

सांख्यिकी अध्ययनशाला

पं. रविशंकर शुक्ल विश्वविद्यालय रायपुर (छ.ग.)

प्रत्र क्रमांक 635 / सांख्यिकी / 2025

रायपुर दिनांक 16.05.2025

To,
The Dy. Registrar (Academic)
PRSU, Raipur

Subject: Regarding Information about new program M. Sc. Data Science
Reference : No. 932 / Aca./ 2025 date: 13/05/2025

Respected sir,

Please find the following are the information about the new program M. Sc. Data Science:

1. Approved Syllabus : Enclosed
2. Number of Seats :15
3. Fee structure :Tution fee Rs. 15,000 per Semester + Lab fee Rs. 600 per Annum + Library fee Rs 240 Per Annum
4. Self-financed program or other information : Self-financed
5. Details of Data Science for Brochure for session 2025-26 = No
(Enclosed the details for session 2025-26)

Thanking you.

In-charge Head

Minad
16-05-2025

[Signature]
PROFESSOR
School of Studies in Statistics
Pt. Pavishankar Shukla university
RAIPUR (C.G.)

*अभी Hold में रखने है
ऑगो 2025 तक में ऑगो 2025 तक
पर 03/05/25
5602
16/7/25 AT 12.08 PM*

To,
[Signature]
12/05/25

15.	सांख्यिकी अध्ययनशाला	एम.एस.सी. डाटा साइंस	Artificial Intelligence (AI) and Machine Learning (ML)	2 वर्ष	4 सेमेस्टर (2 वर्ष)	15	पाठ्यक्रम शुल्क 15,000/ प्रति सेमेस्टर (लैब शुल्क 600/ PA और विभागीय पुस्तकालय शुल्क 240/ PA)
-----	-------------------------	----------------------------	---	--------	----------------------------	----	--

Sl. No.	Name of the SoS	Name of the Course	Specialization	Duration	Number of Seats	Tuition fee
15.	Statistics	M. Sc. Data Science	Artificial Intelligence (AI) and Machine Learning (ML)	4 Semesters (2 Years)	15	Course Fee 15,000/ per semester (Lab Fee 600/PA and Departmental Library Fee 240/ PA)

STATISTICS

This School was initiated in the year 1977. Research focus is on Statistical Quality Control, Reliability Theory, Sampling Theory and Operations Research. It offers postgraduate degree in Statistics, and Data Science. In addition, it also offers research degree Statistics.

M. N. Mead
In-charge Head
16.05.2023

[Signature]
PROFESSOR
School of Studies in Statistics
Pt. Pavishankar Shukla university
RAIPUR (C.G.)

SCHOOL OF STUDIES IN STATISTICS
Pt. Ravishankar Shukla University, Raipur (C.G.)



CURRICULUM & SYLLABI

M. Sc. Data Science
Semester System

Session: 2025-27

Approved by	Board of Studies	Academic Council
Date	17/04/2025	

Minod *BoD* *Dahale* *dh* *Signature* *Signature* *Signature*

M. Sc. Data Science

Program overview:

In today's data driven world, the demand for skilled data scientist is soaring across academia, business and research sectors. The main objective of this program is to develop the skill of Big data analytics, Data Mining and Machine Learning. Students will get an opportunity to apply for data scientist or big data analyst positions in the IT industry. The Master of Science in Data Science is a four-semester program aimed to produce postgraduates with advanced knowledge and understanding of Data Science using statistical tools and computer techniques.

Program Outcomes (POs):

Upon successful completion of the Master of Science in Data Science program, students will be able to:

PO-1	Knowledge: Demonstrate a deep understanding of mathematics, statistics, and computer science to solve complex data science problems.
PO-2	Critical Thinking and Reasoning: Exhibit advanced critical thinking skills by analysing and evaluating, mathematical arguments and
PO-3	Problem Solving: Apply data science concepts and methods to solve problems in real-world contexts and will communicate these solutions effectively.
PO-4	Advanced Analytical and Computational Skills: execute statistical analyses with professional software.
PO-5	Effective Communication: Communicate complex data science concepts and solutions effectively through written reports, and presentations.
PO-6	Social / Interdisciplinary Interaction: Integrate statistical techniques and data science into Interdisciplinary context, collaborating effectively with professionals from other fields to address complex problems.
PO-7	Self-Directed and life-long Learning: Recognize the importance of ongoing technology development and life-long learning in the rapidly evolving field of data science
PO-8	Effective Citizenship, Leadership and Innovation: Lead and innovate in various statistical and computer techniques context, contributing to advancement in the field of data science.
PO-9	Ethics: Demonstrate ethical and responsible conduct in data privacy and integrity in data handling.
PO-10	Further Education or Employment: Engage for further academic pursuits, including Ph. D. Programs in Data Science or related fields. Get employment in academia, research institutions, industry, government and other sectors.
PO-11	Global Perspective: Recognize the global nature of data science and its impact.

Mined *30/10* *20/10* *21/10* *22/10* *23/10* *24/10* *25/10* *26/10* *27/10* *28/10* *29/10* *30/10* *31/10* *1/11* *2/11* *3/11* *4/11* *5/11* *6/11* *7/11* *8/11* *9/11* *10/11* *11/11* *12/11* *13/11* *14/11* *15/11* *16/11* *17/11* *18/11* *19/11* *20/11* *21/11* *22/11* *23/11* *24/11* *25/11* *26/11* *27/11* *28/11* *29/11* *30/11* *1/12* *2/12* *3/12* *4/12* *5/12* *6/12* *7/12* *8/12* *9/12* *10/12* *11/12* *12/12* *13/12* *14/12* *15/12* *16/12* *17/12* *18/12* *19/12* *20/12* *21/12* *22/12* *23/12* *24/12* *25/12* *26/12* *27/12* *28/12* *29/12* *30/12* *31/12* *1/1* *2/1* *3/1* *4/1* *5/1* *6/1* *7/1* *8/1* *9/1* *10/1* *11/1* *12/1* *13/1* *14/1* *15/1* *16/1* *17/1* *18/1* *19/1* *20/1* *21/1* *22/1* *23/1* *24/1* *25/1* *26/1* *27/1* *28/1* *29/1* *30/1* *31/1* *1/2* *2/2* *3/2* *4/2* *5/2* *6/2* *7/2* *8/2* *9/2* *10/2* *11/2* *12/2* *13/2* *14/2* *15/2* *16/2* *17/2* *18/2* *19/2* *20/2* *21/2* *22/2* *23/2* *24/2* *25/2* *26/2* *27/2* *28/2* *29/2* *30/2* *31/2* *1/3* *2/3* *3/3* *4/3* *5/3* *6/3* *7/3* *8/3* *9/3* *10/3* *11/3* *12/3* *13/3* *14/3* *15/3* *16/3* *17/3* *18/3* *19/3* *20/3* *21/3* *22/3* *23/3* *24/3* *25/3* *26/3* *27/3* *28/3* *29/3* *30/3* *31/3* *1/4* *2/4* *3/4* *4/4* *5/4* *6/4* *7/4* *8/4* *9/4* *10/4* *11/4* *12/4* *13/4* *14/4* *15/4* *16/4* *17/4* *18/4* *19/4* *20/4* *21/4* *22/4* *23/4* *24/4* *25/4* *26/4* *27/4* *28/4* *29/4* *30/4* *31/4* *1/5* *2/5* *3/5* *4/5* *5/5* *6/5* *7/5* *8/5* *9/5* *10/5* *11/5* *12/5* *13/5* *14/5* *15/5* *16/5* *17/5* *18/5* *19/5* *20/5* *21/5* *22/5* *23/5* *24/5* *25/5* *26/5* *27/5* *28/5* *29/5* *30/5* *31/5* *1/6* *2/6* *3/6* *4/6* *5/6* *6/6* *7/6* *8/6* *9/6* *10/6* *11/6* *12/6* *13/6* *14/6* *15/6* *16/6* *17/6* *18/6* *19/6* *20/6* *21/6* *22/6* *23/6* *24/6* *25/6* *26/6* *27/6* *28/6* *29/6* *30/6* *31/6* *1/7* *2/7* *3/7* *4/7* *5/7* *6/7* *7/7* *8/7* *9/7* *10/7* *11/7* *12/7* *13/7* *14/7* *15/7* *16/7* *17/7* *18/7* *19/7* *20/7* *21/7* *22/7* *23/7* *24/7* *25/7* *26/7* *27/7* *28/7* *29/7* *30/7* *31/7* *1/8* *2/8* *3/8* *4/8* *5/8* *6/8* *7/8* *8/8* *9/8* *10/8* *11/8* *12/8* *13/8* *14/8* *15/8* *16/8* *17/8* *18/8* *19/8* *20/8* *21/8* *22/8* *23/8* *24/8* *25/8* *26/8* *27/8* *28/8* *29/8* *30/8* *31/8* *1/9* *2/9* *3/9* *4/9* *5/9* *6/9* *7/9* *8/9* *9/9* *10/9* *11/9* *12/9* *13/9* *14/9* *15/9* *16/9* *17/9* *18/9* *19/9* *20/9* *21/9* *22/9* *23/9* *24/9* *25/9* *26/9* *27/9* *28/9* *29/9* *30/9* *31/9* *1/10* *2/10* *3/10* *4/10* *5/10* *6/10* *7/10* *8/10* *9/10* *10/10* *11/10* *12/10* *13/10* *14/10* *15/10* *16/10* *17/10* *18/10* *19/10* *20/10* *21/10* *22/10* *23/10* *24/10* *25/10* *26/10* *27/10* *28/10* *29/10* *30/10* *31/10* *1/11* *2/11* *3/11* *4/11* *5/11* *6/11* *7/11* *8/11* *9/11* *10/11* *11/11* *12/11* *13/11* *14/11* *15/11* *16/11* *17/11* *18/11* *19/11* *20/11* *21/11* *22/11* *23/11* *24/11* *25/11* *26/11* *27/11* *28/11* *29/11* *30/11* *31/11* *1/12* *2/12* *3/12* *4/12* *5/12* *6/12* *7/12* *8/12* *9/12* *10/12* *11/12* *12/12* *13/12* *14/12* *15/12* *16/12* *17/12* *18/12* *19/12* *20/12* *21/12* *22/12* *23/12* *24/12* *25/12* *26/12* *27/12* *28/12* *29/12* *30/12* *31/12* *1/13* *2/13* *3/13* *4/13* *5/13* *6/13* *7/13* *8/13* *9/13* *10/13* *11/13* *12/13* *13/13* *14/13* *15/13* *16/13* *17/13* *18/13* *19/13* *20/13* *21/13* *22/13* *23/13* *24/13* *25/13* *26/13* *27/13* *28/13* *29/13* *30/13* *31/13* *1/14* *2/14* *3/14* *4/14* *5/14* *6/14* *7/14* *8/14* *9/14* *10/14* *11/14* *12/14* *13/14* *14/14* *15/14* *16/14* *17/14* *18/14* *19/14* *20/14* *21/14* *22/14* *23/14* *24/14* *25/14* *26/14* *27/14* *28/14* *29/14* *30/14* *31/14* *1/15* *2/15* *3/15* *4/15* *5/15* *6/15* *7/15* *8/15* *9/15* *10/15* *11/15* *12/15* *13/15* *14/15* *15/15* *16/15* *17/15* *18/15* *19/15* *20/15* *21/15* *22/15* *23/15* *24/15* *25/15* *26/15* *27/15* *28/15* *29/15* *30/15* *31/15* *1/16* *2/16* *3/16* *4/16* *5/16* *6/16* *7/16* *8/16* *9/16* *10/16* *11/16* *12/16* *13/16* *14/16* *15/16* *16/16* *17/16* *18/16* *19/16* *20/16* *21/16* *22/16* *23/16* *24/16* *25/16* *26/16* *27/16* *28/16* *29/16* *30/16* *31/16* *1/17* *2/17* *3/17* *4/17* *5/17* *6/17* *7/17* *8/17* *9/17* *10/17* *11/17* *12/17* *13/17* *14/17* *15/17* *16/17* *17/17* *18/17* *19/17* *20/17* *21/17* *22/17* *23/17* *24/17* *25/17* *26/17* *27/17* *28/17* *29/17* *30/17* *31/17* *1/18* *2/18* *3/18* *4/18* *5/18* *6/18* *7/18* *8/18* *9/18* *10/18* *11/18* *12/18* *13/18* *14/18* *15/18* *16/18* *17/18* *18/18* *19/18* *20/18* *21/18* *22/18* *23/18* *24/18* *25/18* *26/18* *27/18* *28/18* *29/18* *30/18* *31/18* *1/19* *2/19* *3/19* *4/19* *5/19* *6/19* *7/19* *8/19* *9/19* *10/19* *11/19* *12/19* *13/19* *14/19* *15/19* *16/19* *17/19* *18/19* *19/19* *20/19* *21/19* *22/19* *23/19* *24/19* *25/19* *26/19* *27/19* *28/19* *29/19* *30/19* *31/19* *1/20* *2/20* *3/20* *4/20* *5/20* *6/20* *7/20* *8/20* *9/20* *10/20* *11/20* *12/20* *13/20* *14/20* *15/20* *16/20* *17/20* *18/20* *19/20* *20/20* *21/20* *22/20* *23/20* *24/20* *25/20* *26/20* *27/20* *28/20* *29/20* *30/20* *31/20* *1/21* *2/21* *3/21* *4/21* *5/21* *6/21* *7/21* *8/21* *9/21* *10/21* *11/21* *12/21* *13/21* *14/21* *15/21* *16/21* *17/21* *18/21* *19/21* *20/21* *21/21* *22/21* *23/21* *24/21* *25/21* *26/21* *27/21* *28/21* *29/21* *30/21* *31/21* *1/22* *2/22* *3/22* *4/22* *5/22* *6/22* *7/22* *8/22* *9/22* *10/22* *11/22* *12/22* *13/22* *14/22* *15/22* *16/22* *17/22* *18/22* *19/22* *20/22* *21/22* *22/22* *23/22* *24/22* *25/22* *26/22* *27/22* *28/22* *29/22* *30/22* *31/22* *1/23* *2/23* *3/23* *4/23* *5/23* *6/23* *7/23* *8/23* *9/23* *10/23* *11/23* *12/23* *13/23* *14/23* *15/23* *16/23* *17/23* *18/23* *19/23* *20/23* *21/23* *22/23* *23/23* *24/23* *25/23* *26/23* *27/23* *28/23* *29/23* *30/23* *31/23* *1/24* *2/24* *3/24* *4/24* *5/24* *6/24* *7/24* *8/24* *9/24* *10/24* *11/24* *12/24* *13/24* *14/24* *15/24* *16/24* *17/24* *18/24* *19/24* *20/24* *21/24* *22/24* *23/24* *24/24* *25/24* *26/24* *27/24* *28/24* *29/24* *30/24* *31/24* *1/25* *2/25* *3/25* *4/25* *5/25* *6/25* *7/25* *8/25* *9/25* *10/25* *11/25* *12/25* *13/25* *14/25* *15/25* *16/25* *17/25* *18/25* *19/25* *20/25* *21/25* *22/25* *23/25* *24/25* *25/25* *26/25* *27/25* *28/25* *29/25* *30/25* *31/25* *1/26* *2/26* *3/26* *4/26* *5/26* *6/26* *7/26* *8/26* *9/26* *10/26* *11/26* *12/26* *13/26* *14/26* *15/26* *16/26* *17/26* *18/26* *19/26* *20/26* *21/26* *22/26* *23/26* *24/26* *25/26* *26/26* *27/26* *28/26* *29/26* *30/26* *31/26* *1/27* *2/27* *3/27* *4/27* *5/27* *6/27* *7/27* *8/27* *9/27* *10/27* *11/27* *12/27* *13/27* *14/27* *15/27* *16/27* *17/27* *18/27* *19/27* *20/27* *21/27* *22/27* *23/27* *24/27* *25/27* *26/27* *27/27* *28/27* *29/27* *30/27* *31/27* *1/28* *2/28* *3/28* *4/28* *5/28* *6/28* *7/28* *8/28* *9/28* *10/28* *11/28* *12/28* *13/28* *14/28* *15/28* *16/28* *17/28* *18/28* *19/28* *20/28* *21/28* *22/28* *23/28* *24/28* *25/28* *26/28* *27/28* *28/28* *29/28* *30/28* *31/28* *1/29* *2/29* *3/29* *4/29* *5/29* *6/29* *7/29* *8/29* *9/29* *10/29* *11/29* *12/29* *13/29* *14/29* *15/29* *16/29* *17/29* *18/29* *19/29* *20/29* *21/29* *22/29* *23/29* *24/29* *25/29* *26/29* *27/29* *28/29* *29/29* *30/29* *31/29* *1/30* *2/30* *3/30* *4/30* *5/30* *6/30* *7/30* *8/30* *9/30* *10/30* *11/30* *12/30* *13/30* *14/30* *15/30* *16/30* *17/30* *18/30* *19/30* *20/30* *21/30* *22/30* *23/30* *24/30* *25/30* *26/30* *27/30* *28/30* *29/30* *30/30* *31/30* *1/31* *2/31* *3/31* *4/31* *5/31* *6/31* *7/31* *8/31* *9/31* *10/31* *11/31* *12/31* *13/31* *14/31* *15/31* *16/31* *17/31* *18/31* *19/31* *20/31* *21/31* *22/31* *23/31* *24/31* *25/31* *26/31* *27/31* *28/31* *29/31* *30/31* *31/31* *1/32* *2/32* *3/32* *4/32* *5/32* *6/32* *7/32* *8/32* *9/32* *10/32* *11/32* *12/32* *13/32* *14/32* *15/32* *16/32* *17/32* *18/32* *19/32* *20/32* *21/32* *22/32* *23/32* *24/32* *25/32* *26/32* *27/32* *28/32* *29/32* *30/32* *31/32* *1/33* *2/33* *3/33* *4/33* *5/33* *6/33* *7/33* *8/33* *9/33* *10/33* *11/33* *12/33* *13/33* *14/33* *15/33* *16/33* *17/33* *18/33* *19/33* *20/33* *21/33* *22/33* *23/33* *24/33* *25/33* *26/33* *27/33* *28/33* *29/33* *30/33* *31/33* *1/34* *2/34* *3/34* *4/34* *5/34* *6/34</*

At the end of the program, the student will be able to:

PSO1	Understand the fundamental principles, concepts and methods in critical areas of data science and multidisciplinary.
PSO2	Apply statistical and computational methods for data analysis, predictive modeling, and decision-making in various industries.
PSO3	Pursue Research in Challenging areas of data science
PSO4	Develop and implement Artificial intelligences/ machine learning models, and data mining techniques to address complex data-driven problem.
PSO5	Integrate data science tools, cloud computing to design scalable and efficient data pipelines and systems for real-world applications.

M.Sc. (Data Science) Semester Course: 2025-27

Mirad BNP BNP

2

Langle

inf + polio

Dr.

**M.Sc. (Data science)
PROGRAMME STRUCTURE**

Semester	Course Nature	Course Code	Course Title	L	T	P	Hrs/Week	Credits	Marks		
									CIA	ESE	Total
Semester-I	Core	DS110	Principles of Data Science	5	1		6	5	30	70	100
	Core	DS120	Exploratory Data Analysis and Data visualization	5	1		6	5	30	70	100
	Core	DS130	Statistical Method for Data Science	5	1		6	5	30	70	100
	Core	DS140	Programming for Data Science in R	5	1		6	5	30	70	100
	Core	DS150	Lab Course I: Practical based on DS120			6	6	3	30	70	100
	Core	DS160	Lab Course II: Practical based on DS130 & DS140			6	6	3	30	70	100
								26			600
Semester-II	Core	DS210	Multivariate Analysis	5	1		6	5	30	70	100
	Core	DS220	Python Programming for Data Science	5	1		6	5	30	70	100
	Core	DS230	Data Base Management System	5	1		6	5	30	70	100
	Core	DS240	Data Structure and Algorithms	5	1		6	5	30	70	100
	Core	DS250	Lab Course I: Practical based on DS210 and DS220			6	6	3	30	70	100
	Core	DS260	Lab Course II: Practical based on DS230 and DS240			6	6	3	30	70	100
								26			600
Semester-III	Core	DS310	Optimization Techniques for Data Science	5	1		6	5	30	70	100
	Core	DS320	Artificial Intelligence and Machine Learning	5	1		6	5	30	70	100
	EL- I (Select any one)	DS331	High Performance Computing	5	1		6	5	30	70	100
		DS332	Deep Learning								
		DS333	Data Science for Health Care								
	EL -II (Select any one)	DS341	Internet of Things	5	1		6	5	30	70	100
		DS342	Generative AI								
		DS343	Natural Language Processing								
	Core	DS350	Lab Course I: Practical based on DS310			6	6	3	30	70	100
	Core	DS360	Lab Course I: Practical based on DS320			6	6	3	30	70	100
	Core	DS370	Minor Project*				*	2	30	100	100
								28			700
Semester IV	Core	DS410	Major Project Work **/ Research Project Dissertation			5	35	20	200 Sessional Marks	200 Viva voce	400
			Total					100			2300

Mined 3020 Buh 21

 3

 Sub

 John

* Minor project as a internship to be done after completion of II semester exam for duration of 4 to 5 weeks and Evaluation is done in III sem.

**

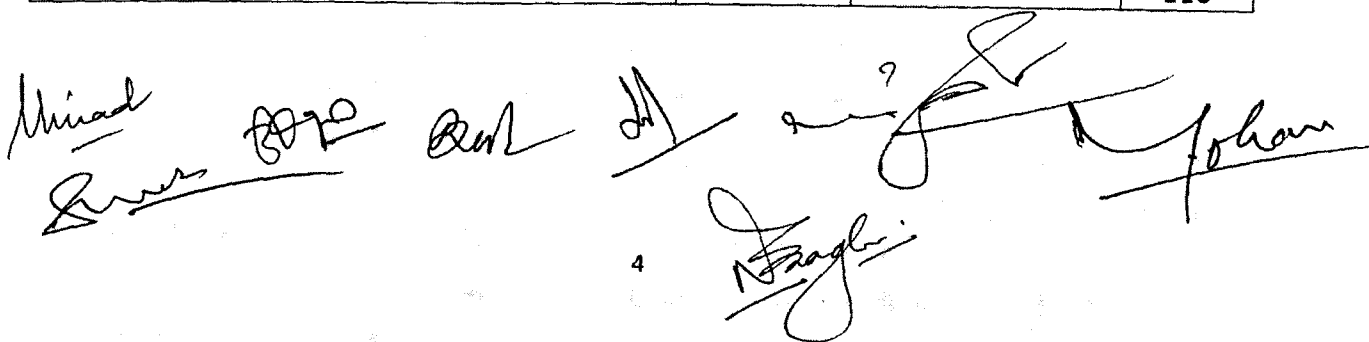
- i. The work done by the students should be enough to justify the duration of the project as 5 to 6 months.
- ii. The certificate Company/Institute must specify the duration of at least five months.
- iii. Students having undergoing project will have to send the confirmation letter from the company/institute within 15 days of joining. This letter will have to consist of the information regarding Company/Institute name, Guide Name, Project Title, Project Starting Date etc.
- iv. Students undergoing major Project should send a progress Report after completion of two months to the department.
- v. The students will have to submit two copies of Major Project Reports after completion of Project work.
- vi. Preferably, independent work should be carried out by each student.
- vii. Major project may be a research Project also.
- viii. Participating in workshops, conferences and seminar or publishing research papers will be given weight age in the research project
- ix. Student should register for any one MOOC course from SWAYAM/NPTEL/RSU LMS under the guidance of a mentor and a certificate of completion must be submitted to the mentor.

Value Added Courses: (offered to the PG students of Data Science):

Sem.	Course Code	Course Title	Course type (T/P)	Hrs/Week	Credits	Marks		
						CIA	ESE	Total
I	DS170	Indian Knowledge System	T	2	2	30	70	100
III	DS380	Intellectual Property Right	T	2	2	30	70	100

M. Sc. DATA SCIENCE program

Specification of Course	Semester	No. of courses	Credits
Core	I-IV	18	90
Elective	III	02	10
# Additional Courses (Qualifying in nature, for Student admitted in School of Studies only)			
Value Added Courses	I and III	02	04
Generic Elective	II-III	02	06
Total			110



 4

SEMESTER I

Course Name : Principles of Data Science

Course Code : DS110

Semester : First

Course Objective: To introduce core concepts in Data Science, to provide strong foundation in application related to Data Science, to understand various techniques used in Data Science

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the foundational concepts of Data Science and Nature of data.
2. Establish the relationship between data dependencies using Statistics.
3. Explain concepts of Artificial Intelligence Roles and skills in Data Science.
4. Demonstrate Data Types, Data Classification.
5. Implement Ethics and Data Science.

UNIT I

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

UNIT II

Frequency distributions: Outliers, relative frequency distributions, cumulative frequency distributions, frequency distributions for nominal data, interpreting distributions, Graphs for quantitative data and qualitative data, typical shapes. Describing data with averages, mode, median, mean, averages for qualitative and ranked data. Describing variability range, variance, standard deviation, degrees of freedom, and interquartile range. Variability for qualitative and ranked data. Describing relationships: correlation and regression.

UNIT III

AI: Cognitive Computing: Learning Perceptions, Terminologies, Machine Learning, Neural Networks, Deep Learning, NLP, Speech Processing, Big Data and AI, Ethics in AI Research, Advanced Application, AI Myth, Data Science Roles Data Scientist, Data Architect, Data Analyst Machine Learning Engineer Skills.

UNIT IV

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

Minad *Boye* *But* *2/2* *and* *John*
Brage

Ethics and Data Science- Doing Good Data Science, Data Ownership, The Five Cs, Implementing the Five Cs, Ethics and Security Training, Developing Guiding Principles, Building Ethics into a Data-Driven Culture, Regulation, Building Our Future, Case Study.

1. David Cielen, Arno D. B. Meysman, and Mohamed Ali (2016): *Introducing Data Science*, Manning Publications.
2. Robert S. Witte and John S. Witte (2017): *Statistics*, 11th edition, Wiley Publications.
3. Joel Grus (2019): *Data Science from Scratch*, 2nd edition, O'Reilly Publisher

1. Francesco Corea (2019): An Introduction to Data. SpringerLillian Pierson, Jake Porway, (2017): Data Science for Dummies, 2nd Edition, John Wiley & Sons, (EBook)
2. SinanOzdemir, and Sunil Kakade (2016): Principles of Data Science, 2nd Edition (EBook)
3. Lillian Pierson and Jake Porway (2017): Data Science for Dummies, 2nd Edition, John Wiley & Sons, (EBook)

Course Objective : The objective of this course is to provide students with a deep understanding of Exploratory Data Analysis (EDA) techniques used to summarize, visualize, and uncover patterns in data. The course emphasizes the use of statistical methods, data visualization tools, to explore data sets, detect anomalies, test assumptions, and generate hypotheses.

1. Recognize Exploratory Data Analysis
2. Explain EDA Assumptions.
3. Derive meaningful insights, and prepare data effectively for further modeling
4. Analysis uniform random numbers and Random walk
5. Generate report using SPSS that for decision making

Introduction: EDA Introduction, What is EDA? EDA vs. Classical & Bayesian, EDA vs. Summary, EDA Goals, The Role of Graphics, An EDA/Graphics Example, General Problem Categories.

Minad Bozo Bur da sin solari
6 Eagle Erur

UNIT II

EDA Assumptions: Underlying Assumptions, Importance, Techniques for Testing Assumptions, Interpretation of 4-Plot, Consequences.

UNIT III

EDA Techniques: Introduction, Analysis Questions, Graphical Techniques: Alphabetical, Graphical Techniques: By Problem Category, Quantitative Techniques, Probability Distributions.

UNIT IV

EDA Case Studies : Case Studies Introduction, Case Studies : Normal random numbers, Uniform random numbers, Random walk, Josephson Junction Cryothermometry, Beam Deflections, Filter Transmittance, Standard Resistor, Heat Flow Meter 1, Airplane Glass Failure Time, Ceramic Strength.

UNIT V

Introduction to different data types, scale of measurements ,classification techniques, basics of SPSS, data entry in SPSS, missing values, multi responses, Data transformation through SPSS; Selection of cases , compute variables, Frequency tables and different types of charts, measures of central tendency, measures of dispersions, distribution of data set, Correlation and Regression using SPSS.

Text Book:

1. Hoang Pham (2023): Handbook of Engineering Statistics, 1st edition, Springer.

Reference Books:

1. Roger D. Peng (2016): Exploratory Data Analysis with R,
2. Scott Murray (2022): Interactive Data Visualization for the Web, 2nd edition, O'Reilly.
3. Jen Stirrup and Ruben oliva Ramos (2017): Advanced Analytics with R and Tableau, Packt Publishing.

Minad *Boze* *Quik* *da* *if* *polian*
Naaghi *Ernes*

Course Name : Statistical Methods for Data Science

Course Code : DS130

Semester : First

Course Objective: The objective of this course is to equip students with a solid foundation in statistical methods essential for data science.

Course Outcome : On successful completion of the course, the student will be able to:

1. Apply descriptive and inferential statistics to analyze data.
2. Explain Linear Regression and its application.
3. Investigate the various Standard Discrete Distributions.
4. Solve real life problem using Normal distribution.
5. Emphasis on understanding hypothesis testing to solve business problem.

UNIT-I

Diagrammatic and representation of data, Frequency distribution, measures of location, dispersion, Skewness, Kurtosis, Simple correlation coefficient, Multiple and Partial Correlation.

UNIT-II

Linear Regression, and their application, Intra class correlation, Correlation ratio. Definition of probability, Additive and multiplicative theorems of probability, Bayes' theorem, Basic distribution function probability mass function, probability density function, joint, marginal and conditional p.m.f.. Random Variables and its mathematical expectations, conditional Expectation.

UNIT-III

Standard Discrete Distributions- Bernoulli, Binomial, Poisson, Geometric. Limiting form of Binomial and Poisson distributions.

UNIT-IV

Standard continuous distributions-Uniform, Exponential, Normal distribution and its properties, standard normal distribution.

UNIT-V

Statistical hypotheses, Type I and II errors, level of significance, degree of freedom, test of significance, concept of p-value. Tests of significance for the parameters of normal distribution (one sample and two sample problems). Chi-square test of goodness of fit and independence of attributes, Homogeneity of populations, sampling distribution of correlation coefficient.

Minad BOP BMR DA [Signature] [Signature] [Signature]
[Signature] [Signature]

Text Books:

1. Dudewicz, E. J. and Mishra, S.N. (1988) : Modern Mathematical Statistics, Wiley, Int'l Student's Edition.
2. Rohatgi, V. K. (1984): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.

Reference Books:

1. Rao C. R. (1973): Linear Statistical Inference and its Applications, 2nd edition, Wiley Eastern.
2. Weather, Burn, C. E. (1961): A first Course in Mathematical Statistics. The English Language Book Society And Cambridge University Press
3. Keany, J. F. (2013) : Mathematics of Statistics Part 1-2, Literary Licensing LLC.
4. Stuart (1994): Kendalls Advanced Theory of Statistics V 1 6e: Distribution Theory, John Wiley & Sons Inc.
5. Mood, A., Gybrill F. and Boes, D. (2017): Introduction to the theory of Statistics, 3rd edition, McGraw Hill Education.
6. Dasgupta B. (2013) : Fundamentals of statistics Vol.I. , World Press Private Ltd.

Course Name : Programming for Data Science in R

Course Code : DS210

Semester : Second

Course Objective: The objective of this course is to equip students with the programming skills necessary to perform data science tasks using R. The course covers the basics of R programming, data manipulation with packages like dplyr and tidyr, data visualization with ggplot2, and statistical analysis techniques. Students will learn how to clean, explore, analyze, and visualize data efficiently using R, and apply these techniques to real-world datasets.

Course Outcome : On successful completion of the course, the student will be able to:

1. Understand the overview of R and to read and write the data.
2. List report with data visualization using R.
3. Solve data science problems and build statistical models for data-driven decision making.
4. Apply Non-parametric tests for various problems.
5. Develop ANOVA (analysis of variance) and to make appropriate Interpretation.

Minad / Bore / Bore / Bore /
 Bore / Bore / Bore / Bore /
 Bore / Bore / Bore / Bore /

[Handwritten signature]

UNIT I

Overview of R , R data types and objects (vector, matrix, data frame, list, array, factor, time series), reading and writing data (both from console and external files) and different types of indexes in R.

UNIT II

Calling external programs in R and linking to data bases. Data visualization using R (both two and three dimensions). Statistical Computing and mathematical computing based on descriptive statistics using R.

UNIT III

Simple correlation, multiple correlation and partial correlation and their significance tests, Spearman rank correlation, Linear regression analysis and multiple regression analysis with assumptions.

UNIT IV

Concept of simple hypothesis test and related error, Tests of Significance Tests of significance for Z, t, F and chi-square statistics for one and two samples. Non-parametric tests- Sign test, Wilcoxon signed rank test, Runs test for randomness, Median test and Mann-Whitney U-test. Kruskal – Wallis test.

UNIT V

Multivariate data representation, one way analysis of variance, two way analysis of variance, numerical integration, root extraction, matrix computations using R, etc.

Text Books:

1. R. A. Thisted (1988): Elements of Statistical Computing. Champman and Hall.
2. Hadley Wickham, Mine Çetinkaya-Rundel, Garrett Golemud: R for Data Science, 2nd, O'Reilly.

Reference Books:

1. Rajaraman, V. (2019): Computer Oriented Numerical Methods, 4th edition, PHI Learning,
2. Sandeep Rakshit (2017): R Programming for beginners, 1st edition, McGraw Hill Education.
3. Michael J. Crawley (2017): The R Book, 2nd edition, Willey.
4. John Chambers (2009): Software for data analysis: Programming with R, 2nd edition, Springer-Verlag New York Inc.
5. Phil Spector (2008) : Data Manipulation with r (use R!), Springer-verlag New York inc,

Mirad *Baro* *Bar* *dl* *sig* *folow*
Shur *Langhi*

SEMESTER II

Course Name : Multivariate Analysis

Course Code : DS210

Semester : First

Course Objective: The objective of this course is to provide students with a comprehensive understanding of multivariate statistical techniques useful to analyze data theoretically as well as using Statistical Software.

Course Outcome: On successful completion of the course, the student will be able to:

1. Apply multivariate methods to real-world datasets,
2. Make decision using Normal distribution.
3. Draw meaningful conclusions for research and decision-making using Fisher-Behran statistic.
4. Interpret complex relationships among variables using correlation.
5. Learn Canonical variables and canonical correlation

UNIT I

Gauss- Markov set-up, Estimability condition, best point estimates/interval estimates of estimable linear parametric functions, Normal equations and Least squares estimates, Gauss-Markov Theorem, Introduction to fixed, mixed and random effects linear models. Analysis of variance for one way and two way classified data with equal number of observations per cells, Analysis of covariance model.

UNIT II

Multivariate Normal Distribution and its properties, Reproductive property, transformation by a vector, singular /non-singular matrix, Maximum likelihood estimators of parameters. Distribution of sample mean vector. Wishart matrix - its distribution and properties, Characteristic function of Wishart distribution, chi-square distribution as a particular case of Wishart distribution.

UNIT III

Distribution of Hotelling's T^2 statistic. Application in tests on mean vector for one and more multivariate normal populations and also on equality of the components of a mean vector in a multivariate normal population, Fisher-Behran statistic, Mahalanobis D^2 Statistic. Properties of Hotelling T^2 distribution.

UNIT IV

Applications and uses of simple correlation coefficient, partial and multiple correlation coefficient. Distribution of sample regression coefficients, Multivariate linear regression model-estimation of parameters, tests of linear hypotheses about regression coefficients. Logistic regression analysis, estimating logistic regression model, testing of significance of the coefficients.

11

UNIT V

Classification and discrimination procedures for discrimination between two multivariate normal populations - sample Discriminant function, probabilities of misclassification and their estimation, classification into more than two multivariate normal populations. Principal components, Dimension reduction, Canonical variables and canonical correlation - definition, use, estimation and computation.

Text Books:

1. Cook, R. D. and Weisberg, S. (1982): Residual and Influence in Regression. Chapman and Hall.
2. Draper, N. R. and Smith, H. (1998): Applied Regression Analysis. 3rd edition, Wiley.

Reference Books:

1. Gunst, R. F. and Mason, R. L. (1980): Regression Analysis and its Applications – A Data Oriented Approach. Marcel and Dekker.
2. Rao, C. R. (1973): Linear Statistical Inference and Its Applications. Wiley Eastern.
3. Weisberg, S. (1985): Applied Linear Regression. Wiley.
4. Anderson, T. W. (1983): An Introduction to multivariate statistical analysis. 2nd Ed.
5. Wiley. Giri, N. C. (1977): Multivariate Statistical inference. Academic Press.
6. Kshirsagar, A. M. (1972): Multivariate Analysis. Marcel Dekker.
7. Morrison, D. F. (1976): Multivariate statistical methods. 2nd Ed. McGraw Hill.
8. Muirhead, R. J. (1982): Aspects of multivariate statistical theory, J. Wiley.
9. Seber, G. A. F. (1984): Multivariate observations. Wiley.
10. Sharma, S. (1996): Applied multivariate techniques. Wiley.
11. Srivastava, M. S. and Khatri, C. G. (1979): An introduction to multivariate statistics. North Holland.
12. Johnson, R. and Wychern (1992): Applied multivariate Statistical analysis, 3rd Ed. Prentice Hall.

Course Name : Python Programming For Data Science

Course Code : DS220

Semester : Second

Course Objective: The objective of this course is to equip students with practical programming skills in Python tailored for data science applications. The course covers fundamental Python concepts, data structures, libraries such as NumPy, Pandas, Matplotlib, and tools for data manipulation, analysis, and visualization.

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the basics of python and structure of Python Program.
2. Write efficient code, work with real-world datasets, and implement data science workflows using Python.

(Handwritten signatures and marks)

12

(Signature: Nagar)

3. Apply Python libraries for Data Science
4. Develop Python programs to clean, analyze, and visualize data,
5. Support data-driven decision-making processes using wrangling data and stretching python's capabilities.

UNIT I

Basics of python, Structure of Python Program, Problem Solving Using Branches and Loops, Functions, Problem Solving using functions, sets, lists, tuples and dictionary.

UNIT II

Introduction to Data Understanding and Preprocessing Knowledge domains of Data Analysis, understanding structured and unstructured data, Data Analysis process, Dataset generation, Importing Dataset: Importing and Exporting Data, Basic Insights from Datasets, Cleaning and Preparing the Data: Identify and Handle Missing Values.

UNIT III

Python libraries for Data Science, Scientific libraries, Basics of Numpy and Scipy Computation on NumPy and Scipy, Basic statistical tools: Pandas- Basics of Pandas, Data cleaning and analysis, Matplotlib and Seaborn, Scikitlearn build machine learning models.

UNIT IV

Model development using machine learning algorithms, Supervised and Unsupervised Learning, Visualizing the data, Choosing the right graph, Creating advanced scatterplots, Visualizing graphs, Understanding the tools: Bar graph, Pie chart, Box plot, Histogram, Line charts.

UNIT V

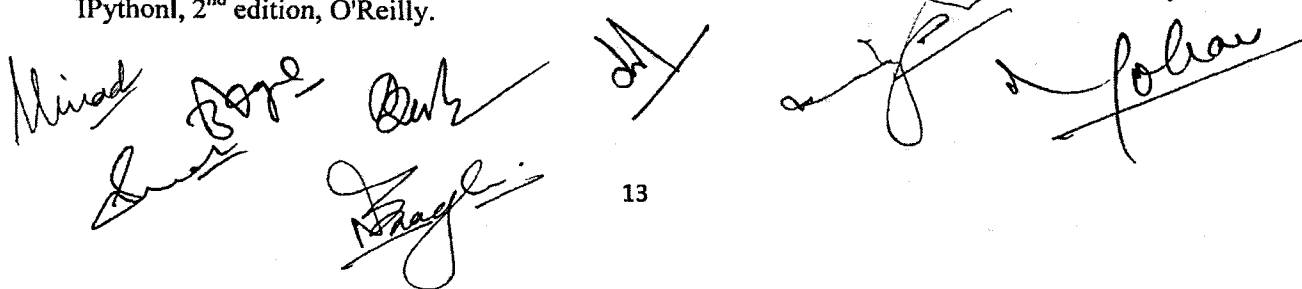
Wrangling data, stretching python's capabilities, Exploring data analysis, Clustering, Preforming Cross, Validation, Selection and Optimization.

Text Books:

1. Wes McKinney (2018) :Python for Data Analysis. 2nd edition, O'Reilly.
2. Reema Thareja (2017): Python Programming: using Problem Solving approach, 1st edition, Oxford University press.
3. Mark Lutz (2013) : learning python, 5th edition, O'Reilly.

Reference Books:

1. David Ascher and Mark Lutz (2000): Learning Python, O'Reilly.
2. Allen Downey, Jeffrey Elkner, and Chris Meyers (2015): Learning with Python, 1st edition, Dreamtech Press.
3. David Taieb (2018): Data Analysis with Python: A Modern Approach“, 1st edition, Packt Publishing.
4. Wes McKinney (2017): Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd edition, O'Reilly.



UNIT – V

Distributed Database Concepts: Distributed databases, structure of distributed databases, trade-offs in distributing the database, transparency and autonomy, distributed query processing, Recovery and Commit Protocols: Recovery in distributed systems, commit protocols, , security and integrity violations. authorization and views , security specification , encryption , Statistical databases

Text Books:

1. Henry F. Korth, and Abraham Silberschatz, Sudarshan(2011): Database system Concepts, 6th edition, McGraw Hill.
2. Raghu Ramakrishnan and Johannes (2003): Database Management Systems, 3rd edition, GehrkeMcGraw Hill.

Reference Books:

1. Bipin C. Desai(1991): An Introduction to data base systems, Galgotia Publications Private Limited,
2. C.J. Date (2004): An Introduction to Database Systems, 8th Edition, Pearson.
3. Rajeev Chopra (2008): Fundamentals of Database Management Systems, S Chand.
4. Ramez E. Shamkant B. Navathe (2016): Fundamentals of Database Systems, 7th Edition, Pearson,

Course Name : Data Structure and Algorithms

Course Code : DS 240

Semester : Second

Course Objective: The objective of this course is to introduce students to fundamental data structures and algorithms essential for efficient problem-solving and software development. The course focuses on the design, implementation, analysis, and application of various data structures such as arrays, linked lists, stacks, queues, trees, graphs, and hash tables. It also covers algorithmic techniques including recursion, sorting, searching, and complexity analysis.

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the role of algorithms in computing.
2. Select and apply appropriate data structures and algorithms to solve computational problems effectively.
3. Explain Red Black Trees and analyze their performance in terms of time and space complexity.
4. Apply dynamic Programming to the LCS Problem.
5. Construct a Priority Search Tree

Minad
2022
2022

2022
2022
2022

2022
2022
2022

UNIT – I

The Role of algorithms in computing – Algorithms, Algorithms as a technology, Getting Started- Insertion Sort, Analyzing algorithms, Designing Algorithms. Growth of functions – Asymptotic notations, Standard notation and common functions. How to create and Analyze the program Program, Arrays - Axiomatization, Ordered lists, Sparse Matrices, Representation of Arrays, Stacks and Queue-Fundamentals, Amazing Problem, Evaluation of Expressions, Multiple Stacks and Queues

UNIT – II

Dictionary Abstract Data Type, Implementation of Dictionaries. Hashing: Review of Hashing, Hash Function, Collision Resolution Techniques in Hashing, Extendible Hashing, Distributed Hashing for Databases, Consistent Hashing Need for Randomizing Data Structures and Algorithms, Search and Update Operations on Skip Lists, Probabilistic Analysis of Skip Lists, Deterministic Skip Lists

UNIT – III

Red Black Trees, 2-3 Trees, B-Trees, Splay Trees, Heaps as priority queues, heap implementation, insertion and deletion operations, binary heaps, Fibonacci heaps, Graphs and Graph Centrality Algorithms.

UNIT – IV

String Operations, Brute-Force Pattern Matching, The Boyer Moore Algorithm, The Knuth-Morris-Pratt Algorithm, Standard Tries, Compressed Tries, Suffix Tries, The Huffman Coding Algorithm, The Longest Common Subsequence Problem (LCS), Applying Dynamic Programming to the LCS Problem.

UNIT – V

One Dimensional Range Searching, Two Dimensional Range Searching, Constructing a Priority Search Tree, Searching Priority Search Trees, Priority Range Trees, Quadrees, k-D Trees. Files – Files Queries, and Sequential Organizations, Index Techniques, File Organizations, Storage management.

Text Books:

1. Ellis Horowitz and Sartaj Sahani (1986): Fundamental of data Structure, Computer science Press. inc, Maryland, USA.
2. Thomas H Cormen et al (2009): Introduction to Algorithms, 3rd Edition, The MIT Press, London England.

Reference Books:

1. Jon Kleinberg and Eva Tardos (2006): Algorithm Design, Pearson, USA.
2. Aho A. V. ,Hopcroft J.E. , and Ullman J. D (1974) : The Design and Analysis of Computer Algorithms , Addison-Wesley Publishing Company.

Mined *Boe* *Bue* *dl* *sign* *sohan*
su *Brage*

10. **Answer: A**—The author states that the "most common" cause of the "dramatic increase in the number of people who are obese" is "the increase in the consumption of fast food."

1. Describe the optimization methods such as linear programming problem.
2. Solve the Transportation Problem and Traveling salesman problem.
3. Analysis neural networks for Industry.
4. Examine gradient descent and Jacobian and multi dimension optimization.
5. Demonstrate RMS in Parametric distribution.

Linear Programming Problems (LPP): Graphical Methods, Simplex Method, Duality, Non-linear Programming: Quadratic, separable, geometric, stochastic programming.

Transportation Problem: Algorithm, method for initial solution, test for optimality. Assignment Problem: mathematical models, solution methods, multiple optimal solution, traveling salesman problem.

Neural networks, linear and nonlinear neural network, set and variables, functions, algorithms, load process, system identification, control system in loss.

Calculus: Limits, Derivative rule, n- order derivative, Jacobian and multi dimension optimization.
Gradient: Output and units, steps, Gradient Descent, graphical implementation, Learning Rate.

Root mean Square (RMS), RMS Function, optimization Error, RMSE, RMS in Parametric distribution, RMS Prop, ADAM optimization in neural networks.

1. J. K. Sharma, (2023): Operations Research: Theory and Applications; 6th edition, Trinity Press.
2. Francois Fleuret (2024): The Little Book of Deep Learning, e-book

Minod ~~Bage~~ Bah ~~DA~~
 Sur ~~Trage~~ 17 ~~foliar~~

Reference Books:

1. Swarup Kanti, P. K and Singh M.M (1985): Operations Research, 6th edition, Sultan chand and sons,
2. Gupta S. C. and V. K. Kapoor (2024): Fundamental of applied Statistics, 12th edition, Sultan chand & sons.
3. Hamdy A. Taha (2019): Operations Research, 10th edition, Pearson.

Course Name : Artificial Intelligence and Machine Learning

Course Code : DS 320

Semester : Third

Course Objective : The objective of this course is to provide students with a comprehensive understanding of the fundamental concepts, techniques, and applications of Artificial Intelligence (AI) and Machine Learning (ML).

Course Outcome : On successful completion of the course, the student will be able to:

1. Explain the basic concepts and scope of AI.
2. Understand the fundamental Concepts of Supervised learning.
3. Demonstrate various clustering, Methodologies; Represent any well-defined problem.
4. Apply AI Searching techniques to solve a problem.
5. Design knowledge base using FOPL and other knowledge representation techniques.

UNIT I:

Introduction to AI and ML: Definition and scope of AI and ML, Types of learning: supervised, unsupervised, reinforcement, History and evolution, AI vs ML vs Deep Learning, Intelligent agents, Agent types and architectures, Applications of AI and ML

UNIT II:

Supervised Learning : Classification and regression tasks, Linear regression, Logistic regression, K-Nearest Neighbours (KNN), Decision Trees, Random Forests, Naive Bayes, Support Vector Machines (SVM), Evaluation metrics: accuracy, precision, recall, F1-score, confusion matrix, Cross-validation, Bias-variance tradeoff, Overfitting and underfitting, Feature scaling and normalization

Minad *DR* *DR* *DR* *DR* *DR* *DR*
DR *DR* *DR* *DR* *DR* *DR* *DR*
DR *DR* *DR* *DR* *DR* *DR* *DR*

UNIT III:

Unsupervised Learning and Dimensionality Reduction: Clustering techniques: K-means, Hierarchical clustering, DBSCAN, Association rule mining: Apriori, FP-Growth, Evaluation of clustering algorithms, Dimensionality reduction: PCA, Feature selection and extraction

UNIT IV:

AI Search and Reasoning :Problem solving as search, Uninformed search: BFS, DFS, Informed search: A*, Greedy, Knowledge representation: Propositional logic, Predicate logic, Rule-based systems, Semantic networks, Ontologies, Reasoning under uncertainty: Bayesian networks, Fuzzy logic

UNIT V:

Neural Networks, Tools and Applications : Perceptron, Multilayer perceptron, Activation functions, Back propagation, Introduction to deep learning: CNN, RNN, Transfer learning, Python for AI/ML, Scikit-learn, Tensor Flow, Keras, Applications: NLP, recommender systems, computer vision, AI in business and healthcare, Ethics in AI, Case studies

Text Books:

1. Tom M Mitchell (2017): Machine Learning, 1st Edition , McGraw Hill.
2. Elaine Rich, Kevin Knight, Shivashankar B. Nair (2017): Artificial Intelligence, 3rd Edition, McGraw Hill Education
3. Ethem Alpaydin (2020): Introduction to Machine Learning, The MIT Press.
4. Dan W. Patterson (2015) : Introduction to Artificial Intelligence and Expert Systems, 1st edition, Pearson Education India.
5. Stuart Russell and Peter Norvig (2022): Artificial Intelligence: A Modern Approach, 4th Edition, Pearson Education.
6. Aurélien Géron (2022): Hands-On Machine Learning with Scikit-Learn, Keras, and Tensor Flow, Concepts, Tools, and Techniques to Build Intelligent Systems, 3rd edition, O'Reilly.

Reference Books:

1. Jude Shavil, Thomas , Dietterich (1990): Readings in Machine Learning, (The Morgan Kaufmann Series in Machine Learning), Morgan Kaufmann.
2. C. M. Bishop (2009): Pattern Recognition and Machine Learning, Springer
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman (2017): The Elements Of Statistical Learning: Data Mining, Inference, And Prediction, 2nd Edition, Springer.
4. J. Han and M. Kamber (2012): 3rd edition, Data Mining tools and Techniques, MKI.

Minad *Boor* *Abh* *Dr* *if* *Polan*
Sus *Pragati*

Course Name : High Performance Computing

Course Code : DS331

Semester : Third

Course Objective : Students will develop a cognitive understanding of parallel processing and hardware architecture of CPU for its implementation.

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the basic principles of Parallel Computing.
2. Explain the concepts of Multiprocessors, Multicomputer, Pipelining etc
3. Become more employable in the area of HPC
4. Pursue research in Computer Architecture and HPC.
5. Apply HPC in various fields

UNIT-I

Introduction - Feng's and Flynn's classification scheme – SISD, SIMD, MISD, MIMD, Multiprocessor and Multicomputer, UMA, NUMA, COMA, NORMA, memory models, parallel computer and its type. Applications of Parallel Computers. Cache Coherence Protocols – Snoopy and Directory Protocols.

UNIT-II

System Interconnect Architecture – Static and Dynamic, Hypercube Interconnection network, multistage interconnection networks-architecture and routing, design consideration, throughput delay, bandwidth. Architecture and routing of 3 stage and 4 stage Banyan Network. Routing and Addition in Hypercube Interconnection network. Performance Metrics and Benchmarks.

UNIT-III

Principle of pipelining-overlapped parallelism, Linear and non-linear pipelining, reservation table, calculation of MAL. Types of Instruction Pipeline. Arithmetic pipeline designs example –Floating point adder, pipelined multiplier

UNIT-IV

Advanced processor Technology – RISC, CISC, VLIW architectures, Hazard detection and resolution, functional organization of instruction in IBM 360/91. Numerical Problems based on CPI, IPC and MIPS.

UNIT-V

Exploring parallelism in program - Parallel Algorithm for Matrix addition and subtraction. Bitonic sort, sorting on linear array processors or odd even sort, PRAM algorithm for addition of numbers or Parallel Reduction. Bernstein's condition, ISO efficiency concept.,

Minad Bore *Bah* *DA* *Signature* *folow*

20

Text Books:

1. Buyya, R. (1999): High Performance Cluster Computing, Volume 1, Architecture and Systems, Prentice.
2. Kai Hwang (2001): Advanced Computer Architecture: Parallelism, Scalability, Programmability, TATA McGraw Hill International Editions

Reference Books:

1. Toby Velte, Anthony Velte, and Robert Elsenpeter (2017): Cloud Computing, A Practical Approach, , McGraw Hill Education.
2. Georg Hager and Gerhard Wellein (2010): Introduction to High Performance Computing for Scientists and Engineers, CRC Press
3. Berman F., Fox G. and A. J. G. Hey (2003): Grid Computing: Making the Global Infrastructure a Reality, Wiley India.

Course Name : Deep Learning

Course Code : DS 332

Semester : Third

Course Objective : The objective of this course is to teach the application of models and their process. Data driven technique for implementation of data cleaning and mining.

Course Outcome : On successful completion of the course, the student will be able to:

1. Understand the Basic function of data learning,
2. Explain the linear Layers and activation functions.
3. Design the Neural networks architecture for industry,
4. Visualization facts from figure using Deep learning networking
5. Solve real world problems using case studies in Deep learning.

UNIT-I

Basic function of data learning, under and over fitting, categories of model, Efficient computation, Training and losses, value of depth, scaling of data. Data Denoising and classification, data cleaning and mining

UNIT-II

Linear layers, activation functions, multi-layer model, Network flow in deep learning, deep learning approach, semi supervised learning, multi task learning.

21

UNIT-III

Deep networks architecture, Convolution Neural Network (CNN), Recurrent neural Network (RNN), Gated Recurrent Units (GRU), Auto encoder (AE), Long-Short Term Memory (LSTM),

UNIT-IV

Generating AI: restrictive machine learning, Gibbs Sampling, Gradient Computational in data networking, Generative Adversarial Network (GAN), LLMs, Finite tuning of deep learning model: Dropout, Hyper Parameter, Hyper Parameter Tuning Model Optimization.

UNIT-V

Case studies in Deep learning, model implementation, model merging in software, Transfer learning, Transfer learning models: Res-Net family, Google Net, VGG Net.

Text Books:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville (2017) : Deep Learning, The MIT Press, Cambridge.
2. Francois Fleuret (2024): The Little Book of Deep Learning, Universite, De Geneve.

Reference Books:

1. Charu C. Agarwal (2018): Neural Networks and Deep Learning, 1st edition, Springer.
2. Gibson Adam, Patterson Josh (2017): Deep Learning, O'Reilly.
3. Nikhil Buduma (2017): Fundamentals of Deep Learning Designing Next-Generation Machine Intelligence Algorithms, O'Reilly.

Course Name : Data Science for Health Care

Course Code : DS333

Semester : Third

Course Objective : This course intends to understand the application of data science in health care domain by applying statistical and machine learning methods to health care data.

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand Health care data and its sources
2. Analyze the data, create model and identify insights from health care data.
3. Explain the clinical trial ethics
4. Apply the Multiple Linear Regression Analysis in health care data
5. Explain the Epidemiological Methods for health care measurement.

Minad *Barz* *Sh* *John*
Sh *Ranger*

UNIT-I

Introduction to Health care data and its sources, types of health care EHR: components, benefits, Introduction to medical Statistics: uses statistical methods (designing and analysis), basic statistical concepts, and variations.

UNIT-II

Data analytics for health care: biomedical image analysis, sensor data analysis, Clinical Text Mining, Mining Biomedical Literature, Biomedical Signal Analysis, Genomic Data Analysis for Personalized Medicine. Clinico-Genomic Data Integration, Application and Practical system for health care: Data Analytics for Pervasive Health, Fraud Detection in Healthcare Data Analytics for Pharmaceutical discoveries

UNIT-III

Clinical Trials Ethics -, Measures of disease frequency and disease burden. Clinical Trials - Goals of Clinical Trials, Phases of Clinical Trials , Randomized Clinical Trails , Classification of Clinical Trials - Randomization: Fixed Allocation, Simple, Blocked, Stratified, Baseline Adaptive and Response Adaptive - Blinding: Single, double

UNIT-IV

Introduction to Multivariate regression methods: Multiple Linear Regression Analysis, Multiple Logistic regression analysis (discriminate analysis, factor analysis, cluster analysis).

UNIT-V

Epidemiological Methods: Evolutions, casual relationships, indices of burden of disease, Prevalence, incidence, crude and specific rate, standardized and adjusted rates, types of studies: cross sectional, case control, cohort studies, Issues in epidemiological Investigations: chance, confounding, bias, interaction. Validity and generalization of results and role of different study design in disease control

Text Books:

1. K.R. Sundaram, S.N. Dwivedi, V. Sreenivas(2016) :Medical Statistics : Principles and methods, 2nd Edition , Wolters Kluwer (India) Pvt. Ltd., New Delhi
2. Sergio Consoli, Milan Petkovic, and Diego Reforgiato Recupero (2019) :Data Science for Health care Methodologies and Application, Springer.

Reference Books:

1. Chandan K. Reddy and Charu C Agarwal, (2015) :Healthcare data analytics, 1st edition, Chapman and Hall/CRC
2. Kleinbaum, D. G., and Klein, M. (2012): Logistic regression: A Self-Learning Text, 3rd edition, Springer – Verlag, NY.
3. Shen-ChungChow , Jen-Pei Liu (1998): Design and analysis of Clinical trials, 1stEdition, Wiley –Inderscience.
4. Daniel, W. W. and Chad L. Cross (2018). Bio-Statistics: A foundation for analysis in the Health Sciences, 11th Edition, John Wiley & Sons, NY.

Course Name : Internet of Things
Course Code : DS 341
Semester : Third

Course Objective : The objective of this course is to provide students with a comprehensive understanding of the Internet of Things (IoT) ecosystem, including the design, development, and deployment of IoT systems. The course covers key topics such as sensors and actuators, communication protocols, IoT architectures, cloud computing, and data analytics.

Course Outcome : On successful completion of the course, the student will be able to:

1. Understand the basic of Internet of Things and its Terminology with Applications,
2. Design and program IOT devices and use real IOT protocols for communications.
3. Generate actionable insights for a variety of applications, including smart homes, healthcare, and industrial automation.
4. Gain hands-on experience with IoT devices, data collection, and real-time monitoring,
5. Explain the challenges related to security, scalability, and privacy in IoT systems.

UNIT I

Introduction to Internet of Things: Origin of Terminology IoT, Applications, Characteristics, Components of IoT, Associated technologies with IoT (M2M, Big Data, Cloud, Smart Grid, IoV, CPS, SDN, 3G/4G/5G), Challenges in IoT.

UNIT II

Connectivity: IoT Network Configurations, Gateway Prefix Allotment, IPv4, IPv6, IPv4 versus IPv6, RPL Data Protocol: Message Queue Telemetry Transport (MQTT), Constrained Application Protocol (CoAP), Advanced message Queuing Protocol (AMQP), Communication Protocols: IEEE 802.15.4, ZWave, Bluetooth, ZigBee, 6LowPAN, HART and Wireless HART, NFC, RFID.

UNIT III

Actuation: Actuator, Actuator Types: Hydraulic Pneumatic, Electrical, Thermal/ Magnetic Mechanical, Soft Actuators, Shape memory polymer (SMP) Types of Motor Actuators: Servo motor, Stepper motor, Hydraulic motor, Solenoid Relay, AC motor Sensing: Definition, Types of sensors, Transducers, Sensors Classes

UNIT IV

Introduction to Arduino Programming: Operators in Arduino, Control Statement, Loops, Integration of Sensors and Actuators with Arduino. Implementation of IoT: Interoperability in IoT, Introduction to NodeMCU (ESP8266), Connectivity of Sensors and Actuators with NodeMCU, Introduction to Python programming, Introduction to Raspberry PI.

UNIT V

Cloud Computing Fundamentals: Recent Trends in Computing, Evolution of Cloud Computing, Evolution of Cloud Computing, Business Advantages, Components Service Models: Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS) Infrastructure as-a-Service (IaaS), Multi-cloud, Inter-

Handwritten signatures and marks:
Several handwritten signatures and initials are present at the bottom of the page, including "Mishra", "Rohit", "Soham", and others. A large, stylized signature "Soham" is prominent on the right side.

cloud, Cloud Computing Service Management and Security, Case studies: Amazon Elastic Compute Cloud (EC2), Microsoft Azure.

Text Books:

1. Arshdeep Bahga and Vijay Madisetti (2015): Internet of Things: A Hands-On Approach, 1st edition, Orient Blackswan Private Limited - New Delhi.
2. Waltenegus Dargie and Christian Poellabauer (2011): Fundamentals of Wireless Sensor Networks: Theory and Practice, Wiley.

Reference Book:

1. Macro Schwart (2016): Internet of Things with Arduino Cookbook, Packt Publishing.
2. Peter Wahar (2015): Learning Internet of things, Packt publishing.

Course Name : Generative AI

Course Code :DS 342

Semester : Third

Course Objective: To provide students with a comprehensive understanding of the theoretical foundations, architectures, and real-world applications of Generative Artificial Intelligence.

Course Outcome : On successful completion of the course, the student will be able to:

1. Understand the foundational principles of AI and neural networks.
2. Analyze and compare different generative model architectures.
3. Apply probabilistic and training techniques in generative AI model design.
4. Evaluate the structure and functioning of large language models and transformers.
5. Analyze the societal, ethical, and business impacts of generative AI technologies.

UNIT-I

Foundations of AI and Neural Networks :History and evolution of AI/ML, Deep learning revolution, Introduction to Transfer learning, History of Neural Natural Language Processing, Structure of Artificial Neural Networks, Steps in Training an Artificial Neural Network: forward pass, loss, Back propagation. Parameters vs Hyper parameters.

UNIT-II

Deep Generative Models: Introduction to Generative AI and its significance, Generative Adversarial Networks (GANs): architecture and objective, Variational Autoencoders (VAEs): latent space, reparameterization, Long Short-Term Memory Networks (LSTMs), Attention Mechanisms.

(Handwritten signatures and marks)

25

UNIT-III

Foundations for Model Development: Basics of probability and statistics relevant to generative models, Data preprocessing techniques (normalization, tokenization, etc.), Model training techniques: optimization algorithms, regularization, early stopping.

UNIT-IV

Large Language Models and Transformers: Transformer architecture: encoder-decoder, self-attention, Pre-training objectives: masked LM, autoregressive LM, Transfer learning strategies in NLP, BERT architecture and variants, Overview of GPT.

UNIT-V

Applications, Trends, and Ethics in Generative AI: Real-world applications: chatbots, content creation, healthcare, design, Ethical concerns: bias, misinformation, deepfakes, plagiarism, Challenges in generative AI: hallucinations, evaluation, interpretability, Emerging trends: multimodal generation, diffusion models, retrieval-augmented generation, Societal and economic impact

TEXT BOOKS

1. Ian Goodfellow, Yoshua Bengio, and Aaron Courville (2016): Deep Learning (Adaptive Computation and Machine Learning series, The MIT Press.
2. Charu C. Aggarwal. (2018): Neural Networks and Deep Learning, 1st edition, Springer.
3. Altaf Rehmani (2024): Generative AI for Everyone: Understanding the Essentials and Applications of This Breakthrough Technology, Kindle Edition.

Reference Books

1. Josh Kalin (2018): Generative Adversarial Networks Cookbook: Packt Publishing,
2. Jesse Sprinter (2023): Generative AI in Software Development: Beyond the Limitations of Traditional Coding.
3. Zhihan Lyu (2024) : Applications of Generative AI, , Springer International Publishing AG.

Course Name : NATURAL LANGUAGE PROCESSING

Course Code : DS 343

Semester : Third

Course Objective : The objective of the course is to study language and the tools that are available to efficiently study and analyse large collections of text. To describe the application based on natural language processing and to show the points of syntactic, semantic and pragmatic processing.

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the fundamental concepts of natural language processing.
2. Understand the text pre-processing and corpora.

Munad *Boo* *Boh* *dh* *sig* *folan*
Rus *26* *Brage*

3. Analyze the words and perform POS tagging.
4. Distinguish between the syntactic and semantic correctness of the natural language.
5. Develop simple language models using NLTK.

UNIT-I

Introduction to various levels (stages) of natural language processing, Ambiguities, varieties and computational challenges in processing natural languages. Introduction to Real life applications of NLP such as spell and grammar checkers, information extraction, information retrieval, question answering, and machine translation.

UNIT-II

Text pre-processing, challenges, tokenization, sentence segmentation, regular expressions, words, text normalization, minimum edit distance, introduction to corpora, corpora analysis. The role of language models. N-gram models. Estimating parameters and smoothing. Evaluating language models.

UNIT-III

Parts of speech and morphology, Inflectional and Derivation Morphology, Morphological Analysis, FSA and Generation using finite state transducers. Introduction to POS tagging, HMM, Viterbi decoding for HMM.

UNIT-IV

Ambiguities in syntactic parsing, context free grammar, CYK parsing, shallow parsing and chunking, dependency parsing, statistical parsing and PCFG, Semantics, Lexical Semantics, Word senses, Relations between senses, Word Sense Disambiguation, Word similarity, WordNet, Thesaurus based word similarity, Thematic Roles, Semantic Role Labelling with CRFs.

UNIT-V

Tokenizing Text and WordNet Basics- Replacing and Correcting Words- Part-of Speech Tagging- Extracting Chunks- Text Classification – Named Entity Recognition.

Text Books:

1. Daniel Jurafsky and James H. Martin, Speech and Language Processing, 2017, 3rd edition, Prentice Hall.
2. Chris Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, 2016, MIT Press.

References Books:

1. James Allen "Natural Language Understanding, 2012, 8th Edition, Pearson Publication.
2. Vajjala, Sowmya, Bodhisattwa Majumder, Anuj Gupta and Harshit Surana. Practical natural language processing: A comprehensive guide to building real world NLP systems, 2020, O'Reilly Media.

Minad

Bore
Qu

Doh

DL

Bagh

Johan

Value Added Courses

Subject : Indian Knowledge System (IKS)

Subject Code : DS170

Semester : First

Course Objective: To learn the important of IKS and its impact on Society.

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the basic and needs of IKS
2. Explain the Introduction to Vedas to society
3. Interpret the Purāṇic repository
4. Implement the National Statistics System and its role in data collection.
5. Know the famous statistician and their contributions to academic and society

UNIT I

Indian Knowledge System (IKS) – An Introduction: What is IKS? Need of IKS. Organizations of IKS. Historicity of IKS, Some salient aspects of IKS.

UNIT II

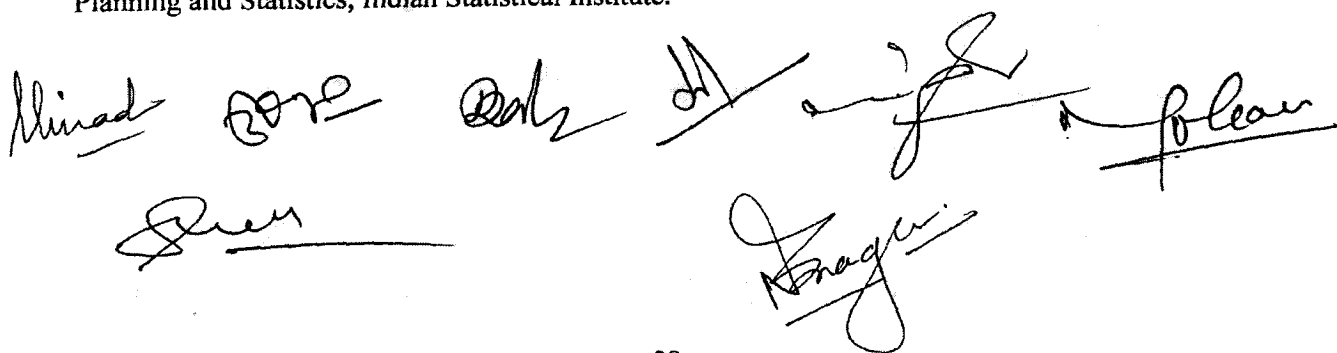
The Vedic Corpus: Introduction to Vedas. A synopsis of the four Vedas. Sub-classification of Vedas. Messages in Vedas. Introduction to Vedāṅgas. Prologue on Śikṣā and Vyākaraṇa. Basics of Nirukta and Chandas. Introduction to Kalpa and Jyotiṣa. Vedic Life: A Distinctive Features.

UNIT III

Wisdom through the Ages: Gateways of ancestral wisdoms. Introduction to Purāṇa. The Purāṇic repository. Issues of interest in Purāṇas. Introduction to Itihāsas. Key messages in Itihāsas. Wisdom through Nīti-śāstras. Wisdom through Subhāṣita.

UNIT IV

Duet Vidya, National Statistics day, International Statistics Day, National Statistics System, Introduction about agencies- NSO and CSO, Directorate of Economics and Statistics. Ministry of Planning and Statistics, Indian Statistical Institute.

A collection of handwritten signatures in black ink, likely representing the faculty members who have approved the course. The signatures are written in a cursive style and are arranged in two rows. The first row contains five signatures, and the second row contains two signatures.

Indian contribution to Statistics, Prasanta Chandra Mahalanobis, C. R. Rao, Devbrata Basu, Samarendra Nath Roy, P. V. Sukhatmi, Raj Chandra Bose, Kanti Lal Mardia, K. C. Sreedhaan Pillai, Pranab K. Sen, B. L. S. Prakash Rao, V. S. Huzurbazar, Jyant Kumar Ghosh.

1. B. Mahadevan, Vinayak Rajat Bhat, R.N. Nagendra Pavana (2022): Introduction to Indian Knowledge System: Concepts and Applications, PHI Learning Pvt. Ltd.

1. K. Kapur A.K. Singh (2005): Indian Knowledge Systems, Vol. 1 & 2. D.K. Print world Pvt. Ltd.,
2. शशिबाला, ओम विकास, अशोक प्रधान (संपादक) (2018), भारती विद्या सार-1, भारतीय विद्या भवन.
3. S. B. Rao (2012): Indian Mathematics and Astronomy: Some Landmarks, Revised 3rd Edition, Bhartiya Vidhya Bhavan.
4. B.S. Yadav and Man Mohan (2010): Ancient Indian Leaps into Mathematics, Birkausher Publication.
5. Dharampal (2000): Indian Science and Technology in the Eighteenth Century, Other India Press.
6. Dharampal (2000): The Beautiful Tree: Indigenous Indian Education in the Eighteenth Century, Other India Press.

Semester : Third

Course Outcome: On successful completion of the course, the student will be able to:

1. Understand the basic of IPR and role of IP in the society.
2. Explain the Categories of IP Patents.
3. Apply the trademarks and Industrial Designs in business.
4. Recall the World Intellectual Property Organization structure.
5. Education and training in intellectual property.

Introduction to IPR- Role of IP in the Economic and Cultural Development of the Society, IP Governance, IP as a Global Indicator of Innovation, Origin of IP, History of IP in India

29

29

UNIT – II

Categories of IP Patents : Conditions for Obtaining a Patent Protection , To Patent or Not to Patent an Invention, Rights Associated with Patents, Enforcement of Patent Rights, Inventions Eligible for Patenting, Non-Patentable Matters, Patent Infringements , Avoid Public Disclosure of an Invention before Patenting, Process of Patenting,

UNIT – III

Copyrights and Related Rights, Trademarks, Industrial Designs, Geographical Indications, Trade Secrets

UNIT – IV

World Intellectual Property Organization- Introduction , Origin of WIPO, Salient Features of WIPO , Main Activities of WIPO, India and WIPO , Access to International IP Filing Systems , Protection of Traditional Knowledge (TK), Technology and Innovation Support Centres (TISC), Access to International Microorganism Deposit System 3.8. Data for International IP Filing System

UNIT – V

Education and training in intellectual property - WIPO e-Learning Centre - A Global Online Platform, Other Online Platforms Offering Courses in IP.

Text Book:

1. Tewari and Bhardwaj (2021): Intellectual Property, A Primer for Academia, Panjab University.

Reference Books:

1. Morris Asimow (1962): Introduction to design, Englewood Cliffs, N.J., Prentice-Hall
2. Mayall, W. H. (1967): Industrial Design for Engineers, McGraw Hill.
3. Debora J. Halbert (2006): Resisting Intellectual Property, 1st ed., Taylor & Francis Ltd.
4. Asha Vijay Durafe, Dhanashree K. zoradmalle (2020): Intellectual Property Rights Kindle Edition, Wiley.

[Handwritten signatures and initials]