.	ES1101	Introduction to Computing*	3	0	2	4	ESC-I
5.	ES1201	Engineering Drawing with CAD Software	2	0	2	3	ESC-II
6.	ES1301	Introduction to Mechatronics	2	0	2	3	ESC-III
					Total Credits	20	

*Common to all branches

Semester II

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	PY1201	Electromagnetic Theory and Quantum Physics	3	0	2	4	BSC-II
2.	MA1301	Probability and Statistics	3	1	0	4	BSC-III
3.	ES1401	Introduction to Electronics & Electrical Engineering	2	0	0	2	ESC-IV
4.	ES1501	Introduction to Manufacturing	2	0	0	2	ESC-V
5.	GS1101 & GS1201	Introduction to Environmental Sciences – I Introduction to Environmental Sciences - II	2 1	0 0	0 2	2 2	GSC-I & GSC-II
6.	HS1101	Communication Skills & Ethics	2	0	2	3	HSM-I
					Total Credits	19	

Second Year

Semester III

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	ES1701	Artificial Intelligence and Machine Learning*	2	0	2	3	ESC-VII
2.	DS1301	Data Structures and Algorithm	3	0	2	4	DCC-I
3.	DS1302	Python for Data Science	3	0	2	4	DCC-II
4.	DS1303	Introduction to IoT	3	0	2	4	DCC-III
5.	DS6301/ DS5301	#Open Elective – I (OE-I) / Minor Specialization Course (MSC)	3	0	2	4	OE-I/MSC
6.	DS7301	Industrial Tour				2	NAC
					Total Credits	21	

*Common to All branches

#Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Open Elective – I (OE-I)#

Sr. No.	Course Code OE-I/MSC	Course Name	L	T	P	Credits	Category
1.	DS6301	Statistical Thinking for Data Science	3	0	2	4	OE-I
2.	DS6302	Discrete Structures	3	1	0	4	OE-I

#Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Semester IV

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	HS2301 / HS2302 / HS2303 / HS2304	Basics of Economics / French- Basic / General Psychology / Sociology	2 2 2 2	1 1 1 1	0 0 0 0	3	HSM-II
2.	DS1401	Analytical Tools for Data Science	3	0	2	4	DCC-IV
3.	DS1402	Operating Systems	3	0	2	4	DCC-V
4.	DS1403	Data Acquisition and Information Retrieval	3	0	2	4	DCC-VI
5.	DS6401/ DS5401	*Open Elective –II (OE-II) / Minor Specialization Course (MSC)	3	0	2	4	OE-II/MSC
6.	PC7401	Proficiency-I				2	Proficiency
					Total Credits	21	

*Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Open Elective – II (OE-II)*

Sr. No.	Course Code OE-II/MSC	Course Name	L	T	P	Credits	Category

1.	DS6401	Introduction to Data Analytics and Visualisation	3	0	2	4	OE-II
2.	DS6402	Software Engineering	3	0	2	4	OE-II

*Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Third Year

Semester V

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS1501	Machine Learning for Data Science	3	0	2	4	DCC-VII
2.	DS1502	Theory of Computation	3	1	0	4	DCC-VIII
3.	DS1503	Computer Networks and Network Analysis	3	0	2	4	DCC-IX
4.	DS1504	Minor Project	0	0	6	3	DCC-X
5. 6.	DS2501 to DS2506	Department Elective Course – I (DEC-I) / Department Elective Course – II (DEC-II)	3 3	0 0	2 2	4 4	DEC-I / DEC-II
					Total Credits	23	

Department Elective Course - I (DEC-I) / Department Elective Course - II (DEC-II)

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2501	Discrete Structures	3	1	0	4	DEC-I/DEC-II
2.	DS2502	Software Engineering	3	0	2	4	DEC-I/DEC-II
3.	DS2503	Database Systems	3	0	2	4	DEC-I/DEC-II
4.	DS2504	Web Information Retrieval and Crawling	3	0	2	4	DEC-I/DEC-II
5.	DS2505	Natural Language Processing	3	0	2	4	DEC-I/DEC-II
6.	DS2506	Data and Information Security	3	0	2	4	DEC-I/DEC-II

Semester VI

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS8601	Internship Training (Optional)	0	0	24	12	Internship
		(OR)					
Students Opting for course work will do department elective (4 credits), open elective (4 credits) and Project Work							
1.	DS2601	Department Elective	3	0	2	4	DE
2.	DS6601 to DS6602	Open Elective*	3	0	2	4	OE
3.	DS7601	Project Work	0	0	8	4	Project
					Total Credits	12	

*Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Department Elective

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2601	Knowledge Representation and Reasoning	3	0	2	4	DE

Open Elective*

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS6601	Optimization Strategies in Big Data Analytics*	3	0	2	4	OE
2.	DS6602	Image Processing and Computer Vision*	3	0	2	4	OE

*Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Final Year

Semester VII

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	HS2701 / HS2702 / HS2703 / HS2704 / HS2705 / HS2706 / HS2707 / HS2708	Anthropology / Appreciation of Art / English Literature / History / Introduction to Art History/ Philosophy – an Introduction / Political Science / Public Administration	2 3 2 3 3 3 2 2	1 0 1 0 0 0 1 1	0 0 0 0 0 0 0 0	3	HSM-III
2. 3.	DS2701 to DS2708	Department Elective Course – III (DEC-III) / Department Elective Course – IV (DEC-IV)	3 3	0 0	2 2	4 4	DEC-III / DEC-IV
4.	DS6701 / DS5701	*Open Elective –III (OE-III) / Minor Specialization Course (MSC)	3	0	2	4	OE-III/MSC
5.	DS6702	*Open Elective –IV (OE-IV)	3	0	2	4	OE-IV
6.	DS7701	Major Project-I	0	0	4	2	Project
					Total Credits	21	

*Open Elective course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Department Elective Course - III (DEC-III) / Department Elective Course - IV (DEC-IV)

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2701	Agile Software Development and Modern Practices	3	0	2	4	DEC-III/DEC-IV
2.	DS2702	Deep Learning	3	0	2	4	DEC-III/DEC-IV
3.	DS2703	Mathematical Techniques for Data Science	3	0	2	4	DEC-III/DEC-IV
4.	DS2704	Graph Theory and Social Network Analysis	3	0	2	4	DEC-III/DEC-IV
5.	DS2705	Advanced Artificial Intelligence	3	0	2	4	DEC-III/DEC-IV
6.	DS2706	Cloud Data Management	3	0	2	4	DEC-III/DEC-IV
7.	DS2707	Mining Massive Data Sets	3	0	2	4	DEC-III/DEC-IV
8.	DS2708	Architectures for Management of Large Data Sets	3	0	2	4	DEC-III/DEC-IV

Open Elective – III (OE-III)*

Sr. No.	Course Code OE-III/MSC	Course Name	L	T	P	Credits	Category
1.	DS6701	Machine Learning for Data Science*	3	0	2	4	OE-III
2.	DS6702	Advanced Artificial Intelligence*	3	0	2	4	OE-III

*OE-III course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Open Elective – IV (OE-IV)*

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS6703	Python for Data Science*	3	0	2	4	OE-IV
2.	DS6704	Data and Information Security*	3	0	2	4	OE-IV

*OE-IV course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Semester VIII

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	HS2801 /	Business Environment and Business Laws /	2	1	0	3	HSM-IV
	HS2802 /	Entrepreneurship and Project Management /	2	1	0		
	HS2803	Financial Management	2	1	0		
2.	DS2801 to	Department Elective Course – V (DEC-V)	3	0	2	4	DEC-V

	DS2805						
3.	DS6801/ DS5801	*Open Elective – V / Minor Specialization Course (MSC)	3	0	2	4	OE-V/MSC
4.	DS6802	*Open Elective –VI	3	0	2	4	OE-VI
5.	DS8801	Discipline				2	Discipline
6.	PC7801	Proficiency-II				2	Proficiency
7.	DS7801	Major Project-II	0	0	8	4	Project
					Total Credits	23	

*Open Elective course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Department Elective Course - V (DEC-V)

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2801	Introduction to Data Analytics and Visualization	3	0	2	4	DEC-V
2.	DS2802	Statistical Thinking for Data Science	3	0	2	4	DEC-V
3.	DS2803	Recommendation Systems	3	0	2	4	DEC-V
4.	DS2804	Image Processing and Computer Vision	3	0	2	4	DEC-V
5.	DS2805	Optimization Strategies in Big Data Analytics	3	0	2	4	DEC-V

***Open Elective – V (OE-V) / Minor Specialization Course (MSC)**

Sr. No.	**Course Code**	**Course Name**	**L**	**T**	**P**	**Credits**	**Category**
1.	DS6801	Introduction to IOT*	3	0	2	4	OE-V

*OE-V course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence).

Open Elective – VI (OE-VI)*

Sr. No.	**Course Code**	**Course Name**	**L**	**T**	**P**	**Credits**	**Category**
1.	DS6802	Management of Large Data Sets	3	0	2	4	OE-VI

*OE-VI course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence).

Detailed Syllabus

Semester V

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS1501	Machine Learning for Data Science	3	0	2	4	DCC-VII
2.	DS1502	Theory of Computation	3	1	0	4	DCC-VIII
3.	DS1503	Computer Networks and Network Analysis	3	0	2	4	DCC-IX
4.	DS1504	Minor Project	0	0	6	3	DCC-X
5. 6.	DS2501 to DS2506	Department Elective Course – I (DEC-I) / Department Elective Course – II (DEC-II)	3 3	0 0	2 2	4 4	DEC-I /DECII
					Total Credits	23	

Department Elective Course - I (DEC-I) / Department Elective Course - II (DEC-II)

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2501	Discrete Structures	3	1	0	4	DEC-I/DEC-II
2.	DS2502	Software Engineering	3	0	2	4	DEC-I/DEC-II
3.	DS2503	Database Systems	3	0	2	4	DEC-I/DEC-II
4.	DS2504	Web Information Retrieval and Crawling	3	0	2	4	DEC-I/DEC-II
5.	DS2505	Natural Language Processing	3	0	2	4	DEC-I/DEC-II
6.	DS2506	Data and Information Security	3	0	2	4	DEC-I/DEC-II

Course Name	**:**	**Machine Learning for Data Science**
Course Code	**:**	**DS1501**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Core Course – VII (DCC-VII)**

Course Objectives :

Students should be able -

- To analyze and apply mathematical concepts in the context of data science and machine learning.
- To explore various data acquisition, cleaning, and preprocessing techniques.
- To develop skills in applying machine learning algorithms and predictive modeling on various datasets.
- To prepare students for lifelong learning in the rapidly evolving field of data science and machine learning.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**Introduction:** Basics of python for data science and machine learning, Types of data: structured data, unstructured data, other data types, the data science process, Explore and visualize data, features, Linear algebra, correlation	**6**
Unit 2	**Data Acquisition and Preprocessing:** Data sources and collection, Data storage and management, Data Preparation: handling missing data, feature encoding, feature scaling, Data Cleaning, Dataset division: train, test and validation sets	**7**
Unit 3	**Machine Learning Fundamentals:** Applications of Machine Learning, Introduction to Machine Learning Techniques: Supervised Learning, Unsupervised Learning and Reinforcement Learning, bias-variance tradeoff, overfitting-underfitting, Evaluation Measures: SSE, R2, confusion matrix, precision, recall, F-Score, ROC-Curve, cross-validation.	**7**
Unit 4	**Regression and Classification:** Simple Linear Regression, Multiple Linear Regression, Logistic Regression, gradient descent algorithm, K-Nearest Neighbor, Support Vector Machine (SVM), Naïve Bayes, Decision Trees.	**12**
Unit 5	**Dimensionality Reduction and Clustering:** Feature selection, Principle Component Analysis, K-means clustering, Hierarchical clustering, Density based clustering.	**10**

List of Experiments:		**Number of Turns**
1	Explore various python libraries used for data handling and machine	**2**

	learning techniques: numpy, pandas, matplotlib, scikit-learn etc.	
2	Implementation of data handling and preprocessing techniques	**2**
3	Implementation of various classification and regression algorithms – Linear and logistic regression, KNN, SVM, Naïve Bayes, Decision Trees	**6**
4	Implementation of Dimensionality Reduction and Clustering techniques: PCA, K-means clustering, Density based clustering algorithms	**4**

Course Outcomes: **After completion of course, students would be able to:**	
1	Explore and apply mathematical and engineering knowledge to analyze complex data science problems.
2	Apply modern tools and techniques to solve data science problems.
3	Apply different machine learning models on various datasets.
4	Analyze the trade-off of various machine learning models for real-life applications.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"Introduction to Data Science" by Jeffrey Stanton, SAGE Publications	2017
2	"Introduction to Machine Learning with Python: A Guide for Data Scientists", Andreas Müller, Sarah Guido, O'Reilly Media	2016
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"The Machine Learning Yearning" by Andrew NG, Deeplearning.ai	2021
2	"Python for Data Analysis" by Wes McKinney, O'Reilly Media	2022
3	"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems", by Aurélien Géron, 3rd Edition, O'Reilly	2022

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Machine Learning Specialization - Stanford University https://www.coursera.org/specializations/machine-learning-introduction	Coursera
2	Data Science Specialization - Johns Hopkins University https://www.coursera.org/specializations/jhu-data-science	Coursera

Course Name	**:**	**THEORY OF COMPUTATION**
Course Code	**:**	**DS1502**
Credits	**:**	**4**
L T P	**:**	**3 1 0**
Type of Course	**:**	**Department Core Course –VIII (DCC-VIII)**

Course Objectives :

Students should be able -

- To understand fundamental mathematical and computational principles that are foundations of computer science.
- To learn about abstract models of computation, finite representations for languages.
- To gain formal understanding of algorithms and procedures.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**AUTOMATA AND LANGUAGE THEORY** Introduction to formal proofs, Finite Automata (FA), Deterministic Finite Automata (DFA), Non-deterministic Finite Automata (NFA), Finite Automata with Epsilon transitions	8
Unit 2	**REGULAR EXPRESSIONS AND LANGUAGES** Regular Expression, FA and Regular Expressions, Proving languages not to be regular, Closure properties of regular languages, Equivalence and minimization of Automata	9
Unit 3	**CONTEXT-FREE GRAMMARS AND LANGUAGES** Context-Free Grammar (CFG), Parse Trees, Ambiguity in grammars and languages, Definition of the Pushdown automata, Languages of a Pushdown Automata, Equivalence of Pushdown automata and CFG Deterministic Pushdown Automata, Normal forms for CFG, Pumping Lemma for CFL, Closure Properties of CFL	10
Unit 4	**COMPUTABILITY THEORY** Turing Machines, Programming Techniques for TM, Variations of TM, Non-Universal TM, Universal TM, decidability, halting problem, reducibility.	8
Unit 5	**UNDECIDABALITY** Complexity classes P, NP, L, NL, PSPACE, BPP and IP, A language that is not Recursively Enumerable (RE), An undecidable problem that is RE, Undecidable problems about Turing Machine, Post's Correspondence Problem	7

List of Experiments:		**Number of Turns**
1	Operations on Strings, Operations on Languages	1

2	Construct the DFA, the transition diagram and transition table	1
3	Convert the regular expressions into their equivalent NFAs	1
4	Convert the NFAs to DFAs, DFA Minimization,	1
5	Arden's theorem, DFA equivalence	1
6	Pumping lemma	1
7	CFG, ambiguity	1
8	CFG simplification, CNF, GNF	2
9	Push Down Automata, CFG to PDA	2
10	Turing Machines, Halting problem	2
11	P, NP, PCP	1

Course Outcomes:	
1	Recognise and comprehend formal reasoning about languages
2	Express computer science problems as mathematical statements
3	Distinguish different computing languages and classify their respective types
4	Develop a competent understanding of the basic concepts of complexity theory

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	J.E. Hopcroft, R. Motwani and J.D. Ullman, "Introduction to Automata Theory, Languages and Computations", second Edition, Pearson Education	2007
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	H.R. Lewis and C.H. Papadimitriou,"Elements of the theory of Computation", Second Edition, Pearson Education	2003
2	Thomas A. Sudkamp," An Introduction to the Theory of Computer Science, Languages and Machines", Third Edition, Pearson Education.	2007
3	J. Martin, "Introduction to Languages and the Theory of computation" Third Edition, Tata Mc Graw Hill.	2007

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Theory of Computation https://onlinecourses.nptel.ac.in/noc19_cs79/preview	Swayam+ IIT Kanpur
2	The complete theory of computation https://www.udemy.com/course/the-complete-theory-of-computation/	Udemy

Course Name	**:**	**Computer Networks and Network Analysis**
Course Code	**:**	**DS1503**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Core Course-IX (DCC-IX)**

Course Objectives :

Students should be able -

- To understand the basic concepts of computer networks
- To explore various applications and their protocols
- To design and implement various transport layer protocols and algorithms
- To analyze and develop various network layer protocols and algorithms
- To evaluate the performance of various data link control protocols and algorithms

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**COMPUTER NETWORKS AND THE INTERNET** What is the Internet; network edge; network core; Delay, Loss and throughput in Packet-Switched Networks; Protocol Layers and their Service Models.	**6**
Unit 2	**APPLICATION LAYER** Principles of Network Applications; The Web and HTTP; File Transfer: FTP; Electronic Mail in the Internet; DNS - The Internet's Directory Service	**8**
Unit 3	**TRANSPORT LAYER** Introduction and Transport-Layer Services; Multiplexing and Demultiplexing; Connectionless Transport: UDP; Connection-Oriented Transport: TCP; TCP Congestion Control.	**10**
Unit 4	**NETWORK LAYER** Introduction; Virtual circuit and datagram networks; What is inside a router; Internet Protocol (IP): Forwarding and Addressing in the Internet; Routing Algorithms; Routing in the Internet	**10**
Unit 5	**DATA LINK LAYER** Introduction to the link layer; Error Detection and Correction Techniques; Multiple Access links and Protocols; Switched local area networks.	**8**

List of Experiments:		**Number of Turns**
1	Understanding and using of commands like ping, ipconfig, hostname, getmac, arp, nslookup, netstat, tracert, route,pathping	**2**
2	Socket Programming – Installing and learning Linux operating system, Understanding the syntax, purpose and use of various functions used in sockets programming – connect(), send(), recv(), bind(), listen(),	**3**

	accept(), sendto(), recvfrom() etc. functions, TCP/IP based echo client server and UDP/IP based echo client server	
3	Practicals based on Wireshark – getting started, HTTP, DNS, TCP, IP, ICMP, Ethernet & ARP, DHCP etc.	**9**

Course Outcomes: At the end of the course, students will be able to:	
1	Understand and explain various components of computer networks
2	Illustrate the various applications and protocols
3	Analyze and develop various transport layer protocols and algorithms
4	Explore and apply network layer protocols and algorithms to design and analyze the networks
5	To evaluate and compare various data link control protocols and algorithms

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	James F. Kurose and Keith W. Ross, "Computer Networking: A top down approach", Pearson Education, 6th edition.	2017
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	A.S. Tanenbaum, "Computer Networks", 5th Edition, PHI	2010
2	G. Keiser, "Local Area Networks", 2nd Edition, TMH	2002
3	D. Bertesekas and R. Gallager, "Data Networks", 2nd Edition, PHI	2010
4	William Stallings, "Data & Computer Communication", PHI, 6th Edition	2002
5	B.A. Forouzan, "Data communications and networking", TMH, 1st ed	2000
6	B.A. Forouzan, "Local Area Networks", TMH.	2001
7	B.A. Forouzan, "TCP/IP Protocol Suite", TMH	2005

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Computer Networks, https://onlinecourses.swayam2.ac.in/cec19_cs07/preview	Swayam
2	Computer Networks and Internet Protocol, https://onlinecourses.nptel.ac.in/noc20_cs23/preview	NPTEL

Course Name	**:**	**DISCRETE STRUCTURES**
Course Code	**:**	**DS2501**
Credits	**:**	**4**
L T P	**:**	**3 1 0**
Type of Course	**:**	**Department Elective Course I / II (DEC-I/DEC-II)**

Course Objectives:

- To develop logical thinking and its application to computer science.
- To reason mathematically about basic data types and structures (such as numbers, sets, graphs, and trees) used in computer algorithms and systems; synthesize elementary proofs, especially proofs by induction.
- To model and analyze computational processes using analytic and combinatorial methods.
- To apply principles of discrete probability to calculate probabilities and expectations of simple random processes.

Total No. of Lectures – 42

Lecture wise breakup		**Number of Lectures**
1	**MATHEMATICAL REASONING** Mathematical reasoning, Propositions, Negation, disjunction and conjunction, Implication and Equivalence, Truth tables, Predicates, Quantifiers, Natural deduction, Rules of Inference, Methods of proofs, Resolution principle, Application to PROLOG.	**8**
2	**SET THEORY** Paradoxes in set theory, Inductive definition of sets and proof by induction, Peano postulates, Relations, Properties of relations, Equivalence Relations and partitions, Partial orderings, POSETs, Linear and well-ordered sets.	**8**
3	**FUNCTIONS** Functions; mappings, Injection and Surjections, Composition of functions, Inverse functions, Special functions, Recursive function theory.	**6**
4	**COMBINATORICS** Elementary combinatorics, Pigeonhole principle, Permutations and Combinations, Counting techniques, Recurrence relations, Solving Linear Recurrence relations, Generating functions.	**9**
5	**GRAPH THEORY** Elements of graph theory, Graph Isomorphism, Euler graph, Hamiltonian path, trees, Tree traversals, Spanning trees, Representation of relations by graphs. **GROUPS, RINGS, FIELDS** Definition and elementary properties of groups, Semigroups, Monoids,	**8**

	Rings, Fields, Vector spaces and lattices.	
6	**DISCRETE PROBABILITY** Introduction, Probability Theory, Bayes' Theorem, Expected Value and Variance, Discrete random variables.	**3**

Course Outcomes: At the end of the course, students will be able to:	
1	Use logical notations to define and reason about fundamental mathematical concepts such as sets, relations, functions, and integers. Analyze logical propositions via truth tables.
2	Synthesize induction hypotheses and simple induction proofs.
3	Calculate numbers of possible outcomes of elementary combinatorial processes such as permutations and combinations.
4	Apply graph theory models of data structures and state machines to solve problems of connectivity and constraint satisfaction, for example, scheduling.
5	Prove elementary properties of modular arithmetic and explain their applications in Computer Science, for example, in cryptography and hashing algorithms.
6	Calculate probabilities and discrete distributions for simple combinatorial processes; calculate expectations.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	K. H. Rosen, Discrete Mathematics and applications, 7th Edition, McGraw Hill	2012

Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**

1	Seymour Lipschutz and Marc Lipson, Schaum's Outline of Discrete Mathematics, 3rd Edition	2010
2	J. L. Mott, A. Kandel, T. P. Baker, Discrete Mathematics for Computer Scientists and Mathematicians, 2nd Edition, Pearson India	2015
3	C. L. Liu and D. P. Mohapatra, Elements of Discrete Mathematics, 4th Edition., McGraw-Hill	2012
4	C. Stein, R. L. Drysdale, K. Bogart, Discrete Mathematics for Computer Scientists, Second edition, Pearson Education Inc.	2011
5	W. K. Grassmann and J. P. Trembnlay, Logic and Discrete Mathematics, A Computer Science Perspective, Prentice Hall Inc	2007
6	M. Litvin and G. Litvin, Coding in Python and Elements of Discrete mathematics, Skylight Publishing	2019
7	A. M. Stavely, Programming and Mathematical Thinking, The New Mexico Tech Press	2014

Equivalent MOOC Courses		
Sr. No.	**Course Link**	**Offered by**
1	https://ocw.mit.edu/courses/6-042j-mathematics-for-computer-science-fall-2010/video_galleries/video-lectures/	MIT Open Courseware
2	https://onlinecourses.nptel.ac.in/noc20_cs82/preview	NPTEL

Course Name	**:**	**SOFTWARE ENGINEERING**
Course Code	**:**	**DS2502**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course I / II (DEC-I/DEC-II)**

Course Objectives:

- Students should understand fundamentals of software engineering.
- They should learn strengths and weaknesses of various software engineering process models used in industrial applications.
- They should be able to construct software that is reasonably easy to understand, design, develop,test and modify.

Total No. of Lectures – 42

Lecture wise breakup		**Number of Lectures**
1	**INTRODUCTION TO SOFTWARE ENGINEERING:** Software, Software Engineering, Myths, Software Process, Work Products,Importance of Software Engineering.	4
2	**SOFTWARE PROCESS MODELS:** Standard for Software Process, Waterfall Model, Prototyping Model, IterativeEnhancement Model, Spiral Model, RAD model, 4th Generation models, Formal Methods, Agile Development, Team Software Process, DevOps, Choosing the best Software Process.	6
3	**REQUIREMENT ENGINEERING:** Software Requirements, Types of Requirements, Requirement Engineering Cycle, Requirements Specification document, Characteristics of Requirements,Requirement verification and validation.	4
4	**SOFTWARE PROJECT MANAGEMENT:** Role of Management in Software Development, Project Estimation Techniques, Staffing, Scheduling, Earned Value Analysis, Software Risks, Software Configuration Management, Software Process and Project metrics.	6
5	**SOFTWARE DESIGN:** Process, Data and Behavioral Modeling, Design Concepts, Modularity, Architectural design, Coupling and Cohesion, Top-down and bottom-up design,Object-oriented Analysis, Function-oriented and Object-Oriented Design approach, Software Design Document.	6
6	**CODING AND TESTING:** Coding styles and documentation, Testing principles, Testing strategies, Black-box and White-box Testing Techniques, Levels of testing -unit, integration, system, regression, Test Plan, Test Cases Specification, Software debugging, Software Maintenance.	6
7	**SOFTWARE QUALITY:** Software Quality Assurance (SQA), SQA tasks, Software amplification and removal, Formal Technical Reviews, Software Quality Factors, ISO 9126, SEICMM, CMMI, Software Reliability. Software Availability.	4

8	**ADVANCED TOPICS IN SOFTWARE ENGINEERING:** Software Process Improvement, Component Based Software Engineering, WebEngineering, Reverse Engineering, Software Engineering challenges of Big Data, Mobile Applications. **COMPUTER AIDED SOFTWARE ENGINEERING:** Computer Aided Software Engineering (CASE) and its Scope, CASE support in Software Life Cycle, Architecture of CASE Environment, Upper CASE andLower CASE, Exposure to CASE tools.	6
List of Experiments:		**Number of Turns**
1	Team Software Project: Requirements Planning, Prepare Models and Designs,Develop and Test the project.	Throughout Semester
2	UML models exercises.	3
3	Programming exercise to understand and develop test cases for Black-boxtesting.	1
4	Programming exercise to understand and develop test cases for White-boxtesting.	1

Course Outcomes: At the end of the course, students will be able to:

1	Understand the concepts of Software Process. Select a suitable Software Process for anyproject.
2	Design and Test the requirements of the software projects.
3	Design and develop projects using concepts like Applet, AWT class, client and server sidescripting.
4	Implement the software development processes activities from requirements to verificationand validation.

Suggested Books:

Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	Roger Pressman, Software Engineering: A Practitioner‟s Approach, 7thEdition, McGraw- Hill	2010
2	Ian Sommerville, Software Engineering, 9th Edition, Addision-Wesley.	2016
3	KK Aggarwal and Yogesh Singh, Software Engineering, 3rd Edition, New AgeInternational	2008
4	Pankaj Jalote, A Concise Introduction to Software Engineering, Springer.	2008

Course Name	**:**	**Database Systems**
Course Code	**:**	**DS2503**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course		**Department Elective Course I / II (DEC-I/DEC-II)**

Course Objectives:

Students should be able -

- To analyze the components/ requirements of DBMS
- To explore the role of storage and file structure in database management.
- To apply their knowledge of normalization theory to the normalization of a database
- To develop various SQL queries related to any given situation.
- To enable them think for transactions and concurrency in database applications in various scenarios

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**INTRODUCTION** Introduction and application of DBMS, Data Independence, Database System Architecture – levels, Mapping, Database users and DBA, Database Languages: DDL, DML. **DATABASE MODELS:** Entity – Relationship model, constraints, keys, Design issues, E-R Diagram, Weak and Strong entity types, Extended E-R features- Generalization, Specialization, Aggregation, Translating E-R model into Relational model. hierarchical and network models	**8**
Unit 2	**RELATIONAL MODE**L: Introduction to relational model, basic structure, Types, Keys, views in a relational database. **SQL:** Fundamentals, basic structure, set operations, aggregate operations, DDL, DML, DCL, nested queries, complex queries, Integrity Constraints, PL/SQL Concepts, triggers	**8**
Unit 3	**RELATIONAL ALGEBRA AND RELATIONAL CALCULUS** Relational Algebra: Fundamental operations, Additional Operations Relational calculus: Tuple Relational calculus, Domain Relational calculus **STORAGE AND FILE STRUCTURE** Overview of physical storage media, magnetic disks, RAID, file organization, organization of records in files, indexing and hashing	**6**
Unit 4	**RELATIONAL DATABASE DESIGN:** Functional Dependencies, Non-loss Decomposition, First, Second, Third Normal Forms, Dependency Preservation, Boyce/Codd Normal Form, Multi-valued Dependencies and Fourth Normal Form, Join Dependencies and Fifth Normal Form, 6th Normal Form.	**8**
Unit 5	**TRANSACTION MANAGEMENT** Transaction concept, transaction state, ACID properties,	**7**

	serializability, Recoverability , Implementation of Isolation , Testing for serializability, **CONCURRENCY CONTROL** Basic Concept , Lock based concurrency control, Time stamping methods, Dealing with Deadlocks, Recovery techniques based on deferred update and recovery techniques based on Immediate Update. Introduction to Checkpoints and Shadow Paging, Introduction to Database Security, Discretionary access control based on granting and revoking privileges	
Unit 6	**DISTRIBUTED DATABASES** Introduction, data mining, Big Data, NoSQL, New SQL, Modern databases based on these concepts.Object oriented databases	**5**

List of Experiments:		**Number of Turns**
1	For a given scenario of database application, Create the required tables using SQL Commands	1
2	Write Sql queries to apply the constraints i.e. Primary Key, Foreign key, UNIQUE to the tables.	2
3	SQL queries for Null values and different clauses	2
4	Usage of SELECT, rename, tuple operations, DELETE etc.	2
5	SQL queries for implementing various String operations, Set operations	2
6	SQL queries for implementing JOINs and types of joins with conditions.	2
7	SQL nested queries for a particular scenario.	1
8	SQL queries to create the views ,triggers	1
9	SQL queries to create indexes and apply on a database.	1

Course Outcomes:	
1	Describe DBMS architecture, physical and logical database designs, database modeling, relational, hierarchical and network models
2	explain basic database storage structures and access techniques such as file organizations, indexing methods, Understand the normalization of databases
3	Analyse /contrast the realtional algebra, calculus , other latest modern databases like data mining, no sql, etc
4	Design , implement, test and debug using Sql ,transaction proceesing, concurrency control mechanism
5	Design a database and Sql queries , triggers etc in various environments.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**

1	"Database System Concepts", Abraham Silberschatz, Henry F. Korth, S. Sudharshan, Tata McGraw Hill,	2006
2	"Fundamentals of Database Systems", Elmsari and Navathe, Pearson Education	2013
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Database Management Systems , Ramakrishnan and Gehrke, McGrawHill	2003
2	"An Introduction to Database Systems", C.J.Date, A.Kannan, S.Swamynathan, Pearson Education	2006
3	J. D. Ullman, "Principles of Database Systems", 2nd Ed., Galgotia Publications	1999

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Database Management Essentials https://www.coursera.org/learn/database-management	coursera
2	Data Base Management System By Prof. Partha Pratim Das https://archive.nptel.ac.in/courses/106/105/106105175/	NPTEL

Course Name	**:**	**Web Information Retrieval and Crawling**
Course Code	**:**	**DS2504**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course I / II (DEC-I/DEC-II)**

Course Objectives:

- To understand information retrieval algorithms
- To identify challenging problems on the Web
- The course will cover both traditional and newly developed algorithms in information retrieval and Web search and their Web applications

Total No. of Lectures – 42

Lecture wise breakup		**Number of Lectures**
1	**Introduction to Web:** How the web works, Web and its upcoming versions: A bird's eye on the evolution and Definition: Web 1.0, Web 2.0, Web 3.0, Web 4.0, Web 5.0, and various applications, Web Crawlers and search engines, Dark and Deep web	**6**
2	**Web IR background:** Web IR vs. traditional IR,Web search history, Web Characterization, Structure of the Web (Bow-tie, scale-free and small world networks, power-law distributions), Deep web and surface web	**10**
3	**Web crawling** Crawler architecture, Various type of crawlers, Politeness policies (robots.txt),. Selection policies (URL characteristics, spider traps, repository freshness), URL normalization,. Sitemap protocol	**10**
4	**Indexing** Boolean retrieval model, Vector space model, Inverted index and implementation issues, Stemming and stop lists,. Synonyms, homographs, and misspellings Term frequency and inverted document frequency	**10**
5	**Ethical Issues in IR :** Privacy, Fairness, Fake news and disinformation, Filter bubble , Viewpoint diversity, Fostering extremism, Internet addiction	**6**

List of Experiments:		**Number of Turns**
1	Implementing basic IR model and making it smart in each session as per the machine learning algorithms	**7**
2	Implementing various types of web crawlers and doing their comparative analysis	**7**

Co**urse Outcomes:** At the end of the course, students will be able to:	
1	Understand web crawling and Information retrieval in details
2	Implement some of the crawlers keeping all the policies in mind.
3	Implement indexer and apply it on real world database
4	Aware about various ethical issues in web crawling and information retrieval.

Suggested Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Introduction to Information Retrieval by Christopher D. Manning, Prabhakar Raghavan and Hinrich Schutze. Cambridge University Press, 2008.	2018
2	Modern Information Retrieval by Ricardo Baeza-Yates and Berthier Ribeiro-Neto Addison-Wesley	2019

Course Name	**:**	**Natural Language Processing**
Course Code	**:**	**DS2505**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course I / II (DEC-I/DEC-II)**

Course Objectives :

Students should be able -

- To explore various language structures and tools that are available for NLP.
- To analyze large collections of text for processing.
- To apply concepts related to language processing and generation.
- To develop tasks related to language processing, information extraction and machine translation.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**INTRODUCTION** Linguistic fundamentals, a computational framework for natural language, description of English or an Indian language in the framework; lexicon, algorithms and data structures for implementation of the framework, the different analysis levels used for NLP (morphological, syntactic, semantic, pragmatic)	**4**
Unit 2	**WORD LEVEL AND SYNTACTIC ANALYSIS** Word Level Analysis: Regular Expressions, Finite-State Automata, Morphological Parsing, Spelling Error Detection and correction, Words and Word classes, Part-of Speech Tagging. Syntactic Analysis: Context-free Grammar, Constituency, Parsing-Probabilistic Parsing. Lexical Resources: World Net, Frame Net, Stemmers, POS Tagger.	**10**
Unit 3	**SEMANTIC ANALYSIS** Semantic Analysis: Meaning Representation, Lexical Semantics, Ambiguity, Word Sense Disambiguation. Discourse Processing: Coreference Resolution, Discourse Coherence and Structure. Knowledge Representation, reasoning.	**10**
Unit 4	**NATURAL LANGUAGE GENERATION** Natural Language Generation (NLG): Architecture of NLG Systems, Generation Tasks and Representations, Application of NLG. Machine Translation approaches, problems in Machine Translation. Characteristics of Indian Languages, Translation involving Indian Languages.	**10**
Unit 5	**INFORMATION RETRIEVAL AND ADVANCED CONCEPTS** Information Retrieval: Design features of Information Retrieval Systems, Classical, Non-classical, Alternative Models of Information Retrieval, Language Models, Transformers.	**8**

List of Experiments:	**Number**

		of Turns
1	Implement program to perform automatic word analysis	2
2	Implement program to perform word generation	2
3	Implement programs related to morphology, N-Grams, N-Grams Smoothing	3
4	Implementation of Hidden Markov Models	2
5	Implement program to build POS Tagger, Chunker	3
6	Case study: language model	2

Course Outcomes:	
1	To understand the challenges of language processing, information extraction and translation.
2	To apply natural language concepts for various language based applications
3	To analyse the language constructs for language processing
4	To evaluate the language based techniques for processing applications
5	To design solutions for natural language processing, information extraction and translation

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	Speech and Language Processing by Daniel Jurafsky and James H Martin., 3rd ed draft. https://web.stanford.edu/~jurafsky/slp3/	2023
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	Natural Language understanding by James Allen, Pearson Education	2008
2	NLP: A Paninian Perspective by Akshar Bharati, Vineet Chaitanya, and Rajeev Sangal, Prentice Hall	1995
3	Meaning and Grammar by G. Chirchia and S. McConnell Ginet, MIT Press	2000
4	Natural Language Processing and Information Retrieval by T.Siddiqui, U.S. Tiwary, OUP.	2008

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Natural Language Processing https://onlinecourses.nptel.ac.in/noc23_cs45/preview	Nptel
2	Natural Language Processing Specialization https://www.coursera.org/specializations/natural-language-processing	coursera

Course Name	**:**	**Data and Information Security**
Course Code	**:**	**DS2506**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course I / II (DEC-I/DEC-II)**

Course Objectives

The main objectives of this course are:	
1	To understand basic design principals of symmetric and asymmetric cryptography and learn how standard cryptanalytic attacks work and thereby how to avoid common design flaws.
2	The student should be able to apply basic number theory in cryptography and will be able to understand the design principals of symmetric and asymmetric cryptography, AES, RSA
3	The course will be emphasizing on the algorithmic complexity and security vs performance trade off.
4.	Learn different perspectives of managing Keys, trust and Privacy

Total No. of Lectures – 42

Course Contents

Lecture wise breakup		**Number of Lectures**
1	**Introduction** - Security mind-set, Computer Security Concepts (CIA), Threats, Attacks Assets, Services and Mechanisms.	3
2	**Foundations** History and evolution of cryptosystems, Classical Encryption Systems, Substitution Ciphers and Transposition Cipher, Block cipher, Stream cipher, Modes of operation, Symmetric and Asymmetric cryptography.	4
3	**Symmetric Key Cryptosystems**: Advanced Encryption Standards Mathematical Theory, Number Theory, Data Encryption Standard: Description of DES, Security of DES, Differential And Linear Cryptanalysis, Design Criteria, DES Variants, DES modes of operation, Finite fields, Galois's field, Multiplicative Inverse, Extended Euclid's algorithm, AES, Other Stream Ciphers	10
4	**Asymmetric key Cryptosystems** Prime Number Generation, Discrete Logarithms in a Finite Field, Chinese Remainder Theorem, DIFFIE-HELLMAN Key Exchange, RSA Key Setup, Encryption and verification, elgamal, and Elliptic Curve Cryptography	8
5	**Key Management, Message Authentication and Digital Signatures** – Key Management, Notion of Trust, Centralised vs Distributed Trust management, Distributed Third Party, Kerberos, PKI, PGP, Certificate Distribution, Management and Revocation	8
6	**Message Authentication and Digital Signatures** Digital Signatures and Certificates, Hash and MAC, Digital Signatures, RSA and DSS based digital Signatures	5
	IDS and Firewalls Intruders, Virus, Worms, Firewalls-need and features of firewall, Types of firewall, Intruder Detection Systems.	4

Lab work

Sr No	Lab Contents	No. of Hours
1	Implementation of key exchange, symmetric and asymmetric algorithms, signature schemes.	7
2	Practical Use of Network Security Tools, Email Header Analysis, Packet sniffing, configuration of network security equipment such as firewall, routers, IDS, Wireless Access Points. Explore security solutions for Web application vulnerabilities.	7

Course Outcomes: At the end of the course, students will be able to:	
1	Understand the cryptographic principles
2	Able to apply the basic rules of public key and symmetric encryption for practical cryptographic problems
4	To evaluate various security issues and understand solutions for it.
5	Given a problem in cryptography, be able to design an algorithm to implement the solution to that problem.

References

Sr No	Book detail	Year of Publication
1	Applied Cryptography protocols, algorithms, and source code in C, Second Edition, Bruce Schneier, John Wiley & Sons	1996 (reference book)
2	Cryptography and Network Security by William Stallings, Prentice Hall. ISBN 0-13-187316-4	2006 (text book)
3	Handbook of Applied Cryptography by Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone, CRC Press	2000 (reference book)
4	Cryptography Theory and Practice by Douglas R. Stinson: 3rd ed., Chapman & Hall/CRC.	2006 (reference book)

Semester VI

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS8601	Internship Training (Optional)	0	0	24	12	Internship
(OR)							
Students Opting for course work will do department elective (4 credits), open elective (4 credits) and Project Work							
1.	DS2601	Department Elective	3	0	2	4	DE
2.	DS6601 to DS6602	Open Elective*	3	0	2	4	OE
3.	DS7601	Project Work	0	0	8	4	Project
Total Credits						12	

*Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Department Elective

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2601	Knowledge Representation and Reasoning	3	0	2	4	DE

Open Elective*

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS6601	Optimization Strategies in Big Data Analytics*	3	0	2	4	OE
2.	DS6602	Image Processing and Computer Vision*	3	0	2	4	OE

*Open Elective course will not be offered to B.Tech. (CSE), B.Tech. CSE(DS) and B.Tech. CSE(AI) students

Course Name	**:**	**Knowledge Representation and Reasoning**
Course Code	**:**	**DS2601**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective (DE)**

Course Objectives :

Students should be able -

- To learn predicate logic, a universal symbolic language to express the basic constructs
- To introduce Logic based concepts and techniques
- To learn theoretical and practical applications in knowledge representation system

Course Contents: **Total No. Lectures: 42**

Sr. No.	Course contents	No. of Lectures
1.	Introduction: Knowledge based systems, The role of logic, semantics, Explicit and implicit beliefs, Knowledge engineering, vocabulary	4
2.	Propositional Logic, Resolution and Horn Clauses, Representing Knowledge in Predicate Logic, Soundness and completeness, Handling Propositional Formulae, Reduction, Belief operators, Characteristic axioms, Non-monotonic logics, Reasoning in Default Logic, Fixpoint semantics for default logic, Properties of default logic	9
3.	Modeling and reasoning: Equivalence of logic programs, Cumulative logics, Well-founded semantics for logic programs, Consequence relations, preferential reasoning, Answer set programming, Abductive reasoning, Qualitative reasoning, Constraint satisfaction	7
4.	Knowledge representation: Ontology languages for the Semantic Web, Rules, Representation of Domain Knowledge, Object-oriented representation, inheritance reasoning, Ontology design patterns, Linked data publishing with ODPs	8
5.	Overview RDF semantics, SPARQL, BGP, Triples, Formalization and representation of knowledge with RDF and OWL	6
6.	Knowledge Representation in the Web: Representing data with HTML, Formalization and representation of information with DTD, XML Schema	8

Lab Work: To follow project based learning approach for the course.

Sr. No.	Lab Contents	No. of Hours
1.	To Translate the sentences from English to first order logic	4
2.	To implement Horn clauses and program positive Horn clauses and negative Horn clauses.	3
3.	To build an ontology and its mappings to query	2
4.	Write RDF document in Turtle syntax	2
5.	Write SPARQL queries using SELECT, CONSTRUCT, ASK, DESCRIBE	3

Course Outcomes:

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

1. Understand the fundamental principles of logic-based Knowledge Representation
2. Model simple application domains in a logic-based language
3. Understand several widely used knowledge representation languages

Bibliography:

Sr. No.	Book Details	Year of Publication
1.	R. J. Brachman, H. J. Levesque, *Knowledge Representation and Reasoning*, Elsevier	2004
2.	Handbook of Knowledge Representation. Frank van Harmelen, Vladimir Lifschitz and Bruce Porter (Eds). Foundations of Artificial Intelligence	2008
3.	Foundations of Semantic Web Technologies. Chapman & Hall/ CRC Textbooks in Computing. Pascal Hitzler, Markus Kroetsch, and Sebastian Rudolph	2009

Course Name	**:**	**Optimization Strategies in Big Data Analytics**
Course Code	**:**	**DS6601**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective (OE)**

Course Objectives :

- Analyze and address challenges in processing vast volumes of diverse big data efficiently.
- Apply diverse optimization techniques to enhance the efficiency of big data processing.
- Implement parallel computing models for effective task distribution in large-scale systems.
- Evaluate scalability concerns and implement resource management strategies for optimized performance.
- Integrate optimization methods seamlessly with machine learning algorithms for efficient data analytics.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	Introduction to Big Data, Definition and Characteristics, Significance in Modern Computing **Challenges in Big Data Processing:** Volume, Velocity, and Variety, Real-time Data Processing Challenges **Types of Optimization:** Time Optimization Strategies, Space Optimization Techniques, Resource Utilization Optimization	**8**
Unit 2	Parallelization Techniques and Models, Task Parallelism vs. Data Parallelism, Parallel Computing Architectures **Distributed Systems:** Concepts and Challenges, Principles of Distributed Computing, Challenges in Coordination and Communication Data and Task Parallelism, Strategies for Distributing Data and Tasks, Achieving Parallelism in Data Processing	**9**
Unit 3	**Optimization Algorithms in Big Data Analytics:** Optimizing Machine Learning Algorithms, Techniques for Enhancing ML Efficiency, Distributed Machine Learning Overview **Streamlining Data Processing Workflows:** Batch Processing vs. Real-time Processing, Workflow Optimization Strategies **Distributed Machine Learning Frameworks:** Overview of Frameworks (e.g., Apache Spark), Implementing Machine Learning at Scale	**9**
Unit 4	**Resource Management and Scalability:** Scalability in Big Data Systems, Horizontal vs. Vertical Scalability, Scaling Storage and Processing Systems	**8**

	Resource Allocation Strategies: Efficient Allocation of Computing Resources, Strategies for Dynamic Resource Management	
Unit 5	**Case Studies and Future Trends:** Real-world Optimization Challenges, Case Studies of Successful Optimization **Lessons Learned and Best Practices**: Lessons from Big Data Optimization Projects, Aalyzing Past Optimization Projects Identifying Common Challenges and Solutions	**8**

List of Experiments:		**Number of Turns**
1	**Benchmarking Big Data Processing:** Set up a benchmarking environment to assess the impact of optimization techniques on the performance of distributed processing systems. Analyze processing speed variations with different data sizes.	**2**
2	**Implementing Parallel Computing Models:** Develop a parallel computing application using Apache Hadoop, comparing task and data parallelism on a dataset. Assess scalability by increasing the number of computing nodes.	**2**
3	**Optimizing Machine Learning Algorithms:** Implement a distributed machine learning algorithm (e.g., Apache Spark MLlib), optimizing strategies to reduce training time and improve model accuracy. Compare performance with and without optimization.	**2**
4	**Resource Management in Distributed Systems:** Set up a Kubernetes cluster, implement dynamic resource allocation for a distributed application, and measure the impact on performance and resource utilization.	**2**
5	**Workflow Optimization in Real-time Processing:** Design and implement a real-time data processing workflow, integrating optimization techniques to enhance data flow efficiency. Evaluate the system's real-time streaming capabilities.	**2**
6	**Case Study Analysis:** Analyze real-world case studies of big data optimization projects, discussing challenges faced, applied optimization techniques, and project outcomes. Relate findings to course concepts	**2**
7	**Future Trends Exploration:** Explore emerging trends in big data optimization (e.g., edge computing, federated learning), discussing potential impacts on large-scale data processing optimization. Present insights in a group discussion or presentation format.	**2**

Course Outcomes:	
1	Students will demonstrate the ability to apply various optimization strategies in the context of big data processing, showcasing a higher-order thinking skill in practical problem-solving.
2	Develop skills in analyzing the performance of distributed systems, employing critical thinking to assess the impact of optimization techniques on processing speed and resource utilization.
3	Students will evaluate the scalability of distributed applications, employing critical analysis to assess how well the system adapts to increasing data sizes and computing nodes.

4	Demonstrate the ability to integrate optimization techniques into real-world scenarios, showcasing creativity in designing and implementing efficient workflows for both batch and real-time processing.
5	Develop a forward-thinking approach by exploring emerging trends in big data optimization, demonstrating the ability to forecast potential impacts on future large-scale data processing environments.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	**Book Title:** "Big Data: A Revolution That Will Transform How We Live, Work, and Think" **Author:** Viktor Mayer-Schönberger and Kenneth Cukier **Publisher:** Eamon Dolan/Mariner Books **Year:** 2013	2013
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"Big Data: A Very Short Introduction" Author: Dawn E. Holmes Publisher: Oxford University Press Year: 2016	2016
2	"Hadoop: The Definitive Guide" Author: Tom White Publisher: O'Reilly Media Year: 2015	2015
3	"Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking" Authors: Foster Provost and Tom Fawcett Publisher: O'Reilly Media Year: 2013	2013

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	**Coursera Course:** "Big Data Specialization" **Platform:** Coursera **Instructor:** University of California, San Diego **Link:** https://www.coursera.org/specializations/big-data	Coursera
2	**edX Course:** "Introduction to Big Data" **Platform:** edX **Institution:** University of California, San Diego **Link:** https://www.edx.org/professional-certificate/introduction-to-big-data	edX

Course Name	**:**	**Image Processing and Computer Vision**
Course Code	**:**	**DS6602**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective (OE)**

Course Objectives :

Students should be able -

- To understand the fundamentals of computer vision and image processing
- To understand various transformations on digital images and color models
- To apply various feature extraction methods and perform image segmentation
- To apply image classification methods for various real life applications

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	Introduction to computer vision and its applications, Image Formation, Sensing and Acquisition. Image Sampling and Quantization, Some Basic Relationships between Pixels, Intensity Transformations and Spatial Filtering, Convolution, Filtering in the Frequency Domain – Fourier Transform.	11
Unit 2	Color Models – Human Color Perception, RGB, CMY, CMYK, HSI, Conversion from RGB to HIS and vice versa. Edge and corner detection, Feature Descriptors - Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HoG), Feature Matching.	11
Unit 3	Image Segmentation and Grouping – Basic Methods, Watershed Algorithm, Segmentation using K-means, Mean Shift, Hough Transform, Fitting Lines and planes. Object tracking – Simple tracking strategies, tracking using matching.	10
Unit 4	Image Classification – Naïve Bayes Classifier, k-nearest neighbours, Support Vector Machine, Neural Networks, Brief Discussion on Convolutional Neural Networks (CNNs). Applications – Salient Object Detection, Face Detection, Optical Character Recognition.	10

List of Experiments:		**Number of Turns**
1	Learn fundamental commands in MATLAB/Python	1
2	Implement various intensity transformations	2
3	Implement spatial filters on digital images	2
4	Implement feature extraction methods	3
5	Implement various image segmentation algorithms	3
6	Implement different classification algorithms	3

Course Outcomes:	
1	Understand the fundamentals of computer vision and image processing
2	Understand various transformations on digital images and color models
3	Apply various feature extraction methods and perform image segmentation

4	Apply image classification methods for various real life applications

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Digital Image processing By Rafael C. Gonzalez and Richard E. Woods-Pearson Education.	Latest Edition
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Computer Vision: Algorithms and Applications, Richard Szeliski, Springer-Verlag	Latest Edition
2	Digital Image Processing by A.K. Jain, PHI	Latest Edtion
3	Practical Computer Vision Applications Using Deep Learning with CNNs: With Detailed Examples in Python Using TensorFlow and Kivy, Apress	Latest Edtion

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	https://nptel.ac.in/courses/108103174	NPTEL
2	https://www.coursera.org/learn/introduction-computer-vision-watson-opencv	Coursera

Semester VII

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	HS2701 / HS2702 / HS2703 / HS2704 / HS2705 / HS2706 / HS2707 / HS2708	Anthropology / Appreciation of Art / English Literature / History / Introduction to Art History/ Philosophy – an Introduction / Political Science / Public Administration	2 3 2 3 3 3 2 2	1 0 1 0 0 1 1 1	0 0 0 0 0 0 0 0	3	HSM-III
2. 3.	DS2701 to DS2708	Department Elective Course – III (DEC-III) / Department Elective Course – IV (DEC-IV)	3 3	0 0	2 2	4 4	DEC-III / DEC-IV
4.	DS6701 / DS5701	*Open Elective –III (OE-III) / Minor Specialization Course (MSC)	3	0	2	4	OE-III/MSC
5.	DS6702	*Open Elective –IV (OE-IV)	3	0	2	4	OE-IV
6.	DS7701	Major Project-I	0	0	4	2	Project
					Total Credits	21	

*Open Elective course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Department Elective Course - III (DEC-III) / Department Elective Course - IV (DEC-IV)

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2701	Agile Software Development and Modern Practices	3	0	2	4	DEC-III/DEC-IV
2.	DS2702	Deep Learning	3	0	2	4	DEC-III/DEC-IV
3.	DS2703	Mathematical Techniques for Data Science	3	0	2	4	DEC-III/DEC-IV
4.	DS2704	Graph Theory and Social Network Analysis	3	0	2	4	DEC-III/DEC-IV
5.	DS2705	Advanced Artificial Intelligence	3	0	2	4	DEC-III/DEC-IV
6.	DS2706	Cloud Data Management	3	0	2	4	DEC-III/DEC-IV
7.	DS2707	Mining Massive Data Sets	3	0	2	4	DEC-III/DEC-IV
8.	DS2708	Architectures for Management of Large Data Sets	3	0	2	4	DEC-III/DEC-IV

Open Elective – III (OE-III)*

Sr. No.	Course Code OE-III/MSC	Course Name	L	T	P	Credits	Category
1.	DS6701	Machine Learning for Data Science*	3	0	2	4	OE-III
2.	DS6702	Advanced Artificial Intelligence*	3	0	2	4	OE-III

*OE-III course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Open Elective – IV (OE-IV)*

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS6703	Python for Data Science*	3	0	2	4	OE-IV
2.	DS6704	Data and Information Security*	3	0	2	4	OE-IV

*OE-IV course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Course Name	**:**	**Agile Software Development and Modern Practices**
Course Code	**:**	**DS2701**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objectives:

- In software problem areas that require exploratory development efforts, those with complex requirements and high levels of change, agile software development practices are highly effective when deployed in a collaborative, people-centered organizational culture.
- This course examines agile methods, including Extreme Programming (XP), Scrum, Lean, Crystal, Dynamic Systems Development Method and Feature-Driven Development to understand how rapid realization of software occurs most effectively.

Course Contents: **Total No. Lectures: 42**

Sr. No.	Course contents	No. of Lectures
1	**Agile Processes:** Lean Production – SCRUM, Crystal, Feature Driven Development- Adaptive Software Development – Extreme Programming: Method Overview – Lifecycle – Work Products, Roles and Practices.	11
2	**Agility and Knowledge Management:** Agile Information Systems – Agile Decision Making – EarlS Schools of KM – Institutional Knowledge Evolution Cycle – Development, Acquisition, Refinement, Distribution, Deployment	11
3	**Leveraging – KM in Software Engineering** – Managing Software Knowledge – Challenges of Migrating to Agile Methodologies – Agile Knowledge Sharing – Role of Story-Cards – Story-Card Maturity Model (SMM).	10
4	**Agile Product Development** – Agile Metrics – Feature Driven Development (FDD) – Financial and Production Metrics in FDD – Agile Approach to Quality Assurance – Test Driven Development -Agile Approach in Global Software Development.	10

Lab Work: To follow project based learning approach for the course.

Sr. No.	Lab contents	No. of Hours
1	Understand the background and driving forces for taking an Agile Approach to Software Development.	4
2	Understand the business value of adopting agile approach.	3
3	Understand agile development practices and its practical implementation and how its different from classical approach	3
4	Apply Design principle and Refactoring to achieve agility	4

Course Outcomes:

At the end of the course, students will be able to:	
1.	Realize the importance of interacting with business stakeholders in determining the requirements for a software system
2.	Perform iterative software development processes: how to plan them, how to execute them.
3.	Point out the impact of social aspects on software development success.

4.	Develop techniques and tools for improving team collaboration and software quality.
5	Perform Software process improvement as an ongoing task for development teams.
6	Show how agile approaches can be scaled up to the enterprise level.

Bibliography:

Sr. No.	Book Detail	Year of Publication
1	David J. Anderson and Eli Schragenheim, Agile Management for Software Engineering: Applying the Theory of Constraints for Business Results, Prentice Hall	2019
2	Hazza and Dubinsky, Agile Software Engineering, Series: Undergraduate Topics in Computer Science, Springer	2019

Course Name	**:**	**DEEP LEARNING**
Course Code	**:**	**DS2702**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objectives:
To introduce the fundamentals of deep learning and the main research activities in this field. To learn architectures and optimization methods for deep neural network training

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
1	**LINEAR ALGEBRA REVIEW AND OPTIMIZATION:** Brief review of concepts from Linear Algebra, Types of errors, bias-variance trade-off, overfitting-underfitting, brief review of concepts from Vector Calculus and optimization, variants of gradient descent, momentum.	4
2	**LOGISTIC REGRESSION:** Basic concepts of regression and classification problems, linear models addressing regression and classification, maximum likelihood, logistic regression classifiers.	2
3	**NEURAL NETWORKS:** Basic concepts of artificial neurons, single and multi-layer perceptrons, perceptron learning algorithm, its convergence proof, different activation functions, softmax cross entropy loss function.	6
4	**CONVNETS:** Basic concepts of Convolutional Neural Networks starting from filtering. Convolution and pooling operation and arithmetic of these, Discussions on famous convnet architectures - AlexNet, ZFNet, VGG, GoogLeNet, ResNet, MobileNet-v1	8
5	**REGULARIZATION, BATCHNORM:** Discussion on regularization, Dropout, Batchnorm, Discussion on detection as classification, region proposals, RCNN architectures	8
6	**RECURRENT NEURAL NETWORKS:** Basic concepts of Recurrent Neural Networks (RNNs), backpropagation through time, Long-Short Term Memory (LSTM) architectures, the problem of exploding and vanishing gradients, and basics of word embedding,	8
7	**Autoencoders:** Autoencoders, Denoising autoencoders, sparse autoencoders, contractive autoencoders	6

Suggested References:

1. Ian Goodfellow, YoshuaBengio, Aaron Courville. Deep Learning, the MIT press, 2016
2. Bengio, Yoshua. " Learning deep architectures for AI." Foundations and trends in Machine Learning 2.1, Now Publishers, 2009

Course Outcomes: After completion of course, students would be able to:

1. Understand the fundamentals of deep learning
2. Compare various deep neural network architectures
3. Apply various deep learning algorithms based on real-world applications.

Lab Assignments:

S. No.	Experiment	No. of Turns
1.	Introduction to python libraries for deep learning: Keras, Tensorflow, Theano, OpenCV	4
2.	Implementation of MultiLayer Perceptron(MLP)	2
3.	Implementation Basic CNN model and various CNN architectures - transfer learning	2
4.	Implementation of Recurrent neural networks	2
5.	Deep Learning based project	4

Course Name	**:**	**Mathematical Techniques for Data Science**
Course Code	**:**	**DS2703**
Credits	**:**	**4**
LTP	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objectives: Students should able

1. To understand the foundational mathematical concepts relevant to data science.
2. To apply linear algebra techniques to analyze and manipulate data.
3. To utilize probability and statistics for data analysis and inference.
4. To apply calculus in the context of data science problems.
5. To understand optimization techniques and their applications in machine learning.

Total No. of Lectures- 42

Lecture Wise Breakup		**Number of Lectures**
1.	**LINEAR ALGEBRA AND MATRIX:** Overview and importance of Mathematical Concepts in data science. Representation of vectors; Linear dependence and independence; vector space and subspaces (definition, examples and concepts of basis); Matrices and their properties (determinants, traces, rank, nullity, etc.); Eigenvalues and eigenvectors with applications to data problems; Matrix factorizations; Inner products; Distance measures; Projections; linear transformations; matrices associated with linear transformations; Least square and minimum normed solutions. **Matrices in Machine Learning Algorithms:** Projection Transformation; Orthogonal Decomposition; Matrix Approximation Techniques: data redundancy: Singular Value Decomposition (SVD); Principal Component Analysis (PCA) and Linear Discriminant Analysis. (LDA)	**12**
2.	**Probability:** Basic concepts of probability and its axioms, conditional probability, total probability, independent events, Bayes' theorem, random variable, Central Limit Theorem, Probability distributions: Binomial, Poisson and Normal. **Statistics:** Introduction to statistics, Central tendencies, Standard Deviation variance, Similarity Measures: Pearson, Cosine and Spearman and Hypothesis testing.	**12**
3.	**Gradient Calculus:** Gradient Calculus: Basic concepts of calculus: partial derivatives, gradient, directional derivatives, Jacobian, Hessian.	**8**
4.	**Optimization:** Unconstrained and Constrained Optimization; Necessary and sufficiency conditions for optima; Gradient descent methods; Constrained optimization, KKT conditions; Introduction to non-gradient techniques; Introduction to least squares optimization; Optimization view of machine learning.	**10**

Project/Lab Work:	**Number of**

Each lab sessions will be aimed to translate the theory lectures into practical implementation through programming paradigms and tools, platforms available in the data science lab.	**Turns 14**

Course Outcomes: At the end of the course, the students should be able to
• Understand and apply Linear Algebra techniques to analyse and manipulate data. • Understand the basic concepts of probability theory and statistics methodologies to conduct data analysis and make informed inferences. • Employ calculus and optimization concepts to solve problems within the context of data science. • Develop critical thinking skills to choose and apply appropriate mathematical techniques in data science projects.

Text Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1.	*"Mathematics for Machine Learning*" by Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong	2020
2.	*Essential Math for Data Science* by Thomas Nield	2022
3.	*Fundamentals Of Mathematical Statistics* by S.C Gupta and V. K Kapoor	2014
MOOC's		
Sr. No.	**URL/Course Name**	**Name**
1.	https://nptel.ac.in/courses/111101004: Introduction to Probability Theory, IIT Bombay	NPTEL

Course Name	**:**	**Graph Theory and Social Network Analysis**
Course Code	**:**	**DS2704**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objectives :

This course aims to give an overview of graph theory, both theoretically and practically, with a focus on social networks.

Students should be able -

- To analyze network data using different methods and packages
- To explore personal network data and input it into social network analysis packages
- To apply social network theory to example data sets
- To develop and test network models
- To choose among social network designs based on research goals

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	Introduction to graphs, graph structures, graph representation, Basic theorems for vertices and edges, special types of simple graphs, bipartite graphs, subgraphs, Representing graphs, adjacency and incidence, Graph isomorphisms, Paths, connectedness and distance, Random walks on graph, Social networks as graphs.	**6**
Unit 2	Overview of social networks – network and society, landmarks of social networks – connectedness, degree distribution, centrality, small world phenomena, cliques and clusters, Social network models – *Erdös* and *Rényi* model, Preferential Attachment Models	**12**
Unit 3	Review of linear algebra, Computing eigen vectors, Eigenvector centralities, PageRank – matrix, score, PageRank centrality, Link Analysis – strong and weak ties, algorithms for link analysis.	**12**
Unit 4	Social network graph analysis, Techniques for graph analytics - Path Analytics, Connectivity Analytics, Community Analytics and Centrality Analytics, Community structures in networks – types of communities, community detection methods, community detection versus community search, SNAP system for large networks analysis and manipulation.	**12**

List of Experiments:		**Number of Turns**
1	Implementation of basic graph algorithms.	**2**
2	Format data for social network analysis and Network data collection using VennMaker.	**2**
3	Creating different types of Social Network Graphs in R	**1**

4	Graph data visualization using Gephi	**2**
5	Analyzing Network Graphs: • Calculating network density • Measuring betweenness • Measuring edge density • Taking random walks in graph • Analyzing graph through cliques Introduction to UCINET package for analysis of social network data	**3**
6	Introduction to NetworkD3 package and its application in creating different network	**2**
7	Generate following random graphs and analyze the pros and cons in the context of large-scale graphs.(1) Erdos-Renyi Model and (2) Albert-Barabasi Model graphs i.e. Preferential Attachment Model	**2**

Course Outcomes: At the end of the Course, the students should be able to	
1	Understand and apply algorithms for analyzing graph structures and properties.
2	Explore and input personal network data into analysis tools.
3	Apply social network theory to analyze and interpret real world datasets.
4	Build models of networks and test them against actual data.
5	Decide on social network designs based on specific research goals and understand the impact of different structures.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994, doi:10.1017/CBO9780511815478.	**1994**
2	J. Scott, Social Network Analysis, London and Beverley Hills, Sage Publications, 1992, doi:10.4135/9781529716597	**4th Edition 2017**
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Tanmoy Chakraborty. *Social Network Analysis*, Wiley	2021
2	Steven Borgatti, Martin Everett and Jeffrey Johnson. *Analyzing Social Networks*, Second Edition, Sage Publications.	2017
3	Tom Valente. *Social Networks and Health: Models, Methods and Applications*, First Edition, Oxford University Press.	2017

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	**https://www.coursera.org/learn/graphs:** Introduction to Graph Theory	Coursera
	https://www.coursera.org/learn/social-network-analysis: Social Network Analysis	Coursera

2	**https://onlinecourses.nptel.ac.in/noc22_cs117/preview:** Social Network Analysis	NPTEL

S.No.	Description	Blooms Level of Taxonomy
1.	Understand and apply algorithms for analyzing graph structures and properties.	Understand, Analze
2.	Explore and input personal network data into analysis tools	Understand, Apply
3.	Apply social network theory to analyze and interpret real world datasets.	Apply, Evaluate
4.	Build models of networks and test them against actual data.	Create, Analyze
5.	Decide on social network designs based on specific research goals and understand the impact of different structures.	Understand, Apply, Evaluate

Course Name	**:**	**ADVANCED ARTIFICIAL INTELLIGENCE**
Course Code	**:**	**DS2705**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objective:
• To cover modern paradigms of AI that go beyond traditional learning

Total No. of Lectures: 42

Lecture wise breakup		**No. of Lectures**
1	**Introduction:** Introduction to AI, expert systems, Utility theory, utility functions	5
2	**Making Decisions:** Decision networks, sequential decision problems, Partially Observable MDPs, Game Theory	9
3	**Reinforcement Learning:** Passive RL, Active RL, Generalization in RL, Policy Search	7
4	**Probabilistic Reasoning over time:** Hidden Markov Models, Kalman Filters	7
5	**Knowledge Representation:** Ontological engineering, Situation Calculus, semantic networks, description logic	6
6	**Planning:** Planning with state space search, Partial-Order Planning, Planning Graphs, Planning with Propositional Logic, hierarchical task network planning, non-deterministic domains, conditional planning, continuous planning, multi-agent planning	8

Total no. of turns: 14

List of Experiments:		**Number of turns**
1	Implement various decision making algorithms	5
2	Implement various Reinforcement Learning approach	2
3	Implement Hidden Markov Models	2
4	Implement various knowledge representation mechanisms	2
5	Implement various planning methodologies	3

Course Outcomes: At the end of the course, students will be able to:	
1	Develop an understanding of modern concepts in AI and where they can be used
2	Design, implement and apply novel AI techniques based on emerging real-world requirements

Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of publicatio n**

1	S. RUSSEL, P. NORVIG, Artificial Intelligence: A Modern Approach, Pearson, 3rd Edition.	2015
2	E. RICH, K. KNIGHT, S. B. NAIR, Artificial Intelligence, McGraw Hill Education, 3rd Edition	2017
3	R.S. SUTTON, A.G. BARTO, Reinforcement Learning: An Introduction, The MIT Press, 2nd Edition	2015

Course Name	**:**	**Cloud Data Management**
Course Code	**:**	**DS2706**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objectives :

Students should be able -

- To explore various data management techniques in cloud environments.
- To analyze various indexing and optimization techniques.
- To analyze various Cloud services for Big Data.
- To develop understanding for data integration, recovery and security methods in cloud environment.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**Introduction to Cloud Computing and Data Management:** Fundamental concepts, advantages and challenges of cloud services, Cloud Computing Principles, Models (IaaS, PaaS, SaaS), deployment types (public, private, hybrid), Major cloud service providers (AWS, Azure, Google Cloud). Introduction to databases and data warehouses. Data Management Models, Different data types, structures, formats, and the principles of data storage. File systems, block storage, object storage. Indexing and optimization techniques.	**8**
Unit 2	**Cloud Database Technologies and Big Data:** Relational databases (Amazon RDS, Amazon S3, Azure SQL Database), NoSQL databases (MongoDB, Cassandra), NewSQL databases. Big data concepts, Hadoop ecosystem, cloud-based big data services (e.g., AWS EMR, Azure HDInsight), real-time analytics for handling large-scale data.	**10**
Unit 3	**Data Integration and ETL in Cloud Environments:** Extract, Transform, Load (ETL) processes in the cloud, Integration tools (e.g., AWS Glue, Azure Data Factory), data pipelines, workflow orchestration techniques for seamless data movement and transformation.	**10**
Unit 4	**Data Backup, Recovery and Disaster Planning:** Data backup, fault tolerance mechanisms, disaster planning in cloud environments, designing robust disaster recovery plans.	**8**
Unit 5	**Data Security and Compliance in the Cloud:** Security best practices, encryption methods, access controls, data governance, compliance standards such as GDPR, HIPAA, methods that impact cloud data management.	**6**

List of Experiments:		**Number of Turns**
1	Project based lab work on accessing cloud services, understanding	14

	available service types for Big Data, integration tools, recovery and security mechanisms.	

Course Outcomes:	
1	To understand the use cases of cloud services.
2.	To analyze the storage mechanisms for organizing and managing data effectively.
3.	To apply best practices for effective big data management.
4.	To evaluate security mechanisms and recovery strategies for application domain.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1.	Building Cloud Data Platforms Solutions, An End-to-End Guide for Designing, Implementing, and Managing Robust Data Solutions in the Cloud, Anouar BEN ZAHRA.	2023
2.	Cloud Database Development and Management, Lee Chao, CRC Press Taylor & Francis Group.	2014
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	Advanced Data Management FOR SQL, NOSQL, CLOUD AND DISTRIBUTED DATABASES, Dr. Lena Wiese, Walter de Gruyter GmbH & Co KG.	2015
2	Web-Scale Data Management for the Cloud, Wolfgang Lehner, Kai-Uwe Sattler, Springer.	2013
3	Cloud Computing Explained: Implementation Handbook for Enterprises, John Rhoton, Recursive Press.	2013

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	GoogleCloud: Modernizing Data Lakes and Data Warehouses with Google Cloud. https://www.edx.org/learn/data-warehouse/google-cloud-modernizing-data-lakes-and-data-warehouses-with-google-cloud	edx
2	Cloud Computing Engineering and Management https://www.mooc-list.com/course/cloud-computing-engineering-and-management-edx	edx

Course Name	**:**	**Mining Massive Datasets**
Course Code	**:**	**DS 2707**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objective(s):

- This course aims to provide students with a comprehensive understanding of advanced techniques and methodologies for extracting valuable insights from large-scale datasets.
- The students will gain proficiency in various aspects of data mining, machine learning, and computational approaches applied to massive datasets.

Total No. Lectures—42

Sr. No.	Course contents	No. of Lectures
1.	**Data Mining:** Introduction, Statistical Modeling, Machine Learning, Computational Approaches to Modeling, Feature Extraction, Statistical Limits on Data Mining, Hash Functions, Indexes, Natural Logarithms, Power Laws.	9
2.	**Map Reduce and the New Software Stack:** Distributed File Systems, Map Reduce, Algorithms Using MapReduce, Extensions to MapReduce, Complexity Theory for MapReduce.	7
3.	**Mining Data Streams:** The Stream Data Model, Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream, Counting Ones in a Window, Decaying Windows.	8
4.	**Frequent Item sets:** The Market-Basket Model, Market Baskets and the A-Priori Algorithm, Handling Larger Datasets in Main Memory, Limited-Pass Algorithms, Counting Frequent Items in a Stream	8
5.	**Clustering:** Introduction to Clustering Techniques, Hierarchical Clustering, K-means Algorithms, The CURE Algorithm, Clustering in Non-Euclidean Spaces, and Clustering for Streams and Parallelism.	5
6.	**Dimensionality Reduction:** Eigen values and Eigenvectors of Symmetric Matrices, Principal-Component Analysis, Singular-Value Decomposition, CUR Decomposition	5

Total Lab Hours: 28

Sr. No.	Title of the Experiment	No. of Hours
1.	Explore a large dataset, identify key features, and perform necessary preprocessing steps to clean and prepare the data for analysis.	2
2.	Apply statistical modeling techniques to analyze the dataset, including measures of central tendency, dispersion, and correlation.	2
3.	Implement and compare different machine learning algorithms on the dataset, evaluating their performance and selecting the most suitable model.	2
4.	Implement computational approaches to model the dataset, such as regression or decision trees, and assess their accuracy and efficiency.	2

5.	Develop a MapReduce program to process and analyze a large dataset, gaining hands-on experience with distributed computing.	2
6.	Implement and experiment with various algorithms using the MapReduce paradigm, understanding their applications and performance characteristics.	2
7.	Explore and implement extensions to MapReduce, such as iterative algorithms or graph processing, to address more complex data mining tasks.	2
8.	Implement algorithms to process and analyze streaming data, addressing challenges such as windowed counting and filtering.	2
9.	Apply the A-Priori algorithm to discover frequent itemsets in a dataset	2
10.	Implement hierarchical clustering and K-means algorithms on datasets	2
11.	Explore and implement clustering techniques suitable for non-Euclidean spaces, such as density-based clustering or spectral clustering.	2
12.	Apply dimensionality reduction techniques, including PCA and SVD, to reduce the feature space and analyze the impact on model performance.	2
13.	Implement CUR decomposition as an alternative dimensionality reduction technique	2
14.	Solve any complex real-world problem, emphasizing the importance of combining methods for comprehensive data mining.	2

Course Outcomes: Students will be able to

1.	Apply their knowledge of statistical modeling, machine learning algorithms, and computational approaches to analyze and interpret massive datasets.
2.	Critically analyze the complexities of MapReduce, distributed file systems, and algorithms used in data mining.
3.	Design and implement solutions for real-world problems in data mining.
4.	Synthesize knowledge from various dimensions of the course, including dimensionality reduction techniques, data stream mining, and advanced clustering algorithms.

Books:

Sr. No.	Name of Book/ Authors/ Publisher	Year of Publication
1.	Mining of Massive Datasets, Second Edition, by by Jure Leskovec, Anand Rajaraman and Jeffrey David Ullman	2016
2.	Data Mining and Machine Learning: Fundamental Concepts and Algorithms by Mohammed J. Zaki, Wagner Meira and Jr, Wagner Meira	2020
3.	Algorithms and Data Structures for Massive Datasets by Dzejla Medjedovic, Emin Tahirovic	2022
4.	Mining Massive Data Sets for Security: Advances in Data Mining, Search, Social Networks and Text Mining, and Their Applications to Security by Françoise Fogelman-Soulié	2008
5.	Handbook of Statistical Analysis and Data Mining Applications by Ken Yale, Robert Nisbet, Gary D. Miner	2017

Available MOOCS:

Sr. No.	Name of Book/ Authors/ Publisher	Provided by
1.	https://online.stanford.edu/courses/soe-ycs0007-mining-massive-data-sets	Stanford
2.	https://onlinecourses.nptel.ac.in/noc21_cs06/preview	NPTEL

Course Name	**:**	**Architectures for Management of Large Data Sets**
Course Code	**:**	**DS2708**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course III / IV (DEC-III/DEC-IV)**

Course Objectives :

1. Analyze and compare various architectural patterns for managing large datasets, focusing on scalability and industry applicability.
2. Comprehend parallel processing models and distributed computing systems for efficient data computation in batch and stream processing.
3. Differentiate between RDBMS, NoSQL databases, and storage systems like Hadoop HDFS for managing vast datasets.
4. Apply data warehousing, OLAP, and machine learning pipelines for real-time decision-making and large-scale data analysis.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**Introduction to Data Management Architectures:** Overview of Large Data Sets and their Characteristics, Architectural Patterns for Data Management, Scalability, Reliability, and Availability in Architectures, Introduction to Distributed Systems and Parallel Computing	**12**
Unit 2	**Database and Storage Architectures:** Relational Database Management Systems (RDBMS) vs. NoSQL Systems, Overview of Storage Architectures: SAN, NAS, and Object Storage, Distributed File Systems: Hadoop HDFS and its Architecture, Replication and Sharding in Database Architectures	**10**
Unit 3	**Processing and Computation Architectures:** Parallel Processing Models: MapReduce and its variations, In-Memory Computing Architectures: Apache Spark and Redis, Stream Processing Architectures: Apache Kafka and Stream Analytics, GPU-Accelerated Computing and High-Performance Computing (HPC) Architectures	**10**
Unit 4	Architectures for Big Data Analytics and Machine Learning: Big Data Analytics Architectures: Data Warehousing and OLAP Cubes, Machine Learning Architectures: Training and Inference Pipelines, Real-time Analytics and Real-time Decision-Making Architectures, Case Studies: Architectures in Cloud Computing Platforms	**10**

List of Experiments:		**Number of Turns**
1	Performance Evaluation of Distributed Storage: Compare Hadoop Distributed File System (HDFS) and object storage systems to assess their scalability and fault tolerance handling large datasets.	**2**
2	Data Replication and Sharding Techniques: Implement data replication and sharding methods in a database system to manage growing data volumes effectively.	**2**

3	Parallel Computing with MapReduce/Spark: Develop programs using MapReduce or Apache Spark to understand parallel computing techniques in distributed architectures.	**2**
4	Machine Learning Pipeline on Large Datasets: Construct and evaluate a machine learning pipeline on distributed data, analyzing the efficiency of various algorithms.	**2**
5	Real-time Stream Processing using Kafka: Design and analyze real-time processing pipelines using Apache Kafka for stream processing architectures.	**2**
6	Data Warehousing Implementation: Implement and evaluate a data warehousing solution to manage structured and unstructured data for analytics purposes.	**2**
7	Scaling Analytics Workloads on Cloud: Explore cloud-based services to scale and manage large data analytics workloads, understanding their capabilities and limitations.	**2**

Course Outcomes:	
1	Understand the pros and cons of the traditional relational databases and need for distributed architecture for data processing of large datasets.
2	Implement map reduce programs to solve data analysis tasks.
3	Explain the need and use cases for emerging architectures such as Spark, Storm, Giraph, Hive etc. and how they differ from Apache Hadoop
4	Understand and explain distributed software architectures, runtime and storage strategies used by Apache Hadoop.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley, 2013	2013
2	V.K. Jain, Big Data and Hadoop, Khanna Book Publishing Company 2020.	2020
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"Designing Data-Intensive Applications" by Martin Kleppmann Publisher: O'Reilly Media, Year: 2017	2017
2	"Hadoop: The Definitive Guide" by Tom White, Publisher: O'Reilly Media, Year: 2015	**2015**
3	"Distributed Systems: Principles and Paradigms" by Andrew S. Tanenbaum and Maarten Van Steen, Publisher: Pearson, Year: 2017	**2017**

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Fundamentals of Software Architecture for Big Data by University of	Coursera

	Colorado Boulder on Coursera Online Link: https://www.coursera.org/learn/software-architecture-for-big-data-fundamentals	
2	Fundamental Course of Data Architecture 2.0 by Dr Jose J Link:https://www.udemy.com/course/data-architecture-big-data-architecture-data-mesh-database-management/	Udemy

Course Name	**:**	**Machine Learning for Data Science**
Course Code	**:**	**DS6701**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of course	**:**	**Open Elective – III (OE-III)**

Course Objectives :

Students should be able -

- To analyze and apply mathematical concepts in the context of data science and machine learning.
- To explore various data acquisition, cleaning, and preprocessing techniques.
- To develop skills in applying machine learning algorithms and predictive modeling on various datasets.
- To prepare students for lifelong learning in the rapidly evolving field of data science and machine learning.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**Introduction:** Basics of python for data science and machine learning, Types of data: structured data, unstructured data, other data types, the data science process, Explore and visualize data, features, Linear algebra, correlation	**6**
Unit 2	**Data Acquisition and Preprocessing:** Data sources and collection, Data storage and management, Data Preparation: handling missing data, feature encoding, feature scaling, Data Cleaning, Dataset division: train, test and validation sets	**7**
Unit 3	**Machine Learning Fundamentals:** Applications of Machine Learning, Introduction to Machine Learning Techniques: Supervised Learning, Unsupervised Learning and Reinforcement Learning, bias-variance tradeoff, overfitting-underfitting, Evaluation Measures: SSE, R2, confusion matrix, precision, recall, F-Score, ROC-Curve, cross-validation.	**7**
Unit 4	**Regression and Classification:** Simple Linear Regression, Multiple Linear Regression, Logistic Regression, gradient descent algorithm, K-Nearest Neighbor, Support Vector Machine (SVM), Naïve Bayes, Decision Trees.	**12**
Unit 5	**Dimensionality Reduction and Clustering:** Feature selection, Principle Component Analysis, K-means clustering, Hierarchical clustering, Density based clustering.	**10**

List of Experiments:		**Number of Turns**
1	Explore various python libraries used for data handling and machine learning techniques: numpy, pandas, matplotlib, scikit-learn etc.	**2**
2	Implementation of data handling and preprocessing techniques	**2**
3	Implementation of various classification and regression algorithms – Linear and logistic regression, KNN, SVM, Naïve Bayes, Decision Trees	**6**
4	Implementation of Dimensionality Reduction and Clustering techniques:	**4**

	PCA, K-means clustering, Density based clustering algorithms	

Course Outcomes: **After completion of course, students would be able to:**	
1	Explore and apply mathematical and engineering knowledge to analyze complex data science problems.
2	Apply modern tools and techniques to solve data science problems.
3	Apply different machine learning models on various datasets.
4	Analyze the trade-off of various machine learning models for real-life applications.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"Introduction to Data Science" by Jeffrey Stanton, SAGE Publications	2017
2	"Introduction to Machine Learning with Python: A Guide for Data Scientists", Andreas Müller, Sarah Guido, O'Reilly Media	2016
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"The Machine Learning Yearning" by Andrew NG, Deeplearning.ai	2021
2	"Python for Data Analysis" by Wes McKinney, O'Reilly Media	2022
3	"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems", by Aurélien Géron, 3rd Edition, O'Reilly	2022

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	Machine Learning Specialization - Stanford University https://www.coursera.org/specializations/machine-learning-introduction	Coursera
2	Data Science Specialization - Johns Hopkins University https://www.coursera.org/specializations/jhu-data-science	Coursera

Course Name	**:**	**ADVANCED ARTIFICIAL INTELLIGENCE**
Course Code	**:**	**DS6702**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective – III (OE-III)**

Course Objective:
• To cover modern paradigms of AI that go beyond traditional learning

Total No. of Lectures: 42

Lecture wise breakup		**No. of Lectures**
1	**Introduction:** Introduction to AI, expert systems, Utility theory, utility functions	5
2	**Making Decisions:** Decision networks, sequential decision problems, Partially Observable MDPs, Game Theory	9
3	**Reinforcement Learning:** Passive RL, Active RL, Generalization in RL, Policy Search	7
4	**Probabilistic Reasoning over time:** Hidden Markov Models, Kalman Filters	7
5	**Knowledge Representation:** Ontological engineering, Situation Calculus, semantic networks, description logic	6
6	**Planning:** Planning with state space search, Partial-Order Planning, Planning Graphs, Planning with Propositional Logic, hierarchical task network planning, non-deterministic domains, conditional planning, continuous planning, multi-agent planning	8

Total no. of turns: 14

List of Experiments:		**Number of turns**
1	Implement various decision making algorithms	5
2	Implement various Reinforcement Learning approach	2
3	Implement Hidden Markov Models	2
4	Implement various knowledge representation mechanisms	2
5	Implement various planning methodologies	3

Course Outcomes: At the end of the course, students will be able to:	
1	Develop an understanding of modern concepts in AI and where they can be used
2	Design, implement and apply novel AI techniques based on emerging real-world requirements

Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of publication**
1	S. RUSSEL, P. NORVIG, Artificial Intelligence: A Modern Approach,	2015

	Pearson, 3rd Edition.	
2	E. RICH, K. KNIGHT, S. B. NAIR, Artificial Intelligence, McGraw Hill Education, 3rd Edition	2017
3	R.S. SUTTON, A.G. BARTO, Reinforcement Learning: An Introduction, The MIT Press, 2nd Edition	2015

Course Name	**:**	**PYTHON FOR DATA SCIENCE**
Course Code	**:**	**DS6703**
Credit	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective – IV (OE-IV)**

Course Objectives:

- Basic process of data science.
- Python and Jupyter notebooks.
- An applied understanding of how to manipulate and analyse uncurated datasets.
- Basic statistical analysis and machine learning methods.
- How to effectively visualize results.

Total No. of Lecture - 42

Lecture wise breakup		**No. of Lectures**
1	**INTRODUCTION:** Introduction to data science, Data acquisition, data pre-processing and preparation, data quality, Tools for data science, significance of python for data science.	**10**
2	**PYTHON PROGRAMMING:** Python Basics, Setting up python, Python Data Structures: lists and tuples, dictionaries, sets; Jupyter, Spyder.	**08**
3	**PYTHON PROGRAMMING FUNDAMENTALS:** Conditions and Branching, Loops, Functions, Objects and Classes, reading files with open, writing files with open.	**10**
4	**PYTHON LIBRARIES FOR DATA SCIENCE:** Numpy, loading data with Pandas, working with and Saving data with Pandas, MatplotLib, Seaborn, Scikit-Learn.	**10**
5	**CASE STUDIES:** Regression, Classification.	**04**

Course Outcomes:	
1	Understanding various process of data science.
2	Design algorithm and code using Jupyter notebook.
3	Experimentation of data analysis using python tools.
4	Visualization results using python libraries.

Project/Lab Work:	**Number of Turns**
Each lab sessions will be aimed to translate the theory lectures into practical implementation through programming paradigms and tools, platforms available in the data science lab.	**14**

Suggested Books:	
1	Python Data Science Handbook: Essential Tools for Working with by Jake VanderPlas, O'Reilly Media, Inc.
2	Python Data Science Handbook, by Jake VanderPlas, O'Reilly Media, Inc.
3	Python for Data Science For Dummies, Book by John Mueller and Luca Massaron, John Wiley & Sons

4	Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, Wes McKinney, Shroff/O'Reilly.

Online MooC course https://www.edx.org/course/python-for-data-science-2

Course Name	**:**	**Data and Information Security**
Course Code	**:**	**DS6704**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective - IV (OE-IV)**

Course Objectives

The main objectives of this course are:	
1	To understand basic design principals of symmetric and asymmetric cryptography and learn how standard cryptanalytic attacks work and thereby how to avoid common design flaws.
2	The student should be able to apply basic number theory in cryptography and will be able to understand the design principals of symmetric and asymmetric cryptography, AES, RSA
3	The course will be emphasizing on the algorithmic complexity and security vs performance trade off.
4.	Learn different perspectives of managing Keys, trust and Privacy

Course Contents

Lecture wise breakup		**Number of Lectures**
1	**Introduction** - Security mind-set, Computer Security Concepts (CIA), Threats, Attacks Assets, Services and Mechanisms.	3
2	**Foundations** History and evolution of cryptosystems, Classical Encryption Systems, Substitution Ciphers and Transposition Cipher, Block cipher, Stream cipher, Modes of operation, Symmetric and Asymmetric cryptography.	4
3	**Symmetric Key Cryptosystems**: Advanced Encryption Standards Mathematical Theory, Number Theory, Data Encryption Standard: Description of DES, Security of DES, Differential And Linear Cryptanalysis, Design Criteria, DES Variants, DES modes of operation, Finite fields, Galois's field, Multiplicative Inverse, Extended Euclid's algorithm, AES, Other Stream Ciphers	10
4	**Asymmetric key Cryptosystems** Prime Number Generation, Discrete Logarithms in a Finite Field, Chinese Remainder Theorem, DIFFIE-HELLMAN Key Exchange, RSA Key Setup, Encryption and verification, elgamal, and Elliptic Curve Cryptography	8
5	**Key Management, Message Authentication and Digital Signatures** – Key Management, Notion of Trust, Centralised vs Distributed Trust management, Distributed Third Party, Kerberos, PKI, PGP, Certificate Distribution, Management and Revocation	8
6	**Message Authentication and Digital Signatures** Digital Signatures and Certificates, Hash and MAC, Digital Signatures, RSA and DSS based digital Signatures	5
	IDS and Firewalls Intruders, Virus, Worms, Firewalls-need and features of firewall, Types of firewall, Intruder Detection Systems.	4

Lab work

Sr No	Lab Contents	No. of Hours
1	Implementation of key exchange, symmetric and asymmetric algorithms, signature schemes.	14
2	Practical Use of Network Security Tools, Email Header Analysis, Packet sniffing, configuration of network security equipment such as firewall, routers, IDS, Wireless Access Points. Explore security solutions for Web application vulnerabilities.	14

Course Outcomes: At the end of the course, students will be able to:	
1	Understand the cryptographic principles
2	Able to apply the basic rules of public key and symmetric encryption for practical cryptographic problems
4	To evaluate various security issues and understand solutions for it.
5	Given a problem in cryptography, be able to design an algorithm to implement the solution to that problem.

References

Sr No	Book detail	Year of Publication
1	Applied Cryptography protocols, algorithms, and source code in C, Second Edition, Bruce Schneier, John Wiley & Sons	1996 (reference book)
2	Cryptography and Network Security by William Stallings, Prentice Hall. ISBN 0-13-187316-4	2006 (text book)
3	Handbook of Applied Cryptography by Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone, CRC Press	2000 (reference book)
4	Cryptography Theory and Practice by Douglas R. Stinson: 3rd ed., Chapman & Hall/CRC.	2006 (reference book)

Semester VIII

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	HS2801 / HS2802 / HS2803	Business Environment and Business Laws / Entrepreneurship and Project Management / Financial Management	2 2 2	1 1 1	0 0 0	3	HSM-IV
2.	DS2801 to DS2805	Department Elective Course – V (DEC-V)	3	0	2	4	DEC-V
3.	DS6801/ DS5801	*Open Elective – V / Minor Specialization Course (MSC)	3	0	2	4	OE-V/MSC
4.	DS6802	*Open Elective –VI	3	0	2	4	OE-VI
5.	DS8801	Discipline				2	Discipline
6.	PC7801	Proficiency-II				2	Proficiency
7.	DS7801	Major Project-II	0	0	8	4	Project
					Total Credits	23	

*Open Elective course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence)

Department Elective Course - V (DEC-V)

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS2801	Introduction to Data Analytics and Visualization	3	0	2	4	DEC-V
2.	DS2802	Statistical Thinking for Data Science	3	0	2	4	DEC-V
3.	DS2803	Recommendation Systems	3	0	2	4	DEC-V
4.	DS2804	Image Processing and Computer Vision	3	0	2	4	DEC-V
5.	DS2805	Optimization Strategies in Big Data Analytics	3	0	2	4	DEC-V

***Open Elective – V (OE-V) / Minor Specialization Course (MSC)**

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS6801	Introduction to IOT*	3	0	2	4	OE-V

*OE-V course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence).

Open Elective – VI (OE-VI)*

Sr. No.	Course Code	Course Name	L	T	P	Credits	Category
1.	DS6802	Management of Large Data Sets	3	0	2	4	OE-VI

*OE-VI course will not be offered to B. Tech. (CSE), B. Tech. CSE (Data Science) and B. Tech. CSE (Artificial Intelligence).

Course Name	**:**	**INTRODUCTION TO DATA ANALYTICS AND VISUALISATION**
Course Code	**:**	**DS2801**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course – V (DEC-V)**

Course Objectives:

- This course is intended to provide a broad overview of data analysis and visualization techniques.
- Students will be given hands-on training in data analytics to build descriptive and predictive models, and validating their models against the actual outcomes.
- The course will provide various techniques and tools for data analysis of noisy, real life data.
- Students will be taught how to perform data wrangling, cleaning, and sampling to get a suitable data set; exploratory data analysis; generating hypotheses and building intuition; prediction or statistical learning; communication – summarizing results through various visualization techniques and providing interpretable summaries.

Total No. of Lectures – 42

Lecture wise breakup		**Number of Lectures**
1.	**INTRODUCTION:** Data science workflow, Automated methods for data collection, Data and Visualization Models, Data wrangling and cleaning, Exploratory data analysis, Dimensionality Reduction.	**10**
2.	**MODELLING AND ANALYSIS** Building and evaluation of models for: Association Analysis, Recommendation Systems, Time-series data, Text Analysis.	**10**
3.	**DATA VISUALIZATION** Visualization Software and Tools, Visualization Design, Multidimensional Data, Graphical Perception, Interaction dynamics for Visual Analysis, Using Space Effectively, Stacked Graphs, Geometry & Aesthetics.	**12**
4.	**GRAPH VISUALIZATION** Networks, Graph Visualization and navigation in information Visualization, mapping & Cartography, Text Visualization	**10**

List of experiments:		**No. of Turns**
1	Learn how to collect data via web-scraping, APIs and data connectors from suitable sources as specified by the instructor	1
2	Perform various types of data cleaning operations on the data collected in previous lab using data exploration, imputation etc.	1
3	Perform dimensionality reduction on a given dataset and create various visualizations like histograms, scatter-plots, etc.	2
4	Perform association analysis on a given dataset and evaluate its	2

	accuracy.	
5	Build a recommendation system on a given dataset and evaluate its accuracy.	2
6	Build a time-series model on a given dataset and evaluate its accuracy.	2
7	Build cartographic visualization for multiple datasets involving various countries of the world; states and districts in India etc.	2
8	Perform text mining on a set of documents and visualize the most important words in a visualization such as word cloud.	2

Course Outcomes: At the end of the course, students will be able to:

1	Explain and demonstrate various techniques for automatic data collection, data cleaning and exploration using visualizations.
2	Implement data collection, data cleaning and exploration techniques in a programming language.
3	Understand and apply modeling and analysis techniques for various types of datasets including e-commerce transactions, review datasets, time series datasets, text documents etc.
4	Evaluate different models and their strengths and weakness for a given dataset and task.
5	Select methods and create effective visualizations to explain the artifacts in the data, distributions of attributes, relationships between the attributes, efficacy of the models and predictions generated by it
6	Become proficient in data analysis tasks involving real-life datasets with noise.

Suggested Books:

Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Skiena, Steven S, The Data Science Design Manual, CRC press.	2017
2	Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, Introduction to Data Mining (Second Edition).	2021
3	V.K. Jain, Data Science and Analytics (with Python, R and SPSS Programming), Khanna Book Publishing Company.	2019
4	V.K. Jain, Big Data and Hadoop, Khanna Book Publishing Company.	2022
5	Tamara Munzner, “Visualization Analysis and Design”, A K Peters/CRC Press; 1st edition	2014
6	Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley.	2013
7	Matthew O. Ward, Georges Grinstein, Daniel Keim, “Interactive Data Visualization: Foundations, Techniques, and Applications”, 2nd Edition, CRC press.	2015

COURSE Name	**:**	**STATISTICAL THINKING FOR DATA SCIENCE**
Course Code	**:**	**DS2802**
Credits	**:**	**4**
LTP	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course – V (DEC-V)**

Course Objectives:
The objective of the course is to make students learn about:
• Data collection, analysis and inference followed by data classification.
• Conditional Probability -How to judge the probability of an event and Bayesian modelling.
• Basics of Linear Regression Model and its variants.
• Data Visualization- Standard tools for graphical visualization predicting patterns.

Total No. of Lecture - 42

Lecture wise breakup		**No. of Lectures**
1	**INTRODUCTION:** Introduction to Data Science, Examples of Statistical Thinking, Numerical Data, Summary Statistics: Population to Sampled Data, Different Types of Biases, Introduction to Probability, Introduction to Statistical Inference.	**10**
2	**STATISTICAL THINKING:** Association and Dependence, Association and Causation, Conditional, Probability and Bayes Rule, Simpsons Paradox, Confounding, Introduction to Linear Regression, Special Regression Models.	**12**
3	**EXPLORATORY DATA ANALYSIS AND VISUALIZATION:** Statistical Graphics and Data Visualization, Graphs of Data, Graphs of Fitted Models, Graphs to Check Fitted Models, Principles of graphics, Data Visualization tools.	**10**
4	**INTRODUCTION TO BAYESIAN MODELING:** Bayesian Inference: combining models and data in a forecasting problem, Bayesian hierarchical modeling for studying public opinion, Bayesian modeling for Big Data.	**10**

Project/Lab Work: Each lab sessions will be aimed to translate the theory lectures into practical implementation through programming paradigms and tools, platforms available in the data science lab. Tool: R studio	**14**

Course Outcomes:
• Development of design proficiency with statistical analysis of data. • Understanding Statistical analysis with open source software. • Apply data science concepts and methods to solve problems in real-world contexts and will communicate these solutions effectively. • Students can carry out standard data visualization and formal inference procedures and can comment on the results.

Suggested Books:	
1	Probability and Statistics for Engineers, Walpole, Pearson.
2	Think Stats: Exploratory Data Analysis, Allen B. Downey, O'Reilly.
3	Data Analytics, Anil Maheshwari, McGrawHill.

Course Name	**:**	**Recommendation Systems**
Course Code	**:**	**DS2803**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course – V (DEC-V)**

Course Objectives:

- To learn techniques for making recommendations, including non-personalized, content-based, and collaborative filtering
- To automate a variety of choice-making strategies with the goal of providing affordable, personal, and high-quality recommendations

Course Contents: **Total No. Lectures: 42**

Sr. No.	Course contents	No. of Lectures
1.	**Introduction** Overview of Information Retrieval, Retrieval Models, Search and Filtering Techniques: Relevance Feedback, User Profiles, Recommender system functions, Matrix operations, covariance matrices, Understanding ratings, Impact of Bad Ratings, Applications of recommendation systems, Issues with recommender system	08
2.	**Content-based Filtering** High level architecture of content-based systems, Advantages and drawbacks of content based filtering, Item profiles, Discovering features of documents, pre-processing and feature extraction, Obtaining item features from tags, Methods for learning user profiles, Similarity based retrieval, Classification algorithms.	08
3.	**Collaborative Filtering** User-based recommendation, Item-based recommendation, Model based approaches, Matrix factorization, Attacks on collaborative recommender systems.	08
4.	**Hybrid approaches** Opportunities for hybridization, Monolithic hybridization design: Feature combination, Feature augmentation, Parallelized hybridization design: Weighted, Switching, Mixed, Pipelined hybridization design: Cascade Meta-level, Limitations of hybridization strategies	08
5.	**Evaluating Recommender System** Introduction, General properties of evaluation research, Evaluation designs: Accuracy, Coverage, confidence, novelty, diversity, scalability, serendipity, Evaluation on historical datasets, Offline evaluations.	06
6.	Trust-Based Recommendation, Recommending for Groups, Context-Aware Recommendation, Cross-domain Recommendations	04

Lab Work: To follow project based learning approach for the course.

Sr. No.	Lab contents	No. of Turns
1.	Program to explore n**on-personalized data and features**	2
2.	Program for rating conversion from normal data	2
3.	**To compute** Mean Rating, Rating Count , correlation and **generate recommendations**	2
4.	To compute Pearson correlation and cosine similarity for given dataset	2
5.	Write program to that make predictions for the test item-user pair and use user-based or item-based Collaborative Filtering	4
6.	To evaluate recommendations on the basis of accuracy and other metrics	2

Course Outcomes:

At the end of the course, students will be able to:	
1.	Design recommendation system for a particular application domain
2.	Evaluate recommender systems on the basis of metrics such as accuracy, rank accuracy, diversity, product coverage, and serendipity

Bibliography:

Sr. No.	Book Detail	Year of Publication
1.	Recommender Systems: An Introduction, Cambridge University Press 1st ed. Jannach D., Zanker M. and FelFering A.	2011
2.	Recommender Systems: The Textbook, Springer (2016), 1st ed. Charu C. Aggarwal	2016
3.	Recommender Systems Handbook, Springer(2011), 1st ed. Ricci F., Rokach L., Shapira D., Kantor B.P	2015
4.	Recommender Systems For Learning, Springer (2013), 1st ed. Manouselis N., Drachsler H., Verbert K., Duval E.	2013

Course Name	**:**	**Image Processing and Computer Vision**
Course Code	**:**	**DS2804**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course – V (DEC-V)**

Course Objectives :

Students should be able -

- To understand the fundamentals of computer vision and image processing
- To understand various transformations on digital images and color models
- To apply various feature extraction methods and perform image segmentation
- To apply image classification methods for various real life applications

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	Introduction to computer vision and its applications, Image Formation, Sensing and Acquisition. Image Sampling and Quantization, Some Basic Relationships between Pixels, Intensity Transformations and Spatial Filtering, Convolution, Filtering in the Frequency Domain – Fourier Transform.	11
Unit 2	Color Models – Human Color Perception, RGB, CMY, CMYK, HSI, Conversion from RGB to HIS and vice versa. Edge and corner detection, Feature Descriptors - Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HoG), Feature Matching.	11
Unit 3	Image Segmentation and Grouping – Basic Methods, Watershed Algorithm, Segmentation using K-means, Mean Shift, Hough Transform, Fitting Lines and planes. Object tracking – Simple tracking strategies, tracking using matching.	10
Unit 4	Image Classification – Naïve Bayes Classifier, k-nearest neighbours, Support Vector Machine, Neural Networks, Brief Discussion on Convolutional Neural Networks (CNNs). Applications – Salient Object Detection, Face Detection, Optical Character Recognition.	10

List of Experiments:		**Number of Turns**
1	Learn fundamental commands in MATLAB/Python	1
2	Implement various intensity transformations	2
3	Implement spatial filters on digital images	2
4	Implement feature extraction methods	3
5	Implement various image segmentation algorithms	3
6	Implement different classification algorithms	3

Course Outcomes:	
1	Understand the fundamentals of computer vision and image processing
2	Understand various transformations on digital images and color models
3	Apply various feature extraction methods and perform image segmentation

4	Apply image classification methods for various real life applications

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Digital Image processing By Rafael C. Gonzalez and Richard E. Woods- Pearson Education.	Latest Edition
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication/ Reprint**
1	Computer Vision: Algorithms and Applications, Richard Szeliski, Springer-Verlag	Latest Edition
2	Digital Image Processing by A.K. Jain, PHI	Latest Edtion
3	Practical Computer Vision Applications Using Deep Learning with CNNs: With Detailed Examples in Python Using TensorFlow and Kivy, Apress	Latest Edtion

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	https://nptel.ac.in/courses/108103174	NPTEL
2	https://www.coursera.org/learn/introduction-computer-vision-watson-opencv	Coursera

Course Name	**:**	**Optimization Strategies in Big Data Analytics**
Course Code	**:**	**DS2805**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Department Elective Course – V (DEC-V)**

Course Objectives :

- Analyze and address challenges in processing vast volumes of diverse big data efficiently.
- Apply diverse optimization techniques to enhance the efficiency of big data processing.
- Implement parallel computing models for effective task distribution in large-scale systems.
- Evaluate scalability concerns and implement resource management strategies for optimized performance.
- Integrate optimization methods seamlessly with machine learning algorithms for efficient data analytics.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	Introduction to Big Data, Definition and Characteristics, Significance in Modern Computing **Challenges in Big Data Processing:** Volume, Velocity, and Variety, Real-time Data Processing Challenges **Types of Optimization:** Time Optimization Strategies, Space Optimization Techniques, Resource Utilization Optimization	**8**
Unit 2	Parallelization Techniques and Models, Task Parallelism vs. Data Parallelism, Parallel Computing Architectures **Distributed Systems:** Concepts and Challenges, Principles of Distributed Computing, Challenges in Coordination and Communication Data and Task Parallelism, Strategies for Distributing Data and Tasks, Achieving Parallelism in Data Processing	**9**
Unit 3	**Optimization Algorithms in Big Data Analytics:** Optimizing Machine Learning Algorithms, Techniques for Enhancing ML Efficiency, Distributed Machine Learning Overview **Streamlining Data Processing Workflows:** Batch Processing vs. Real-time Processing, Workflow Optimization Strategies **Distributed Machine Learning Frameworks:** Overview of Frameworks (e.g., Apache Spark), Implementing Machine Learning at Scale	**9**
Unit 4	**Resource Management and Scalability:** Scalability in Big Data Systems, Horizontal vs. Vertical Scalability, Scaling Storage and Processing Systems	**8**

	Resource Allocation Strategies: Efficient Allocation of Computing Resources, Strategies for Dynamic Resource Management	
Unit 5	**Case Studies and Future Trends:** Real-world Optimization Challenges, Case Studies of Successful Optimization **Lessons Learned and Best Practices**: Lessons from Big Data Optimization Projects, Aalyzing Past Optimization Projects Identifying Common Challenges and Solutions	**8**

List of Experiments:		**Number of Turns**
1	**Benchmarking Big Data Processing:** Set up a benchmarking environment to assess the impact of optimization techniques on the performance of distributed processing systems. Analyze processing speed variations with different data sizes.	**2**
2	**Implementing Parallel Computing Models:** Develop a parallel computing application using Apache Hadoop, comparing task and data parallelism on a dataset. Assess scalability by increasing the number of computing nodes.	**2**
3	**Optimizing Machine Learning Algorithms:** Implement a distributed machine learning algorithm (e.g., Apache Spark MLlib), optimizing strategies to reduce training time and improve model accuracy. Compare performance with and without optimization.	**2**
4	**Resource Management in Distributed Systems:** Set up a Kubernetes cluster, implement dynamic resource allocation for a distributed application, and measure the impact on performance and resource utilization.	**2**
5	**Workflow Optimization in Real-time Processing:** Design and implement a real-time data processing workflow, integrating optimization techniques to enhance data flow efficiency. Evaluate the system's real-time streaming capabilities.	**2**
6	**Case Study Analysis:** Analyze real-world case studies of big data optimization projects, discussing challenges faced, applied optimization techniques, and project outcomes. Relate findings to course concepts	**2**
7	**Future Trends Exploration:** Explore emerging trends in big data optimization (e.g., edge computing, federated learning), discussing potential impacts on large-scale data processing optimization. Present insights in a group discussion or presentation format.	**2**

Course Outcomes:	
1	Students will demonstrate the ability to apply various optimization strategies in the context of big data processing, showcasing a higher-order thinking skill in practical problem-solving.
2	Develop skills in analyzing the performance of distributed systems, employing critical thinking to assess the impact of optimization techniques on processing speed and resource utilization.
3	Students will evaluate the scalability of distributed applications, employing critical analysis to assess how well the system adapts to increasing data sizes and computing nodes.

4	Demonstrate the ability to integrate optimization techniques into real-world scenarios, showcasing creativity in designing and implementing efficient workflows for both batch and real-time processing.
5	Develop a forward-thinking approach by exploring emerging trends in big data optimization, demonstrating the ability to forecast potential impacts on future large-scale data processing environments.

Text Book:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	**Book Title:** "Big Data: A Revolution That Will Transform How We Live, Work, and Think" **Author:** Viktor Mayer-Schönberger and Kenneth Cukier **Publisher:** Eamon Dolan/Mariner Books **Year:** 2013	2013
Reference Books:		
Sr. No.	**Name of Book/ Authors/ Publisher**	**Year of Publication / Reprint**
1	"Big Data: A Very Short Introduction" Author: Dawn E. Holmes Publisher: Oxford University Press Year: 2016	2016
2	"Hadoop: The Definitive Guide" Author: Tom White Publisher: O'Reilly Media Year: 2015	2015
3	"Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking" Authors: Foster Provost and Tom Fawcett Publisher: O'Reilly Media Year: 2013	2013

Equivalent MOOCs courses

Sr. No.	**Course Links**	**Offered by**
1	**Coursera Course:** "Big Data Specialization" **Platform:** Coursera **Instructor:** University of California, San Diego **Link:** https://www.coursera.org/specializations/big-data	Coursera
2	**edX Course:** "Introduction to Big Data" **Platform:** edX **Institution:** University of California, San Diego **Link:** https://www.edx.org/professional-certificate/introduction-to-big-data	edX

Course Name	**:**	**INTRODUCTION TO IOT**
Course Code	**:**	**DS6801**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective – V (OE-V)**

Course objectives
The main objectives of this course are:
• Understanding of core technology, applications, sensors used and IOT architecturealong with the industry perspective.
• Principles and operations of different types of sensors commonly used on mobileplatform will be taught in a manner that by the end of the course the students will be able to design and implement real time solutions using IOT.

Total No. of Lectures – 42

Sr. No.	**Course contents**	**No. of Lectures**
1.	**INTRODUCTION TO IoT:** What is IoT, how does it work? Difference between Embedded device and IoT device, Properties of IoT device, IoT Ecosystem, IoT Decision Framework, IoT Solution Architecture Models, Major IoT Boards in Market, Privacy issues in IoT.	3
2.	**SETTIN GETTING UP RASPBERRY PI/ARUDUINO TO CREATE SOLUTIONS:** Explore Raspberry Pi, setting up Raspberry Pi, showing working of Raspberry Pi using Secure Shell (SSH) Client and Team Viewer, Understand Sensing actions, Understand Actuators and Microelectromechanical Systems (MEMS).	7
3.	**COMMUNICATION PROTOCOLS IN IoT:** Types of wireless communication, Major wireless Short-range communication devices, properties, comparison of these devices (Bluetooth, Wireless Fidelity (Wi-Fi), ZigBee, Low-power Wireless Personal Area Network (6LoWPAN)), Major wireless Long-range communication devices, properties, comparison of these devices (Cellular IoT, Low-Power Wide-Area Network (LPWAN).	6
4.	**IoT APPLICATIONS:** Industrial Internet 4.0, Applications such as: Smart Homes, Wearables,Smart City, Smart Grids, Connected Car, Connected digital health, telehealth, telemedicine), smart retail.	7
5.	**SENSORS:** Applications of various sensors: Google Maps, Waze, WhatsApp, Ola Positioning sensors: encoders and accelerometers, Image sensors: cameras Global positioning sensors: Global Positioning System (GPS), Global Navigation Satellite System (GLONASS), Indian Regional	6

	Navigation Satellite System (IRNSS), Galileo and indoor localization systems, Motion & Orientation Sensors: Accelerometer, Magnetometer, Proximity Sensor, Gyroscope, Calibration, - noise modelling and characterization, and - noise filtering and sensor data processing, Privacy & Security.	
6.	**CLOUD COMPUTING SERVICES:** Introduction to GCP, AWS, Microsoft Azure, IoT Integration of IoT services with the cloud, IoT Data Analytics in the cloud.	6
7.	Fog Computing, Smart Cities and Smart Homes, Connected Vehicles,Smart Grid, Industrial IoT, Case Study: Agriculture, Healthcare, Activity Monitoring.	7

Lab Work:

Sr. No.	**Lab contents**	**Number of Turns**
1.	Project based lab work: Design and build systems that will use sensors, communication protocols and actuators.	14

Course Outcomes:

At the end of the course, students will be able to:	
1.	Understand the concept of IOT.
2.	Study IOT architecture and applications in various fields.
3.	Study the security and privacy issues in IOT.
4.	Understand various applications of sensor in Industrial, healthcare, commercial,and building automation.

Bibliography:

Sr. No.	**Book Detail**	**Year of Publication**
1.	Vijay Madisetti and ArshdeepBahga, “Internet of Things (A Hands-on- Approach)”, VPT, 1st Edition.	2014
2.	Francis daCosta, “Rethinking the Internet of Things: A Scalable Approach to Connecting Everything”, Apress Publications, 1st Edition.	2013
3.	CunoPfister, “Getting Started with the Internet of Things”, OReilly Media.	2011
4.	Kyung, C.-M., Yasuura, H., Liu, Y., Lin, Y.-L., Smart Sensors and Systems, Springer International Publishing.	2015

Course Name	**:**	**MANAGEMENT OF LARGE DATA SETS**
Course Code	**:**	**DS6802**
Credits	**:**	**4**
L T P	**:**	**3 0 2**
Type of Course	**:**	**Open Elective – VI (OE-VI)**

Course Objectives :

Students should be able -

- To provide a comprehensive overview of the data evolution and data-intensive computing at scale
- To introduce various open-source tools for managing large data storage, processing and analytics
- Evaluate relational, Hadoop, and noSQL database tooling as to understand their underlying similarities and necessary differences.
- Identify the most suitable database systems for a specific application's data storage requirements.

Total No. of Lectures – 42

Lecture wise breakup		**No. of Lectures**
Unit 1	**Introduction** Relational vs Non-Relational Data, Need of managing of large data sets, The Big Data Dimensions (3V's, 5V's), Data Replication, Data Rebalancing, Data Partitioning, Transactions, Concurrency Control, Performance, Security, noSQL, SQL Server index types, Combining Relational and noSQL concepts in SQL Server, Challenges in managing large data sets	**6**
Unit 2	**The MapReduce Programming Model** Data Parallel Problems, Algorithms Using MapReduce, HADOOP, HDFS, MapReduce and YARN, Pig, Hive, HCatalog	**9**
Unit 3	**Distributed Databases** Distributed Indexing, noSQL, MongoDB and Redis, Cassandra, Kakfa and KSQL	**9**
Unit 4	**Graph Management** Graph and Network Datasets use-cases, Graph Database, Graph storage and Processing, Neo4j, Managing and Mining Social-Network Graphs	**9**
Unit 5	**Large-Scale Machine Learning** Perceptrons, The Winnow Algorithm, Parallel Implementation of Perceptrons , Support-Vector Machines, Parallel Implementation of SVM, Learning from Nearest Neighbors, Dealing with High-Dimensional Euclidean Data	**9**

List of Experiments:		Number of Turns
1	Study Parallel and Distributed Computing for Data-Intensive Applications	2
2	Perform Data Replication, Data Rebalancing, Data Partitioning	2
3	Execute Distributed Indexing	2
4	Perform MapReduce Algorithmic Design Patterns	2
5	Perform Graph Storage and Data Lookups using Neo4j	2
6	Combining Relational and noSQL concepts in SQL Server	2
7	Perform Best Practices for Cassandra Data Modeling	2

- Instructor may change the order of the experiment

Course Outcomes:	
1	Describe the multiple dimensions and challenges involved in storing, processing modelling and managing large data.
2	Comprehend which principles are to be applied depending on the context along with their potential implications and trade-offs.
3	Understand the capabilities and pitfalls of large data storage models towards relational storage models when applied on structured and unstructured data.
4	Acknowledge how to model, adapt and extend data analysis techniques

Text Book:		
Sr. No.	Name of Book/ Authors/ Publisher	Year of Publication / Reprint
1	Martin Kleppmann, "Designing DataIntensive Applications", O'Reilly , ISBN: 978-1-44937332-0	2017
Reference Books:		
Sr. No.	Name of Book/ Authors/ Publisher	Year of Publication / Reprint
1	Jimmy Lin and Chris Dyer "Data-Intensive Text Processing with MapReduce" Morgan and Claypool ISBN: 978-1-60845342-9	2010
2	Rajkumar Buyya and Rodrigo N. Calheiros and Amir Vahid Dastjerdi "Big Data: Principles and Paradigms" Morgan Kaufmann	2016
3	Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman "**Mining of Massive Datasets**" Cambridge University Press http://www.mmds.org/#book	2014

Equivalent MOOCs courses

Sr. No.	Course Links	Offered by
1	Managing your big dataset https://www.coursera.org/lecture/data-driven-astronomy/managing-your-big-datasets-wp7qt?utm_source=link&utm_medium=page_share&utm_content=vlp&utm_campaign=top_button	Cousera
2	Introduction to Big Data with Spark and Hadoop https://www.coursera.org/learn/introduction-to-big-data-with-spark-hadoop	Cousera+IBM