No. 6-38/2025(M.Sc.Data Science)HPU (Acad.)
Himachal Pradesh University, Summer Hill, Shimla-5
(NAAC Accredited "A" Grade University)
"Academic Branch",

Dated: 10 JUN

To

1. The Dean, Faculty of Physical Science, HPU, Shimla-5
2. The Controller of Examinations, HPU, Shimla-5.
3. The D.R. Exam. (PG) HPU, Shimla-5.
4. The D.R. Eval./Re-Eval./Conduct, HPU, Shimla-5.
5. The D. R. Secrecy, HPU, Shimla-5. **(with 2 spare copies.)**
6. The S.O. Exam (M.Sc. Data Science) HPU, Shimla-5.
7. The Librarian, HPU Main Library, Shimla-5.
8. The Incharge, Computer Centre, Examination Wing (PG), HPU, Shimla-5.

Subject: ***Complimentary copy of*** **Plan, Scheme and syllabi for 2 Years M.Sc. Data Science**

Sir/Madam,

I am sending herewith a complimentary copy **Plan, Scheme and syllabi for 2 Years M.Sc. Data Science** of duly approved by the Standing Committee of Academic Council in its meeting held on 27.05.2025 vide item No. 8 (ii), on the recommendations of the concerned Board of Studies (PG) and Faculty w.e.f. 2025-26 as per annexure.

Yours faithfully,

Rajender 10.6

Deputy Registrar (Acad.)
HP University Shimla-5.
Dated: 10 JUN 2025

Endst. No. Even
Copy to:

1. The Chairman, Deptt. of Data Science, HPU, Shimla-5 for information and **send the soft copy in PDF format to web Admin, HPU, Shimla-5 immediately.**
2. The Web Admin, HPU, Shimla-5, with the request to upload this letter with syllabus on the website after receiving the soft copy from the department.
3. The Dealing Assistant Meeting (Acad.), HPU, Shimla-5, for information.
4. Guard file.

Rajender 10.6.

Deputy Registrar (Acad.)

Annexure-1

Himachal Pradesh University

NAAC Accredited "A" Grade University

Gyan Path, Summer Hill,

Shimla -171005

Plan, Scheme, and Syllabus

For the Programme

Master of Science (M.Sc.)

in

Data Science

(Effective from the session 2025-26)

Department of Data Science & Artificial Intelligence

Faculty of Physical Sciences

Himachal Pradesh University, Shimla - 5

List of Contents

		Page
1.	Introduction	3
2.	Programme Details:	4
	a) Programme	
	b) Duration	
	c) Eligibility	
	d) Fee Structure	
	e) Student Intake	
	f) Mode of Admission	
3.	Examinations	7
4.	Minimum Criteria to Award the Degree	10
5.	Honorarium	10
6.	Scheme and Syllabus	11
7.	Detailed Syllabus	18

1. Introduction

Data science is amongst the hottest fields of the 21st century that will impact all segments of daily life by 2025. The recent development in Data science is bringing significant social and economic benefits to the world. As our daily lives are seamlessly integrating more and more data-driven applications, the roleof data analytics and artificial intelligence becomes increasingly important in transforming organizations, industries and society in general. Using Data Science techniques,digital machines can analyse and learn from big datasets and discover more efficient ways to do complex tasks; thereby, making intelligent decisions with much higher accuracy and speedthan human beings. Thus, the need of the hour is to integrate the power of Data science to every business to boost the global economy by transforming business models across all sectors like science, engineering, banking, sales, finance, marketing, construction, manufacturing, healthcare, travel, hospitality, leisure, environmentalmonitoring, logistics etc. Hence, the field of Data Science has a potential to employ a large quantum of human resources and serve countries all over the globe.

As per World Economic Forum, Data Scientists and Analysts will become the number one emerging role in the world and thus will find job opportunities which are expected to rise appreciably in the years to come. So, academic institutions must take timely initiatives to offer academic programmes to equip aspirants with requisite job oriented skills and training and thereby contribute to meet the global industrial demand of workforce and diversity in AI and Data Science. Further, in higher education institutions, there is an urgent need for increased collaboration between industry and academia, through creation of channels of communication between faculty and industry, to promote exchange of ideas and expertise. Various avenues of collaboration need to be explored, including workshops, incentives for guest lectures by professionals and institutional arrangements for regular design of courses/curricula in collaboration with the Private Sector Units (PSU).

Master's Programme in Data science proposed in this report, andto be started by Himachal Pradesh University under the Faculty of Physical Sciences, shall beone of such academic platforms, which caters to impart most advanced knowledge, methods and processes to exploit data science-based solutions to real-world problems. After the completion of this course, the students may have career opportunities with exceptional prospective fields of healthcare, business, e-Commerce, social networking companies, climatology, biotechnology, genetics and other important areas.

2. Programme Details:

a. Programme : M.Sc. in Data Science

b. Duration : Two (02) Years Programme divided into four (04) Semesters

c. Eligibility : Any Engineering Graduate

Or

Any other graduate with Mathematics/ Computer/IT/Applications/Statistics as Major Subjects in all three years.

With 50% aggregate marks in qualifying exam (45% for SC/ST/PWD)

d. Fee Structure :

Subsidized Seats: ₹40,000/- (Rupees Forty Thousand Only inclusive of ₹5,000/- of equipment fee) annually.

Non-Subsidized Seats: ₹60,000/- (Rupees Sixty Thousand Only inclusive of ₹5,000/- of equipment fee) annually.

e. Student Intake :

Total Seats:

Subsidized Seats:	**15+02+02+03*** (Supernumerary 02 seats Reserved for Single Child who is a Girl + Supernumerary 02 seats Reserved for J&K migrants + 03* seats for EWS of H.P.)
Non-Subsidized Seats:	**15+03** (Supernumerary 03 seats for Wards of H.P. University Employees)

Total Seats in M.Sc. Artificial Intelligence Course:
15+02+02+03*+15+03=40

There are 10% additional seats reserved for the Economically Weaker Section (EWS) category for admission and all are the subsidized seats.
If these seats remain vacant then these seats neither be filled with other category nor will carry forward.
Supernumerary seats as per university norms for foreign National who apply through ICCR are also available as per H.P. University Rules.

f. Mode of Admission: On the basis of Merit of Entrance Test

The admission to this course will be made on the basis of merit of the Entrance Examination(written test) conducted by H.P. University.

Written test	**100 Marks**
Duration of test	**1:30 Hours**

The written test shall include the following three sections:

Sr. No	Contents	Marks
1.	ICT Awareness	60
2.	Mathematics	20
3.	General Logic Ability & Aptitude	20
Total		**100**

The minimum qualifying marks in the Entrance Examination (written test) for subsidized as well as non-subsidized seats will be 35%.

Age Limit:

There is no age bar as such.

RESERVATION:

I. Subsidized Seats

a. 15% and 7.5% of the seats shall be reserved for bonafide Himachali Scheduled Castes and Scheduled Tribes candidates respectively who have passed their qualifying examination from Himachal Pradesh University or from any other University established by law in India which is equivalent to the qualifying examination of H.P. University.

b. The remaining seats shall be filled as under

   i. 25% seats shall be open for all the candidates irrespective of the institution from where they have passed their qualifying examination.

   ii. 75% of the seats shall be filled out of the candidates who have passed their qualifying examination from Himachal Pradesh University or H. P. Krishi Vishva Vidyalaya or Dr. Y. S. Parmar University of Horticulture and Forestry or Himachal Pradesh Technical University or Central University of Himachal Pradesh or the candidates who are Himachal Pradesh domicile irrespective of passing qualifying examination from any other university established by law in India which is equivalent to the qualifying examination of Himachal Pradesh University, subject to the following reservation: 5% of

the seats shall be reserved for Physically Handicapped candidates with a minimum of 40% disability and who are Himachal Pradesh Domicile.

iii. Two supernumerary seats shall be reserved for Himachali Bonafide only single child who is a girl. In this respect, an affidavit issued by a competent authority in original shall have to be submitted by the candidate at the time of counselling.

iv. There are 10% additional seats reserved for the Economically Weaker Section (EWS) category for admission. If these seats remain vacant then these seats neither be filled with other category nor will carry forward.

v. Two supernumerary seats shall be reserved for the students of the state of Jammu and Kashmir Migrants.

vi. Supernumerary seats as per university norms for foreign National who apply through ICCR are also available as per H.P. University Rules.

II. Non-Subsidized Seats

a. 15% and 7.5% of the seats shall be reserved for bonafide Himachali Scheduled Caste and Scheduled Tribe candidates respectively who have passed their qualifying examination from Himachal Pradesh University or from any other University established by law in India which is equivalent to the qualifying examination of H.P. University.

b. The remaining seats shall be filled as under

i. 25% seats shall be open to all the candidates irrespective of the institution from where they have passed their qualifying examination.

ii. 75% of the seats shall be filled out of the candidates who have passed their qualifying examination from Himachal Pradesh University or H. P. Krishi Vishva Vidyalaya or Dr. Y. S. Parmar University of Horticulture and Forestry or Himachal Pradesh Technical University or Central University of Himachal Pradesh or the candidates who are Himachal Pradesh domicile irrespective of passing qualifying examination from any other university established by law in India which is equivalent to the qualifying examination of Himachal Pradesh University.

iii. Three supernumerary seats shall be reserved for wards of H.P. University Employee.

III. Reservation Roster

The reserved seats shall be worked out on the basis of the existing 120-point roster.

1PwD	26SC	51	76	101PwD
2	27ST	52SC	77	102
3	28	53ST	78SC	103
4	29	54	79SP	104
5	30	55	80ST	105
6	31	56	81CUL	106SC
7SC	32	57	82 (PwD)	107ST
8	33SC	58	83	108
9	34	59SC	84	109
10	35	60SP	85SC	110
11	36	61CUL	86	111
12SC	37	62 (PwD)	87	112
13ST	38CUL	63	88	113SC
14	39SC	64	89	114
15	40SP	65SC	90	115
16	41ST	66	91SC	116
17	42(PwD)	67ST	92	117SP
18SC	43	68	93ST	118CUL
19SP	44	69	94	119SC
20CUL	45	70	95	120ST
21PwD	46SC	71	96	
22	47	72SC	97SC	
23	48	73	98CUL	
24	49	74	99SP	
25	50	75	100	

(PwD-5%) (CUL-5%) (SP-5%) (SC-15%) (ST-7.5%)

PwD : Person with Disability; CUL : Cultural; SP : Sports; SC : Scheduled Caste and ST : Scheduled Tribe.

3. Examinations:

As the degree is spanned over two years and distributed into four semesters, the learning outcomes shall be assessed after every semester. The assessment of the students shall consist of the following components:

Sr. No.	Assessment Component
1	Semester End External Examinations (Theory)
2	Internal Assessment (Theory)
3	Semester End External Examinations (Practical)
4	Internal Assessment (Practical)
5	Project Work Evaluation

Note: The concerned subject teacher shall submit the marks of Internal Assessment (Theory), Internal Assessment (Practical), and Semester End Examinations (Practical) to theChairman/Head of the Department (as per the schedule mentioned in Academic Calendar) in triplicate - first copy for Examination Branch of HPU, second copy as an Office copy of the Department, and the third copy may be retained by the subject teacher.

Internal assessment will be given on the basis of class tests (best of 2 in a semester), seminars,

surprise quizzes, class participation and regularity of the student in the class, be evaluated by the department and the award list shall be sent to the examination branch by the Chairman/HOD.

In the third semester, the Chairman/Head of the Department will assign a guide/supervisor to each candidate for his/her project work. The candidate shall be required to maintain his/herproject diary (logbook) of work in the organization/department. Each student will be requiredto give at least two seminars on his/her project work . Each student is required to submit three copies of his/her project reports in the department after completion of the project work which will be evaluated by an external examiner.

Further, the project work (final semester) will be jointly evaluated by an internal guide and external examiner.

Scheme of Examination:

English shall be the medium of instruction and examination. The pass marks in each course shall be 40% in each written paper and in the internal assessment separately, and 40% in viva-voce, project work and semester course and 50% in the aggregate subject to the conditions that aggregate shall be determined at the end of the examination. Other rules shall be as per the rules of the university.

Theory Papers:

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

Practical Examination

Each paper will be of 75 marks (50 marks for practical exam and 25 marks for internal assessment) and duration of each paper will be 3 hours. The marks awarded by the teacher on account of internal assessment in relation to theory/practical paper as mentioned above shall be submitted to the office

of Chairman.

Conduct:

Practical exam will be conducted by the external examiner from the panel submitted by Chairman, Department of Data Science and Artificial Intelligence, Himachal Pradesh University and duly approved by the university authority/evaluation branch, Himachal Pradesh University, Shimla.

Project Work:

In 2nd year (fourth semester) the student has to develop one software project, which will be evaluated by the external examiner from the panel submitted by Chairman, Department of Data Science and Artificial Intelligence, Himachal Pradesh University, and duly approved by the university authority/evaluation branch, Himachal Pradesh University, Shimla on the following basis:

System Development Project:

Seminar (2)	50 marks
Log Book & Interim Report	50 marks
Internship/Project Report	100 marks
Viva-Voce	250 marks
Total	**450 marks**
OR	
Seminar (2)	50 marks
Log Book & Interim Report	50 marks
Elective 4	100 marks
Project Report	100 marks
Viva-Voce	150 marks
Total	**450 Marks**

In fourth semester, the Chairman/Head of the Department will assign a guide/supervisor, to each candidate for his/her project work. The candidate shall be required to maintain his/her project diary (logbook) of work in the organization. Each student will be required to give at least two seminars on his/her project work. Each student is required to submit three copies of his/her project reports in the Department after completion of the project work which will be evaluated by external examiner.

4. Minimum Criteria to Award the Degree

Rules regarding the minimum criteria to award a Degree of Master of Science in Data Science shall remain the same as applicable in other Master of Science programmes run by HPU.

5. Honorarium

1. All the teachers shall be paid an honorarium on lecture basis as per university norms for all the lectures taken for engaging theory as well as practical classes in the MSc Data Science course in the department.
2. Remuneration for evaluation of answer-scripts(Theory) exam of M.Sc. Data Science @ ₹25/- per answer-script subject to minimum of ₹150/-
3. **Remuneration charges for the conduct of practical/Viva-voce for External Examiners:**
    i. Sitting charges per session per day for the answer sheets of practical examination @ ₹25/- per answer-sheet.
    ii. Sitting charges per session per day for practical- as per university norms
    iii. Project evaluation-₹100 per student
    iv. TA/DA - As per the University norms
4. All the expenditure incurred with respect to the honorarium etc. shall be met out of the funds generated through the running of M.Sc. Data Science course.

Master of Science

in

Data Science

M. Sc. (Data Science)

Scheme

&

Syllabus

Credit Based System

Effective from

Academic Session 2025-2026

PROGRAMME OUTCOMES (POs)	
PO1	Knowledge: Capable of demonstrating comprehensive disciplinary knowledge gained during course of study.
PO2	Research Aptitude: Capability to ask relevant/appropriate questions for identifying, formulating and analyzing the research problems and to draw conclusions from the analysis.
PO3	Problem Investigation & Solving: a) Ability of critical thinking, analytical reasoning, and research-based knowledge including design of experiments, analysis, and interpretation of data to provide conclusions, and b) Capability of applying knowledge to solve scientific and other problems.
PO4	Individual/Team Work and Modern Tool Usage: Capable to learn and work effectively as an individual, and as a member or leader in diverse teams, in multidisciplinary settings with or without the developed capabilities of modern tools.
PO5	Science and Society: Ability to apply reasoning to assess the different issues related to society and the consequent responsibilities relevant to the professional scientific practices.
PO6	Life-Long Learning: Motivation to take higher studies and aptitude to apply knowledge and skills that are necessary forparticipating in learning activities throughout life.
PO7	Ethics: Capability to identify and apply ethical issues related to one's work, avoid unethical behaviour such as fabrication of data, committing plagiarism and unbiased truthful actions in all aspects of work.
PO8	Project Management: Ability to demonstrate knowledge and understanding of the scientific principles and apply these to manage projects.

PROGRAMME SPECIFIC OUTCOMES (PSOs)	
PSO1	Demonstrate expertise in foundational AI concepts, including probability, statistics, data structures, and Python programming.
PSO2	Apply machine learning, deep learning, and NLP techniques to solve real-world problems, demonstrating proficiency in experimentation and analysis.
PSO3	Design and implement robust AI systems for tasks within computer vision, image processing, and optimization.
PSO4	Analyze large datasets using big data tools, create compelling data visualizations, and extract actionable insights.
PSO5	Utilize computational intelligence techniques like genetic algorithms and explore applications of AI in robotics.

Abbreviations Used:

L	LECTURES
T	TUTORIALS
P	PRACTICALS
C	CREDITS
H	HOURS
EE	EXTERNAL EXAMINATIONS
IA	INTERNAL ASSESSMENT
WAP	WRITE A PROGRAM
PO	PROGRAM OUTCOME
CO	COURSE OUTCOMES
PSO	PROGRAM SPECIFIC OUTCOMES

Semester-I

Sr. No	Course Code	Course Title	Contact Hrs/week				C	Semester End Marks	
			L	T	P	H		EE	IA
1.	MSDS-101	Probability and Statistics	4	0	0	60	4	75	25
2.	MSDS-102	Data Structures and Algorithms	4	0	0	60	4	75	25
3.	MSDS-103	Python Programming	4	0	0	60	4	75	25
4.	MSDS-104	Introduction to Data Science & Artificial Intelligence	4	0	0	60	4	75	25
5.	MSDS- 105	Data Mining	4	0	0	60	4	75	25
6.	MSDS-151	Data Structures and Algorithms using C - Lab	0	0	4	60	2	50	25
7.	MSDS-152	Python Programming - Lab	0	0	4	60	2	50	25
TOTAL							24	475	175
								Total = 650	

Semester-II

Sr. No	Course Code	Course Title	Contact Hrs/week				C	Semester End Marks	
			L	T	P	H		EE	IA
1.	MSDS-201	Machine Learning	4	0	0	60	4	75	25
2.	MSDS-202	Big Data Analytics and Data Visualization	4	0	0	60	4	75	25
3.	MSDS-203	Introduction to Computer Vision and Image Processing	4	0	0	60	4	75	25
4.	MSDS-204	IoT Based Data Analytics	3	0	0	45	3	75	25
5.	MSDS-EL	Elective 1	3	0	0	45	3	75	25
6.	MSDS-251	Big Data Analytics & Visualisation - Lab	0	0	4	60	2	50	25
7.	MSDS-252	Machine Learning - Lab	0	0	4	60	2	50	25
TOTAL							22	475	175
								Total = 650	

Semester-III

Sr. No	Course Code	Course Title	Contact Hrs/week				C	Semester End Marks	
			L	T	P	H		EE	IA
1.	MSDS-301	Deep Learning	4	0	0	60	4	75	25

2.	MSDS-302	Distributed and Cloud Computing	**4**	**0**	**0**	**60**	**4**	75	25
3.	MSDS-303	Natural Language Processing	**4**	**0**	**0**	**60**	**4**	75	25
4.	MSDS-EL	Elective 2	**3**	**0**	**0**	**45**	**3**	75	25
5.	MSDS-EL	Elective 3	**3**	**0**	**0**	**45**	**3**	75	25
6.	MSDS-351	Deep Learning- Lab	**0**	**0**	**4**	**60**	**2**	50	25
7.	MSDS-352	Natural Language Processing - Lab	**0**	**0**	**4**	**60**	**2**	50	25
TOTAL							**22**	**475**	**175**
								Total = 650	

Semester-IV

Sr. N.	Course Code	Course Title	Evaluation Components	C	Semester End Marks	
					EE	IA
1.	MSDS-401	**Project Work**	Seminar (2)	**2**		50
			Log Book & Interim Report	**2**		50
			Internship/Project Report	**4**		100
			Viva-Voce	**10**	250	
			Total	**18**	**250**	**200**
					Total =450	
			OR			
			Seminar (2)	**2**		50
			Log Book & Interim Report	**2**		50
			Elective 4	**4**	75	25
			Internship/Project Report	**4**		100
			Viva-Voce	**6**	150	
TOTAL				**18**	**225**	**225**
					Total =450	

Total Credits : 24+22+22+18 = **86**
Total Maximum Marks : 650+650+650+450 = **2400**

LIST OF ELECTIVE COURSES

NAME OF COURSE	COURSE CODE
1. ELECTIVE 1	
a) Database Management System	MSDS-EL-211
b) Information Retrieval	MSDS-EL-212
c) Software Engineering	MSDS-EL-213
2. ELECTIVE 2	
a) Data Engineering	MSDS-EL-311
b) Soft Computing	MSDS-EL-312
c) Complex Network Analysis	MSDS-EL-313
3. ELECTIVE 3	
a) Time Series Analysis and Forecasting Techniques	MSDS-EL-321
b) Multi-Criteria Decision Making	MSDS-EL-322
c) Queuing Theory and Network Analysis	MSDS-EL-323
4. ELECTIVE 4	
a) Predictive Analysis	MSDS-EL-411
b) Social Media Analytics	MSDS-EL-412
c) Deep Reinforcement Learning	MSDS-EL-413

DETAILED SYLLABUS

SEMESTER-I

Name of the Course	Probability And Statistics		
Course Code	MSDS-101	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	Max Marks: 75	Min Pass Marks: 40%	Max. Time: 3 Hrs
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			Max Marks: 25 Min Marks: 10

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Calculate and apply probabilities using fundamental concepts and various distributions.
CO2	Analyze data sets using descriptive statistics, central tendency, dispersion, and measures of shape.
CO3	Conduct hypothesis tests for making statistical inferences about populations.
CO4	Apply correlation and regression techniques to model relationships between variables.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-101 Probability and Statistics

UNIT-I

Introduction: Probability, Bayes Theorem, Conditional expectation and variance, mean, median, mode and standard deviation, Random Variables & Probability Distributions: Discrete and continuous random variables - distribution function and its properties , Joint Probability Law, probability mass function and probability density function - discrete and continuous probability distributions – Bernoulli Distribution, Binomial, Geometric, Poisson,Uniform, Exponential and Normal distributions, Cumulative distribution function.

UNIT-II

Types of data: primary and secondary data - classification and representation of data - formation of frequency distribution - various measures of central tendency, dispersion - and their merits and

demerits - concept of skewness and kurtosis. Sampling, analysis of sample data - Empirical Distributions, Sampling from a Population Estimation, confidence intervals, point estimation- Maximum Likelihood.

UNIT-III

Testing of Hypothesis I: Concept of large and small samples – Tests concerning a single population mean for known σ – equality of two means for known σ- Test for single variance – Test for equality of two variance for normal population – Tests for single proportion – Tests of equality of two proportions for the normal population.

Test of Hypothesis: Z, t, Chi-Square & F-test. ANOVA & Designs of Experiments--Single, Two factor ANOVA, Factorials ANOVA models.

UNIT-IV

Correlation and Curve Fitting: Correlation coefficient and regression - rank correlation - curve fitting by least square methods, fitting a straight line, parabola, power curve and exponential curves. (no derivation, numerical problems only) Correlation & Regression Models-- linear regression methods, Ridge regression, LASSO, univariate and Multivariate Linear Regression, probabilistic interpretation, Regularization, Logistic regression, locally weighted regression.

Text Books:

1. Gupta, S.C. and V. K. Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand & Sons,New Delhi, 12th edition, 2020.
2. T. Veerarajan, Probability, Statistics and Random Processes, Tata McGraw Hill, 3rd edition, 2017.
3. Robert V Hogg, Elliot A Tannis and Dale L.Zimmerman, Probability and Statistical Inference,10th edition, Pearson publishers,2021

Reference Books:

1. Douglas C. Montgomery and George C. Runger, Applied Statistics and Probability for Engineers,Third Edition, John Wiley & Sons Inc., 6th edition, 2016.
2. Ronald E. Walpole, Raymond H Myres, Sharon. L. Myres and Kying Ye, Probability and Statisticsfor Engineers and Scientists, 9th edition, Pearson Education, 2010.
3. Richard Arnold Johnson, Irwin Miller, John E. Freund , Miller & Freund's Probability andStatistics for Engineers, 8th edition, Prentice Hall, 2011.
4. Goon, A.M., M. K. Gupta and B. Das Gupta Fundamentals of Statistics- Vol. I, World Press Ltd,Kolkata, 2013.
5. Brian Caffo, Statistical Inference for Data Science, Learnpub, 2016.
6. Hogg, R.V. and A. Craig, Introduction to Mathematical Statistics,7th edition, Pearson Education, 2012..
7. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to StatisticalLearning with Applications in R, Springer 2017.
8. Dr. P. Kandaswamy, Dr. K. Thilagavathy and Dr. K. Gunavathy, Probability and Queuing Theory,Revised edition, S. Chand Publishing, 2013.

Name of the Course	Data Structures and Algorithms		
Course Code	**MSDS-102**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Analyze algorithms using asymptotic notation and assess the efficiency of different data structures.
CO2	Implement fundamental data structures and perform associated operations.
CO3	Implement fundamental data structures and perform associated operations.
CO4	Design and implement algorithms using advanced techniques like divide-and-conquer, greedy methods, and dynamic programming.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-102 Data Structures and Algorithms

UNIT-I

Introduction to Data Structures and algorithms: Primitive and Composite data types, Classification of Data Structures, algorithm, complexity of algorithms, analyzing algorithms, designing algorithms, asymptotic notations.

Arrays, and their operations (insertion, deletion, traversing), Linked Lists (traversal, insertion, deletion), and type (linear, circular, doubly linked, inverted).

UNIT-II

Stacks & Queues: Representation of Stacks, Stack Operations, Application of stacks (converting arithmetic expression from infix notation to polish and their subsequent evaluation, recursion), Queues, Operations on Queues, Circular Queues, Dequeue, Priority Queues.

Searching & Sorting: Linear Search, Binary Search, Selection Sort, Insertion Sort, Bubble Sort. Implementation of these searching and sorting through algorithms.

UNIT-III

Trees and graphs: binary trees and their types, representation in memory, Threaded Binary Trees, Binary Search Trees and Operations, AVL Trees, heap, M-way Search Trees, B-Trees, B+ Trees, hashing, Graph representation and traversal (BFS and DFS).Divide and Conquer: The General Method, Merge Sort, Quick Sort.

UNIT-IV

Greedy Algorithms: General Method, Knapsack problem, Job sequencing with deadlines, Minimum Spanning Trees, Kruskal's Algorithm, Prim's Algorithm, and Dijkstra's Single Source Shortest Path Algorithm, Dynamic Programming: The General Method, All Pairs Shortest Paths, 0/1 Knapsack, Traveling Salesperson Problem.

Text Book:

1. Seymour Lipschutz, "Data Structures", McGraw Hill Education, Revised edition, 2014.
2. T.H. Cormen, C.E. Leiserson, R.L. Rivest and C. Stein, Introduction to Algorithms, Prentice-Hall of India, 4th edition, 2022.

Reference Books:

1. Parag H. Dave, Himanshu B. Dave, Design and Analysis of Algorithms, Pearson Education, 3rd edition, 2021.
2. Jean Paul Tremblay & Paul G. Sorenson, "An Introduction to Data Structures with Applications", Tata McGraw Hill Publications,1984.

Name of the Course	Python Programming		
Course Code	MSDS-103	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	Max Marks: 75	Min Pass Marks: 40%	Max. Time: 3 Hrs
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			Max Marks: 25 Min Marks: 10

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Execute Python code using basic syntax, data types, functions, and control flow structures.
CO2	Apply numpy to create arrays, manipulate data, and perform vectorized computations for data analysis.
CO3	Utilize pandas to create and manipulate dataframes and Series, perform descriptive analysis, and visualize data.
CO4	Clean and prepare data for analysis, including handling missing values, transforming data, and identifying outliers.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-103 Python Programming

UNIT-I

Python Programming: Basics of Python Programming: Features of Python, History and Future of Python, Writing and executing Python program, Literal constants, variables and identifiers, Built-in Data types and their Methods: Strings, List, Tuples, Dictionary, Set, Input operation, Comments, Reserved words, Indentation, Operators and expressions, Expressions in Python, Math and Random number functions. User defined functions - function arguments & its types.

UNIT-II

Decision Control Statements, Functions and Modules: Decision Control Statements: Decision control statements, Selection/conditional branching Statements: if, if-else, nested if, if-elif-else statements. Basic loop Structures/Iterative statements: while loop, for loop, selecting appropriate loop. Nested loops, the break, continue, pass, else statement used with loops.

Need for functions, Function: definition, call, variable scope and lifetime, the return statement. Defining functions, Lambda or anonymous function, documentation string, good programming practices. Introduction to modules, Introduction to packages in Python, Introduction to standard library modules. Exception Handling: Built-in Exceptions, Handling Exceptions, Exception with Arguments, User-defined Exceptions.

UNIT-III

Working with Numpy Basics, Data Cleaning and Preparation: Arrays and Vectorized Computation- The NumPy ndarray - Creating ndarrays - Data Types for ndarrays - Arithmetic with NumPy Arrays- Basic Indexing and Slicing - Boolean Indexing-Transposing Arrays and Swapping Axes. Universal Functions: Fast Element-Wise Array Functions- Mathematical and Statistical Methods-Sorting Unique and Other Set Logic.

Data Cleaning and Preparation: Handling Missing Data - Data Transformation: Removing Duplicates, Transforming Data Using a Function or Mapping, Replacing Values, Detecting and Filtering Outliers String Manipulation: Vectorized String Functions in pandas. Plotting with pandas: Line Plots, Bar Plots, Histograms and Density Plots, Scatter or Point Plots.

UNIT-IV

Introductions to Pandas Data Structures: Series, Data Frame, Essential Functionality: Dropping Entries Indexing, Selection, and Filtering- Function Application and Mapping- Sorting and Ranking. Summarizing and Computing Descriptive Statistics- Unique Values, Value Counts, and Membership. Reading and Writing Data in Text Format. Frequency tables. Simple Aggregation in Pandas GroupBy: Split, Apply, Combine.

Text Books:

1. Mark Lutz, Learning Python, O'Reilly, 5th Edition, June 2013
2. Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and I Python, O'Reilly, 3rd Edition, August 2022.

Reference Books:

1. Y. Daniel Liang, Introduction to Programming using Python, Pearson, 2012
2. Wesley J. Chun, Core Python Applications Programming,3rd Edition, Pearson,2012.

3. R. Nageswara Rao, “Core Python Programming”, Dreamtech Press; Third edition,2021
4. Reema Thareja, “Python Programming Using Problem Solving Approach”, 2nd Edition,Oxford University Press,2019.
5. Jake Vander Plas ,Python Data Science Handbook O’Reilly, 2nd Edition ,2022.

Name of the Course	Introduction to Data Science and Artificial Intelligence		
Course Code	**MSDS-104**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain fundamental concepts of data science, including its importance, workflow, and the difference between structured and unstructured data.
CO2	Apply basic SQL and data science platforms for data manipulation, analysis, and visualization.
CO3	Describe the foundations of artificial intelligence, including key definitions, applications, and the principles of intelligent agents.
CO4	Identify and apply various AI problem-solving and search techniques and understand knowledge representation methods.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-104 Introduction to Data Science and Artificial Intelligence

UNIT-I

Introduction to Data Science: Importance of Data Science, Need for Data Science**,** Data Science Process**,** business intelligence and data science, components of data science, tools and skills needed.

Data: structured data, unstructured data, challenges with unstructured data, Data pre-processing: data cleaning, data integration, data transformation, data reduction, data discretization.

D**ata Modeling and Analytics**: data science methodology, analytics for data science, data analytics life cycle, data discovery, data preparation, model planning, model building, communicate results, operationalization.

UNIT-II

Platforms for Data Science: SQL tools for data science, basics statistics with SQL, data munging with SQL, preparing data for analytics tool, advanced NOSQL for data science, document databases for data science, wide-column databases for data science.
Platforms for Data Science: data science tool python: basics of python for data science, python libraries: data frames manipulation with pandas and numpy, exploration data analysis with python, python IDEs for data science, **data science tool R**: reading and getting data into R, writing data into files, scan() function, built in data sets, ordered and unordered factors, array and matrices, creating an array, accessing elements in an array, array manipulation, matrices, creating a matrix, matrix transpose, Data visualization.

UNIT-III

Introduction to AI: Introduction to Artificial Intelligence, various definitions of AI, AI Applications and Techniques, Turing Test and Reasoning - forward & backward chaining.
Intelligent Agents: Introduction to Intelligent Agents, Rational Agent, their structure, reflex, model-based, goal-based, and utility-based agents, behavior and environment in which a particular agent operates, Expert System Architectures and Development, Applications of Expert System, Representing and using domain knowledge, Expert System Shells, Knowledge Acquisition, Different types of uncertainty - degree of belief and degree of truth.

UNIT-IV

Problem Solving and Search Techniques: Problem Characteristics, Production Systems, Control Strategies, Breadth First Search, Depth First Search, iterative deepening, uniform cost search, Hill climbing and its Variations, simulated annealing, genetic algorithm search; Heuristics Search Techniques: Best First Search, A* algorithm, AO* algorithm, Minmax & game trees, refining minmax, Alpha – Beta pruning, Constraint Satisfaction Problem, Means-End Analysis.
Knowledge Representation: Introduction to First Order Predicate Calculus, Resolution Principle, Unification, Semantic Nets, Conceptual Dependencies, semantic networks, Frames system, Production Rules, Conceptual Graphs, Ontologies.

Text Books:

1. S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th edition, Pearson Education, 2022
2. Elaine Rich and Kelvin Knight, Artificial Intelligence, 3rd edition, Tata McGraw Hill , 2017
3. Sanjeev J. Wagh, Manisha S. Bhende, Anuradha D. Thakare, Fundamentals of Data Science, CRC 2021

Reference Books:

1. Michael Wooldridge, An Introduction to MultiAgent Systems, 2nd edition, John Wiley & Sons, 2009.
2. Tom Markiewicz, Josh Zheng, Getting Started with Artificial Intelligence,2nd edition, O'Reilly, 2020.
3. Ivan Bratko, Prolog Programming for Artificial Intelligence, Addison-Wesley, Pearson Education, 4th edition, 2011.
4. Fabio Luigi Bellifemine, Giovanni Caire, Dominic Greenwood, Developing Multi-Agent Systems with JADE, Wiley Series in Agent Technology, John Wiley & Sons, 2007.
5. Charu C. Aggarwal, Artificial Intelligence, Springer, 2021.

Name of the Course	Data Mining		
Course Code	**MSDS-105**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr. Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs.**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Understand the fundamentals of data mining and knowledge discovery processes.
CO2	Implement clustering algorithms like k-means, hierarchical, and DBSCAN methods.
CO3	Apply sequential pattern mining algorithms to discover patterns in time-series data.
CO4	Utilize big data tools such as Hadoop for mining large-scale datasets.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each unit and will carry 27 marks out of the total marks of the semester end examination for the course. Sections A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester-end examination for the course.

For candidates: Candidates must attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-105 Data Mining

UNIT-I

Foundations of Data Mining: Definition of data mining, the knowledge discovery process, applications in real-world domains, challenges in mining various types of data (structured, semi-structured, unstructured), data pre-processing techniques, data cleaning, data integration, transformation, reduction, feature selection, dimensionality reduction methods (PCA, SVD), outlier detection methods, introduction to data warehousing, OLAP systems.

UNIT-II

Clustering and Association Rule Mining: Clustering techniques from **"Data Mining: Concepts and Techniques"**, partitioning (k-means, k-medoids), hierarchical clustering, density-based clustering (DBSCAN), grid-based methods, evaluation of clustering results, association rule mining, Apriori algorithm, FP-Growth algorithm, correlation analysis in association rule mining, multi-level and multi-dimensional association rule mining.

UNIT-III

Web, Text, and Sequential Pattern Mining: Web mining techniques, including content, structure, and usage mining from **"Mining the Web",** introduction to text mining from **"Introduction to Data Mining",** pre-processing (tokenization, stemming, stop word removal), mining sequential patterns (GSP, Prefix Span), mining time-series data, applications in real-world domains (e.g., e-commerce, social media).

UNIT-VI

Advanced-Data Mining and Big Data Analytics: Spatial data mining from **"Data Mining: Concepts and Techniques"**, mining multimedia data, time-series data mining, anomaly detection techniques, big data mining with Hadoop and Map Reduce, ethical considerations in data mining (privacy, bias, security), case studies from healthcare, finance, and marketing, hands-on experience with data mining tools (Weka, R, Python libraries).

Text Books:

1. Jiawei Han, Micheline Kamber, Jian Pei, *Data Mining: Concepts and Techniques*, 4th Edition, 2022, Morgan Kaufmann.
2. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, 2nd Edition, 2018, Pearson.
3. Soumen Chakrabarti, Mining the Web: Discovering Knowledge from Hypertext Data, 1st Edition, 2003, Morgan Kaufmann.

Reference Books:

1. Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
2. Tan, P.-N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Pearson Addison-Wesley.
3. David J. Hand, Heikki Mannila, Padhraic Smyth, *Principles of Data Mining*, 1st Edition, 2001, MIT Press
4. Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of Data Mining*. MIT Press.
5. Chakrabarti, S. (2003). *Mining the Web: Discovering Knowledge from Hypertext Data*. Morgan Kaufmann.

Name of the Course	**Data Structures and Algorithms using C- Lab**		
Course Code	**MSDS-151**	Credits-2	L-0, T-0, P-4
Lectures to be Delivered	**60 hours of Lab Sessions**		
Semester End Examination	**Max Marks: 50**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Implement fundamental algorithms to solve problems like finding prime/Armstrong numbers, calculating factorials, and sorting.
CO2	Demonstrate the use of arrays, pointers, and data structures like stacks and queues for various operations.
CO3	Design and implement linked lists along with operations like insertion and deletion.
CO4	Construct and traverse binary trees using different traversal methods

MSDS-151 Data Structures and Algorithms using C -Lab

1. Write a program to find whether the entered number is Armstrong or not.

2. Write a program to find whether the entered number is prime or not.

3. Write a program to find the factorial of any number.

4. Swap the value of two variables using call by value & by reference.

5. Write a program to a) add a number to the given array and b) Delete a number from the given array.

6. Write a program using pointers to find the smallest number in an array.

7. Write a program to calculate the sum of all digits on a number.

8. Write a program to multiply two matrices.

9. Write a program to Count the total number of vowels and consonants in a string.

10. Implement the Selection Sort

11. Implement Bubble Sort

12. Implement Insertion Sort

13. Implement the Linear Search
14. Implement Binary Search.
15. Implement the following:
    a. Inserting a node into the Linked List (First Node, Last Node, and nth Node)
    b. Deleting a node from the Linked List (First Node, Last Node, and nth Node)
16. Implement Stack and its operations using arrays and Linked List.
17. Implement Queue and its operations arrays and Linked List.
18. Implement circular queue.
19. Implement binary tree and traverse it using in-order, pre-order, and post-order.
20. Implement queue by using stacks.

Name of the Course	Python Programming-Lab		
Course Code	**MSDS-152**	Credits-2	L-0, T-0, P-4
Lectures to be Delivered	**60 hours of Lab Sessions**		
Semester End Examination	**Max Marks: 50**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Work with Python's fundamental data types and perform various operations on them.
CO2	Apply programming concepts to solve computational problems.
CO3	Utilize numpy effectively for array creation, manipulation, and mathematical computations.
CO4	Create and manipulate Pandas dataframes, perform data exploration, cleaning, and analysis.

MSDS-152 Python Programming-Lab

1. Write a program to demonstrate different number data types in Python.
2. Write a program to perform different Arithmetic Operations on numbers in Python.
3. Write a program to create, concatenate and print a string and accessing sub-string from a given string.
4. Write a python script to print the current date in the following format “Sun May 29 02:26:23 IST 2024” **.**
5. Write a program to create, append, and remove lists in python, demonstrate working with tuples, dictionaries in python**.**
6. Write a Python program to convert temperatures to and from Celsius, Fahrenheit. [Formula : c/5 = f-32/9]**.**
7. Write a Python script that prints prime numbers less than 20.
8. Write a python program to find factorial of a number using Recursion.
9. Write a program that accepts the lengths of three sides of a triangle as inputs. The program output should indicate whether or not the triangle is a right triangle (Recall from the Pythagorean Theorem that in a right triangle, the square of one side equals the sum of the squares of the other two sides)**.**

10. Write a python program to define a module to find Fibonacci Numbers and import the module to another program.
11. Write a Python function that prints out the first n rows of Pascal's triangle.
12. Write a Python program to square and cube every number in a given list of integers using Lambda.
13. Write a Python program to create a lambda function that adds 15 to a given number passed in as an argument, also create a lambda function that multiplies argument x with argument y and prints the result.

Sample Output:
25
48

14. Write a program Utilize NumPy functions to create arrays from lists, tuples, or built-in functions.
15. Write a program Access array attributes like shape, dimensions, and data type, and use methods like reshape, flatten, and transpose.
16. Write a program Access elements and sub-arrays of NumPy arrays using indexing and slicing techniques.
17. Write a program Concatenate multiple arrays along different axes or split arrays into smaller ones.
18. Write a program Stack arrays horizontally and vertically or perform broadcasting operations**.**
19. Create Pandas DataFrame from dictionaries, lists, or NumPy arrays.
20. Read and write data from various file formats such as CSV, Excel, JSON, SQL, etc., into Pandas DataFrame.
21. Use methods like **`head()`**, **`tail()`**, **`info()`**, **`describe()`** to view and inspect the DataFrame.
22. Use methods like **`loc[]`**, **`iloc[]`**, and boolean indexing to select rows and columns of the DataFrame.
23. Perform time series operations like resampling, shifting, and rolling window calculations.
24. Group data using **`groupby()`** and perform aggregation operations like **`sum()`**, **`mean()`**, **`count()`.**

SEMESTER-II

Name of the Course	Machine Learning		
Course Code	**MSDS-201**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Distinguish between various types of machine learning and explain key ML concepts.
CO2	Design and evaluate classification model using appropriate performance metrics.
CO3	Implement regression techniques and understand how gradient descent is used for optimization.
CO4	Apply ensemble methods to improve model performance.
CO5	Perform dimensionality reduction and unsupervised clustering for data analysis.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-201 Machine Learning

UNIT-l

The Machine Learning Landscape: What is Machine learning and why use machine learning, types of machine learning systems, Supervised and Unsupervised learning, Instants based vs Models based, Challenges in Machine Learning, Testing and Validation.

Classification: MNIST, Training a binary classifier, performance measures: measuring accuracy using cross validation, confusion matrix, precision and recall, precision/recall trade off, the ROC curve, multiclass classification, error analysis, multilevel classification, multioutput classification.

UNIT-II

Training Models using Regression: Linear Regression, Gradient Descent: Batch, Stochastic, Mini-Batch, Polynomial Regression, Learning Curves, Regularized Linear Models: Ridge Regression, Lasso Regression, Elastic Net, Early Stopping, Logistic Regression: Estimating Probabilities, Training and Cost Function, Decision Boundaries, Softmax Regression.

Support Vector Machines: Linear SVM, Softmargin Classification, Non Linear Classification: Polynomial Kernel, Adding Similarity Features, Gaussian RBF Kernel, Computational Complexity, SVM Regression.

Decision Trees: Training and Visualizing a Decision Tree, Making Predictions, Estimating Class Probabilities, The CART Training Algorithm, Computational Complexity, Gini Impurtiy or Entropy, Regularization Hyperparameters, Regression.

UNIT-III

Ensemble Learning and Random Forests: Voting Classifiers, Bagging and Pasting: Bagging and Pasting in Scikit-Learn, Out of Bag Evaluation, Random Patches and Random Subspaces, Random Forest:Extra –Trees, Feature Importance, Boosting:AdaBoost, Gradient Boosting, Stacking.

Dimensionality Reduction: The Curse of Dimensionality, Main Approaches for Dimensionality Reduction: Projection, Manifold Learning, PCA: Preserving the Variance, Principal Components, Projecting Down tod Dimensions, Using Scikit-Learn, Explained Variance Ratio, Choosing the Right Number of Dimensions, PCA for Compression, Randomized PCA, Incremental PCA, Kernel PCA: Selecting a Kernel and Tuning Hyperparameters, LLE.

UNIT-IV

Unsupervised Learning Techniques: Clustering: K-Means, Limit of K-means, Using clustering for image segmentation, Using Clustering for Preprocessing, Using Clustering for Semi-Supervised Learning, DBSCAN, Other Clustering Algorithms, Gaussian Mixtures: Anomaly Detection using Gaussian Mixtures, Selecting the Number of Clusters, Bayesian Gaussian Mixture Models, Other Anomaly Detection and Novelty Detection Algorithms.

Text Books:

1. Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd Edition, O'Reilly Media, 2022.
2. Alpaydin, Ethem. Introduction to Machine Learning, Prentice Hall Indian Learning Pvt. Ltd.,4th edition,2020.

Reference Books:

1. James Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning: With Applications in R. Germany, Springer New York, 2nd edition, 2021.
2. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 4th edition,2020.

3. Murphy, Kevin P. Machine Learning: A Probabilistic Perspective. Prentice Hall Indian Learning Pvt. Ltd.,2nd edition,2022.

4. Simon O. Haykin, Neural Networks and Learning Machines, Pearson Education,3rd edition, 2009.

Name of the Course	Big Data Analytics and Data Visualization		
Course Code	**MSDS-202**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain fundamental concepts of big data, its characteristics, data storage, and the Hadoop Ecosystem for managing big data.
CO2	Design and implement nosql database solutions for big data applications.
CO3	Utilize tools like mapreduce, Hive, and Spark to perform data analysis and apply ETL processes on big data sets.
CO4	Create effective data visualizations using Python libraries to extract insights and communicate data stories.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-202 Big Data Analytics and Data Visualization

UNIT-I

Introduction: Introduction to Big Data Analytic, Big Data, Scalability and Parallel Processing, Designing Data Architecture, Data Sources, Quality, Pre-Processing and Storing, Data storage and Analysis, Big Data Analytics Applications and Case Studies.

Introduction to Hadoop: Hadoop and its Ecosystem, Hadoop Distributed File System, Hadoop yarn, Hadoop Ecosystem Tools.

NoSQL Big Data Management: Introduction to NoSQL, NoSQL Data Store, NoSQL Data Architecture Patterns, NoSQL to Manage Big Data, Shared-Nothing Architecture for Big Data Tasks; MongoDB Database.

UNIT-II

MapReduce, Hive and Pig: Introduction, MapReduce Map Tasks, Reduce Tasks, and MapReduce Execution, Composing MapReduce for Calculations and Algorithms, Hive, HiveQL, Pig.

Spark and Big Data Analytics: Introduction, Spark, Introduction to Data Analysis with Spark, Downloading Spark, and Programming using RDDs and MLIB, Data ETL (Extract, Transform and Load) Process, Introduction to Analytics, Reporting and Visualizing.

UNIT-III

Introduction to Data Visualization with Python: Acquiring and Visualizing Data, Heavyweight Scraping with Scrapy , Simultaneous acquisition and visualization, Applications of Data Visualization, Keys factors of Data Visualization. Exploring the Visual Data Spectrum: Charting Primitives (Data Points, Line Charts, Bar Charts, Pie Charts, Area Charts), Pairwise data, Matplotlib Statistical distributions: (hist(x), boxplot(x), errorbar, violinplot, eventplot, hist2d, hexbin, ecdf), Gridded data, Irregularly gridded data: (tricontour, tricontourf, tripcolor, triplot), 3D and volumetric data, Lines, bars and markers, Subplots, axes and figures, Pie and Polar charts, Text, labels, and annotations, shapes and collections, Exploring Advanced Visualizations (Candlestick Charts, Bubble Charts, Surface Charts, Map Charts, Infographics).

Visualizing Data with Matplotlib: Introduction, Architecture, Elements, High Level Plotting-Historical Background, Pyplot and Object-Oriented Matplotlib, Starting an Interactive Session, Interactive Plotting with Pyplot's Global State: Configuring Matplotlib, Setting the Figure's Size, Points, Not Pixels, Labels and Legends, Titles and Axes Labels, Saving your Charts, Figures and Object-oriented Matplotlib, Plot Types, Matplotlib, NetworkX, , New Styles in Matplotlib, Seaborn; Data Analysis - SciPy & Seaborn.

UNIT-IV

Visualizing Data with D3: Imagining a Nobel Visualization, Building a Visualization: HTML skeleton, CSS styling, The JavaScript Engine, Introducing D3: framing the problem, working with selections, adding DOM elements, Leveraging D3, Measuring up with D3's Scales, unleashing the power of D3 with data binding, the enter method, accessing the bound data, the update pattern, axes and labels, transitions, Visualizing Individual Prizes, Mapping with D3: Available Maps, D3's Mapping Data formats, D3 Geo, projections, and paths, putting elements together, updating the map, adding value indicators.

Text Books:

1. Raj Kamal and Preeti Saxena, BIG DATA ANALYTICS: Introduction to Hadoop, Spark, and Machine-Learning, McGraw-Hill Education, 2019.
2. Kyran Dale, Data Visualization with Python and JavaScript, O'Reilly, 2nd Edition, 2022.

3. Scott Murray, Interactive Data Visualization for Web, O'Reilly,2nd edition, 2017.

Reference Books:

1. Jon Raasch, Graham Murray, Vadim Ogievetsky, Joseph Lowery, “JavaScript and jQuery for Data Analysis and Visualization”, WROX,2014.
2. Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley, 2013.
3. P. J. Sadalage, M. Fowler, NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, Addison-Wesley Professional, 2012.
4. Ritchie S. King, Visual story telling with D3” ,Pearson, 2015.
5. Ben Fry, Visualizing data: Exploring and explaining data with the processing environment, O'Reilly, 2007.
6. A Julie Steele and Noah Iliinsky, Designing Data Visualizations: Representing Informational Relationships, O'Reilly, 2011.
7. Andy Kirk, Data Visualization: A Successful Design Process, Packt, 2012.
8. Nathan Yau, Data Points: Visualization that means something, Wiley, 2013.

Name of the Course	Introduction to Computer Vision and Image Processing		
Course Code	**MSDS-203**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60(1 Hr Each) (L=60 , T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Apply fundamental image processing techniques, including transformations, spatial/frequency domain filtering, and noise reduction.
CO2	Utilize image compression techniques and understand the principles of digital watermarking.
CO3	Implement morphological operations and perform image segmentation.
CO4	Explain concepts in computer vision, including image formation, motion representation, 3D reconstruction, and object recognition.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-203 Introduction to Computer Vision and Image Processing

UNIT-I

Introduction and Digital Image Fundamentals: Applications of digital image processing, Fundamental steps in digital image processing, Image sensing and acquisition, image sampling and quantization, basic relationships between pixel, Gray level transformations.

Image enhancement in the spatial domain and frequency domain: Histogram Processing, local enhancement, image subtraction, image averaging, Fundamentals of Spatial Filtering, smoothing and sharpening spatial filters, Discrete Fourier transformation, filtering in the frequency domain,

image smoothing using Lowpass Frequency Domain filters and image sharpening using Highpass filters, Noise Models, Restoration in the presence of Noise Only- Spatial filtering, Periodic Noise Reduction using Frequency Domain Filtering, Image Reconstruction from Projections.

UNIT-II

Image Compression and Watermarking: Fundamentals, Huffman Coding, Golomb Coding, Arithmetic Coding, LZW Coding, Run-length Coding, Symbol-Based Coding, Bit-plane Coding, Block Transform Coding, Predictive Coding.

Morphological image processing: Erosion and dilation, opening and closing, hit-or-miss transformation, some basic morphological algorithms, Morphological Reconstruction.

Image segmentation: Point, line and edge detection, gradient operator, edge linking and boundary detection, thresholding, region-based segmentation, representation schemes like chain codes, polygonal approximations, boundary segments, skeleton of a region, boundary descriptor.

UNIT-III

Introduction to Computer Vision and motion representation: Image Formation, Feature based alignment, 2D and 3D feature based alignment, Pose estimation, geometric intrinsic calibration, Triangulation, two frame structure from motion, factorization, bundle adjustment, Translation alignment, parametric motion, spline based motion, optical flow, layered motion.

Computation Photography: Photometric calibration, high dynamic range imaging, super-resolution and blur removal, image matting and composing, Texture analysis and synthesis.

UNIT-IV

3D Reconstruction and Image-based rendering: Shape from X, Active rangefinding, Surface representations, point-based representations, volumetric representations, model-based reconstruction, recovering texture maps and albedos, layered depth images, light fields and lumigraphs.

Recognition: object detection, face recognition, instance recognition, category recognition, context and scene understanding, recognition databases and test sets.

Text Books:

1. Rafael C. Gonzalez and Richard E.Woods, Digital Image Processing (4th edition), Prentice–Hall of India, 2016.
2. Richard Szeliski, Computer Vision: Algorithms and Applications (2nd edition),Springer,2022.

Reference Books:

1. Bernd Jahne, Digital Image Processing, (7th edition), Springer, 2022.
2. M.A. Joshi, Digital Image Processing: An Algorithmic Approach (2nd edition), Prentice-Hall of India,2021

3. B. Chandra and D.D. Majumder, Digital Image Processing and Analysis, Prentice-Hall of India, 2011.
4. Reinhard Klette, Concise Computer Vision: An Introduction into Theory and Algorithms, Springer,4th edition, 2022.
5. E.R. Davies, Computer Vision: Principles, Algorithms, Applications, Learning(5th edition), 2017.

Name of the Course	IoT Based Data Analytics		
Course Code	**MSDS-204**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45(1 Hr Each) (L=45 , T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Understand the vision of IoT from a global context.
CO2	Analyze various protocols of IoT.
CO3	Determine the Market perspective of IoT
CO4	Apply IoT in various applications like Industrial, Commercial Building Automation, Manufacturing, etc.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-204 IoT Based Data Analytics

UNIT-I

IoT Fundamentals: Genesis of IoT, IoT and Digitization, IoT Impact, Connected Roadways, Connected Factory, Smart Connected Buildings, Smart Creatures, Convergence of IT and OT, IoT Challenges.

IoT Network Architecture and Design: Drivers Behind New Network Architectures, Scale, Security, Constrained Devices and Networks, Data, Legacy Device Support, Comparing IoT Architectures, IoT Data Management and Compute Stack, Fog Computing. Edge Computing, The Hierarchy of Edge, Fog, and Cloud.

UNIT-II

Smart Objects: The "Things" in IoT, Sensors, Actuators, and Smart Objects, Sensors, Actuators, Micro-Electro-Mechanical Systems (MEMS), Smart Objects, Smart Objects: A Definition, Trends in Smart Objects, Sensor Networks, Wireless Sensor Networks (WSNs), Communication Protocols for Wireless Sensor Networks.
IoT Access Technologies: IEEE 802.15.4, IEEE 802.15.4g and 802.15.4e, IEEE 1901.2a, IEEE 802.11ah, LoRaWAN, LTE Cat 0, LTE-M, NB-IoT Standardization and Alliances, Physical Layer, MAC Layer, Topology, Security, Competitive Technologies.

UNIT-III

IP as the IoT Network Layer: The Business Case for IP, The Key Advantages of Internet Protocol, Adoption or Adaptation of the Internet Protocol, The Need for Optimization, Constrained Nodes, Constrained Networks, IP Versions, Optimizing IP for IoT, From 6LoWPAN to 6Lo, Header Compression, Fragmentation, Mesh Addressing, Mesh-Under Versus Mesh-Over Routing, 6Lo Working Group, 6TiSCH, RPL.

UNIT-IV

Data and Analytics for IoT: An Introduction to Data Analytics for IoT, Structured Versus Unstructured Data, Data in Motion Versus Data at Rest, IoT Data Analytics Overview, IoT Data Analytics Challenges, Machine Learning, Machine Learning Overview, Supervised Learning, Unsupervised Learning, Neural Networks, Machine Learning and Getting Intelligence from Big Data, Predictive Analytics.
Securing IoT: A Brief History of OT Security, Common Challenges in OT Security, Erosion of Network Architecture, Pervasive Legacy Systems, Insecure Operational Protocols, Modbus, DNP3 (Distributed Network Protocol), ICCP (Inter-Control Center Communications Protocol), OPC (OLE for Process Control).

Text Books:

1. David Hanes, Gonzalo Salgueiro, Patrick Grossetete, Rob Barton and Jerome Henry, "IoT Fundamentals: Networking Technologies, Protocols and Use Cases for Internet of Things" , Cisco Press, 2017.
2. Raj Kamal, "Internet of Things – Architecture and Design", McGraw Hill, 1/e, 2017.

Reference Books :

1. Vijay Madisetti, Arshdeep Bahga, "Internet of Things: A Hands-On Approach". 2016.
2. Olivier Hersent, David Boswarthick, Omar Elloumi , "The Internet of Things – Key applications and Protocols" , Wiley, 2012.
3. Dieter Uckelmann, Mark Harrison, Michahelles, Florian (Eds), "Architecting the Internet of Things" , Springer, 2011.
4. Michael Margolis, Arduino Cookbook, "Recipes to Begin, Expand, and Enhance Your Projects" , O'Reilly Media, 2/e, 2011.

Name of the Course	Big Data Analytics and Data Visualization Lab		
Course Code	**MSDS-251**	Credits-2	L-0, T-0, P-4
Lectures to be Delivered	**60 hours of Lab Sessions**		
Semester End Examination	**Max Marks: 50**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Set up a Big Data environment, ingest data into HDFS, and explore storage formats.
CO2	Preprocess and clean large datasets for analysis using Big Data tools.
CO3	Implement machine learning algorithms on Big Data sets and evaluate model performance.
CO4	Create interactive visualizations to explore and communicate insights from large datasets.
CO5	Analyze Big Data scalability, optimize code, and evaluate performance trade-offs.

MSDS-251 Big Data Analytics and Data Visualization Lab

1: Data Ingestion and Storage

- Set up a Hadoop cluster or use a cloud-based Big Data service.
- Ingest a large dataset (e.g., CSV or JSON format) into the Hadoop Distributed File System (HDFS).
- Explore different storage formats (e.g., Parquet, ORC) and analyze the impact on storage size and query performance.
- Write MapReduce or Spark code to perform basic data transformations and store the processed data in a NoSQL database (e.g., MongoDB).

2: Data Preprocessing and Cleaning

- Choose a large dataset from a public repository (e.g., Kaggle).
- Apply data cleaning techniques such as handling missing values, outliers, and noise.
- Use Apache Spark or a similar framework to parallelize the data processing tasks.
- Split the dataset into training and testing sets for future analysis.

3: Data Analysis and Mining

- Select a real-world dataset related to a specific domain (e.g., customer churn prediction, fraud detection).
- Implement a machine learning algorithm (e.g., decision tree, logistic regression) using Apache Mahout or a similar library.
- Split the dataset into training, validation, and testing sets.
- Evaluate the performance of the model using appropriate performance metrics (e .g., accuracy, precision, recall).

4: Data Visualization

- Choose a large dataset with multiple dimensions and attributes.
- Use a visualization tool like Tableau or D3.js to create interactive visualizations.
- Explore different visualization techniques, such as histograms, scatter plots, and heatmaps.
- Present your findings and insights in a visually appealing and informative manner.

5: Scalability and Performance Optimization

- Scale up your analysis from Assignment 3 to handle larger datasets.
- Evaluate the performance of the Big Data platform in terms of processing time and resource utilization.
- Identify bottlenecks and optimize the code (e.g., using data partitioning, caching) to improve performance.
- Compare the performance of different hardware configurations or cloud-based services for Big Data analytics.

6: Case Study and Application

- Choose an industry or domain of interest (e.g., healthcare, finance, marketing).
- Identify a specific problem or challenge in that domain that can be addressed using Big Data analytics.
- Gather relevant datasets and perform exploratory data analysis.
- Apply appropriate data analysis and mining techniques to derive insights and solutions.
- Present your findings and recommendations in a comprehensive report or presentation.

Name of the Course	Machine Learning Lab		
Course Code	**MSDS-252**	Credits-2	L-0, T-0, P-4
Lectures to be Delivered	**60 hours of Lab Sessions**		
Semester End Examination	**Max Marks: 50**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Prepare data for machine learning: Handle missing values, encode categorical variables, and explore data using scatter plots.
CO2	Implement linear regression for prediction tasks.
CO3	Apply classification algorithms: Naive Bayes, Decision Trees, Support Vector Machines.
CO4	Utilize dimensionality reduction techniques and clustering algorithms.

MSDS-252 Machine Learning Lab

1. Write a Python program to Prepare a Scatter Plot (Use Forge Dataset / Iris Dataset).
2. Write a Python program to find all null values in a given data set and remove them.
3. Write a Python program the Categorical values in numeric format for a given dataset.
4. Write a Python program to implement simple Linear Regression for predicting House-price.
5. Write a Python program to implement multiple Linear Regression for a given dataset.
6. Write a Python program to implement Polynomial Regression for the given dataset.
7. Write a Python program to Implement Naïve Bayes.
8. Write a Python program to Implement Decision Tree whether or not to play tennis.
9. Write a Python program to implement linear SVM.
10. Write a Python program to transform data with Principal Component Analysis (PCA)
11. Write a Python program to implement the k-nearest Neighbors ML algorithm to build prediction model (Use Forge Dataset).
12. Write a Python program to implement the k-means algorithm on a synthetic dataset.
13. Write a Python program to implement Agglomerative clustering on a synthetic dataset.

ELECTIVE-1

Name of the Course	Database Management System		
Course Code	**MSDS-EL-211**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain core database concepts, advantages of a DBMS, data models, and DBMS architecture.
CO2	Design databases using the Entity-Relationship model and normalize them to optimize structure.
CO3	Construct and execute complex SQL queries including the use of joins, views, and indexes.
CO4	Understand transaction processing, concurrency control, and database recovery techniques to ensure data integrity.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-211 Database Management System

UNIT-I

Basic Concepts: File Systems vs. DMBS, Characteristics of the Data Base Approach, Abstraction and Data Integration, Database users, Advantages and Disadvantages of a DBMS.

Data Base Systems Concepts and Architecture: Schema and Instances, DBMS architecture and Data Independence, database languages and Interfaces, DBMS functions and component modules, Centralized and Client/Server Architectures for DBMS, Data Models.

Entity Relationship Model: Entity Types, Entity Sets, Attributes & keys, Relationships, Relationships Types, Roles and Structural Constraints, Design issues, E-R Diagrams, Design of an E-R Database Schema, Reduction of an E-R schema to Tables.

UNIT-II

Relational Data Model: Relational model concepts, Integrity constraints over Relations, Relational Algebra– Basic Operations, Relational Calculus, Codd Rules.

SQL: Data Definition and Data Types, Components of SQL: DDL, DML, and DCL, Schema Change Statement in SQL, Views, Joins& Queries in SQL, Specifying Constraints & Indexes in SQL, Database Triggers, SQL Injection.

UNIT-III

Relational Data Base Management System: RDBMS, Basic structure, Date Base Structure & its manipulation in an RDBMS, Storage Organization. Conventional Data Models: An Overview of Network and Hierarchical Data Models.

Relational Data Base Design: Functional Dependencies, Decomposition, Normal forms based on primary keys (1 NF, 2 NF, 3 NF, & BCNF), Multi-valued Dependencies, 4 NF, Join dependencies, 5 NF, Algorithms for Query Processing and Optimization.

UNIT-IV

Transaction Processing Concepts: Introduction to Transaction Processing, Transaction & System Concepts, Properties of Transaction, Schedules and Recoverability, Serializability of Schedules. **Concurrency Control Techniques:** Locking Techniques, Time stamp ordering, Multi-version Techniques, Optimistic Techniques, Granularity of Data items.

Recovery Techniques: Recovery concepts, Recovery and Atomicity, Recovery Algorithm, ARIES, Recovery in Multi database Systems, Remote Backup Systems.

Text Books:

1. Elmasri & Navathe, Fundamentals of Database systems, Pearson Education,7th edition, 2017.
2. Korth & Silberschatz, Database System Concept, McGraw Hill International Edition,7th edition, 2019.

Reference Books:

1. Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom, Database Systems: The Complete Book,Pearson,3rd edition, 2014.
2. Ivan Bayross, SQL, PL/SQL- The Program Language of ORACLE, BPB Publication,3rd revised edition, 2010.
3. Raghu Ramakrishnan& Johannes Gehrke, Database Management Systems, Mcgraw Hill,3rd edition,2003.

4. Peter Rob, Carlos Colonel, Database system Design, Implementation, and Measurement, Cengage Learning,13th edition,2019.

5. Alexis Leon & Mathews Leon: Database Management System, Leon Vikas Publication,2008.

Name of the Course	Information Retrieval		
Course Code	**MSDS-EL-212**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain fundamental concepts of information retrieval
CO2	Apply probabilistic and vector space models for document representation and retrieval.
CO3	Describe modern search engine architectures, including indexing, crawling, and ranking techniques.
CO4	Implement components of a search engine, demonstrating practical understanding of indexing, query processing, and ranking techniques.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-212 Information Retrieval

UNIT- 1

Introduction to Information Retrieval: Scope, goals, and applications. Historical perspectives of IR, Boolean Retrieval Model: Set theory concepts, inverted indices, query processing.

Information Needs and Relevance: Defining user queries, the concept of relevance, challenges in its measurement, Basic Evaluation Measures: Precision, recall, F-measure, mean average precision.

UNIT -II

Probabilistic Retrieval Models: Understanding uncertainty in relevance, ranking documents based on probability, the Binary Independence Model, Vector Space Model and tf-idf Weighting: Representing documents and queries as vectors, cosine similarity, term frequency-inverse document frequency.

Term Weighting Schemes: Variations on tf-idf, normalization, advanced weighting for better retrieval, Latent Semantic Indexing (LSI): Dimensionality reduction, uncovering hidden relationships between terms and documents.

UNIT- III

Indexing: Efficient index construction techniques, compression methods, handling dynamic updates, Web Crawling: Building a web crawler, focused crawling, challenges of web-scale data collection.

Search Engine Architecture: Components, query processing pipelines, distributed and parallel retrieval, Link Analysis and Ranking: PageRank algorithm, HITS, combating link spam.

UNIT -IV

Advanced Evaluation Techniques: Test collections, A/B testing, interleaved comparisons, beyond precision/recall, User Interfaces and Visualization: Search result presentation, search engine result pages (SERPs), visualizing retrieval results.

Text Classification and Clustering: Document categorization, clustering techniques for search result organization, Emerging Trends: Personalized search, question answering systems, conversational IR.

Text Book:

1. Stefan Buettcher, Charles L.A. Clarke, Gordon V. Cormack, Information Retrieval: Implementing and Evaluating Search Engines, MIT Press, 2016.

Reference Books:

1. Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze,**"Introduction to Information Retrieval"** , Cambridge University Press, 2008.
2. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, **"Modern Information Retrieval**" The Concepts and Technology Behind Search" , ACM Press, 1999.

Name of the Course	Software Engineering		
Course Code	**MSDS-EL-213**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain software engineering principles, analyze software development lifecycles, and select appropriate models.
CO2	Perform requirements analysis and create software specifications.
CO3	Design software architectures applying functional and object-oriented principles.
CO4	Utilize software maintenance, configuration management, and the importance of software certification.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-213 Software Engineering

UNIT- I

Introduction: Software Engineering, Changing nature of Software, Software Myths, Terminologies, Role of management in software development Software Process and desired Characteristics, Software Life Cycle Models: Build & Fix Model, Water Fall Model, Incremental Process Model, Evolutionary Process Models, Unified Process, Comparison of Models, Other Software Processes, Selection of a Model Software Requirements Analysis & Specifications: Requirements Engineering, Types of Requirements, Feasibility Studies, Requirements Elicitation, Requirements - Analysis Documentation, Validation and Management.

UNIT - II

Software Architecture: Its Role, Views, Component & Connector View and its architecture style, Architecture Vs Design, Deployment View & Performance Analysis, Documentation, Evaluation Software Project Planning: Size estimation, Cost Estimation, COCOMO, COCOMO – II, Software Risk Management.

UNIT - III

Function Oriented Design: Design principles, Module level Concepts, Notation & Specification, Structured Design Methodology, Verification Object-Oriented Design: OO Analysis & Design, OO Concepts, Design Concepts, UML – Class Diagram, Sequence & Collaboration Diagram, Other diagrams & Capabilities, Design Methodology – Dynamic and Functional Modeling, Internal Classes & Operations Detailed Design: PDL, Logic/Algorithm Design, State Modeling of Classes, Verification – Design Walkthroughs, Critical Design Review, Consistency Checkers.

UNIT- IV

Coding: Programming Principles & Guidelines, Coding Process, Refactoring, Verification Software Metrics: What & Why, Token Count, Data Structure Metrics, Information Flow Metrics, Object-Oriented Metrics, Use Case Oriented Metrics, Web Engineering Project Metrics, Metric Analysis Software Maintenance & Certification: Maintenance, Maintenance Process and Models, Estimation of Maintenance Costs, Regression Testing, Reverse Engineering, Software Re-engineering, Configuration Management, Documentation, Requirements of Certification, Types.

Text Books:

1. K.K. Aggrawal and Yogesh Singh, “Software Engineering”, 3rd Edition, New Age International (P) Ltd, 2008.
2. Pankaj Jalote, “An Integrated Approach to Software Engineering”, 3rd Edition, Narosa Publishing House, 2005.

Reference Books:

1. Pressman, R.S., “Software Engineering – A Practitioner's Approach”, Third Edition, McGraw Hills, 2008.

SEMESTER-III

Name of the Course	Deep Learning		
Course Code	**MSDS-301**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain the fundamentals of deep learning, including the biological inspiration, mathematical foundations, and architectures.
CO2	Design and train feedforward neural networks, apply regularization, and optimize for deep models.
CO3	Implement convolutional neural networks and apply them to image and text classification.
CO4	Build recurrent neural networks including advanced architectures, for sequence modeling and encoder-decoder tasks.
CO5	Utilize autoencoders for representation learning and dimensionality reduction.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-301 Deep Learning

UNIT-I

Introduction: Historical context and motivation for deep learning, biological neurons, neural networks, artificial neural networks ,linear perceptron, perceptron learning algorithm.

Applied maths and ML basics: scalars, vectors, matrices and tensors, multiplying matrices and vectors, special kinds of matrices and vectors, principal components analysis, probability distribution, conditional probability, independence and conditional independence, expectation, variance and covariance.

UNIT-II

Neural Networks: Feedforward neural networks, deep networks, regularizing a deep network, model exploration, and hyperparameter tuning.
Convolution Neural Networks: Introduction to convolution neural networks: stacking, striding and pooling, applications like image, and text classification.

UNIT-III

Sequence Modeling: Recurrent Nets: Unfolding computational graphs, recurrent neural networks (RNNs), bidirectional RNNs, encoder-decoder sequence to sequence architectures, deep recurrent networks, recursive neural networks, echo state networks, LSTM and other gated RNNs.
Autoencoders: Undercomplete autoencoders, regularized autoencoders, sparse autoencoders, denoising autoencoders, representational power, layer, size, and depth of autoencoders, stochastic encoders and decoders.

UNIT-IV

Structured probabilistic Models for Deep Learning: The challenge of unstructured modeling, using graphs to describe model structure, sampling from graphical models, Advantage of structured modeling, learning about dependencies, inference and approximate inference, the deep learning approach to structured probabilistic models.

Text Books:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2016.
2. Jeff Heaton, Deep Learning and Neural Networks, Heaton Research Inc, 2015

Reference Books:

1. Mindy L Hall, Deep Learning, VDM Verlag, 2011.
2. Li Deng ,Dong Yu, Deep Learning: Methods and Applications , Now Publishers Inc, 2014.

Name of the Course	Distributed and Cloud Computing		
Course Code	MSDS-302	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60(1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Understand distributed system fundamentals and communication models effectively.
CO2	Comprehend cloud computing models and virtualization techniques thoroughly.
CO3	Manage cloud infrastructure and data solutions for applications efficiently.
CO4	Analyze advanced topics and industry applications in cloud computing.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-302 Distributed and Cloud Computing

UNIT-I

Fundamentals of Distributed Systems: Introduction to Distributed Systems: Characteristics and Challenges, Architectures: Client-Server, Peer-to-Peer, Hybrid, Communication Models: Sockets, Remote Procedure Calls (RPC), Synchronization: Mutual Exclusion, Locks, Fault Tolerance: Failover Mechanisms, Redundancy, Clock Synchronization: Logical Clocks.

UNIT-II

Cloud Computing Basics and Virtualization: Overview of Cloud Computing, Cloud Service Models: IaaS, PaaS, SaaS, Virtualization Techniques: VMs, Containers, Hypervisors, Resource Management and Scheduling: Resource Provisioning, Auto-Scaling, Cloud Networking: Virtualized Networking, Software-Defined Networking (SDN), Advanced Virtualization: Container Orchestration with Kubernetes.

Basic Industry Application: AWS, Microsoft Azure, Google Cloud Platform, IBM Cloud.

UNIT-III

Cloud Infrastructure and Data Management : Cloud Data Centers: Data Center Architecture, Server Racks, Power and Cooling, Data Storage Solutions: NoSQL Databases, Object Storage, Key-Value Stores, Distributed File Systems: HDFS, Ceph, Big Data Technologies: Hadoop, MapReduce, Apache Spark, Data Management in IoT: IoT Data Streams, Storage Challenges, Distributed Machine Learning: Training Algorithms in Cloud Environments

UNIT-IV

Advanced Topics in Distributed and Cloud Computing: Security and Privacy: Security Threats, Data Encryption, Edge Computing: Edge Nodes, Fog Computing, IoT Cloud Integration, Federated Learning: Privacy-Preserving Machine Learning, Blockchain Technology: Decentralization, Smart Contracts, Distributed AI: Collaboration of AI Agents, Serverless Computing: Functions as a Service (FaaS), Cloud Infrastructure for IoT: Architecture and Management.

Basic Industry Application: Serverless Architectures (AWS Lambda, Azure Functions), Edge AI (NVIDIA Jetson), Blockchain for Supply Chain Management.

Textbooks:

1. Andrew S. Tanenbaum and Maarten Van Steen, Distributed Systems: Principles and Paradigms, 4th Edition, 2024, Prentice Hall.
2. Thomas Erl, Cloud Computing: Concepts, Technology & Architecture, 1st Edition, 2013, Prentice Hall.
3. Kai Hwang, Geoffrey Fox, and Jack Dongarra - Distributed and Cloud Computing: From Parallel Processing to the Internet of Things, 2nd Edition, 2017, Morgan Kaufmann.

Reference Books:

1. Rob Kloss, Distributed Computing: Principles and Applications, 1st Edition, 2015, Springer.
2. Raghunath Nambiar and Kenjiro Taura, Cloud Computing: Theory and Practice, 1st Edition, 2015, Morgan Kaufmann.
3. M. S. S. S. N. R. Bhargav, Cloud Computing: Principles and Paradigms, 1st Edition, 2016, Wiley.

Name of the Course	Natural Language Processing		
Course Code	**MSDS-303**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain core NLP concepts, levels of linguistic analysis, and applications of NLP.
CO2	Design grammars for natural language, utilize parsing techniques bottom-up, and address ambiguity.
CO3	Understand principles of semantic interpretation, create logical forms, and address thematic roles in language.
CO4	Build language models, implement machine translation systems, and apply NLP to tasks like information retrieval.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-303 Natural Language Processing

UNIT-I

Introduction to Natural language Processing: The Study of Language, Applications of NLP, Evaluating Language Understanding Systems, Different Levels of Language Analysis, Representations and Understanding, Organization of Natural language Understanding Systems, Linguistic Background: An outline of English Syntax, Elements of Simple Noun Phrases, The Elements of Simple Sentences, Prepositional Phrases, Embedded Sentences, Complements, Adjective Phrases.

UNIT-II

Grammars and Parsing: Grammars and Parsing- Top- Down and Bottom-Up Parsers, Transition Network Grammars, Feature Systems and Augmented Grammars, Morphological Analysis and the Lexicon, Parsing with Features, Augmented Transition Networks.

Grammars for Natural Language: Grammars for Natural Language, Movement Phenomenon in Language, Handling questions in Context Free Grammars, Hold Mechanisms in ATNs, Gap Threading, Human Preferences in Parsing, Shift Reduce Parsers, Deterministic Parsers.

UNIT-III

Semantic Interpretation: Semantic & Logical form, Word senses & ambiguity, The basic logical form language, Encoding ambiguity in the logical Form, Verbs & States in logical form, Thematic roles, Speech acts & embedded sentences, Defining semantics structure model theory.

Language Modeling :Introduction, n-Gram Models, Language model Evaluation, Parameter Estimation, Language Model Adaption, Types of Language Models, Language-Specific Modeling Problems, Multilingual and Crosslingual Language Modeling, Neural Networks and Neural Language Models, Feedforward networks for NLP, RNNs as Lannguage Models, RNNs for other NLP tasks, The LSTM, Transformers as Language Models.

UNIT-IV

Machine Translation and Encoder-Decoder Models: Introduction, Problems of Machine Translation, Is Machine Translation Possible, Machine Translation Evaluation, Word Alignment, Phrase-Based Models, Tree-Based Models, Linguistic Challenges, Language Divergences and Typology, The Encoder-Decoder Model, Beam Search, Encoder-Decoder with RNNs and Transformers, MT Systems.

Multilingual Information Retrieval: Introduction, Document Preprocessing, Monolingual Information Retrieval, CLIR, MLIR, Evaluation in Information Retrieval, Tools, Software and Resources.

Combining Natural Language Processing Engines: Introduction, Desired Attributes of Architectures for Aggregating Speech and NLP Engines, Architectures for Aggregation, Case Studies, Lessons Learned.

Text Books:

1. James Allen, Natural Language Understanding, 2nd Edition, 2003, Pearson Education.
2. Daniel Jurafsky and James H., Speech and Language Processing, 3rd Edition, Martin Prentice Hall,2024.
3. Daniel M.Bikel and Imed Zitouni , Multilingual Natural Language Processing Applications : From Theory To Practice, IBM Press,2012.

Reference Books:

1. Charniack, Eugene, Statistical Language Learning, MIT Press, 1993.
2. Akshar Bharathi,Vineet chaitanya,Natural Language Processing, A paninian perspective, Prentice –Hall of India, 2008.
3. Jurafsky, Dan and Martin, James, Speech and Language Processing, 2nd Edition, Pearson Education, 2014.
4. Manning, Christopher and Henrich, Schutze, Foundations of Statistical Natural Language Processing, MIT Press, 1999.

Name of the Course	Deep Learning Lab		
Course Code	**MSDS-351**	Credits-2	L-0, T-0, P-4
Lectures to be Delivered	**60 hours of Lab Sessions**		
Semester End Examination	**Max Marks: 50**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Build basic and multi-layer neural networks for classification.
CO2	Apply gradient descent and backpropagation for training.
CO3	Optimize deep learning models.
CO4	Implement convolutional neural networks for image tasks.
CO5	Use rnns with LSTM for sequence modeling.

MSDS-351 Deep Learning Lab

1. Implement a basic feedforward neural network using Python and a deep learning framework like TensorFlow or Keras. Train the network on a simple dataset, such as the XOR problem, and analyze the model's performance.
2. Build a multi-layer perceptron (MLP) to classify images from the MNIST dataset. Experiment with different architectures and activation functions to achieve higher accuracy.
3. Implement gradient descent and backpropagation algorithms from scratch in Python to train a simple neural network. Compare the performance with the same model trained using a deep learning framework.
4. Explore the impact of different optimization algorithms (e.g., Adam, RMSprop, SGD) and learning rates on the convergence of a deep neural network.
5. Implement a simple autoencoder using Python and TensorFlow/Keras to perform dimensionality reduction on a dataset.
6. Build a simple feedforward neural network to classify handwritten digits from the MNIST dataset.
7. Experiment with different architectures, activation functions, and optimization algorithms to achieve higher accuracy.

8. Implement a multi-layer perceptron (MLP) neural network from scratch using Python and NumPy. Train the network on a synthetic dataset and visualize the decision boundaries.
9. Design and train a convolutional neural network (CNN) to classify images from the CIFAR-10 dataset. Evaluate the model's performance and analyze common misclassifications.
10. Implement a recurrent neural network (RNN) with Long ShortTerm Memory (LSTM) cells to perform sentiment analysis on a text dataset. Analyze the model's predictions and discuss the importance of sequence modeling.

Name of the Course	Natural Language Processing Lab		
Course Code	**MSDS-352**	Credits-2	L-0, T-0, P-4
Lectures to be Delivered	60 Hrs of Lab Session		
Semester End Examination	**Max Marks: 50**	**Min Pass Marks: 40%**	**Max. Time: 3 hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Pre-process text data (cleaning, normalization) using libraries like NLTK and spaCy.
CO2	Implement bag-of-words and TF-IDF for document representation and explore spelling correction techniques.
CO3	Implement bag-of-words and TF-IDF for document representation and explore spelling correction techniques.
CO4	Apply constituency/dependency parsing, explore word sense disambiguation, and develop sentiment analysis systems.

Natural Language Processing Lab

List of practical:

1. **Text Preprocessing Program**: Write a program that cleans text data from a file. This could involve tasks like removing punctuation, converting to lowercase, removing stop words, and performing stemming/lemmatization. (Libraries: NLTK, spaCy)
2. **Bag-of-Words and TF-IDF:**
   Implement functions to convert text documents into two numerical representations:
   - Bag-of-Words (BoW): Represents a document as a dictionary where keys are words and values are their frequency in the document.
   - TF-IDF (Term Frequency-Inverse Document Frequency): Assigns weights to words based on their importance in a document collection.
3. **Spelling Correction:**
   - Develop an NLP system that can correct spelling errors in a given text.
   - Explore approaches like edit distance, n-grams, and context-based correction.
4. **Language Modeling:**
   - Create a language model using n-gram models or recurrent neural networks(RNNs).
   - Train the model on a large corpus and evaluate its performance.
5. **Part-of-Speech(POS) Tagging**
   - Write a program to perform POS tagging on sentences.
6. **Constituency and Dependency Parsing:**
   - Implement constituency parsing(e.g. Using CYK algorithm) and dependency parsing (e.g. using transition-based parsers).

7. **Lexical Semantics:**
    - Investigate word sense disambiguation techniques.
    - Build a system that disambiguates word meanings based on context.
8. **Sentiment Analysis:**Build a sentiment analysis classifier to determine if a piece of text expresses positive, negative, or neutral sentiment.
   Here's the process:
    - Collect a sentiment analysis dataset where each text snippet is labeled as positive, negative, or neutral.
    - Train a machine learning model (e.g., Naive Bayes) on the labeled data.
    - Use the trained model to predict sentiment on new, unseen text data.
    - Evaluate the model's performance using metrics like accuracy, precision, and recall.
      Python Libraries: NLTK, spaCy, scikit-learn
9. **Extractive Text Summarization**:
    - Implement an extractive summarization algorithm that identifies key sentences based on features like sentence position, word frequency, or sentence length.
    - Utilize libraries like NLTK's LexRank algorithm or spaCy's summarization functionality
10. **Text Classification:**
    - Build a text classifier (e.g. sentiment analysis, spam detection) using machine learning algorithms (e.g. Naive Bayes, SVM, neural networks)
    - Train and evaluate the classifier on labeled datasets.

Reference Books:

1. Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python (1st ed.). O'Reilly Media.
2. Jurafsky, D., & Martin, J. H. (2021). Speech and language processing (3rd ed.). Pearson.+
3. Goldberg, Y. (2017). Neural network methods for natural language processing (1st ed.). Morgan Kaufmann Publishers

Libraries :

1. NLTK, spaCy, scikit-learn, TensorFlow, Keras(for RNNs).

ELECTIVE-2

Name of the Course	Data Engineering		
Course Code	**MSDS-EL-311**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain the role of data engineering, the data lifecycle, and design data warehouses and data lakes.
CO2	Design etl processes, perform data cleaning, and apply feature engineering techniques for machine learning.
CO3	Utilize distributed systems and understand big data processing and stream processing concepts.
CO4	Work with cloud data platforms, implement data security measures, and understand data governance principles.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-311 Data Engineering

UNIT- I

Role of Data Engineering: Its relationship to data science and AI, critical components of data engineering within an organization, Data Lifecycle: Data generation, collection, ingestion, storage, processing, analysis, visualization, and the ethical considerations surrounding them.

Data Warehousing and Data Lakes: Concepts, design principles, OLAP vs. OLTP, data lake architectures, Databases: Relational (SQL) databases, NoSQL databases (document, columnar, graph), and choosing the right databases for different needs.

UNIT- II

ETL Processes: Extract, Transform, and Load concepts, design patterns, and best practices.

Data Cleaning and Quality: Techniques for identifying and addressing missing values, outliers, errors, inconsistencies, and bias.

Data Wrangling Tools: Introduction to popular tools like Python (Pandas), Spark, R (dplyr, tidyr), or SQL, Feature Engineering: Techniques for creating, selecting, and transforming features to enhance machine learning models.

UNIT - III

Big Data Fundamentals: Defining characteristics of Big Data (Volume, Velocity, Variety), implications for data engineering, Distributed Systems: MapReduce, Hadoop, Spark, architectures, and use cases.

Stream Processing: Concepts of real-time data, processing technologies (Kafka, Flink, Storm), and applications, Data Pipelines: Building robust, fault-tolerant pipelines, workflow orchestration (Airflow, Luigi).

UNIT- IV

Cloud Data Platforms: AWS (S3, Redshift, Glue), Azure (Data Lake Storage, Synapse),Google Cloud (BigQuery, Dataflow), Data Security and Privacy: Encryption, masking, access control, compliance with regulations (GDPR, CCPA).

Data Governance: Establishing data quality rules, lineage, metadata management, and data ownership, DevOps for Data Engineering: CI/CD principles, infrastructure as code (Terraform, etc.), and monitoring.

Text books:

1. Paul Crickard ,Data Engineering with Python, Packt Publishing, 2020.
2. Liang Zhao, Sherif Sakr, Anna Liu, Athman Bouguettaya,Cloud Data Management, SpringerLink, 2014.

Reference book:

1. Martin Kleppmann , Designing Data-Intensive Applications, O'Reilly Media, Inc. , 2017.

Name of the Course	Soft computing		
Course Code	MSDS-EL-312	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	Max Marks: 75	Min Pass Marks: 40%	Max. Time: 3 Hrs
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			Max Marks: 25 Min Marks: 10

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain the principles of soft computing, its advantages, and differentiate it from traditional computing.
CO2	Design and implement fuzzy logic systems, understanding membership functions, operators, and inference for control applications.
CO3	Explain artificial neural network architectures, learning algorithms, and their applications.
CO4	Apply genetic algorithms to solve optimization problems, understanding the principles of selection, crossover, and mutation.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-312 Soft computing

UNIT - I

Introduction to Soft Computing: Definition and Scope of Soft Computing, Applications of Soft Computing in Various Fields, Advantages of Soft Computing over Traditional Computing (e.g., handling uncertainty, learning and adaptation, robustness to errors),Overview of Fuzzy Logic, Neural Networks, and Evolutionary Computation.

UNIT - II

Fuzzy Logic Systems: Fundamentals of Fuzzy Set Theory, Membership Functions and Fuzzy Operators, Fuzzy Reasoning and Inference, Applications of Fuzzy Logic Control Systems (e.g., temperature control systems, robotics),Design and Implementation of Fuzzy Logic Systems .

UNIT- III

Artificial Neural Networks: Introduction to Artificial Neural Networks (ANNs),Structure of Neural Networks: Neurons, Layers, Connections, Learning Algorithms for Neural Networks (e.g., Perceptrons, Backpropagation),Different Network Architectures for Various Tasks (e.g., feedforward, recurrent),Optional: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs),Applications of Neural Networks in Pattern Recognition, Prediction, and Optimization.

UNIT - IV

Evolutionary Computation: Introduction to Evolutionary Computation, Principles of Genetic Algorithms: Selection, Crossover, Mutation, Applications of Evolutionary Computation in Optimization Problems (e.g., scheduling, resource allocation),Hybrid Soft Computing Approaches: Integration of Fuzzy Logic, Neural Networks, and Evolutionary Computation.

Text Books:

1. S. N. Sivanandam and S. Sumathi , Principles of Soft Computing, Alpha Science International Ltd., , 3rd edition , 2018 .
2. Dr. Nilakshi Jain, Artificial Intelligence Making a System Intelligent, wiley, 2019.

Reference Books:

1. Timothy J. Ross, "Fuzzy Logic with Engineering Applications", 4th edition, Wiley, 2017.
2. George J. Klir and Bo Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications", Prentice Hall, 1995.

Name of the Course	Complex Networks Analysis		
Course Code	**MSDS-EL-313**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45(1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Represent and analyze networks using graph theory, centrality measures, and community detection algorithms.
CO2	Model network formation, diffusion processes, and assess network resilience.
CO3	Analyze the dynamics of temporal networks.
CO4	Apply network science concepts to solve problems across various domains.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-313 Complex Networks Analysis

UNIT- I

Graph Theory Foundations: Degree distributions, path lengths, connected components, clustering coefficients, Random Network Models: Erdős–Rényi models, small-world networks, Watts-Strogatz models, preferential attachment models (Barabási–Albert model).

UNIT- II

Centrality Measures: Degree centrality, Eigenvector centrality, Betweenness centrality, Closeness centrality, PageRank, Community Detection: Modularity-based methods, hierarchical clustering, spectral clustering, label propagation.

Influence and Diffusion: Diffusion models (threshold models, cascade models), influence maximization, identifying influential nodes.

UNIT - III

Network Evolution Models: Node/edge addition/removal, network growth models.

Resilience and Robustness: Vulnerability analysis, cascading failures, robustness metrics.

Temporal Networks: Representations, metrics, dynamics on temporal networks.

UNIT- IV

Social Networks: Network formation, information diffusion, community dynamics, social influence, Biological Networks: Protein interaction networks, gene regulatory networks, metabolic networks, Infrastructure Networks: Transportation networks, power grids, communication networks.

Text Books:

1. Mark Newman ,"Networks: An Introduction", Oxford University Press , 2010 .
2. Albert-László Barabási , "Network Science", Cambridge University Press , July 2016 .

ELECTIVE-3

Name of the Course	Time Series Analysis And Forecasting Techniques		
Course Code	**MSDS-EL-321**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Analyze time series data, decompose components, explore smoothing techniques, and assess stationarity.
CO2	Build and understand time series models: Autoregressive , Moving Average, ARMA, and ARIMA processes.
CO3	Select appropriate forecasting models, apply model selection techniques, and evaluate forecast accuracy.
CO4	Utilize transfer function models, intervention analysis, and spectral analysis for advanced time series modeling.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-321 Time Series Analysis and Forecasting Techniques

UNIT-I

Exploratory Analysis of Time Series: Graphical display, classical decomposition model, Components and various decompositions of Time Series Models-Numerical description of Time Series: Stationarity, Auto covariance and Autocorrelation functions - Data transformations - Methods of estimation –Trend, Seasonal and exponential.

Smoothing Techniques: Moving Averages: Simple, centered, double and weighted moving averages; single and double exponential smoothing – Holt's and winter's methods - Exponential smoothing techniques for series with trend and seasonality-Basic evaluation of exponential smoothing.

UNIT-II

Stationary Models: Time series data, Trend, seasonality, cycles and residuals, Stationary, White noise processes, Autoregressive (AR), Moving Average (MA), Autoregressive and Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) processes, Choice of AR and MA periods.

Non-Stationary Time Series Models: Tests for Non-stationarity: Random walk –random walk with drift –Trend stationary –General Unit Root Tests: Dickey Fuller Test, Augmented Dickey Fuller Test. ARIMA Models: Basic formulation of the ARIMA Model and their statistical properties - Autocorrelation function (ACF), Partial autocorrelation function (PACF) and their standard errors.

UNIT-III

Forecasting: Nature of Forecasting – Forecasting methods- qualitative and quantitative methods – Steps involved in stochastic model building – Forecasting model evaluation. Model selection techniques: AIC, BIC and AICC – Forecasting model monitoring.

Transfer Function And Intervention Analysis: Transfer function models- Transfer function – noise models; Cross correlation function; Model specification; Forecasting with Transfer function – noise models; Intervention analysis.

UNIT-IV

Spectral Analysis: Spectral density function (s. d. f.) and its properties, s. d. f. of AR, MA and ARMA processes, Fourier transformation and periodogram.

Contemporary Issues: Research and Analytical problems on various applications of the sampling Techniques.

Text Books:

2. Douglas C. Montgomery, Cheryl L. Jennings, Murat Kulahci, Introduction to Time Series Analysis and Forecasting, Second Ed., Wiley, 2016**.**
3. George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung, Time Series Analysis: Forecasting and Control, Fifth Ed., Wiley, 2016.

Reference Books:

1. Brockwell, P. J., & Davis, R. A., Introduction to time series and forecasting, Third Ed., Springer, 2016.
2. Terence C. Mills, Applied Time Series Analysis: A Practical Guide to Modeling and Forecasting, Academic Press, 2019.

Name of the Course	Multicriteria Decision Making		
Course Code	**MSDS-EL-322**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Apply foundational MCDM methods to structure decision problems, define criteria, and model preferences.
CO2	Address uncertainty in decision-making using fuzzy sets, rough sets, and probabilistic techniques.
CO3	Utilize advanced MCDM methods for complex decisions and facilitate group decision-making processes.
CO4	Integrate MCDM principles into data science and AI applications for feature selection, resource allocation, model selection, and recommendation systems.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-322 Multicriteria Decision Making

UNIT - I

Key Decision Elements: Problem framing, defining objectives/criteria, identifying alternatives, preference modeling, uncertainty, Preference Structures: Ordering techniques, utility functions, trade-offs, and their elicitation.

Classical MCDM Methods: Weighted sum, weighted product, AHP (Analytic Hierarchy Process),

ELECTRE methods.

UNIT- II

Fuzzy Sets in MCDM: Representing preferences and criteria using fuzzy sets, fuzzy aggregation operators, Rough Sets: Dealing with qualitative data, handling inconsistency.

Probabilistic MCDM: Stochastic dominance, techniques involving probability distributions and Monte Carlo simulation, Handling Uncertainties in Preferences: Sensitivity analysis, robustness analysis.

UNIT - III

Outranking Methods: PROMETHEE, ELECTRE-III, focus on preference modeling over strict ranking, Goal Programming: Optimization under multiple conflicting goals, preemptive and non-preemptive approaches.

Group Decision Making: Aggregation of preferences and consensus building, Interactive MCDM: Incorporating the decision-maker's evolving preferences in the decision-making process

UNIT- IV

Feature Selection and Ranking: Using MCDM techniques for dimensionality reduction in data science, Resource Allocation: Efficient distribution of computational and data resources using MCDM frameworks.

Decision Support Systems: Designing intelligent systems incorporating MCDM principles. Algorithm and Model Selection: Evaluating and choosing machine learning approaches based on multiple performance metrics, Recommender Systems: Building systems considering multiple user preferences and attributes.

Text Books:

1. Salvatore Greco, Matthias Ehrgott, José Rui Figueira , "Multiple Criteria Decision Analysis: State of the Art Surveys", Springer New York, 2016.
2. Valerie Belton and Theodor Stewart , "Multiple Criteria Decision Making" , Springer New York, 2002.

Name of the Course	Queuing Theory and Network Analysis		
Course Code	**MSDS-EL-323**	Credits-3	L-3, T-0, P-0
Lectures to be Delivered	45 (1 Hr Each) (L=45, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Apply fundamental queuing theory concepts to model and analyze single-server and multi-server systems.
CO2	Analyze advanced queuing models, including priority queues and bulk service, and utilize simulation techniques.
CO3	Explain network fundamentals, topologies, protocols, routing, switching, and analyze network performance metrics.
CO4	Apply queuing theory and network analysis to solve problems.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-323 Queuing Theory and Network Analysis

UNIT -I

Basic Terminology and Concepts: Arrivals, departures, service times, queues, servers, utilization, Little's Law, Classification of Queuing Systems: Kendall's notation, single-server vs. multi-server queues, finite vs. infinite capacity.

Poisson Process and Exponential Distribution: Their role in modeling arrivals and service times, Simple Queuing Models: M/M/1, M/M/c, M/G/1 queues, derivation of performance measures (queue length, waiting time).

UNIT- II

Priority Queues: Preemptive and non-preemptive disciplines, analysis of waiting times.

Bulk Service Queues: Queueing models with batch arrivals and/or batch service.

Queueing Networks: Jackson networks, open and closed networks, product-form solutions, Simulation of Queuing Systems: Discrete-event simulation techniques, generation of random variates, analysis of simulation output.

UNIT- III

Network Fundamentals: Network topologies (star, bus, ring), OSI Model, TCP/IP protocol suite, Routing and Switching: Routing algorithms (distance vector, link-state), switching techniques (circuit, packet).

Network Performance Analysis: Metrics (throughput, delay, jitter), traffic flow modeling.

Network Security: Firewalls, intrusion detection systems, encryption techniques.

UNIT -IV

Resource Allocation and Scheduling: Server farms, cloud computing, task scheduling in distributed systems, Web Traffic Analysis and Modeling: Understanding user behavior, predicting traffic patterns, website design optimization.

Network Optimization in Machine Learning: Distributed training, communication bottlenecks in deep learning models, Social Network Analysis: Centrality measures, community detection, information diffusion.

Text Books:

1. U. Narayan Bhat ,An Introduction to Queuing Systems: From Theory to Applications, Birkhäuser, 2015.
2. Mark Newman ,Networks: An Introduction" , Oxford University Press, 2010.

Reference Books:

1. Albert-László Barabási , “Network Science" , Cambridge University Press, 2016.
2. Mor Harchol-Balter. , Performance Modeling and Design of Computer Systems: Queueing Theory in Action" , Cambridge University Press, 2013. .
3. Frank Kelly and Elena Yudovina , Stochastic Networks: Theory and Applications", Cambridge University Press, 2014.

SEMESTER-IV

Name of the Course	**PROJECT WORK**		
Course Code	**MSDS-401**	Credits-18	Internship/Project
Project/Internship outside HPU			
Semester End Examination	External Marks: 250		Min. Pass Marks:100
Internal Assessment (based on Internship/Project Report 50%, Seminar 25% and 25% Log Book & Interim Report)		Max. Marks:200	Min. Pass Marks:80
OR			
Project/Internship inside HPU			
Semester End Examination	External Marks: 150		Min. Pass Marks:60
End Semester Examination Elective-4	External Marks: 75		Min. Pass Marks:30
End Semester Examination Elective-4	Internal Assessment: 25		Min. Pass Marks:10
Internal Assessment (based on Project/Internship Report 50%, Seminar 25% and 25% Log Book & Interim Report)		Max. Marks:200	Min. Pass Marks:80

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Utilize the methods and techniques of system development.
CO2	Apply best practices for effective project management.
CO3	Analyze the real life problems in order to develop optimum solutions.
CO4	Evaluation of project deliverables.
CO5	Creation of team spirit and profession ethics.

The fourth semester of MSc Data Science primarily focuses on practical experience and applied knowledge through project work, internships, and electives. Students will engage in independent projects, and the evaluation will be based on several components, including seminars, reports, and viva-voce exams.

In the Third semester, the Chairman/Head of the Department will assign a guide/supervisor to each candidate for his/her Project Work. The candidate shall be required to maintain his/her Project diary (logbook) of work in the Organization / Department. Each student will be required to give at least two seminars on his/her project work (one in the mid semester and other in the end of the semester), dates to be decided by the department. Each student is required to submit three copies of his/her project reports in the Department after completion of the project work which will be evaluated by

an external examiner.

Further, the Project work (final semester) will be jointly evaluated by an internal guide and external examiner. Internal assessment will be given on Internship/Project Report, Seminar and Log Book & Interim Report, be evaluated by the department.

1. **Project/Internship Work outside HPU:**

    - Students are required to undertake a significant project during this semester.
    - **Seminar:** Students will present their project progress through two seminars, which will account for 50 marks.
    - **Log Book & Interim Report:** Regular tracking of project work will be assessed through logbooks and interim reports, worth another 50 marks.
    - **Project/Internship Report:** After completing their project, students will submit a project/internship report, accounting for 100 marks.
    - **Viva-Voce**: After completing their project, students will participate in a viva-voce, accounting for 250 marks.
2. **Elective Path :**
    - Alternatively, students can choose to work on an elective (Elective 4) instead of a project/ internship outside HPU. The elective carries 100 marks.
    - **Seminar:** Students will present their project progress through two seminars, which will account for 50 marks.
    - **Log Book & Interim Report:** Regular tracking of project work will be assessed through logbooks and interim reports, worth another 50 marks.
    - In this pathway, students will still need to submit a project/internship report, which will account for 100 marks.
    - **Viva-Voce**: After completing their project, students will participate in a viva-voce, accounting for 150 marks.
    - and complete a viva-voce for 250 marks.

Total Marks Distribution:

- The entire semester is evaluated out of 450 marks.
    - For students undertaking project work and internship: **250** marks (EE) + 2**00** marks (IA) = **450** total marks.
    - For students opting for the elective path: **225** marks (EE) + **225** marks (IA) = **450** total marks.

ELECTIVE-4

Name of the Course	Predictive Analysis		
Course Code	**MSDS-EL-411**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Build and interpret simple and multiple linear regression models for prediction.
CO2	Diagnose and address common regression problems.
CO3	Extend linear regression using Generalized Linear Models.
CO4	Apply model validation techniques to assess model performance.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and will carry 27 marks out of the total marks of the semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-411 Predictive Analysis

UNIT-II

Simple Regression Analysis: Introduction to a linear and nonlinear model. Ordinary Least Square methods. Simple linear regression model, using simple regression to describe a linear relationship. Fitting a linear trend to time series data, Validating simple regression model using t, F and p test. Developing confidence interval. Precautions in interpreting regression results.

Multiple Regression Analysis: Concept of Multiple regression model to describe a linear relationship, Assessing the fit of the regression line, inferences from multiple regression analysis, problem of overfitting of a model, comparing two regression model, prediction with multiple regression equation.

UNIT-II

Fitting Curves and Model Adequacy Checking: Introduction, fitting curvilinear relationship, residual analysis, PRESS statistics, detection and treatment of outliers, lack of fit of the regression model, test of lack of fit, Problem of autocorrelation and heteroscedasticity. Estimation of pure errors from near neighbors.

Transformation techniques: Introduction, variance stabilizing transformations, transformations to linearize the model, BoxCox methods, transformations on the repressor's variables, Generalized and weighted least squares, Some practical applications.

UNIT-III

Multicollinearity: Introduction, sources of multicollinearity, effects of multicollinearity. Multicollinearity diagnostics: examination of correlation matrix, variance Inflation factors (VIF), Eigen system analysis of X1X. Methods of dealing with Multicollinearity: collecting additional data, model , re-specification, and ridge regression.

Generalized Linear Models: link functions and linear predictors, parameter estimation and inference in the GLM, prediction and estimation with the GLM, Residual Analysis, and concept of over dispersion.

UNIT-IV

Model building and Nonlinear Regression: Variable selection, model building, model misspecification. Model validation techniques: Analysis of model coefficients, and predicted values, data splitting method. Nonlinear regression model, nonlinear least squares, transformation to linear model, parameter estimation in nonlinear system, statistical inference in nonlinear regression.

Contemporary issues: Research and Analytical problems on various applications of the regression analysis and predictive modeling.

Text Books

[1]. Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining," Introduction to Linear Regression Analysis "Wiley India Pvt. Ltd 6th edition 2021.

[2]. Norman R. Draper, Harry Smith," Applied Regression Analysis" Wiley India Pvt. Ltd 3rd edition 2016.

Reference Books

[1]. Johnson, R A., Wichern, D. W," Applied Multivariate Statistical Analysis", PHI learning, 6th edition, 2021.

[2]. Iain Pardoe, "Applied Regression Modeling", John Wiley and Sons, 3rd edition,2020.

Name of the Course	Social Media Analytics		
Course Code	**MSDS-EL-412**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60(1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Apply graph theory concepts to understand social network structures and employ tools for social network analysis.
CO2	Collect and preprocess social media data from various platforms, addressing ethical concerns.
CO3	Perform exploratory data analysis on social media datasets, using text analysis and visualization techniques.
CO4	Apply machine learning to social media data for tasks like sentiment analysis, trend prediction, and recommendation systems.
CO5	Analyze case studies on social media marketing, social movements, and computational social science to understand real-world applications.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-412 Social Media Analytics

UNIT-I

Introduction to Social Networks: Graph theory basics, network structures (directed/undirected, weighted), common social network properties (centrality, clustering, etc.), Social Media Platforms and Data - Overview of major platforms (Twitter, Facebook, Instagram, Reddit, etc.), understanding their data structures, formats (JSON, XML), and APIs for data access.

Data Collection and Preprocessing: Web scraping, using social media APIs, ethical considerations, dealing with messy data (cleaning, normalization).

UNIT-II

Social Network Analysis Techniques: Measures of centrality (degree, betweenness, eigenvector), community detection algorithms, network evolution analysis.

Content Analysis: Text preprocessing (tokenization, stemming, lemmatization), basic sentiment analysis, topic modeling (LDA, NMF), Visualization for Social Media - Network visualization tools (Gephi, NetworkX), geographic visualization, temporal analysis techniques.

UNIT- III

Opinion Mining and Sentiment Analysis: Advanced sentiment classification, lexicon-based and deep learning-based approaches (LSTM, BERT).

Trend Detection and Prediction: Time series analysis, anomaly detection, social signal forecasting, Recommendation Systems: Collaborative filtering, content-based recommendations in the context of social media.

UNIT - IV

Social Media Marketing: Influencer identification, campaign analysis, virality prediction.

Social Movements and Online Behavior: Protest and mobilization analysis, information diffusion, detecting bots and disinformation.

Computational Social Science: Utilizing social media data to study social phenomena, ethical considerations surrounding social media research.

Textbooks:

1. Matthew A. Russell, Mikhail Klassen, "Mining the Social Web" , O'Reilly Media, 3rd Edition, 2019.
2. Nathan Danneman & Richard Heimann ,"Social Media Mining with R" , Packt Publishing, 2014
3. Steven Bird, Ewan Klein, and Edward Loper , "Natural Language Processing with Python" , O'Reilly Media, 2009.

Name of the Course	Deep Reinforcement Learning		
Course Code	**MSDS-EL-413**	Credits-4	L-4, T-0, P-0
Lectures to be Delivered	60 (1 Hr Each) (L=60, T=0 for each semester)		
Semester End Examination	**Max Marks: 75**	**Min Pass Marks: 40%**	**Max. Time: 3 Hrs**
Internal Assessment (based on sessional test (2) 50%, Tutorials/Assignments 30%, Quiz/Seminar 10%, Attendance 10%)			**Max Marks: 25** **Min Marks: 10**

Course Outcomes (COs)	At the end of this course, the student will be able to:
CO1	Explain reinforcement learning principles, model problems as Markov Decision Processes, and understand deep Q-Networks.
CO2	Apply policy gradient methods to optimize policies, train agents in openai Gym, and implement advanced actor-critic algorithms.
CO3	Utilize evolutionary algorithms for reinforcement learning, understand their advantages, and solve problems.
CO4	Design distributional DQN architectures, address sparse rewards with curiosity-driven exploration, and extend RL to multi-agent scenarios.

INSTRUCTIONS

For Paper Setters: The question paper will consist of five sections A, B, C, D & E. Section E will be compulsory, it will consist of a single question with 9 subparts of short answer type, which will cover the entire syllabus, ensuring only 2-3 questions from each units and semester end examination for the course. Section A, B, C & D will have two questions from the respective sections of the syllabus and each question will carry 12 marks out of the total marks of the semester end examination for the course.

For candidates: Candidates are required to attempt five questions in all selecting one question from each of the sections A, B, C & D of the question paper and all the subparts of the questions in Section E. Use of non-programmable calculators is allowed.

MSDS-EL-413 Deep Reinforcement Learning

UNIT-I

Introduction to reinforcement learning: Reinforcement learning, dynamic programming versus Monte Carlo , reinforcement learning framework, why deep reinforcement learning, our didactic tool: String diagrams.

Modeling reinforcement learning problems: Markov decision processes, String diagrams and our teaching methods, solving the multi-arm bandit, applying bandits to optimize ad placements, building networks with PyTorch, solving contextual bandits, the Markov property, predicting future rewards: value and policy functions.

Predicting the best states and actions: DeepQ-networks, navigating with Q-learning, Preventing catastrophic forgetting: Experience replay, improving stability with a target network.

UNIT-II

Learning to pick the best policy: policy gradient methods, policy function using neural networks, reinforcing good actions: the policy gradient algorithm, working with OPENAI Gym, the REINFORCE algorithm, distributed training, advantage actor-critic, N-step actor-critic.

Evolutionary Algorithms: a different approach to reinforcement learning, reinforcement learning with evolution strategies, a genetic algorithm for CartPole, pros and cons of evolutionary algorithms, evolutionary algorithms as a scalable alternative.

UNIT-III

Distributional DON: Probability and statistics revisit, The Bellman equation , Distributional Q-learning, Comparing probability distributions, Dist-DQN on simulated data, Using distributional Q-learning to play Freeway.

Curiosity-driven exploration: Tackling sparse rewards with predictive coding, Inverse dynamics prediction, Setting up Super Mario Bros, Preprocessing and the Q-network, Setting up the Q-network and policy function, Intrinsic curiosity module, Alternative intrinsic reward mechanism.

UNIT-IV

Multi-agent reinforcement learning: From one to many agents, Neighborhood Q-learning, the ID Ising model, Mean field Q-learning and the 2D Ising model, Mixed cooperative-competitive games.

Interpretable reinforcement learning: Attention and relational models, Machine learning interpretability with attention and relational biases, relational reasoning with attention, Implementing self-attention for MNIST, Multi-head attention and relational DQN, Double Q-learning, Training and attention visualization.

Text Book:

1. Alexander Zai, Brandon Brown, Deep Reinforcement Learning in Action, Simon and Schuster, 2020.

Reference Books:

1. Richard S. Sutton and Andrew G. Barto, “Reinforcement Learning: An Introduction" MIT press, 2018.
2. Maxim Lapan , “Deep Reinforcement Learning Hands-On" , Packt Publishing Limited, 2018.

DEPARTMENT OF DATA SCIENCE AND ARTIFICIAL INTELLIGENCE

HIMACHAL PRADESH UNIVERSITY SHIMLA

Syllabus for the Entrance Exam of M.Sc. Data Science and M.Sc. Artificial Intelligence

Sr. No	Contents	Marks
1.	ICT Awareness	60
2.	Mathematics	20
3.	General Logic Ability & Aptitude	20
Total		**100**

The minimum qualifying marks in the Entrance Examination (written test) for subsidized as well as non-subsidized seats will be as per university norms.

Detailed Syllabus

1. ICT Awareness

10 questions worth 1 mark each *will have to be set from each of the following six sections, for a total of* ***60 questions****.*

1. **Fundamentals of Computers**: Introduction to computers, types of computers, components of a computer system (hardware and software), basics of computer architecture, input and output devices, memory hierarchy and types of memory, types of software (system software, application software), binary and other number systems, evolution of computing generations, boot operation and fundamentals of digital logic (AND, OR, NOT gates).
2. **Operating System**: Functions of operating systems, types of operating systems (batch, time-sharing, distributed, real-time), basics of process management (process creation, scheduling algorithms, inter-process communication), memory management techniques (paging and segmentation), file systems and file management, basics of scheduling algorithms, deadlock prevention and handling, types of user interfaces (CLI, GUI), Linux and Windows operating system basics.
3. **Computer Networking and Data Communication**: Basics of computer networks, types of networks (LAN, MAN, WAN), network topologies, OSI and TCP/IP models, data encapsulation and packet switching, protocols (IP, TCP, UDP, HTTP, FTP, SMTP), types of transmission media (coaxial, fiber optic, wireless), network devices (router, switch, hub), error detection and correction techniques, data link and network layer functionalities, network security concepts (encryption, firewalls, VPNs), IPv4 and IPv6 addressing.
4. **Database Management System (DBMS)**: Basics of databases, types of databases (relational, NoSQL, hierarchical, object-oriented), entity-relationship (ER) modeling, relational database design, normalization (1NF, 2NF, 3NF, BCNF), SQL basics (queries, subqueries, joins, aggregation), database indexing, basics of transactions and concurrency control, ACID properties, introduction to stored procedures and triggers.
5. **Data Structures**: Types of data structures (arrays, linked lists, stacks, queues, trees, graphs), algorithm complexity (Big O), searching and sorting algorithms (linear search, binary search, bubble sort, selection sort, merge sort, quicksort), operations on data structures

(insertion, deletion, traversal), basics of recursion ,hashing and hash tables, binary trees, binary search trees.

6. **Software Engineering**: Introduction to software development lifecycle (SDLC), software development models (waterfall, agile, V-model), requirement analysis and specifications, software design concepts (modularity, cohesion, coupling, data flow diagrams), object-oriented analysis and design (OOAD), software testing methods (unit testing, integration testing, system testing, acceptance testing), maintenance and evolution, software project management basics (cost estimation, scheduling, risk analysis), introduction to software quality assurance, software metrics.

2. Mathematics

2 questions worth 1 mark each *will have to be set from each of the following 10 sections, for a total of* ***20 questions***

1. **Number System**: Types of numbers (natural, whole, integers, rational, irrational, prime), divisibility rules, factors and multiples, HCF and LCM, and basic operations.
2. **Fractions and Decimals**: Conversions between fractions and decimals, basic operations, and comparisons.
3. **Percentages**: Concepts of percentages, percentage change, increase and decrease, and applications in problem-solving.
4. **Ratios and Proportion**: Understanding ratio, proportion, direct and inverse variation, and their applications.
5. **Averages**: Calculation of averages, weighted averages, and applications in real-life scenarios.
6. **Simple and Compound Interest**: Basics of interest calculations, differences between simple and compound interest, and applications in finance problems.
7. **Profit, Loss, and Discounts**: Calculating profit and loss, percentage profit/loss, marked price, cost price, and discounts.
8. **Time and Work**: Basics of work and time, problems on work efficiency, combined work, and problems involving pipes and cisterns.
9. **Time, Speed, and Distance**: Concepts of speed, distance, and time, relative speed, and problems on trains, boats, and streams.
10. **Differentiation and Integration**: Basics of differentiation (simple rules, chain rule, product rule), basic integration (antiderivatives), applications in finding tangents, areas under curves, and solving problems related to rates of change.

3. General Logic Ability and Aptitude

2 questions worth 1 mark each *will have to be set from each of the following 10 sections, for a total of* ***20 questions****.*

1. **Analogy**: Identifying relationships between words, numbers, or shapes and selecting the correct analogy.
2. **Classification**: Grouping similar items or figures based on certain characteristics, and identifying the odd one out.

3. **Series Completion**: Finding the next number, letter, or shape in a given series based on a pattern.
4. **Coding-Decoding**: Understanding and interpreting codes, where letters or numbers are substituted for other letters or numbers, and solving related puzzles.
5. **Blood Relations**: Solving problems based on familial relationships, determining relationships between different family members.
6. **Direction Sense**: Problems involving directions (north, south, east, west), distance, and the relative positions of objects or people.
7. **Syllogism**: Understanding statements and conclusions, and solving problems based on logical deductions using Venn diagrams.
8. **Puzzle Test**: Solving problems involving arrangement or sequencing based on given conditions (e.g., seating arrangement, order of events).
9. **Logical Venn Diagrams**: Solving problems based on the relationships between different sets of objects, with the help of Venn diagrams.
10. **Statement and Assumptions**: Interpreting and solving problems involving given statements and assumptions, determining whether the assumption follows from the statement.

SAMPLE QUESTION PAPER

Below is a **sample question paper** with **2 multiple-choice questions from each section** Each question includes **four answer options**.

1. ICT Awareness

Fundamentals of Computers

- Which of the following is an input device?
  A. Printer
  B. Monitor
  C. Keyboard
  D. Speaker
- The primary memory of a computer is:
  A. Hard Disk
  B. RAM
  C. CD-ROM
  D. USB Drive

Operating System

- Which of the following is *not* a function of an operating system?
  A. Managing memory
  B. Creating web pages
  C. Handling files
  D. Process management

- Linux is an example of a:
  A. System software
  B. Application software
  C. Utility software
  D. Firmware

Computer Networking and Data Communication

- Which of the following is a network device used to connect different networks?
  A. Switch
  B. Hub
  C. Router
  D. Modem
- In the OSI model, the function of data encryption is performed at the:
  A. Physical layer
  B. Network layer
  C. Transport layer
  D. Presentation layer

Database Management System (DBMS)

- Which of these is a property of a relational database?
  A. Uses object inheritance
  B. Organized in tables
  C. Uses HTML tags
  D. Stores data in a flat file
- The command used to retrieve data from a database is:
  A. UPDATE
  B. SELECT
  C. DELETE
  D. INSERT

Data Structures

- A linear data structure where elements are inserted at one end and removed from the other is called:
  A. Stack
  B. Array
  C. Queue
  D. Graph
- The time complexity of binary search in a sorted array is:
  A. O(n)
  B. O(log n)
  C. $O(n^2)$
  D. O(1)

Software Engineering

- In which software model is each phase completed before the next one begins?
  A. Agile
  B. Waterfall
  C. Spiral
  D. V-Model
- What does SDLC stand for?
  A. Software Data Learning Cycle
  B. Structured Design Logic Code
  C. Software Development Life Cycle
  D. System Debugging Lifecycle

2. Mathematics

Number System

- Which of the following is a prime number?
  A. 21
  B. 17
  C. 33
  D. 27
- HCF of 18 and 24 is:
  A. 2
  B. 3
  C. 6
  D. 12

Fractions and Decimals

- 0.25 is equal to:
  A. 1/2
  B. 1/4
  C. 3/4
  D. 1/8
- 2/5 + 3/10 equals:
  A. 5/10
  B. 4/5
  C. 7/10
  D. 1

Percentages

- 25% of 160 is:
  A. 40
  B. 60
  C. 20
  D. 80
- A number increases from 80 to 100. The percentage increase is:
  A. 20%
  B. 25%

C. 18%
D. 22%

Ratios and Proportion

- If A : B = 2 : 3, and B : C = 4 : 5, then A : C is:
  A. 8 : 15
  B. 2 : 5
  C. 3 : 10
  D. 4 : 9
- If 5 pens cost Rs. 60, then 8 pens will cost:
  A. Rs. 90
  B. Rs. 96
  C. Rs. 100
  D. Rs. 120

Averages

- The average of 4, 6, and 10 is:
  A. 6
  B. 7
  C. 8
  D. 9
- The average of five numbers is 20. The sum is:
  A. 80
  B. 100
  C. 120
  D. 140

Simple and Compound Interest

- Simple interest on Rs. 1000 at 5% for 2 years is:
  A. Rs. 50
  B. Rs. 100
  C. Rs. 200
  D. Rs. 75
- The compound interest on Rs. 1000 at 10% per annum for 2 years is:
  A. Rs. 200
  B. Rs. 210
  C. Rs. 220
  D. Rs. 100

Profit, Loss, and Discounts

- A person bought an item for Rs. 200 and sold it for Rs. 250. The profit percentage is:
  A. 20%
  B. 25%
  C. 30%
  D. 50%

- If the marked price is Rs. 500 and the discount is 10%, selling price is:
  A. Rs. 450
  B. Rs. 400
  C. Rs. 480
  D. Rs. 470

Time and Work

- If A can do a job in 6 days and B in 12 days, together they can finish it in:
  A. 2 days
  B. 3 days
  C. 4 days
  D. 5 days
- Two taps can fill a tank in 12 and 15 minutes. Together, they take:
  A. 6.5 min
  B. 6 min
  C. 7 min
  D. 7.5 min

Time, Speed, and Distance

- A car travels 60 km in 1.5 hours. Its speed is:
  A. 40 km/h
  B. 50 km/h
  C. 60 km/h
  D. 80 km/h
- A train 100 meters long crosses a pole in 10 seconds. Its speed is:
  A. 10 m/s
  B. 20 m/s
  C. 100 m/s
  D. 15 m/s

Differentiation and Integration

- Derivative of x^2 is:
  A. x
  B. 2x
  C. x^3
  D. $3x^2$
- Integral of 2x dx is:
  A. $x^2 + C$
  B. $x^3 + C$
  C. $2x^2 + C$
  D. x^2

3. General Logic Ability and Aptitude

Analogy

- Book : Reading :: Fork : ?
  A. Drawing
  B. Writing
  C. Stirring
  D. Eating
- Moon : Satellite :: Earth : ?
  A. Star
  B. Planet
  C. Galaxy
  D. Orbit

Classification

- Which one is different from the rest?
  A. Apple
  B. Banana
  C. Carrot
  D. Mango
- Find the odd number:
  A. 2
  B. 3
  C. 5
  D. 9

Series Completion

- 2, 4, 8, 16, ?
  A. 18
  B. 24
  C. 32
  D. 30
- A, C, E, G, ?
  A. H
  B. I
  C. J
  D. K

Coding-Decoding

- If CAT = 24, then DOG = ?
  A. 26
  B. 30
  C. 28
  D. 22

- In a certain code, FLOW is written as GMPX. How is RAIN written?
  A. SBLM
  B. SBJO
  C. QZHM
  D. RAKO

Blood Relations

- If A is B's mother and B is C's father, what is A to C?
  A. Aunt
  B. Sister
  C. Grandmother
  D. Cousin
- If X is the brother of Y and Y is the sister of Z, then how is X related to Z?
  A. Cousin
  B. Brother
  C. Uncle
  D. Cannot be determined

Direction Sense

- A man walks north, then turns right, then right again. In which direction is he now?
  A. South
  B. East
  C. West
  D. North
- If east becomes north, what does north become?
  A. West
  B. East
  C. South
  D. South-East

Syllogism

- All pens are blue. Some blue things are round. Conclusion: Some pens are round. Is it:
  A. True
  B. False
  C. Can't say
  D. Both A and B
- No cats are dogs. All dogs are animals. So, some animals are not cats. Is this:
  A. True
  B. False
  C. Can't say
  D. None

Puzzle Test

- Four friends sit in a row. A is left to B but right to C. Who is in the middle?
  A. A

B. B
C. C
D. Can't be determined

- If Monday is the first day of the month, what day will the 15th be?
  A. Tuesday
  B. Wednesday
  C. Monday
  D. Sunday

Logical Venn Diagrams

- Which group best represents: Men, Fathers, Engineers?
  A. All are separate
  B. Men ⊃ Fathers ⊃ Engineers
  C. Engineers ⊂ Fathers ⊂ Men
  D. Engineers ⊃ Men ⊃ Fathers
- Choose the best Venn diagram: Doctors, Males, Humans.
  A. All overlapping circles
  B. Non-overlapping
  C. One inside another
  D. Two inside one

Statement and Assumptions

- Statement: "Use eco-friendly bags." Assumption:
  A. Plastic bags are harmful
  B. All people use plastic
  C. Bags are free
  D. Eco-friendly bags are expensive
- Statement: "Join the evening yoga class." Assumption:
  A. People are free in the evening
  B. Everyone loves yoga
  C. Yoga is for old people
  D. The class is free