PONDICHERRY UNIVERSITY
(A CENTRAL UNIVERSITY)

B.Sc. Data Science (Honors)
B.Sc. Data Science (Honors with Research)

REGULATIONS, CURRICULUM & SYLLABUS
(For Affiliated Colleges)

(Under the National Education Policy - NEP 2020)
Effective from the Academic Year 2023 - 2024

Revised in June 2024

Board of Studies (BOS) - Computer Science

S.No.	Name	Affiliation	BoS
1	Dr. S.K.V. Jayakumar	Dept. of CS, Pondicherry University	Chairman
2	Dr. M. Punithavalli	Dept. of CA, Bharathiar University	External Members
3	Dr. P.S. Grover	Dept. of CS, Delhi University	
4	Dr. M. Sornam	Dept. of CS, University of Madras	Special Invitee
5	Dr. R. Subramanian	Dept. of CS, Pondicherry University	Members of Board
6	Dr. T. Chithralekha	Dept. of CS, Pondicherry University	
7	Dr. S. Bhuvaneswari	Dept. of CS, Pondicherry University	
8	Dr. R. Sunitha	Dept. of CS, Pondicherry University	
9	The Head, Dept. of CS	Dept. of CS, Achariya Arts & Sci. College	
10	The Head, Dept. of CS	Dept. of CS, Idhaya College of Arts & Sci.	
11	The Head, Dept. of CS	Dept. of CS, Avvaiyar Govt. College	
12	The Head, Dept. of CS	Dept. of CS, Don Bosco College	
13	The Head, Dept. of CS	Dept. of CS, PSV College of Arts & Sci.	
14	The Head, Dept. of CS	Dept. of CS, Rajiv Gandhi Arts & Science	

NEP Committee

S.No.	Name	Designation	Affiliation
1	Prof. S. K.V. Jayakumar	Professor and Head	Dept. of Computer Science, Pondicherry University
2	Prof. T. Chithralekha	Professor	
3	Prof. P. Sujatha	Professor	
4	Dr. R. Sunitha	Associate Professor	
5	Dr. V. Uma	Associate Professor	
6	Dr. K. S. Kuppusamy	Associate Professor	
7	Dr. M. Sathya	Assistant Professor	

Syllabus Revision Committee (Data Science)

S.No.	Name	Designation	Affiliation
1	Dr. M. Nandhini	Professor	Dept. of Computer Science, Pondicherry University
2	Dr. P. Shanthi Bala	Professor	
3	Dr. R. Sunitha	Associate Professor	

Table of Contents

1. PREAMBLE & PROGRAMME OUTCOMES

1.1. Preamble

In today's rapidly evolving world, the importance of data and its analysis has never been more profound. The field of Data Science stands at the forefront of this data revolution, driving innovations, solving complex problems, and shaping the future of industries across the globe. Recognizing the profound significance of Data Science in modern society, there is an impending need to introduce an undergraduate program in Data Science.

Data Science, at its core, is the discipline that empowers individuals and organizations to harness the power of data for informed decision-making. It blends mathematics, statistics, computer science, and domain expertise to extract valuable insights, predict future trends, and optimize various processes. This field has already permeated every sector, from healthcare and finance to marketing and environmental sciences, and its impact continues to expand.

The significance of including an undergraduate program in Data Science within educational framework is multifold. Firstly, it addresses the growing demand for professionals with specialized skills in data analysis, machine learning, and artificial intelligence. Secondly, an undergraduate program in Data Science fosters interdisciplinary learning, as students gain proficiency in mathematics, statistics, programming, and data ethics. This interdisciplinary approach encourages a holistic understanding of complex problems and enhances critical thinking and problem-solving skills. Thirdly, this program in Data Science aligns with the global imperative of cultivating a workforce that can contribute to scientific research, economic growth, and societal welfare.

Moreover, by offering an undergraduate program in Data Science, students are empowered to participate actively in the data-driven society, facilitating data literacy and digital citizenship. This, in turn, enhances their career opportunities and prepares them for a future where data will continue to play an increasingly central role.

In conclusion, this program embodies the objective of equipping students with the knowledge, skills, and ethical values necessary to navigate the complexities of the data-driven world, ensuring a brighter future for individuals and society.

The present Curriculum Framework for B.Sc (Data Science) degree is intended to facilitate the students to achieve the following.

- To provide an understanding and knowledge of the basic theory of Computer Science and Information Technology with good foundation on theory, systems and applications such as algorithms, data structures, data handling, data communication and computation

- To offer a strong foundation in Data Science, including statistical analysis, machine learning, data analysis, and applications, along with practical skills that are immediately applicable in the industry, aligning with NEP's skill-based learning approach.

- To encourage students to engage in cross-disciplinary learning, promote a holistic understanding of data science's real-world applications and connect it with other domains, fostering a multidisciplinary approach as outlined in the NEP.

- To encourage students to undertake projects, collaborate on research, and contribute to developing cutting-edge data science solutions.

1.2 Programme Outcomes

Upon completion of the Bachelor of Science (B.Sc.) programme in Data Science, students will demonstrate the following outcomes:

UG Certificate Level

- Solid Understanding of fundamental concepts in computer science and data science
- Demonstrate basic skills in problem-solving and programming.

UG Diploma Level

- Gain knowledge about statistics, mathematics required to perform data analysis
- Possess practical skills to perform programming and data analysis
- Familiarized with computer science concepts.

UG Degree Level

- Use data science tools, to tackle real-world data challenges effectively.
- Become adept at applying data science techniques in various domains, fostering a multidisciplinary approach to problem-solving and decision-making.

UG Degree with Honors / Honors with Research

- Acquire ability to conduct data-driven research and innovation in data science, contributing to the advancement of the field.
- Excel in effectively communicating complex data-driven insights through reports, visualizations, and presentations.

2. DEFINITIONS

Terms used in the NEP Regulations shall have the meaning assigned to them as given below unless the context otherwise requires:

A. Credit: A credit is the number of hours of instruction required per week for the given subject in a given semester of 16-18 weeks. One credit is equivalent to 15 hours of teaching (lecture or tutorial) or 30 hours of practice/field work/community engagement and service per Semester.

B. Academic Year: Means the year starting on 1st day of July and ends on the 30th day of June in the succeeding year.

C. Residence time: Means the time a student spends for attending classes in the College/Institution (either Online/Offline) as a full-time student and enrolled in any Academic programme of the Institution.

D. Semester: Means 18 weeks (90 Working days) of teaching-learning sessions of which two weeks shall be set apart for examinations and evaluation;

E. Grade: Means a letter grade assigned to a student in a Course for his/her performance at academic sessions as denoted in symbols of: O(Outstanding), A+(Excellent), A (Very good), B+ (Good), B (Above average), C (average), P (Pass) F (Fail) and Ab (Absent) with a numeric value of O=10, A+=9, A=8, B+=7, B=6, C=5 P=4, and F=0, Ab=0;

F. Grade Point Average (GPA): Means an average of the Grades secured by a student in all courses in a given academic session duly weighted by the number of credits associated to each of the courses.

G. Cumulative GPA (CGPA): is the weighted average of all courses the student has taken in a given Programme.

H. A common Course: Means the set of courses that all student s who are admitted to any Programme of the University are required to study; these courses include, Languages (English- modern Indian languages), NEP specific courses- viz. Understanding India, Environmental sciences/Education, Health and Well being/Yoga, Digital & Technological solutions;

I. Major Discipline: Means the core subjects mandatory for the programme, Major discipline may be a single discipline or interdisciplinary/ multidisciplinary courses. Eg. B.Sc. (Physics) or B.Sc. (Physics, Maths and Chemistry).

J. Minor Discipline: Means courses which are specific to the specializations in Computer Science.

K. Credit Requirement: For a Degree/Diploma/Certificate Programme means the minimum number of credits that a student shall accumulate to achieve the status of being qualified to receive the said Degree, Diploma/Certificate as the case may be.

L. Exit option: Means the option exercised by the students, to leave the Programme at the end of any given Academic year.

M. Lateral entry: Means a student being admitted into an ongoing Programme of the University otherwise than in the 1st year of the programme.

N. Vocational Studies/Education: This refers to set of activities for participation in an approved project or practical or lab, practices of application of scientific theories, studio activities involving students in creative artistic activities, workshop-based activities, field-based shop-floor learning, and Community engagement services, etc. (These courses are expected to enable students to incorporate the learned skills in daily life and start up entrepreneurship.)

O. Skill-based learning/project: This refers to activities designed to understand the different socio-economic contexts, first-hand understanding of the policies, regulations, organizational structures, processes, and programmes that guide the development process.

P. Work-based internship: Means structured internships with Software Companies, Research and Higher Educational Institution Laboratories, Corporate offices, etc. which will further improve employability.

3. DURATION, ELIGIBILITY & AWARD OF UG DEGREE/DIPLOMA/CERTIFICATE

3.1. Duration of the Programme

The duration of the UG programme is 4 years or 8 semesters. Students who desire to undergo a Three-year UG Programme will be allowed to exit after completion of the 3rd year. If a student wants to leave after the completion of the first or second year, the student will be given a UG Certificate or UG Diploma, respectively, provided they secure the prescribed number of credits (as given in table below).

3.2. Eligibility

Senior Secondary School Leaving Certificate or Higher Secondary (12th Grade) Certificate obtained after successful completion of Grade 12 or equivalent stage of education corresponding to Level-4 (Levels in NHEQF). For detailed eligibility conditions , refer the Admissions and Lateral Section below.

3.3. Awarding of UG Certificate, UG Diploma and Degrees Nomenclature

Four years B.Sc. Degree Programme shall have options for earning a Certificate / Diploma / UG Degree / UG Degree (Honors) / UG Degree (Honors with Research) based on the exit option exercised by the candidates.

3.3.1. UG Certificate

Students who opt to exit after completion of the first year (2 Semesters) and have earned a minimum of 40 credits will be awarded a UG certificate in Data Science if, in addition, they complete work based vocational course / internship of 4 credits during the summer vacation of the first year.

3.3.2. UG Diploma

Students who opt to exit after completion of the second year (4 Semesters) and have earned a minimum of 80 credits will be awarded the UG diploma in Data Science if, in addition, they complete work based vocational course / internship of 4 credits during the summer vacation of the second year.

3.3.3. Three-year UG Degree

Students who wish to discontinue after the 3-year (6 Semesters) UG programme will be awarded a UG Degree in Data Science after successful completion of three years, earning a minimum of 120 credits and satisfying the minimum credit requirements as mentioned in Table1 below.

3.3.4. Four-year UG Degree (Honors)

A four-year UG Honors degree in Data Science will be awarded to those who complete a four-year (8 Semesters) degree programme, earning a minimum of 160 credits and have satisfied the credit requirements as mentioned in Table1.

3.3.5. Four-year UG Degree (Honors with Research)

Students who secure a minimum of 7.5 CGPA in the first six semesters and wish to undertake research at the undergraduate level can choose a research stream in the fourth year. They should do a research project or dissertation under the guidance of a faculty member of the University.

The research project/dissertation will be in the major discipline, Data Science. The students who secure a minimum of 160 credits, including 12 credits from a research project/dissertation, will be awarded UG Degree in Data Science (Honors with Research).

3.3.6. Programme overview

As per the guidelines of NEP, students are mandated to complete 120 credits to complete a basic Bachelor's Degree in 3 years. With an additional 40 credits of course work one can pursue 4th Year Honors or Honors with Research Degree. The UG Programme will consist of the following categories of courses and the minimum credit requirements for 3-year UG and 4-year UG(Honors) or UG (Honors with Research) programmes are given below.

Table 1: Breakup of Credits and Courses – Minimum Requirements

S.No.	Component	3 Year UG	4 Year UG (Honors/ Honors with research)
1	Major Disciplinary - Computer Science	60 Credits (15 Courses of 4 credits)	80 Credits (20 Courses of 4 credits)
2	Minor Disciplinary – Specialization Courses	24 Credits (6 Courses of 4 Credits)	32 Credits (8 Courses of 4 credits)
3	Multi-Disciplinary Courses	9 Credits (3 courses of 3 credits)	9 Credits (3 courses of 3 credits)
4	Ability Enhancement Courses	8 Credits (4 courses of 2 credits)	8 Credits (4 courses of 2 credits)
5	Skill Enhancement Course – On the chosen Specialization	9 Credits (3 courses of 3 credits)	9 Credits (3courses of 3 credits)
6	Value-added courses	8 Credits (4 courses of 2 credits)	8 Credits (4 courses of 2 credits)
7	Summer internship	4 Credits (Included in Major courses of 60 credits)	4 Credits (Included in Major courses of 80 credits)
8	Community engagement and service	2 Credits (1 Field based Course)	2 Credits (1 Field based Course)
9	Research Dissertation Project	-	12 Credits
	Total	**120**	**160**

Note: Honors students not undertaking research will do 3 courses for 12 credits in lieu of a Research Project / Dissertation.

3.3.7. Degree and Nomenclature

Candidates who complete Eight semesters and earn a minimum of 160 credits and have satisfied the credit requirements as mentioned in the table below will be awarded either of the following degrees.

- B.Sc. Data Science*
- B.Sc. Data Science (Honors) #
- B.Sc. Data Science (Honors with Research) ##

* for candidates who wish to exit at the end of the third year with 120 credits earned and satisfied the other minimum requirements given in 3.3.9.

for candidates who complete 3 theory courses (MJD 21, MJD 22, and MJD 23) instead of the research project work in the Eighth Semester.

for candidates who complete a research project work in the Eighth Semester

3.3.8. Degree with Specialization

Out of the above said 160 credits (Table1), the candidates shall earn 103 credits (83 credits out of 120 credits in the case of 3 year UG) from the Hardcore courses (Major Disciplinary, Multi-disciplinary, Ability Enhancement, Value added Courses and Community Engagement and Service) and the remaining 57 credits (37 credits in the case of 3 year UG) shall be earned from the subjects they choose to study from the list of softcore courses. These 57 credits shall be earned through studying the specialization courses in Minor Disciplinary – Specialization Courses, Skill Enhancement Courses in all the semesters and the Research Project or the Courses the candidates choose to study in the Eighth Semester. The Programme Structure is detailed in the following figure.

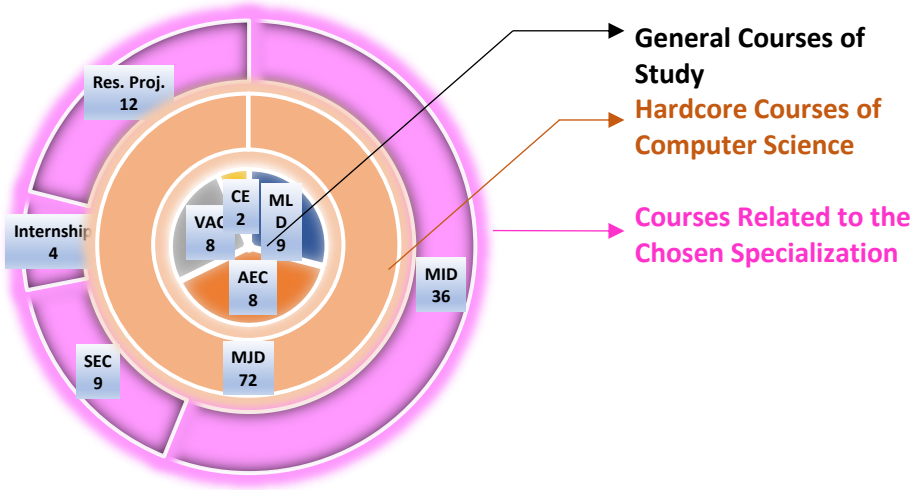

Figure : Programme Structure with credit breakup

3.3.9. Exit Options and Nomenclature of Certificate, Diploma

Candidates can exercise the following exit options and obtain the said certificate or diploma or degree, if the minimum required credits are earned and other conditions are met. Students exercising the option of exit at the end of 2nd semester or 4th semester need to have completed an internship for atleast 8 weeks along with the necessary credit requirements to qualify for the relevant certificate or

diploma. In any case, every student, whenever exit (or complete the 4 year programme), should have completed atleast one internship for a minimum period of 8 weeks.

Exit after 2nd Semester: Certificate in Data Science will be awarded for candidates who exit the course at the end of 2nd semester and earned a minimum of 40 credits and have completed a Summer Internship of 4 credits for a minimum period of 8 weeks, during the summer vacation post 2nd semester.

Exit after 4th Semester: Diploma in Data Science will be awarded for candidates who exit the course at the end of 4th semester and earned a minimum of 80 credits and have completed a Summer Internship of 4 credits for a minimum period of 8 weeks, during the summer vacation post 4th semester.

Exit after 6th Semester: UG Degree in Data Science (B.Sc. (DS)) will be awarded for candidates who exit the course at the end of 6th semester and earned a minimum of 120 credits and have completed a Summer Internship of 4 credits for a minimum period of 8 weeks, during the summer vacation post 4th semester.

Exit after	Credits and other requirements	Awards
2nd Semester	Min: 40 Credits & Internship	Certificate in Data Science
4th Semester	Min: 80 Credits & Internship	Diploma in Data Science
6th Semester	Min: 120 Credits & Internship	B.Sc. Data Science

4. STRUCTURE OF THE UNDERGRADUATE PROGRAMME

This B.Sc Honors programme offered in the affiliated colleges shall confirm to the structure specified hereunder. As per the decided programme mandate, the students to complete 120 credits to complete a basic Bachelor's Degree in 3 years. With an additional 40 credits of course work one can pursue 4th Year Honors or Honors with Research Degree. The UG Programme will consist of the categories of courses and the minimum credit requirements for 3-year UG and 4-year UG(Honors) or UG (Honors with Research) programmes as given in Table 1.

4.1. Types of Courses

Hardcore Courses	Softcore Courses (Specialization specific)
Major Disciplinary – Data Science	Minor Disciplinary
Multi-Disciplinary Courses	Skill Enhancement Courses
Ability Enhancement Courses	Summer Internship
Value Added Courses	Research Dissertation Project
Community Engagement and Service	

4.2. Description of Courses

The following are the types of courses in this programme:

4.2.1. Major Discipline: 60 Credits - 3 Year UG, & 72 Credits - 4 Year UG

Major discipline here means to Data Science. Students should secure the prescribed number of credits (not less than 50% of the total credits) through core courses in the major discipline. The major discipline would provide the opportunity for a student to pursue in-depth study of a particular subject or discipline. A student may choose to change the major discipline within the broad discipline at the end of the second semester provided all the prerequisites of the respective degree programme are fulfilled.

4.2.2. Minor Discipline / Specialization: 24 Credits - 3 Year UG, & 40 Credits - 4 Year UG

Minor discipline helps a student to gain a broader understanding beyond the major discipline.

4.2.3. Multidisciplinary courses (MD): 9 Credits

All undergraduate students must pursue 9 credits worth of courses in such Multi-disciplinary areas/Courses from NEP defined subjects. Colleges may identify any 3 multiple disciplinary streams listed below based on availability of resources and manpower.

a) Natural Sciences
b) Physical Sciences
c) Mathematics & Statistics
d) Computer Science/Applications
e) Data Analysis
f) Social Sciences
g) Humanities
h) Commerce & Management
i) Library Science
j) Media Sciences, etc.

Students are expected to learn basic/introductory courses designed by other departments for this purpose. Colleges may list any 3 introductory courses (one each in Natural Sciences, Physical Sciences, Humanities) for uniform adoption of all UG students.

4.2.4. Ability Enhancement Courses (AEC): 8 credits

All Undergraduate (UG) students are mandated to complete at least 8 Credits worth of Courses which focus on Communication and Linguistic skills, Critical reading, writing skills. These courses are expected to enhance the ability in articulation and presentation of their thoughts at workplace. Colleges may design these ability enhancement courses tuned to the requirements of given major discipline. Eg. A course in Business Communication is more appropriate in place of literature/prose/poetry.

Ability Enhancement Course	
I. English Language	**II. Indian Language (two courses)**
a. English Language & Literature - 1 and 2	a. Indian language & Literature - 1 and 2
b. Functional English - 1 and 2	b. Functional language - 1 and 2
c. Communicative English - 1 and 2	c. Communicative language - 1 and 2

4.2.5. Skill Enhancement Courses (SEC): 9 credits

These courses are aimed at imparting practical skills, hands-on training, soft skills, and other skills to enhance the employability of students. Courses are designed as per the students' needs with the available resources. Students can choose these courses from the list of courses offered in the chosen specialization as said in 4.2.2. Colleges may also outsource the Skill Enhancement Courses to AICTE approved agencies for conducting short term Training Workshops, Skill India initiatives of GOI and approved Trades by Skill development of corporation are to be considered.

4.2.6. Value-Added Courses (VAC) Common to All UG Students: 8 credits

Under NEP, the UGC has proposed for 6 to 8 credits worth of common courses which are likely to add value to overall knowledge base of the students. These courses include:

a) Understanding India
b) Environmental Science/Education, Higher Order Thinking
c) Digital and Technological solutions
d) Health &Wellness, Yoga Education, Sports & Fitness, Universal Human Values

The course structure and coverage of topics are suggested by UGC in its draft documents, colleges/UG Boards of Studies may design the methodology for conducting these value added courses.

4.2.7. Summer Internship: 4 Credits

All students will undergo Internships / Apprenticeships in a firm, industry, or organization or Training in labs with faculty and researchers in their own or other Higher Education Institutions / Research institutions during the summer term. Students will be provided with opportunities for internships to actively engage with the practical side of their learning. Such Summer Internships is to be conducted in between 4th and 5th Semester. A review report and award of grade based on work-based learning by students is to be recorded during the 5th Semester. Students who exercise the option of exit at the end of 1st and 2nd year need to do internship as specified in the respective section. Summer internship shall be conducted for a minimum of 8 weeks.

4.2.8. Community Engagement and Service: 2 Credits

The curricular component of 'Community Engagement and Service' seeks to expose students to the socio-economic issues in society so that the theoretical learnings can be supplemented by actual life experiences to generate solutions to real-life problems. This can be part of summer term activity or part of a major or minor course. Community Engagement shall be conducted for a minimum of 2 weeks.

4.2.9. Research Project / Dissertation: 12 Credits

Students choosing a 4 Year Bachelor's degree (Honors with Research) are required to take up research projects under the guidance of a faculty member. The students are expected to complete the Research Project in the eighth semester.

4.2.10. Audit courses: 0 credits

Audit courses offered do not carry any credits. Evaluation will be based on continuous assessment. Students may be given a Pass or Fail (P/F) based on the assessment that may consist of class tests, homework assignments, and/or any other innovative assessment methodology suitable to the expected learning outcome, as determined by the faculty in charge of the course of study.

4.3. Levels of the Courses

Course codes are based on the academic rigor. The first four letters of the course code indicate the department/Centre, followed by the academic rigor level code in digits (For e.g., Comp 201). The coding structure follows:

4.3.1. 0-99: Pre-requisite courses

It is required to undertake an introductory course which will be a pass or fail course with no credits. It will replace the existing informal way of offering bridge courses that are conducted in some of the colleges/ universities.

4.3.2. 100-199: Foundation or introductory courses

These are courses which are intended for students to gain an understanding and basic knowledge about the subjects and help decide the subject or discipline of interest. These courses generally would focus on foundational theories, concepts, perspectives, principles, methods, and procedures of critical thinking in order to provide a broad basis for taking up more advanced courses.

4.3.3. 200-299: Intermediate-level courses including subject-specific courses

These courses are intended to meet the credit requirements for minor or major areas of learning. These courses can be part of a major and can be pre-requisite courses for advanced-level major courses.

4.3.4. 300-399: Higher-level Courses

These courses are required for majoring in a disciplinary/interdisciplinary area of study for the award of a degree.

4.3.5. 400-499: Advanced Courses

These courses which would include lecture courses with practicum, seminar-based course, term papers, research methodology, advanced laboratory experiments/software training, research projects, hands-on-training, internship/apprenticeship projects at the undergraduate level or first year post-graduate theoretical and practical courses.

4.4. Credit-hours for different types of courses

A three-credit lecture course in a semester means three one-hour lectures per week with each one-hour lecture counted as one credit. One credit for tutorial work means one hour of engagement per

week. A one-credit course in practicum or lab work, community engagement and services, and fieldwork in a semester mean two-hour engagement per week.

The Faculty to Student Ratio in all the practical / laboratory classes shall be maintained at 1:25.

In a semester of 15 weeks duration, a one-credit practicum in a course is equivalent to 30 hours of engagement. A one-credit of Seminar or Internship or Studio activities or Field practice/projects /community engagement and service means two-hour engagements per week. Accordingly, in a semester of 15 weeks duration, one credit in these courses is equivalent to 30 hours of engagement.

4.4.1. Pedagogical Styles

In order to achieve the expected Learning outcomes, UGC Framework has specified different Pedagogical approaches for different courses at undergraduate level. These approaches include:

a) Lecture course b) Tutorial course

c) Practice cum or laboratory courses d) Seminar Course

e) Internship course f) Studio activity-based course

g) Field practicing h) Project work courses

i) Community engagement and service course

The details of these different types of Pedagogical methods are as follows:

Table 2: Pedagogical Approaches

COURSE TYPES	APPROACH
Lecture Courses	Regular classroom lectures by qualified / experienced Expert Teachers • These Lectures may also include classroom discussion, demonstrations, case analysis • Use of Models, Audio-Visual contents, Documentaries, PPTs may supplement.
Tutorial Courses	Problem solving Exercise classes guided discussion, supplementary readings vocational training, etc.
Practical / Lab work	Practical Lab activity with Theoretical support Mini projects, Activity based engagement, Program executions, Data processing and presentation exercise.
Seminar Course	A course requiring student to design and participate in discussions, Group Discussions, Elocution and Debate, Oral Communication Paper presentations, Poster Presentation, Role play participation, Quiz competitions, Business plan preparation/presentation, etc.

Internship course	Courses requiring students to *Learn by Doing* in the workplace external to the educational Institutions. Internships involve working in Software Companies, Research and Higher Educational Institution Laboratories, Corporate Offices, etc. All Internships should be properly guided and inducted for focused learning.
Research Project	Students need to study and analyze the recent research publications from indexed/peer reviewed journals in their area of specialization. Outcome of the study and analysis need to be presented as a thesis or research report with necessary experimental results.

4.5. Semester-wise Break: for courses of 3 year UG and 4 Year UG(Hons) Degree programmes

Incorporating the focus of NEP in terms of different categories of courses and award of Certificates, Diplomas and Degrees during different stages of 4 year Degree programmes, a template for Semester-wise course work was designed by the UGC and presented in para 5.3 of Curriculum Framework . Salient features of it are as follows:

- All courses shall carry specified number of credits.
- Every Semester shall have a minimum of 20 credits worth of courses.
- Credits for a course shall be decided on the basis of number of Contact hours of the teaching in a classroom.
- One credit means one hour of Teaching in the case of a Theory subject and at least 2 hours of conducting Practical in the case of Lab subjects.
- All Major and Minor disciplinary Courses shall have 4 credits with 6 hours of work load (including 2 hours of tutorials)
- Language courses, ability enhancement, skill enhancement and value added common course also will have 2 hours of hands on training.
- Progress of Learning is measured in terms of credits earned by the students by successfully completion the course.
- Students can exercise his/her choice for exiting the course at the end of every Academic year.
- Graduate attributes listed by UGC shall be the focus of Teaching-Learning process.
- Semester I and II shall focus on introductory courses/subjects in Major/Minor disciplines and shall focus on providing knowledge in Multidisciplinary areas, skill enhancement and ability enhancement courses.
- Semester III and IV shall focus on Core disciplinary courses with a focus on building strong foundation in the given Discipline.

- Semester V and VI shall focus on providing in-depth knowledge and skills required for taking up a career in the given discipline.

- Semester VII and VIII shall focus on Advanced knowledge and shall direct the students to take up socially relevant projects/Research works newer applications of the knowledge.

5. ADMISSION ELIGIBILITY, LATERAL ENTRY

5.1 Admission Eligibility

The candidates for admission to this programme shall be required to have passed 10+2 / 10+3 system of examinations or equivalent with Mathematics / Business Mathematics / Computer Science / Computer Applications / Informatics Practices / or Equivalent as one of the subjects of study.

Students shall be admitted to this programme based on admissions criteria fixed by the University / Government of Puducherry from time to time.

5.2 Admissions by Lateral Entry

In this programme, where admission was carried out adopting approved procedures in preceding years, subject to availability, lateral entry admission shall be permitted, subject to:

Candidates seeking entry at the second, third and fourth year, should meet the necessary eligibility criteria with respect to the certificate / diploma / degree they possess, with necessary minimum credits banked in the Academic Bank of Credits (ABC). Such students who get admitted in later years, other than first year will be guided by the following clauses:

- That the University shall notify the admission process and number of vacancies open for lateral entry.
- That the Lateral entrants shall be admitted only after such transparent screening process and such procedure that the University may prescribe from time to time. University may prescribe different methods of screening for different programmes depending on the circumstances prevailing in each case.
- Lateral entry shall be permissible only in the beginning of years 2, 3, 4 of the Under Graduate / Honors programme; provided that the students seeking lateral entry shall have obtained the minimum pass marks / grades fixed by the University in their previous academic years.

6. EVALUATION

All Credit courses are evaluated for 100 marks. Internal Assessment component is for 25 marks and the End Semester University exam is for 75 marks. In case of Practicals, Project work etc., it is 50 marks for Internal and 50 marks for End-Semester Exams.

6.1. Category of Courses

There are three category of courses as shown in 6.2. Category A, theory courses with lecture hours and tutorials are evaluated for an Internal assessment component of 25 Marks and End Semester University Exam for 75 Marks.

Category A	Theory Courses with Lecture hours and hours allotted for Tutorials wherever required.
Category B	Practical Courses with only Practical hours or Laboratory hours. Laboratory Courses, Internships, Research Project Works and other courses allotted only with practical hours in the curriculum shall be under this category.
Category C	Theory & Practice combined Courses where Lecture and Practical hours allotted.

6.2. Learning Assessment

<table>
<tr><th>Course Types</th><th>Internal Assessment</th><th>End Semester Assessment</th></tr>
<tr><td>Category A

IA: 25 Marks
EA: 75 Marks</td><td>25 Marks

<table><tr><th>Evaluation Component</th><th>Marks</th></tr><tr><td>I. Mid Semester Exam (one)</td><td>20</td></tr><tr><td>II. Percentage of Attendance</td><td>05</td></tr><tr><td>Total</td><td>25</td></tr></table></td><td>75 Marks

(Evaluation Details given in Table 3)</td></tr>
<tr><td>Category B

IA: 50 Marks
EA: 50 Marks</td><td>50 Marks

For Practical / Internship Courses<table><tr><th>Evaluation Component</th><th>Marks</th></tr><tr><td>I. Weekly Observation Book / Report</td><td>15</td></tr><tr><td>II. Practical Record / Internship Report</td><td>15</td></tr><tr><td>III. Model Practical Exam</td><td>15</td></tr><tr><td>IV. Percentage of Attendance</td><td>05</td></tr><tr><td>Total</td><td>50</td></tr></table>For Research Project Work Course<table><tr><th>Evaluation Component</th><th>Marks</th></tr><tr><td>I. Monthly Review (3 Reviews – 10 Marks each)</td><td>30</td></tr><tr><td>II. Project Report</td><td>10</td></tr><tr><td>III. Project Work</td><td>10</td></tr><tr><td>Total</td><td>50</td></tr></table></td><td>50 Marks

(Evaluation Details given in Table 3)</td></tr>
<tr><td>Category C

IA: 25 Marks
EA: 75 Marks</td><td>25 Marks

<table><tr><th>Evaluation Component</th><th>Marks</th></tr><tr><td>I. Mid Semester Exam (one) - Theory</td><td>10</td></tr><tr><td>II. Observation Book, Record Book</td><td>10</td></tr><tr><td>III. Percentage of Attendance</td><td>05</td></tr><tr><td>Total</td><td>25</td></tr></table></td><td>75 Marks

(Evaluation Details given in Table 3)</td></tr>
</table>

6.3. Marks for Attendance

Attendance %	Marks
Below 75%	0
75% - 80%	1
81% - 85%	2
86% - 90%	3
91% - 95%	4
96% - 100%	5

6.4. Internal Test Scheme

Principal of the College schedules the Mid-Semester Exam for all courses during 8/9th week of start of classes. All faculty members are expected to conduct this Mid-Semester exam for 1½ hour duration and evaluate, upload the marks to Controller of Examinations of University. Colleges need to preserve the answer books of Mid-Semester exams until declaration of results by the University.

6.5. End Semester University Exam

Controller of Examinations (COE) of Pondicherry University schedules the End-Semester exams for all three categories of courses. **For Category C courses, theory and practical exams will be conducted separately by the Controller of Examinations of Pondicherry University.**

A detailed Exam Time Table shall be circulated to all Colleges atleast 15 days before the start of exams mostly during 15/16th week of the Semester. Question Papers shall be set externally based on BOS approved syllabus. All students who have a minimum of 70% attendance are eligible to attend the end-semester exams. The breakup of end semester marks is as given below.

6.6. Break up of End semester marks
(All End Semester Exams shall be conducted by the Pondicherry University)

The question paper shall be set as per the Bloom's Taxonomy. Table3 below gives the details of evaluation methods for Category A, B and C courses. Various levels along with their description and sample questions are as follows:

Knowledge: Recall or remember previously learned information.

Example: List the basic data types in Python

Comprehension: Demonstrate understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating the main ideas.

Example: Explain how a stack data structure works.

Application: Apply knowledge and concepts to solve problems in new situations. Use learned information in a different context.

Example: Write a Python program to solve the deadlock problem.

Analysis: Break down information into parts and examine the relationships between the parts. Identify motives or causes.

Example: Analyse the efficiency of two sorting algorithms and compare their advantages and disadvantages.

Synthesis: Create a new whole by combining elements in novel ways. Use creativity to produce something original.

Example: Design a web application that can generate a time table of a school.

Table 3: End Semester Assessment examination details for all three categories of courses

Course Components	Marks	Duration
Category A. Theory subjects: Sec A: 10 Questions of 2 Marks each (20 Marks) *(Knowledge: 3, Comprehension: 2, Application: 3, Analysis:2)* Sec B: 5 out of 7 Questions of 5 Marks each (25 Marks) *(Knowledge: 1, Comprehension: 2, Application: 1, Analysis:3)* Sec C: 2 Either/OR choice questions of 15 Marks each (30 Marks) *(Application: 2 Analysis:2* **Questions from all units of Syllabus equally distributed.**	75 Marks	3 Hours
Category B. Skill Enhancement / Practical Courses Based on Practical Exams conducted by COE of University **Internship / Research Project Work:** Presentation of the work / Report / Viva-voce examinations	50 Marks	3 Hours --
Category C. Theory Subjects with Practical Components: **i. Theory Component:** Sec A: 5 Questions of 2 Marks each (10 Marks) *(Knowledge: 3, Comprehension: 2, Application: 3, Analysis:2)* Sec B: 5 out of 7 Questions of 4 Marks each (20 Marks) *(Comprehension: 2, Application: 3, Analysis:2)* Sec C: 2 Either or type questions of 10 Marks each (20 Marks) *(Analysis / Synthesis)* **Questions from all units of Syllabus equally distributed.** **ii. Practical Component:** Based on Practical Exams conducted by COE of University **The examination shall be conducted for 50 Marks and reduced to 25 Marks.** **Total Marks: 75 (Theory: 50 Marks + Practical: 25 Marks)**	50 Marks 25 Marks	3 Hours 3 Hours

7. CONSOLIDATION OF MARKS, PASSING MINIMUM AND ARREAR EXAM

Controller of Examinations of the University consolidates the Internal Assessment marks uploaded by the Colleges and marks secured by students in end-semester examination. The total marks will be converted into letter grades as shown in the table below.

7.1. Passing Minimum

As per NEP Regulations, the passing minimum is 50% marks (IA + End semester put together). However, Pondicherry University considers 40% marks as pass during first 3 years of study and students who secured less than 50 will be awarded 'P' (Pass Grade).

7.2. Arrear Exam

A student who failed to secure 50% marks in aggregate is declared as Failed. Failed students are eligible to take up supplementary examination by registering to the failed course in the following Semester. All other candidates who failed due to shortage of attendance, those who are seeking to improve the grade shall repeat the course.

8. LETTER GRADES AND RANGE OF MARKS

Total Marks secured by a student in each subject shall be converted into a letter grade. UGC Framework has suggested a Country wide uniform letter grades for all UG courses.

8.1. Letter Grades

The following Table shows the seven letter grades and corresponding meaning and the grade points for calculation of CGPA.

Letter Grade	Grade Point
O (outstanding)	10
A+ (Excellent)	9
A (Very good)	8
B+ (Good)	7
B (Above average)	6
C (Average)	5
P (Pass)	4
F (Fail)	0
Ab (Absent)	0

In order to work out the above letter grades, the marks secured by a student (Total of Internal Assessment and End Semester Assessment) would be categorized for relative grading.

8.2. Range of Marks for each letter grades

The ranges of marks for each grade would be worked as follows:

Highest marks in the given subject	X
Cut of marks for grading purpose	50 Marks
Passing mark (for 3 year UG)	40 Marks
Number of grades G (Excl. P grade)	Grades: O, A+, A, B+, B, C, Hence, **G = 6**
Range of marks	K
K = (X – 50) / G	

The following table gives the range of marks and letter grades. According to K value, one of the following grading scheme will be followed.

(i) If K ≥ 5, then the grades shall be awarded as given in the following table.

Range of Marks in %	Letter Grade Points for	Grade Points for
X to (X-K) + 1	O	10
(X-K) to (X-2K) + 1	A+	9
(X-2K) to (X-3K) + 1	A	8
(X-3K) to (X-4K) + 1	B+	7
(X-4K) to (X-5K) + 1	B	6
(X-5K) to 50	C	5
40 – 49	P	4
Below 40	F	0
Absent (Lack of Attendance)	Ab	0

(ii) If K< 5, then the grades shall be awarded as given in the following table.

Range of Marks in %	Letter Grade Points for	Grade Points for
80-100	O	10
71-79	A+	9
66-70	A	8
61-65	B+	7
56-60	B	6
50-55	C	5
40-49	P	4
Below 40	F	0
Absent (lack of attendance)	Ab	0

9. CALCULATION OF SGPA AND CGPA

Semester Grade Point Average (SGPA) is calculated by taking a weighted average of all grade points secured by a candidate from all subjects registered by him/her in the given Semester. The weights being the number of credits that each subject carries.

Cumulative Grade Point Average (CGPA) shall be calculated as the weighted average of credits that course carries and the value of Grade points averaged for all subjects.

9.1. Procedure of computation of SGPA and CGPA

The following procedure shall be followed to compute the Semester Grade Point Average (SGPA) and Cumulative Grade Point Average (CGPA):

The SGPA is the ratio of the sum of the product of the number of credits with the grade points scored by a student in all the courses taken by a student and the sum of the number of credits of all the courses undergone by a student, i.e. $SGPA\ (Si) = \Sigma(Ci \times Gi) / \Sigma Ci$

Where Ci is the number of credits of the i^{th} course and Gi is the grade point scored by the student in the i^{th} course.

9.2. Example for Computation of SGPA where candidate has not failed in any course.

Semester	Course	Credit	Letter Grade	Grade Point	Credit Point (Credit x Grade)
I	Course 1	3	A	8	3 X 8 = 24
I	Course 2	4	B+	7	4 X 7 = 28
I	Course 3	3	B	6	3 X 6 = 18
I	Course 4	3	O	10	3 X 10 = 30
I	Course 5	3	C	5	3 X 5 = 15
I	Course 6	4	B	6	4 X 6 = 24
		20			139
	SGPA				139/20=6.95

9.3. Example for Computation of SGPA where candidate has failed in one course.

Semester	Course	Credit	Letter Grade	Grade Point	Credit Point (Credit x Grade)
I	Course 1	3	A	8	3 X 8 = 24
I	Course 2	4	B+	7	4 X 7 = 28
I	Course 3	3	B	6	3 X 6 = 18
I	Course 4	3	O	10	3 X 10 = 30
I	Course 5	3	C	5	3 X 5 = 15
I	Course 6	4	F	0	4 X 0 = 00
		20			115
	SGPA				115/20=5.75

9.4. Example for Computation of SGPA where candidate has failed in two courses.

Semester	Course	Credit	Letter Grade	Grade point	Credit Point (Credit x Grade)
I	Course 1	3	A	8	3 X 8 = 24
I	Course 2	4	B+	7	4 X 7 = 28
I	Course 3	3	F	0	3 X 0 = 00
I	Course 4	3	B	6	3 X 6 = 18
I	Course 5	3	C	5	3 X 5 = 15
I	Course 6	4	F	0	4 X 0 = 00
		20			85
	SGPA				85/20=4.25

The CGPA shall also be calculated in similar way as shown in examples (i), (ii) and (iii) of SGPA for all subjects taken by the students in all the semesters. However, if any student fails more than once in the same subject, then while calculating CGPA, the credit and grade point related to the subject in which the student fails in multiple attempts will be restricted to one time only. The SGPA and CGPA shall be rounded off to 2 decimal points and reported in the transcripts.

In case of audit courses offered, the students may be given (P) or (F) grade without any credits. This may be indicated in the mark sheet. Audit courses will not be considered towards the calculation of CGPA.

10. DECLARATION OF RESULTS

Controller of Examinations (COE) of the University shall declare the results of given UG programme following the CGPA secured by students by the end of 6th Semester and 8th Semester.

Pass Classes:

Range of CGPA	Result
9.0 -10.0	First Class with distinction
6.0 - 8.99	First Class
5.0 - 5.99	Second Class
4.0 - 4.99	Pass Class

11. MINIMUM CREDIT REQUIREMENTS

S.No.	Components	3-year UG			4-year UG (Honors / Honors With research)		
		Credits	Courses	Cr/Course	Credits	Courses	Cr/Course
1	Major Disciplinary/ Interdisciplinary Courses	56	14	4	76	19	4
2	Minor Disciplinary/ Interdisciplinary Courses	24	6	4	32	8	4
3	Multi-Disciplinary Courses	9	3	3	9	3	3
4	Ability Enhancement Courses	8	4	2	8	4	2
5	Skill Enhancement Courses	9	3	3	9	3	3
6	Value-added courses	8	4	2	8	4	2
7	Summer Internship (MJD 11)	4	1	4	4	1	4
8	Community Engagement and Service	2	1	2	2	1	2
9	Research Project/Dissertation	--	--	--	12	Project or 3 Courses##	
Total		**120**			**160**		

##**Note:** Honors students not undertaking research will do 3 courses for 12 credits in lieu of a research project/Dissertation.

- MJD: Major Disciplinary (Compulsory – Hardcore Subjects)
- MID: Minor Disciplinary (Specialization Specific – Softcore Subjects)
- MLD: Multi-Disciplinary
- AEC: Ability Enhancement Courses
- SEC: Skill Enhancement Courses
- VAC: Value Added Courses

12. COURSE CODE

- Course code : 7 Characters: 4 Alphabets and 3 Digits. Ex: ABCD123
- Alphabets : 1st and 2nd Alphabets: Major domain
  3rd and 4th Alphabets: Specialization
- Digits : 1st Digit: Levels (100, 200, 300, 400)
  2nd and 3rd Digits: Serial number of the courses in the given year

Example: CSDS312: Computer Science Data Science, Level - 300, Serial number of the course in the given year – 12.

CURRICULUM

FIRST SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 1	CSDS101	Digital Logic Fundamentals	H	4	3		2
2	MID 1	CSDS102	Foundations of Data Science-I	S	4	3		2
3	MLD 1		One course from the MLD streams (Table 10)	H	3	4		
4	AEC 1		English I / Modern Indian Languages I	H	2	4		
5	SEC 1	CSDS103/ CSDS104	S.No. 1 or 2 from Table 7	S	3	2		2
6	VAC 1		Understanding India	H	2	4		
7	VAC 2		Environmental Science/ Education, Higher Order Thinking	H	2	4		
				Total	20	30 Hours		

SECOND SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 2	CSDS105	Problem Solving & Programming Fundamentals	H	4	3		2
2	MID 2	CSDS106	Foundations of Data Science-II	S	4	3		2
3	MLD 2		One course from the MLD streams except the stream chosen in MLD1 (Table 10)	H	3	4		
4	AEC 2		English I / Modern Indian Languages I	H	2	4		
5	SEC 2	CSDS107/ CSDS108	S.No. 3 or 4 from Table 7	S	3	2		2
6	VAC 3		Health & Wellness/Yoga Education/Universal Human Values	H	2	2		2
7	VAC 4	COMS100	Digital Technologies	H	2	4		
				Total	20	30 Hours		

THIRD SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 3	CSDS201	Object Oriented Programming	H	4	3		2
2	MJD 4	CSDS202	Data Structures	H	4	3		2
3	MID 3	CSDS203	Probability and Statistics	S	4	3		2
4	MLD 3		One course from the MLD streams except the stream chosen in MLD1 and MLD2 (Table 10)	H	3	4		
5	AEC 3		English II / Modern Indian Languages II	H	2	4		

6	SEC 3	CSDS204/ CSDS205	S.No. 5 or 6 from Table 7	S	3	2		2
				Total	20	27 Hours		

FOURTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 5	CSDS206	Computer System Architecture	H	4	3		2
2	MJD 6	CSDS207	Design and Analysis of Algorithms	H	4	3		2
3	MJD 7	CSDS208	Database Management Systems	H	4	3		2
4	MID 4	CSDS209	Applied Regression Analysis	S	4	3		2
5	AEC 4		English II / Modern Indian Languages II	H	2	4		
6	CES 1	CSDS210	Community Engagement and Service	H	2			6
				Total	20	30 Hours		

FIFTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 8	CSDS301	Operating Systems	H	4	3		2
2	MJD 9	CSDS302	Mathematical Foundations of Computer Science	H	4	4	1	
3	MJD 10	CSDS303	Computer Networks	H	4	3		2
4	MID 5	CSDS304	Artificial Intelligence	S	4	3		2
5	MJD 11	CSDS305	Summer Internship	H	4			6
				Total	20	26 Hours		

SIXTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 12	CSDS306	Management Strategies and Concepts	H	4	5		
2	MJD 13	CSDS307	Software Engineering Theory and Practice	H	4	3		2
3	MJD 14	CSDS308	Distributed Systems	H	4	3		2
4	MJD 15	CSDS309	Operations Research	H	4	4	1	
5	MID 6	CSDS310/ CSDS311	Any one course from Table 1	S	4	3		2
				Total	20	25 Hours		

SEVENTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 16	CSDS401	Web Engineering	H	4	3		2

2	MJD 17	CSDS402	System Modelling & Simulation	H	4	3		2
3	MJD 18	CSDS403	Wireless Communication Networks	H	4	3		2
4	MID 7	CSDS404/ CSDS405	Any one course from Table 2	S	4	3		2
5	MID 8	CSDS406/ CSDS407	Any one course from Table 3	S	4	3		2
				Total	20	25 Hours		

EIGHTH SEMESTER – B.Sc. Data Science (Honors)								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 19	CSDS408/ CSDS409	Any one course from Table 4	S	4	3		2
2	MJD 20	CSDS410/ CSDS411	Any one course from Table 5	S	4	3		2
3	MJD 21	CSDS412	Deep Learning	H	4	3		2
4	MJD 22	CSDS413	Time Series Analysis	H	4	3		2
5	MJD 23	CSDS414	Natural Language Processing	H	4	3		2
				Total	20	25 Hours		

EIGHTH SEMESTER – B.Sc. Data Science (Honors with Research)								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 19	CSDS408/ CSDS409	Any one course from Table 4	S	4	3		2
2	MJD 20	CSDS410/ CSDS411	Any one course from Table 5	S	4	3		2
3	MJD 21	CSDS415	Research Project	H	4			5
4	MJD 22	CSDS416	Project Report	H	4			5
5	MJD 23	CSDS417	Project Viva-voce	H	4			5
				Total	20	25 Hours		

Table 1: MID 6 – SIXTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MID 6	CSDS310	Massive Data Management	S	4	3		2
2	MID 6	CSDS311	Hadoop Eco System	S	4	3		2

Table 2: MID 7 – SEVENTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P

1	MID 7	CSDS404	Big Data Analytics	S	4	3		2
2	MID 7	CSDS405	Predictive Analytics	S	4	3		2

Table 3: MID 8 – SEVENTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MID 8	CSDS406	Data Mining	S	4	3		2
2	MID 8	CSDS407	Text and Speech Analytics	S	4	3		2

Table 4: MJD 19 – EIGHTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 19	CSDS408	Machine Learning	S	4	3		2
2	MJD 19	CSDS409	Health Care Analytics	S	4	3		2

Table 5: MJD 20 – EIGHTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 20	CSDS410	Business Analytics	S	4	3		2
2	MJD 20	CSDS411	Social Network Analysis	S	4	3		2

Table 6: MJD 21 / MJD 22 / MJD 23 – EIGHTH SEMESTER								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	MJD 21	CSDS412	Deep Learning	S	4	3		2
2	MJD 22	CSDS413	Time Series Analysis	S	4	3		2
3	MJD 23	CSDS414	Natural Language Processing	S	4	3		2

Table 7: List of Skill Enhancement Courses/ SEC 1 / SEC 2 / SEC 3 – I / II / III SEMESTERS								
S.No	Component	Course Code	Title of the Course	H/S	Credits	Hours/Week		
						L	T	P
1	SEC 1	CSDS103	Python Programming	S	3	3		2
2	SEC 1	CSDS104	R Programming	S	3	3		2
3	SEC 2	CSDS107	Exploratory Data Analysis	S	3	3		2
4	SEC 2	CSDS108	Data Wrangling with R	S	3	3		2
5	SEC 3	CSDS204	Interactive Data Visualization	S	3	3		2
6	SEC 3	CSDS205	Financial Data Analytics	S	3	3		2

Table 8: List of Major Disciplinary Courses				
S.No	**Component**	**Course Code**	**Title of the Course**	**H/S**
1.	**MJD 1**	**CSDS101**	**Digital Logic Fundamentals**	**H**
2.	**MJD 2**	**CSDS105**	**Problem Solving & Programming Fundamentals**	**H**
3.	**MJD 3**	**CSDS201**	**Object Oriented Programming**	**H**
4.	**MJD 4**	**CSDS202**	**Data Structures**	**H**
5.	**MJD 5**	**CSDS206**	**Computer System Architecture**	**H**
6.	**MJD 6**	**CSDS207**	**Design and Analysis of Algorithms**	**H**
7.	**MJD 7**	**CSDS208**	**Database Management Systems**	**H**
8.	**MJD 8**	**CSDS301**	**Operating Systems**	**H**
9.	**MJD 9**	**CSDS302**	**Mathematical Foundations of Computer Science**	**H**
10.	**MJD 10**	**CSDS303**	**Computer Networks**	**H**
11.	**MJD 11**	**CSDS305**	**Summer Internship**	**H**
12.	**MJD 12**	**CSDS306**	**Management Strategies and Concepts**	**H**
13.	**MJD 13**	**CSDS307**	**Software Engineering Theory and Practice**	**H**
14.	**MJD 14**	**CSDS308**	**Distributed Systems**	**H**
15.	**MJD 15**	**CSDS309**	**Operations Research**	**H**
16.	**MJD 16**	**CSDS401**	**Web Engineering**	**H**
17.	**MJD 17**	**CSDS402**	**System Modeling and Simulation**	**H**
18.	**MJD 18**	**CSDS403**	**Wireless Communication Network**	**H**
19.	**MJD 19**	**CSDS408 / CSDS409**	**Machine Learning / Health Care Analytics**	**S**
20.	**MJD 20**	**CSDS410 / CSDS411**	**Business Analytics / Social Network Analysis**	**S**

Table 9: List of Minor Disciplinary Courses				
S.No	**Component**	**Course Code**	**Title of the Course**	**H/S**
1.	**MID 1**	**CSDS102**	**Foundations of Data Science - I**	**S**
2.	**MID 2**	**CSDS106**	**Foundations of Data Science - II**	**S**
3.	**MID 3**	**CSDS203**	**Probability and Statistics**	**S**
4.	**MID 4**	**CSDS209**	**Applied Regression Analysis**	**S**
5.	**MID 5**	**CSDS304**	**Artificial Intelligence**	**S**
6.	**MID 6**	**CSDS310 / CSDS311**	**Massive Data Management / Hadoop Eco System**	**S**
7.	**MID 7**	**CSDS404/ CSDS405**	**Big Data Analytics / Predictive Analytics**	**S**
8.	**MID 8**	**CSDS406 / CSDS407**	**Data Mining / Text and Speech Analytics**	**S**

*Table 10: MLD 1 / MLD 2 / MLD 3 in Sem 1 / Sem 2 / Sem 3			
Streams	**Course Code**	**Title of the Course**	**H/ S**
Natural Science		Biology	**H**
		Botany	**H**
		Zoology	**H**
		Biotechnology	**H**
		Biochemistry	**H**
Physical Sciences		Chemistry	**H**
		Physics	**H**
		Biophysics	**H**
		Astronomy	**H**
		Astrophysics	**H**
		Earth and Environmental Sciences	**H**
Mathematics & Statistics		STATA	**H**
		SPSS	**H**
		Tally	**H**
Computer Science	**COMS101**	Introduction to Python Programming	**H**
	COMS102	Foundations of Information Technology	**H**
Social Sciences		Political Sciences	**H**
		History	**H**
		Social work	**H**
		Sociology	**H**
Humanities		Anthropology	**H**
		Psychology	**H**
		Economics	**H**
Commerce & Management		Business Management	**H**
		Accountancy	**H**
		Finance	**H**
		Financial Institutions	**H**
Media Sciences		Journalism	**H**
		Mass Media	**H**
		Communication	**H**

Table 11: List of Ability Enhancement Courses				
S.No	**Component**	**Course Code**	**Title of the Course**	**H/S**
1.	**AEC 1**		**English I / Modern Indian Languages I**	**H**
2.	**AEC 2**		**English I / Modern Indian Languages I**	**H**
3.	**AEC 3**		**English II / Modern Indian Languages II**	**H**
4.	**AEC 4**		**English II / Modern Indian Languages II**	**H**

Table 12: List of Skill Enhancement Courses				
S.No	**Component**	**Course Code**	**Title of the Course**	**H/S**
1	**SEC 1**	**CSDS103**	**Python Programming**	**S**
2	**SEC 1**	**CSDS104**	**R Programming**	**S**
3	**SEC 2**	**CSDS107**	**Exploratory Data Analysis**	**S**
4	**SEC 2**	**CSDS108**	**Data Wrangling with R**	**S**
5	**SEC 3**	**CSDS204**	**Interactive Data Visualization**	**S**
6	**SEC 3**	**CSDS205**	**Financial Data Analytics**	**S**

Table 13: List of Value-Added Courses				
S.No	**Component**	**Course Code**	**Title of the Course**	**H/S**
1.	**VAC 1**		**Understanding India**	**H**
2.	**VAC 2**		**Environmental Science/ Education, Higher Order Thinking**	**H**
3.	**VAC 3**		**Health & Wellness / Yoga Education / Universal Human Values**	**H**
4.	**VAC 4**	**COMS100**	**Digital Technologies**	**H**

Table 14: Project (WP/ Internship)				
S.No	**Component**	**Course Code**	**Title of the Course**	**H/S**
1.	**CES 1**	**CSDS210**	**Community Engagement and Service**	**H**

SYLLABUS

SEMESTER I

Year	I	Course Code: CSDS101 Course Title: Digital Logic Fundamentals	Credits	4
Sem.	I		Hours	75
			Category	C
Course Prerequisites, if any	NIL			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Understand and describe the principles of digital systems and binary number operations • Apply Karnaugh mapping to simplify Boolean expressions and optimize digital circuits • Analyze and design basic combinational circuits using various digital components • Synthesize and evaluate synchronous sequential circuits using storage elements and HDL • Design and implement various types of registers and counters using HDL			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Digital Systems and Binary Numbers** Digital Systems – Binary Numbers – Number-Base Conversions – Octal and Hexadecimal Numbers – Complements of Numbers – Signed Binary Numbers – Binary Codes – Binary Storage and Registers – Binary Logic – Axiomatic Definition of Boolean Algebra – Basic Theorems and Properties of Boolean Algebra – Boolean Functions – Canonical and Standard Forms – Other Logic Operations – Digital Logic Gates – Integrated Circuits	9
Unit II	**Gate-Level Minimization** Introduction – The Map Method – Four-Variable K-Map – Product-of- Sums Simplification – Don't-Care Conditions – NAND and NOR Implementation – Other Two-Level Implementations – Exclusive-OR Function – Hardware Description Language	9
Unit III	**Combinational Logic** Introduction – Combinational Circuits – Analysis Procedure – Design Procedure – Binary Adder–Subtractor – Decimal Adder – Binary Multiplier – Magnitude Comparator – Decoders – Encoders – Multiplexers – HDL Models of Combinational Circuits	9
Unit IV	**Synchronous Sequential Logic** Introduction – Sequential Circuits – Storage Elements – Latches – Flip-Flops – Analysis of Clocked Sequential Circuits - Synthesizable HDL Models of Sequential Circuits – State Reduction and Assignment – Design Procedure	9
Unit V	**Registers and Counters** Registers – Shift Registers – Ripple Counters – Synchronous Counters – Other Counters – HDL for Registers and Counters	9

Practical Component		
Exercises	1. Binary to Decimal and vice-versa 2. Decimal to Hexadecimal and Vice-Versa 3. Digital Logic Gates 4. Simplification of Boolean Functions 5. Combinational Logic Circuits i. Code Converters ii. Arithmetic (Adders, Subtractors,Multipliers, Comparators) iii. Data Handling (Multiplexers, Demultiplexers, Encoders & Decoders) 6. Combinational Logic Circuit Design 7. Binary Adder-Subtractor Simulation 8. Decimal Adder Simulation 9. Binary Multiplier Simulation 10. Sequential Circuit Storage Elements: Flip-Flop Simulation	30
Recommended Learning Resources		
Print Resources	1. M. Morris Mano, Michael D. Ciletti, Digital design with an Introduction to the Verilog HDL, Pearson, 6th Edition, 2018. 2. M. Rafiquzzaman, Fundamentals of Digital Logic and Microcomputer Design, John Wiley & Sons, Inc. 5th Edition, 2009	
Syllabus Design: Dr. M.Sathya, Assistant Professor, PUDoCS		

Year	I	Course Code: CSDS102 Course Title : Foundations of Data Science - I	Credits	4
Sem.	I		Hours	75
			Category	C
Course Prerequisites, if any	Basic problem solving skills			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Comprehend the fundamental concepts of data science. • Solve problems in Probability and Statistics • Solve fundamental problems in Matrices and comprehend the application of matrices • Use databases for structured and unstructured data • Summarize the steps in the Data Science life cycle			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Need for Data Science – Data Science Process – Business Intelligence and Data Science – Prerequisites for a Data Scientist. Exploratory Data Analysis – Data Format, Types of EDA – Univariate, multivariate non-graphical, graphical EDA	9
Unit II	**Probability and Statistics** Understanding Probability – Math – Binomial distribution Data – descriptive vs Inferential – Population , Sample, Bias – Descriptive statistics – Inferential Statistics – T Distributions	9
Unit III	**Linear Algebra** Vector – Linear Transformations – Matrix multiplication – Determinants – Types of matrices – system of equation and inverse matrices – Eigen vector , Eigen values	9
Unit IV	**Databases for Data Science** Structured Query Language (SQL): Data Munging, Filtering, Joins, Aggregation, Window Functions, Ordered Data, No-SQL, Document Databases, Wide-column Databases and Graphical Databases.	9
Unit V	**Data Science Methodology** Analytics for Data Science – Examples of Data Analytics – Data Analytics Lifecycle: Data Discovery, Data Preparation, Model Planning, Model Building.	9
	Practical Component	
Exercises	1. Download, install NumPy, SciPy and pandas in Python. 2. Build a data frame using pandas from a csv file. 3. Write a program for finding the frequency, Mean, Median, Mode, Variance, and Standard Deviation of data using Python pandas data-frame. 4. Plot a graph for probability distribution using Python (Normal Distribution). 5. Perform data analysis using SciPy . 6. Create a database and establish relationships between tables. 7. Create view to extract details from two or more tables.	30

Recommended Learning Resources	
Print Resources	1. Sanjeev Wagh, Manisha Bhende, Anuradha Thakare, Fundamentals of Data Science , CRC Press, 1st Edition, 2022. 2. Howard J.Seltman, Experimental Design and Analysis , CMU, 2018. 3. Thomas Nield, Essential Math for Data Science , O'Reilly Media Inc., 2022.
Syllabus Design: Dr.R.Sunitha, Associate Professor, PUDoCS	

Year	I	**Course Code: CSDS103** **Course Title : Python Programming**	**Credits**	3
Sem.	I		**Hours**	60
			Category	B
Course Prerequisites, if any	Basic problem solving skills			
Internal Assessment Marks: 50	End Semester Marks: 50		Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Understand the basics of writing Python code • Implement programs using lists, tuples and dictionaries • Understand the use of control structures • Ability to write programs using packages • Understand file manipulation			

Unit No.	**Course Content**	**Hours**
	Theory Component	
Unit I	**Introduction** Introduction to Python – Advantages of using Python – Executing Python Programs – Python's Core data types – Numeric Types – String Fundamentals.	6
Unit II	**Lists, Tuples, Dictionaries** Lists: list operations, list slices, list methods, list loop, mutability, aliasing, cloning lists, list parameters; Tuples: tuple assignment, tuple as return value; Dictionaries: operations and methods; advanced list processing – list comprehension.	6
Unit III	**Control Flow, Functions, Modules** Python Statements: Assignments – Expressions – If condition – While and For Loops. Functions: Definition, Calls – Scopes – Arguments – Recursive Functions– Functional Programming tools. Classes and Object Oriented programming with Python Modules and Packages: Purpose, using packages– Exception Handling with Python.	6
Unit IV	**Packages** Packages: NumPy, Pandas, Scikit learn Machine learning with Python – Cleaning up, Wrangling, Analysis Visualization - Matplotlib package – Plotting Graphs.	6
Unit V	**File Handling** Files and exception: text files, reading and writing files, format operator; command line arguments, errors and exceptions, handling exceptions	6

Practical Component		
Exercises	1. Exchange the values of two variables 2. Finding minimum among n variables 3. Perform Simple sorting 4. Generate Students marks statement 5. Find square root, GCD, exponentiation 6. Sum the array of numbers 7. Perform linear search, binary search 8. Perform Matrix operations using NumPy 9. Perform Dataframe operations using Pandas 10. Use Matplotlib on dataset and visualise 11. Perform Word count, copy file operations	30
Recommended Learning Resources		
Print Resources	1. Mark Lutz, Learning Python , O'Reilly, 5th Edition, 2013. 2. Daniel Liang, Introduction to Programming using Python , Pearson, 1st Edition, 2021. 3. Wes Mc Kinney, Python for Data Analysis , O'Reilly Media, 2012. 4. Tim Hall and J-P Stacey, Python 3 for Absolute Beginners , Apress, 1st Edition, 2009. 5. Magnus Lie Hetland, Beginning Python: From Novice to Professional , Apress, 2nd Edition, 2005.	
Syllabus Design: Dr. Uma, Associate Professor, PUDoCS		

Year	I	**Course Code: CSDS104** **Course Title : R Programming**	**Credits**	3
Sem.	I		**Hours**	60
			Category	**B**
Course Prerequisites, if any	Basic problem solving skills			
Internal Assessment Marks : 50	End Semester Marks : 50		Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Demonstrate proficiency in R programming basics, including installation, opening, saving, and editing R code. • Perform basic math operations, assign objects, and manipulate vectors in R. • Analyze and manipulate matrices and arrays. • Create and manipulate lists and data frames. • Read and write files in R, handling R-ready datasets and external data files.			

Unit No.	**Course Content**	**Hours**
	Theory Component	
Unit I	**Introduction** R Installation – opening – Saving and Editing – Conventions **Number, Arithmetic, assignment & Vectors** R for Basic Math – Assigning Objects – Vectors.	6
Unit II	**Matrices and Arrays** Defining a Matrix – Subsetting – Matrix Operations & Algebra – Multidimensional Arrays.	6
Unit III	**Non-numeric Values :** Logical Values – Characters - Factors **Lists and Data Frames:** Objects – Data Frames.	6
Unit IV	**Special Values , Classes, and Coercion** Some special values – Understanding Types, classes and Coercion **Basic Plotting** Using Plot with coordinate Vectors – Graphical Parameters – Adding Plots, lines, and Text – ggplot2.	6
Unit V	**Reading and Writing Files** R-Ready Data sets – Reading in External data files – Writing out Data files and Plots – Adhoc Object R/W.	6

<table>
<tr><th colspan="3">Practical Component</th></tr>
<tr><td>Exercises</td><td>1. Practice Installing , opening and saving files in R.
2. Create and store a vector that contains
• A sequence of integers.
• A threefold repetition of a real valu
• Numbers divisible by 2.
3. Create a matrix and find the number of entries in each row which are greater than ‘n’
4. Write a program to Add, Multiply two matrices
5. Write a program to transpose and find the inverse of a matrix.
6. Store a vector with 15 values as an object. Identify those equal to 6, those greater than or equal to 6, those less than 6 + 2, those not equal to 6
7. With the Weight (kg), height (cm) and Sex data of 10 students, create a plot of weight on the x-axis and height on the y-axis. Use different point characters or colors to distinguish between males and females and provide a matching legend. Label the axes and give the plot a title.
8. Demonstrate Visualization using ggplot2</td><td>30</td></tr>
<tr><th colspan="3">Recommended Learning Resources</th></tr>
<tr><td>Print Resources</td><td colspan="2">1. Tilman M.Davies, The Book of R: A First Course in Programming and Statistics , No Starch press, 2016.
2. Bradley C. Boehmke, Data wrangling with R , Springer, 2016.
3. Andrea de Vries, Joris Meys, R programming for Dummies , 2nd Edition, Wiley, 2016.</td></tr>
<tr><td colspan="3">Syllabus Design: Dr.R.Sunitha, Associate Professor, PUDoCS</td></tr>
</table>

SEMESTER II

Year	I	Course Code: CSDS105 Course Title: Problem Solving & Programming Fundamentals	Credits	4
Sem.	II		Hours	75
			Category	C
Course Prerequisites, if any	NIL			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Analyze problems and develop top-down designs. • Write, compile, and debug basic programs. • Implement logic with conditionals and loops. • Manipulate arrays of various dimensions. • Design and implement functions with recursion.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction to Computer Problem-Solving** Problem-solving Aspect – Top-down Design – Implementationof Algorithms – Program Verification – Efficiency of Algorithms – Analysis of Algorithms	9
Unit II	**Basic Programming Constructs** Basic Data types (Numerical, String) – Variables – Expressions – I/O statements – Compile and Run – Debugging	9
Unit III	**Decision Making – Branching & Looping** Decision making – Relational Operators – Conditional statement, Looping Statements – Nested loops – Infinite loops – Switch Statements	9
Unit IV	**Array Techniques** Array Manipulation – Different operations – One dimensional Array – Two-dimensional Array – Multi-dimensional Array – Character – Arrays and Strings	9
Unit V	**Modular solutions** Introduction to Functions – Importance of Design of Functions – Arguments – Parameters – Return Values – Local and Global Scope – Recursion	9
	Practical Component	
Exercises	1. Program to array counting, array order reversal & find the maximum number in a set. 2. Program for removal of duplicates from an ordered array and to partition an array. 3. Program to find the k^{th} smallest element. 4. Program to exchange the values of two variables without using a third variable. 5. Program that takes a list of numbers as input and counts the total number of elements in the list. 6. Program to compute the factorial of a given integer. 7. Program to compute the sine of an angle (in degrees) using a series expansion. 9. Program to generate the Fibonacci sequence up to aspecified limit. 10. Program that converts a number from one base to another	30

Recommended Learning Resources	
Print Resources	1. R. G. Dromey, How to Solve it by Computer, Pearson Education India, 13th Edition, 2013. 2. Allen B. Downey, Think Python: How to Think like a Computer Scientist, O'Reilly Publishers, 3rd Edition, 2020.
Syllabus Design: Dr. M.Sathya, Assistant Professor, PUDoCS	

Year	I	Course Code: CSDS106 Course Title : Foundations of Data Science - II	Credits	4
Sem.	II		Hours	75
			Category	C
Course Prerequisites, if any	Foundations of Data Science - I			
Internal Assessment Marks : 25	End Semester Marks : 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Understand the concepts of regression analysis • Explore Data Science Tools • Explain the fundamental machine learning algorithms • Apply Data Analytics in Text Mining • Perform Data Visualization			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Regression Analysis** Linear – Logistic – Multinomial logistic regression – Time Series Models	9
Unit II	**Machine Learning** Introduction to Decision Trees- Naïve Bayes- Support Vector Machines- Nearest Neighbour learning- Clustering- Confusion Matrix.	9
Unit III	**Data Analytics on Text** Text Mining - Text Analytics - Natural Language Processing- Major Components - Stages - Statistical Processing - Applications	9
Unit IV	**Data Science Tools** Python : Basics, Library – R : Reading and Getting Data – MATLAB : Work flow , Importing Data	9
Unit V	**Data Visualization using Tableu** Introduction - Dimensions, measures, descriptive statistics - basic charts - Dashboard Design and principles - Special charts – Integrate Tableau with Google sheet.	9
	Practical Component	
Exercises	1. Program to calculate regression coefficient 2. Program to count word frequency 3. Install NLTK library and perform simple text processing and analysis tasks 4. Program to process the text (Identifying stop words, Stemming, Lemmatizing). 5. Practice plotting different charts using Tableu	30
	Recommended Learning Resources	
Print Resources	1. Sanjeev Wagh, Manisha Bhende, Anuradha Thakare, Fundamentals of Data Science, CRC Press, 1st Edition, 2022.	
Syllabus Design: Dr.R.Sunitha, Associate Professor, PUDoCS		

Year	I	**Course Code: CSDS107** **Course Title : Exploratory Data Analysis**	**Credits**	3
Sem.	II		**Hours**	60
			Category	B
Course Prerequisites, if any	Python Programming			
Internal Assessment Marks : 50	End Semester Marks : 50	Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Perform data loading, transformation, and preliminary analysis for real-world data • Create charts and graphs to effectively communicate and interpret patterns in data during Exploratory Data Analysis. • Apply advanced statistical measures to describe and interpret datasets, including measures of central tendency and dispersion • Critically evaluate and draw meaningful conclusions from the analysis results. • Demonstrate proficiency in handling time series datasets and performing Time Series Analysis (TSA) using Python.			

Unit No.	**Course Content**	**Hours**
	Theory Component	
Unit I	**Introduction** Understanding Data Science – Significance of EDA – Making sense of Data – Comparing EDA with classical and Bayesian analysis – software tools.	6
Unit II	**Visual aids for EDA** Line – Bar charts – Scatter Plot – Area Plot – Pie – Table – Polar chart – Histogram – Lollipop **EDA with Personal Email** Technical requirements –Loading – Transformation -Data Analysis	6
Unit III	**Data Transformation** Managing Database – Transformation Techniques – Benefits **Descriptive Statistics** Understanding statistics – Measures of central tendency – Measures of dispersion.	6
Unit IV	**Grouping Datasets** Understanding groupby() – Groupby mechanics – Data aggregation – Pivot tables – Cross-tabulations. **Correlation** Understanding correlation – Types of analysis – Multivariate analysis using Titanic dataset.	6
Unit V	**Model Development and Evaluation** Hypothesis Testing and Regression, Model Development and Evaluation **EDA on Wine Quality Data Analysis.** Disclosing – Red Wine Analysis	6

Practical Component		
Exercises	1. Download, Install and practice opensource tools for EDA – Eg. WEKA. 2. Visualize the data using various graphs . 3. Perform histogram analysis using NumPy, Matplotlib, pandas. 4. Program to generate different charts and plots. 5. Program to generate pivot using groupby() method. 6. Perform Time Series analysis and test with with a predictive model. 7. Program to identify the correlation of the features/parameters in the Titanic Dataset. 8. Perform EDA on Wine Data . 9. Demonstrate different visualizations based on Exercise 7.	30
Recommended Learning Resources		
Print Resources	1. Suresh Kumar Mukhiya, Usman Ahmed, Hands-On Exploratory Data Analysis with Python , PACKT Publishing , 2020.	
Syllabus Design: Dr.R.Sunitha, Associate Professor, PUDoCS		

Year	I	**Course Code: CSDS108** **Course Title : Data Wrangling with R**	**Credits**	3
Sem.	II		**Hours**	60
			Category	B
Course Prerequisites, if any	Foundations of Data Science, R programming			
Internal Assessment Marks : 50	End Semester Marks : 50		Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Demonstrate the ability to write and execute R code efficiently for data manipulation. • Apply data wrangling skills to various datasets, understanding the data generation process, interpreting different data types, and effectively using data to address analytical queries. • Utilize core functions of dplyr for efficient data manipulation, sequential operations, grouping, and joining of data frames • Access and integrate data from databases and web APIs using R • Design and create interactive visualization with ggplot2			

Unit No.	**Course Content**	**Hours**
	Theory Component	
Unit I	**Introduction** Setting the computer – Command line **Managing Projects** Version Control – Markdown	6
Unit II	**R Fundamentals** Introduction – Functions – Vectors – Lists	6
Unit III	**Data Wrangling** Understanding Data - Data Frames - Manipulating Data With Dplyr - Reshaping Data with Tidyr	6
Unit IV	**Data Wrangling** Accessing Databases – Accessing Web APIs	6
Unit V	**Data Visualization** Designing Data Visualizations - Creating Visualizations with ggplot2 – Interactive Visualization in R	6
	Practical Component	
Exercises	1. Practice version control using GitHub 2. Perform operations using Dataframes 3. Perform sequential operations using dplyr 4. Practice reshaping educational statistics using tidyr 5. Download, install, practice Plotly, Rbokeh, Leaflet Package 6. Perform interactive visualization Using Seattle dataset 7. Demonstrate visualization using Plotly, Rbokeh, and Leaflet Packages	30

Recommended Learning Resources	
Print Resources	1. Michael Freeman and Joel Ross, Programming Skills for Data Science: Start Writing code to Wrangle, Analyze, and Visualize Data with R , Addison-Wesley, 2018.
Syllabus Design: Dr.R.Sunitha, Associate Professor, PUDoCS	

VALUE ADDED COURSES

Year	I	Course Code: COMS100 Course Title: Digital Technologies	Credits	2
Sem.	II		Hours	45
			Category	C
Course Prerequisites, if any	Nil			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Get introduced to the digital systems and its building blocks. • Understand how the Digital Communication happens and to Learn the advantages and disadvantages including Cybersecurity. • Learn the day to day digital activities and the initiatives on Digital India. • Acquire knowledge on current Technologies and Trends in Digital Space. • Explore the applications on the state of the art in Digital Technologies.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction**: Digital Systems - Information & Communication Technology - ICT Tools Computer Architecture – Software – Hardware - Operating System - Algorithms - Flowcharts.	7
Unit II	**Communication Systems:** Transmission Media - Computer Networks – Internet - Web Browsers - Search Engines - Messaging, Email - Social Media – Online Ethics. **Cybersecurity:** Threats, Significance, Challenges, Precautions, Safety Measures. Cyber Crime Awareness.	7
Unit III	**Digital India & e-Governance:** Initiatives - Unified Payment Interface - Aadhar online services - Credit / Debit Cards - e-Wallets – Mobile and Internet Banking – NEFT / RTGS / IMPS - Online Payments & PoS – Digital Accessibility	7
Unit IV	**Emerging Technologies & Applications: (Basic introduction only).** Overview of Artificial Intelligence, Cloud Computing, Big Data, Internet of Things, Virtual Reality, 5G, 3D Printing.	7
Unit V	**Case Studies:** Any one case study on the emerging technologies and report submission by the candidates.	7
	Practical Component	
Practice	1. Operating System Installation and configuration 2. Application Software Installation and configuration 3. Hardware understanding and minor troubleshooting 4. Networking, cabling, configuration	10

Recommended Learning Resources	
Print Resources	1. Pramod Kumar, Anuradha Tomar, R. Sharmila, Emerging Technologies in Computing - Theory, Practice, and Advances , Chapman and Hall / CRC, 1st Edition, 2021, 2. V Rajaraman, Introduction to Information Technology , PHI, 3rd Edition, 2018. 3. E Balagurusamy, Fundamentals of Computers , Tata Mc GrawHill, 2nd Edition, 2011. 4. Behrouz A. Forouzan, Data Communications and Networking , McGraw Hill, 4th Edition, 2007. 5. Rajkumar Buvya, James Broberg, and Andrzej Gosciniski, Cloud Computing- Principals and Paradigms , Wiley, 2011. 6. Stuart Russel and Peter Norvig, Artificial Intelligence - A Modern Approach , Pearson Education, 3rd Edition, 2010. 7. Samuel Greengard, Internet of Things , MIT Press, 2015. 8. C.S.V. Murthy, E- Commerce (Concept - Models - Strategies) , Himalaya Publishing House, 2015. 9. Hurwith, Nugent Halper, Kaufman, Big Data for Dummies , Wiley & Sons, 1st Edition, 2013.
Syllabus Design: Dr.S.K.V.Jayakumar, Professor, PUDoCS	

SEMESTER III

Year	II	Course Code: CSDS201 Course Title: Object-Oriented Programming	Credits	4
Sem.	III		Hours	75
			Category	C
Course Prerequisites, if any	Basic Programming knowledge			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	● Understand the principles of OOP and the concept of class and objects. ● Apply the concept of Object initialization and overloading. ● Understand the concept of inheritance and reusability. ● Understand file operations and exception handling. • Apply OOP to design and implement solutions to real-world problems.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Principles of Object Oriented Programming (OOP)** Object Oriented Programming Paradigm-Basic Concepts of OOP-Benefits of OOP - Application of OOP - Simple C++ program - Compiling and Linking	9
Unit II	**Classes and Objects** Specifying class - Member functions - Nesting of Member functions - Access specifier - Static Data members and functions - Arrays within a Class - Arrays of Objects - Objects as Arguments - Returning Objects - Friend Function	9
Unit III	**Object Initialization and Overloading** Types of Constructors - Dynamic Initialization of Objects - Destructors Operator overloading - function Overloading - Manipulation of Strings	9
Unit IV	**Inheritance** Derived Classes - Types of inheritance - Virtual Base Classes - Abstract Classes - Pointers to Derived Classes - Virtual base class - Method Overriding - Pure Virtual Functions	9
Unit V	**File operations and Exception handling** Classes for File Operations - File Modes - opening and closing a File - Basics of Exception Handling - Try-Catch block - Case Studies on Real Time Applications	9
	Practical Component	
Exercises	1. Program to Read and Print Number Input from the User. 2. Program using a class and objects 3. Program to demonstrate the usage of a constructor and destructor in a class 4. Program to overload + operator to add two complex numbers. 5. Program to demonstrate the usage of function overloading. 6. Program to display employee information using multiple inheritance. 7. Program to demonstrate multilevel inheritance. 8. Program to copy a file from one location to another location.	30

Recommended Learning Resources	
Print Resources	1. E Balagurusamy, Object Oriented Programming with C++ , Tata McGraw Hill, 7th Edition, 2020.
Syllabus Design: Dr. T. Vengattaraman, Associate Professor, PUDoCS	

Year	II	Course Code: CSDS202 Course Title: Data Structures	Credits	4
Sem.	III		Hours	75
			Category	C
Course Prerequisites, if any	Introductory knowledge about Computing			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Learn basic terminologies of linear and nonlinear data structures and algorithms • Understand the concept of polynomial addition and sparse matrices using arrays • Apply linked lists to solve problems related to stacks, queues, and sparse matrices • Understand the operations and traversals of binary trees • Apply graph algorithms to solve problems like topological sorting and finding minimum cost spanning trees			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Basic terminologies – Linear and Nonlinear data structures – Algorithm: Definition – Pseudo code – Analysis – Design Techniques	7
Unit II	**Arrays, Stacks and Queues** Representation – Polynomial Addition – Sparse Matrices – Multidimensional Arrays – Stack ADT – Operations – Evaluation of Expressions – Queue ADT – Operations – Application – Multiple Stacks and Queues	11
Unit III	**Lists** Singly Linked Lists – Linked Stacks and Queues – Operations – Circularly Linked Lists – Equivalence Relations – Sparse Matrices – Doubly Linked Lists	9
Unit IV	**Trees** Basic Terminologies – Binary trees – Representation, Operations, Traversals, Types – Applications of Trees	9
Unit V	**Graphs** Basic Terminologies – Representation, Operations, Traversals – Applications - Shortest path problem, Topological sorting, Minimum Cost Spanning trees	9
	Practical Component	
Exercises	1. Searching Algorithms - Sequential, Binary and Fibonacci search algorithms 2. Evaluation of arithmetic expressions 3. Stack, Queue, Circular queue, priority queue 4. Singly Linked List, Doubly Linked List, Circular Linked List 5. Tree Traversal techniques 6. Graph Traversal techniques 7. Dijkstra's Algorithm to obtain the shortest paths	30

Recommended Learning Resources		
Print Resources	1. Ellis Horowitz, Sartaj Sahni, Susan Anderson-Freed, Fundamentals of Data Structures in C, 2nd Edition, 2008. 2. Debasis Samanta, Classic Data Structures, Prentice-Hall of India, 2nd Edition, 2009. 3. Dinesh P Mehta & Sartaj Sahni, Handbook of Data Structures and Applications, Chapman and Hall, 2nd Edition, 2020.	
Syllabus Design Dr M.Sathya, Assistant Professor, PUDoCS		

Year	II	Course Code: CSDS203 Course Title : Probability and Statistics	Credits	4
Sem.	III		Hours	75
			Category	C
Course Prerequisites, if any	NA			
Internal Assessment Marks : 25	End Semester Marks : 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Understand the concepts of probability, Random variables • Understand the concepts of mean, variance • Apply correlation and apply in Data Science • Understand the importance of statistics in data science • Understand the concept of regression			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Probability** Introduction - Sample Space - Events - Counting Sample Points - Probability of Event - Additive Rules - Conditional Probability - Independence and Product Rule - Bayes Rule	9
Unit II	**Random Variables** Concept - Discrete probability distributions - Continuous probability distributions - Joint Probability distributions - Mean - Variance - Covariance - Chebyshev's Theorem	9
Unit III	**Probability Distributions** Binominal and multinomial distributions - Hypergeometric distributions - Negative binomial and geometric distributions - Poisson Distribution - Continuous Uniform distribution -Normal distribution- Chi-squared distribution	9
Unit IV	**Estimation Problems** Statistical Inference - Classical Methods of Estimation - Estimating the Mean - Standard Mean - Prediction Intervals - Tolerance Limits - Paired observations - Proportion and variance Estimation - Maximum Likelihood Estimation - Test of Hypotheses	9
Unit V	**Regression and Correlation** Linear Regression - Least squares and Fitted model - Estimators - Inferences - Prediction - Analysis of Variance Approach – Data Plots and Transformation - Correlation	9
	Practical Component	
Exercises	1. Implement Bayes' Theorem 2. Simulate random and Continuous variable 3. Implement Binomial distribution 4. Implement Poisson Distribution 5. Implement Correlation 6. Implement ANOVA Test	30

Recommended Learning Resources	
Print Resources	1. R.E.Walpole, R.H.Myers, S.L.Myers, Keying Ye, Probability and Statistics for Engineers and Scientists, Prentice Hall, 9th Edition, 2012 2. Hogg, R.V. , Mc Kean J W and Craig, A.T.,Introduction to Mathematical Statistics, Pearson, 6th Edition, 2021
Syllabus Design Dr. P.Shanthi Bala, Professor, PUDoCS	

Year	II	**Course Code: CSDS204** **Course Title : Interactive Data Visualization**	**Credits**	3
Sem.	III		**Hours**	60
			Category	B
Course Prerequisites, if any	Foundations of Data Science			
Internal Assessment Marks : 50	End Semester Marks : 50		Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	**Students will be able to** • Understand the concepts of Data visualization • Understand the process of building effective visualization • Apply the Python Libraries to build visualization • Analyze various visualization techniques • Create effective visualization			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Definition - History - Relationship between visualization and other fields - Visualization Process - Role of cognition - Scatterplot - Visualization Foundations	6
Unit II	**Spatial Visualization Techniques** Introduction - Visualization Techniques for Spatial Data - One Dimensional Data - Two Dimensional Data - Three Dimensional Data - Dynamic Data Visualization Techniques for Geospatial Data - Point Data, Line Data, Area Data	6
Unit III	**Temporal Visualization Techniques** Visualization Techniques for Time oriented Data - Visualization Techniques for Multivariate Data - Point, Line and Region based Techniques- Combination of techniques - Visualization Techniques for Trees, Graphs and Networks	6
Unit IV	**Static Visualization** Introduction - Scatter plots - Hexagonal binning plots - Contour plots - Line - Heatmaps - Histogram - Violin - Box plots- From static to Interactive Visualization - Static vs Interactive - Interactive scatter plots - Interactive plots in altair	6
Unit V	**Interactive Visualization** Introduction - Interactive Visualization of data across Time - Temporal Data, Visualization of temporal data, Interactive temporal Visualization, Interactive Visualization of geographical data - Choropleth maps - Bubble plots	6

Practical Component		
Exercises	1. Read data from files using pandas 2. Write a data frame to a file using python 3. Create simple plot to visualize a distribution of variables using python 4. Create Box plot and calculate mean price distribution using python 5. Create static scatter plot using python 6. Create static hexagonal binning plot using python 7. Create linkage in static heatmaps using python 8. Create interactive data visualization with Bokeh 9. Create interactive visualization of Temporal data 10. Create interactive visualization of geographical data	30
Recommended Learning Resources		
Print Resources	1. Matthew Ward, Georges Grinstein, Daniel Keim, Interactive Data Visualization – Foundations, Techniques and Applications, CRC Press, 2nd Edition, 2015 . 2. Abha B, Sharath C G, Shubhangi H, Anshu K, Interactive Data Visualization with Python, Packt Publishing, 2nd Edition, 2020.	
Syllabus Design Dr. P.Shanthi Bala, Professor, PUDoCS		

Year	II	Course Code: CSDS205 Course Title : Financial Data Analytics	Credits	3
Sem.	III		Hours	60
			Category	B
Course Prerequisites, if any	NA			
Internal Assessment Marks : 50	End Semester Marks : 50	Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Approach and analyze any financial data. • Differentiate between various time series models. • Perform cross-validation of various financial models developed. • Forecast future observations on financial data.			
Unit No.	**Course Content**		**Hours**	
Unit I	**Financial Data and their Properties** Asset Returns – Bond Yields and Prices – Implied Volatility – Examples and Visualization of financial data – Multivariate returns.		6	
Unit II	**Linear, Seasonal, Long Memory Models for Financial Time Series** Simple autoregressive models – Simple moving average models – Simple ARMA models – Unit Root non-stationarity – Exponential smoothing.		6	
Unit III	**Linear, Seasonal, Long Memory Models For Financial Time Series** Seasonal models – Regression models with time series errors – Long memory models. **Case Studies Of Linear Time Series** Weekly Regular Gasoline Price.		6	
Unit IV	**Asset Volatility , Volatility Models And Applications** Characteristics of Volatility – Structure of a model – Model Building - Testing for ARCH Effect – ARCH Model – GARCH Model – GARCH-M Model – Exponential Garch Model – Threshold GARCH model – Stochastic volatility model – alternative approaches. Garch Volatility Term structure .		6	
Unit V	**High-Frequency Financial Data** Nonsynchronous trading – Bid-ask spread of trading prices – Empirical characteristics of trading data – Models for price changes. **Value At Risk** Risk measure and Coherence – Risk metrics –Extreme value approach to Value at Risk – Peak over thresholds.		6	
Practical Component				
Exercises	1. Given a simple daily return of a concern as data, implement and execute a R program to compute the sample mean, standard deviation, skewness, excess kurtosis, minimum and maximum of each simple return series. 2. Consider the daily range (daily high–daily low) of Apple stock from January 2, 2007 to December 23, 2011. Obtain the data by the package quantmod. Compute the first 100 lags of ACF of the series. Check for evidence of long-range dependence. If the range series has a long memory, build an ARMA model for the data. 3. Consider the 30-year conventional mortgage rates from April 1971 to November 2011. Build a pure time series model for the monthly mortgage rate. Perform model checking and find the fitted model.		30	

	4. Use the quantmod package to obtain the daily prices of Apple stock from January 2, 2007, to November 30, 2011. Use an ARMA–GARCH model to obtain the daily volatility of the stock. Compare the three volatility series.	
Recommended Learning Resources		
Print Resources	1. Ruey S. Tsay, An Introduction to Analysis of Financial Data with R, Wiley, 2013. 2. Ruey S. Tsay, Analysis of Financial Time Series,Wiley Series in Probability and Statistics, 3rd Edition, 2010.	
Syllabus Design: Dr.R.Sunitha, Associate Professor, PUDoCS		

SEMESTER IV

Year	II	Course Code: CSDS206 Course Title: Computer System Architecture	Credits	4
Sem.	IV		Hours	75
			Category	C
Course Prerequisites, if any	Fundamentals of Computers			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Understand the concept of digital electronics and logic circuits. • Working with binary and arithmetic operations. • Understand the organization of CPU and working principles. • Understand the Input-Output organization in a computer. • Understand the Memory organization in a computer			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Digital Logic Circuits** Digital Computers-Logic Gates-Boolean Algebra-Map Simplification-Combinational Circuits-Flip-Flops-Sequential Circuits – Digital Components	9
Unit II	**Data Representation and Transfer** Datatypes-Complements-Fixed-Point Representation-Floating Point Representation - Register Transfer-Bus and Memory Transfer-Arithmetic, Logic and Shift Microoperations	9
Unit III	**CPU Organization** Register and Stack – Instruction Format – Addressing Modes – Data Transfer and Manipulation – Program Control – RISC – Basics of Pipelining	9
Unit IV	**Input-Output Organization** Peripheral devices – I/O Interface – Asynchronous data transfer – Models of transfer – Priority Interrupt – DMA – Serial Communication	9
Unit V	**Memory Organization** Memory Hierarchy- Main Memory-Auxiliary Memory-Associative Memory-Cache Memory-Virtual Memory-Memory Management Hardware	9
	Practical Component	
Exercises	1. Simplify Boolean expressions using Karnaugh maps. 2. Design a combinational circuit. 3. Implement Logical Left and Right Shifts 4. Understand different data types and how to calculate complements. 5. Evaluate performance improvement through instruction level parallelism. 6. Analyze the effect of cache performance on system performance. 7. Understand the impact of memory hierarchy on access time.	30

Recommended Learning Resources	
Print Resources	1. Morris Mano, Computer System Architecture , Pearson Education, 3rd Edition, 2017
Syllabus Design: Dr. Sukhvinder Singh, Assistant Professor, PUDoCS	

Year	II	Course Code: CSDS207 Course Title: Design and Analysis of Algorithms	Credits	4
Sem.	IV		Hours	75
			Category	C
Course Prerequisites , if any	Basic Knowledge in Data Structures and Programming.			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Analyze the efficiency of algorithms and compare their performance using appropriate metrics. • Understand the general approach of Brute Force and Divide and Conquer algorithms. • Understand the principles of the Greedy Method in algorithm design. • Understand the principles of Dynamic Programming. • Understand the principles of Backtracking and Branch & Bound strategies			
Unit No.	**Course Content**		**Hours**	
	Theory Component			
Unit I	**Introduction** Notation of Algorithm - Analysis of Algorithm Efficiency - Asymptotic Notations and Basic Efficiency classes - Mathematical Analysis of Non-Recursive and recursive Algorithms		9	
Unit II	**Divide and Conquer** Brute Force and Divide and conquer - Binary Search – Finding the maximum and minimum – merge sort - quick sort		9	
Unit III	**Greedy Method** General method - Knapsack problem - Job Sequencing - Spanning Trees - Prims's Algorithm and Kruskal's Algorithm		9	
Unit IV	**Dynamic Programming** General method - Principle of Optimality - Multistage Graphs - 0/1 Knapsack - Travelling Salesman Problem		9	
Unit V	**Backtracking, Branch & Bound** Backtracking - General Method – 8 - Queen Problem - Sum of Subsets - Hamiltonian Cycles Branch and Bound: Introduction FIFO Solution – LC Branch and Bound – 0/1 Knapsack		9	
	Practical Component			
Exercises	1. Write recursive and iterative algorithms and analyze their time complexities using Big-O notation. 2. Implement and compare the efficiency of sorting algorithms (e.g., bubble sort, quicksort) on different input sizes. 3. Implement merge sort and analyze its time complexity with different input sizes. 4. Implement a greedy algorithm for the knapsack problem and analyze its efficiency. 5. Implement Prim's algorithm for finding the minimum cost spanning tree. 6. Implement Kruskal's algorithm for finding the minimum cost spanning tree and compare with the results of exercise 6 7. Solve the 0/1 knapsack problem using dynamic programming and analyze the time complexity.		30	

	8. Implement a backtracking solution for the subset sum problem and analyze its efficiency.	
Recommended Learning Resources		
Print Resources	1. Horowitz, E. and Sahani, S, Fundamentals of Computer Algorithms, Universities Press, 2nd Edition, 2008. 2. S.Sridar, Design and Analysis of Algorithms, Oxford University Press, 2014	
Syllabus Design: Dr. T. Vengattaraman, Associate Professor, PUDoCS		

Year	II	Course Code: CSDS208 Course Title: Database Management Systems	Credits	4
			Hours	75
Sem.	IV		Category	C
Course Prerequisites, if any	Knowledge of data structures and file-handling			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	● Understand the fundamentals of relational model. ● Design real time applications using database query language (SQL) ● Familiarize with the different kinds of PL/SQL objects. ● Understand the various database applications using the Relational model, ER model and EER model. ● Construct and normalize conceptual data models.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction to Relational model** Structure of relational database, Database schema, Keys, Schema diagram, Relational Query language, Relational Algebra.	9
Unit II	**Introduction to SQL** SQL data definition, basic structure of SQL Queries, set operations, null values, aggregate functions, nested subqueries	9
Unit III	**Intermediate and advanced SQL** Join expressions, views, transaction, integrity constraints, functions and procedures, triggers.	9
Unit IV	**Database design using ER model** The Entity-Relationship model, complex attributes, mapping cardinalities, primary key, removing redundant attributes in entity sets, reducing ER diagrams to relational schemas, extended ER features.	9
Unit V	**Relational database design** Decomposition using functional dependencies, normal forms, functional dependency theory, algorithms for decomposition using functional dependencies, decomposition using multivalued dependencies.	9

Practical Component		
Exercises	1. Implement the DDL commands using SQL. 2. Implement the DML commands. 3. Implement the DDL constraints, DCL, and TCL commands. 4. Implement various built functions and aggregate functions. 5. Implement the various join operations. 6. Implement the various nested subqueries. 7. Creation and manipulation of Views. 8. Practice the basics of PL/SQL [control structures]. 9. Create the functions and procedures using PL/SQL. 10. Create the Triggers using PL/SQL.	30
Recommended Learning Resources		
Print Resources	1. Abraham Silberschatz , Henry F. Korth Sudarsan, Database System Concepts, McGraw Hill , 7th Edition, 2021. 2. Brumm.B, Beginning Oracle SQL for Oracle Database 18c: From Novice to Pofessional, Apress, 1st Edition, 2019. 3. Kevin Loney, Bob Bryla , Oracle Database 12c The Complete Reference, McGraw Hill education, 1st Edition, 2013.	
Syllabus Design: Dr. S.L.Jayalakshmi, Assistant Professor, PUDoCS.		

Year	II	Course Code: CSDS209 Course Title : Applied Regression Analysis	Credits	4
Sem.	IV		Hours	75
			Category	C
Course Prerequisites, if any	Fundamental knowledge of Statistics			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory) : 03 hrs Duration of ESA (Practical) : 03 hrs		
Course Outcomes	• Understand the basics of regression analysis • Build and fit simple and multiple linear regression models • Validate the modelling assumptions with formal tests • Diagnose the model violations • Model qualitative data			

Unit No.	Course Content	Hours
	Theory Component	
UNIT I	**Introduction** Regression Analysis - Publicly Available Data Sets - Applications - Steps in Regression Analysis	9
UNIT II	**Simple Linear Regression** Introduction - Covariance and Correlation Coefficient - The Simple Linear Regression Model - Parameter Estimation - Tests of Hypotheses - Confidence Intervals - Predictions - Measuring Quality of Fit - Regression Line Through the Origin - Trivial Regression Models	9
UNIT III	**Multiple Linear Regression** Introduction - Description of the Data and Model - Supervisor Performance Data - Parameter Estimation - Interpretations of Regression - Centering and Scaling - Properties of the Least Squares Estimators - Multiple Correlation Coefficient - Inference for Individual Regression Coefficients - Tests of Hypotheses in a Linear Model	9
UNIT IV	**Regression Diagnostics: Detection of Model Violations** Introduction - Standard Regression Assumptions - Various Types of Residuals - Graphical Methods - Graphs Before and After Fitting a Model	9
UNIT V	**Logistic Regression** Introduction - Modeling Qualitative Data - The Logit Model - Estimating Probability of Bankruptcies - Logistic Regression Diagnostics	9
	Practical Component	
Exercises	1. Implement simple Linear Regression 2. Implement simple Multiple Regression 3. Visualize Underfitting and Over Fitting in a Regression 4. Perform Regression Diagnostics using Graphical Methods 5. Implement Logistic Regression	30

Recommended Learning References		
Print Resources	1. Samprit Chatterjee and Ali S. Hadi, Regression Analysis by Example, John Wiley and Sons, 5th Edition , 2013. 2. Samprit Chatterjee and Jeffrey S. Simonoff, Handbook of Regression Analysis, John Wiley and Sons, 1st Edition, 2013	
Syllabus Design : Dr. M.Nandhini, Professor, PUDoCS		

SEMESTER V

Year	III	Course Code: CSDS301	Credits	4
			Hours	75
Sem.	V	Course Title: OPERATING SYSTEMS	Category	C

Course Prerequisites, if any	Knowledge of computers & computer organization.	
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.
Course Outcomes	● Understand the basic concepts of Operating System and Process. ● Learn the various mechanisms of CPU scheduling, process synchronization and deadlocks. ● Understand how the memory is utilized. ● Analyze various File System methods and Disk scheduling algorithms. ● Evaluate system structures in various operating systems, such as Linux and Windows and identify similarities and differences.	

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Overview and Process management** Introduction: Operating System Structures - Operating systems services - System calls. Process Management**:** Process Concept – process scheduling-operation on processes-Inter process communications- Threads.	9
Unit II	**Scheduling algorithms and Process Synchronization** CPU Scheduling: Basic Concepts – Scheduling Algorithms. Process Synchronization**:** Critical Section problem – Semaphores - Classical problems of synchronization- Monitors. Deadlock: Deadlock Characterization-Deadlock Handling-Deadlock Prevention-Deadlock Avoidance- Deadlock Detection-Deadlock Recovery.	9
Unit III	**Memory Management** Main Memory: Contiguous Memory Allocation-Paging- Structure of the Page Table –Swapping. Virtual Memory: Demand Paging-Page Replacement-Thrashing.	9
Unit IV	**Storage Management** Mass Storage structure: Overview- HDD (Disk) Scheduling – storage management- RAID Structure. File Systems: File concepts – Access methods – Directory Structure – File Protection –File system Implementation- File-System Structure-File-System Operations- Allocation methods.	9
Unit V	**Case Studies** The Linux system: Design principles – kernel modules-process management- Scheduling- Memory Management-Linux File System. Windows Operating system: Systems components-Windows File System.	9
	Practical Component	
Exercises	1. Practice file handling utilities, Process utilities, Disk utilities, and Networking commands.	30

	2. Program to implement various system call operations. 3. Program to demonstrate various File management operations. 4. Program to simulate CPU scheduling algorithms: FCFS, SJF, Round Robin, and priority. 5. Program to simulate Intra & Inter – Process Communication (IPC) techniques: Pipes, Messages Queues, and Shared Memory. 6. Program to simulate solutions to Classical Process Synchronization Problems: Dining Philosophers, Producer – Consumer, Readers – Writers. 7. Program to simulate Bankers Algorithm for Deadlock Avoidance. 8. Program to simulate Page Replacement Algorithms: FIFO, Optimal, LRU. 9. Program to simulate implementation of HDD Scheduling Algorithms: FCFS, SCAN, C-SCAN. 10. Case study on Linux and Windows Operating systems features and prepare a report on the same	
Recommended Learning Resources		
Print Resources	1. Abraham Silberschatz Peter B Galvin, G. Gagne, Operating Systems Concepts, Addison Wesley, 10th Edition, 2018. 2. William Stallings, Operating Systems: Internals and Design Principles, Prentice Hall, 10th Edition, 2021.	
Syllabus Design: Dr. S.L.Jayalakshmi, Assistant Professor, PUDoCS		

Year	III	Course Code: CSDS302 Course Title: Mathematical Foundations of Computer Science	Credits	4
Sem.	V		Hours	75
			Category	A
Course Prerequisites, if any	Basic knowledge in mathematics			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs.		
Course Outcomes	• Understand logical statement structures • Apply operations in problem-solving • Analyze integer representations and congruences • Understand counting principles. • Evaluate combinatorial solutions.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Logic and Proofs** Propositional Logic – Predicates and Quantifiers – Rules of Inference - Proofs – Methods and Strategy	15
Unit II	**Basic Structures** Sets – Functions – Sequences and Summations– Matrices Relations – properties – representation	15
Unit III	**Number Theory** Divisibility and Modular Arithmetic – Integer Representations and Algorithms – Primes and Greatest Common Divisors – Congruences	15
Unit IV	**Induction and Recursion** Mathematical Induction - Strong Induction and Well Ordering - Recursive Definitions and Structural Induction	15
Unit V	**Counting** Basics – Pigeonhole principle – Permutations and Combinations – Binomial Coefficients	15
	Recommended Learning Resources	
Print Resources	1. Kenneth H. Rosen, Discrete Mathematics and its Applications , McGraw Hill, 7th Edition, 2017. 2. Trembley. J.P and Manohar. R., Discrete Mathematical Structures with Applications to Computer Science , Tata McGraw Hill, 2020	
Syllabus Design Dr. M.Sathya, Assistant Professor, PUDoCS		

<table>
<tr><td>Year</td><td>III</td><td colspan="2" rowspan="2">Course Code: CSDS303
Course Title: Computer Networks</td><td>Credits</td><td>4</td></tr>
<tr><td>Sem.</td><td>V</td><td>Hours</td><td>75</td></tr>
<tr><td colspan="4"></td><td>Category</td><td>C</td></tr>
<tr><td>Course Prerequisites, if any</td><td colspan="5">Fundamentals of Computers</td></tr>
<tr><td>Internal Assessment Marks: 25</td><td colspan="2">End Semester Marks: 75</td><td colspan="3">Duration of ESA (Theory): 03 hrs.
Duration of ESA (Practical): 03 hrs.</td></tr>
<tr><td>Course Outcomes</td><td colspan="5">• Learn the basics of Network topology
• Learn about the various physical network media
• Understand the functionalities of all the network layers
• Familiarize the protocols of different layers
• Implement various network protocols</td></tr>
<tr><td>Unit No.</td><td colspan="4">Course Content</td><td>Hours</td></tr>
<tr><td colspan="6">Theory Component</td></tr>
<tr><td>Unit I</td><td colspan="4">Introduction
Introduction to Networks – Topology - Network Architecture - Reference Models - Transmission Media-Multiplexing -Switching</td><td>9</td></tr>
<tr><td>Unit II</td><td colspan="4">Data link layer
Design Issues - Error Detection and Correction - Elementary Data - Link Protocols - Sliding window Protocols</td><td>9</td></tr>
<tr><td>Unit III</td><td colspan="4">Network Layer
Design Issues - Routing - Logical Addressing - IP Working - IPV4 Vs IPV6 - Address Mapping - delivery - Forwarding and routing</td><td>9</td></tr>
<tr><td>Unit IV</td><td colspan="4">Transport Layer
The Transport Service - Service provided to the Upper Layers - Flow Control & Buffering - TCP Congestion Control - UDP - TCP Vs UDP</td><td>9</td></tr>
<tr><td>Unit V</td><td colspan="4">Application layer
Domain Naming System - DNS Namespace - Resource Records - Name Servers - Electronic mail - Messages Format, Message Transfer</td><td>9</td></tr>
<tr><td colspan="6">Practical Component</td></tr>
<tr><td>Exercises</td><td colspan="4">1. Implementation of Basic Chat
2. Implementation of Multiple User Chat
3. Implementation of File Transmission
4. Implementation of Simple Mailing Application
5. Implementation of Client Server Application
6. Given IP address and subnet mask, Computation of (i)Subnet addresses (ii) Number of hosts in each subnet (iii IP addresses of hosts in each subnet
7. Implementation of Error Detection / Error Correction Techniques
8. Implementation of socket program Remote Procedure Call
9. Implementation of any 1 routing protocol
10. Implementation of congestion control protocol</td><td></td></tr>
</table>

Recommended Learning Resources	
Print Resources	1. Andrew S. Tanenbaum, David J. Wetherall, Computer Networks, Prentice Hall Publisher, 5th Edition, 2022. 2. Larry L. Peterson, Bruce S. Davie, Computer Networks: A Systems Approach, Morgan Kaufmann Publishers Inc., 5th Edition, 2015. 3. James F. Kurose, Keith W. Ross, Computer Networking - A Top-Down Approach Featuring the Internet, Pearson Education, 7th Edition,2022.
Syllabus Design: Dr. G. Krishnapriya, Assistant Professor, PUDoCS	

Year	III	**Course Code: CSDS304** **Course Title: Artificial Intelligence**	**Credits**	4
Sem.	V		**Hours**	75
			Category	C
Course Prerequisites, if any	Basic Programming Skills			
Internal Assessment Marks : 25	End Semester Marks : 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	● Familiarize with the diverse traits of a problem-solving agent ● Explore methods for tackling problems amidst different constraints ● Implement AI techniques in various applications ● Grasp the distinct models of learning ● Develop an expert system			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Foundation and History of AI - Intelligent Agents - Agents and Environments - The Concept of Rationality - Nature of Environments - Structure of Agents - Problem Solving Agents - Examples	9
Unit II	**Searching** Searching for Solutions, Uniformed Search Strategies - Heuristics Search Strategies - Local Search Algorithms and Optimization Problems- Hill Climbing- Simulated Annealing- Local Beam Search- Genetic Algorithms - Optimal Decisions in Games - Alpha–Beta Pruning	9
Unit III	**Agents** Logical Agents- Knowledge-Based Agents- The Wumpus World- Logic- Propositional Logic - Propositional Theorem Proving - Effective Propositional Model Checking - Agents Based on Propositional Logic	9
Unit IV	**First Order Logic** Introduction- Syntax and Semantics - Inference - Propositional Vs. First-Order Inference - Unification and Lifting - Forward Chaining - Backward Chaining- Resolution.	9
Unit V	**Learning** Forms of Learning- Supervised Learning- Learning Decision Trees- Hypothesis- Theory of Learning - Prolog - Programs - Data Objects	9
	Practical Component	
Exercises	1. Implement Breadth First Search 2. Implement Depth First Search 3. Implement Tic-Tac-Toe game 4. Implement 8-Puzzle problem 5. Implement Water-Jug problem 6. Implement Monkey Banana Problem 7. Implement Alpha-Beta Pruning 8. Develop an expert system using Prolog	30

Recommended Learning Resources	
Print Resources	1. S. Russell and P. Norvig, Artificial Intelligence – A Modern Approach, Pearson Education, 4th Edition, 2022. 2. Max Bramer, Logic Programming with Prolog, Springer, 2013.
Syllabus Design: Dr. P.Shanthi Bala, Professor, PUDoCS	

SEMESTER VI

Year	III	Course Code:CSDS306 Course Title: Management Strategies and Concepts	Credits	4
			Hours	75
Sem.	VI		Category	A
Course Prerequisites, if any	- Nil -			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs.		
Course Outcomes	● Understand the fundamentals of Management Theories. ● Learn the management & communication Process Concepts ● Analyse the performance of decentralized and centralized organizational structures ● Analyse the different leadership styles and their effects on team performance and organizational culture. ● Evaluate the effectiveness of the strategies in enhancing productivity and efficiency.			
Unit No.	**Course Content**		**Hours**	
Theory Component				
Unit I	**Management Theories** Science Theory and Practice - Management and Society: Social Responsibility and Ethics. The nature and purpose of planning - objectives – Strategies Policies and planning premises.		15	
Unit II	**Decision Making** Process of decision making- organizing- Nature and purpose of organizing – Basics of departmentalization - Line/Staff Authority and Decentralization - Effective Organizing and organizational structure & culture.		15	
Unit III	**Human Resource Management & Selection** Staffing-Manpower planning - Recruitment & Selection-Performance appraisal and career strategy - Organizational development.		15	
Unit IV	**Managing the Human factor** Motivation - Leadership –Communication		15	
Unit V	**The System & Process of Controlling** Control techniques and Information Technology - Productivity and Operations Management - Overall and Preventive Control - Towards a Unified, Global management theory.		15	
Recommended Learning Resources				
Print Resources	1. Herald Knootz and Heinz Weihrich, Essentials of Management, McGraw-Hill Publishing Company, 11th Edition, 2020. 2. Fred R. David and Forest R. David, Strategic Management: Concepts and Cases, Prentice Hall India Learning Private Limited, 16th Edition, 2020.			
Syllabus Design: Dr. S.L.Jayalakshmi, Assistant Professor, PUDoCS				

Year	III	Course Code: CSDS307	Credits	4
Sem.	VI	Course Title: Software Engineering Theory and Practice	Hours	75
			Category	C
Course Prerequisites, if any	● Basic knowledge of programming and information systems			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Understand the fundamental concepts of design thinking • Analyze and document the software requirements • Apply appropriate software engineering design concepts to develop software. • Apply software testing strategies • Understand and consider the significance of security in software development process			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction to Design Thinking** Design process - Traditional design - Design thinking - Existing sample design projects - Study on designs around us - Compositions/structure of a design - Innovative design - Breaking of patterns - Reframe existing design problems - Principles of creativity Empathy - Customer Needs - Insight-leaving from the lives of others/standing on the shoes of others - Observation.	9
Unit II	**Software Engineering and Software Requirements** Defining software engineering, Software life cycle models, Selection of a life cycle model - Requirements engineering, Types of requirements, Feasibility studies, Requirements elicitation, Requirement analysis, Requirement documentation, Requirement validation.	9
Unit III	**Software Project Planning** Size estimation, Cost estimation, Models, Constructive cost model, Software risk management, Software design, Modularity, Strategy of design, Function oriented design, Object oriented design.	9
Unit IV	**Testing Strategies** A strategic approach to software testing, Test strategies for conventional software, Black-Box and White-Box testing, Validation testing, System testing, The art of Debugging.	9
Unit V	**Secure Software Engineering** Introduction - The problem – Software assurance and software security – Threats to software security – Software insecurity – Benefits of detecting software security defects early – Managing secure software development – Defining Properties – Influencing the security properties of software – To assert and specify desired security properties.	9
	Practical Component	
Exercises	1. Conceptualize a novel app that will help to save: a) Energy b) Water	30

	c) Food 2. Apply the phases of Software Development Life Cycle for the following applications and develop the same : a) Library Management System b) Hospital Management System 3. Design the above two systems with security features and implement the same.	
Recommended Learning Resources		
Print Resources	1. Tim Brown, "Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation", 1st edition, HarperCollins Publishers Ltd, 2019. 2. Roger S. Pressman, Bruce Maxim, "Software Engineering, A Practitioner's Approach", 9th edition, McGraw Hill International Edition, 2023. 3. Julia H. Allen, "Software Security Engineering : A Guide for Project Managers", 1st edition, 2008.	
Syllabus Design: Dr. T. Chithralekha, Professor, PUDoCS ***Dr. G. Krishnapriya, Assistant Professor, PUDoCS***		

Year	III	Course Code: CSDS308 Course Title : Distributed Systems	Credits	4
Sem.	VI		Hours	75
			Category	C
Course Prerequisites , if any	Basic knowledge in Operating Systems and Computer Networks			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practicals) : 03 hrs.		
Course Outcomes	● Learn about Distributed Systems ● Understand state-of-the-art distributed system ● Design and develop Client/Server Applications. ● Setup fault tolerance and replication servers. ● Design and implement CORBA and DCOM.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction to distributed systems** Definition – Goals – Hardware and Software Concepts – Client/Server Model Communication – Layered Protocols RPC –Remote Object Invocation – Message Oriented Communication.	9
Unit II	**Client-Server** Client Server and Naming Entity – Threads- Client Server – Code Migration – S/W Agents – Naming Entity – Location Mobile Entity.	9
Unit III	**Synchronization** Distributed Transactions - Synchronization – Clock Synchronization – Logical Clocks – Global States – Election Algorithms – Mutual Exclusion – Distributed Transaction Consistency and Replication–Data Centric Consistency–Fault Tolerance – Distributed Commit –Recovery	9
Unit IV	**Distributed Objects** Distributed Object Database System - CORBA – DCOM – GLOBE.	9
Unit V	**Distributed File System** Introduction – Distributed Document based System – WWW – Distributed Coordination based System – JINI.	9
	Practical Component	
Exercises	1. Perform arithmetic operations using RMI. 2. Calculate simple and compound interest using RMI.	30

	3. Implementation of ATM using RMI. 4. Implementation of Telephone Directory using RMI. 5. Implementation of Quiz Server using Servlets. 6. Implementation of an Online Shopping System using servlets. 7. Implementation of matrimonial System using servlets. 8. Implementation of servlet based Airline Reservation system. 9. Create a Word Document with text using DCOM and Visual Basic.	

Recommended Learning Resources	
Print Resources	1. Andrew S. Tanenbaum, Maarten van Steer, Distributed Systems Principles and Paradigms, Prentice Hall India, 3rd edition, 2017. 2. George Couloursis, Jean Dollomore and Tim Kinderberg, Distributed Systems - Concepts and Design, Addison-Wesley, 5th Edition 2011.
Syllabus Design: Dr.T.Sivakumar, Assistant Professor, PUDoCS	

Year	III	Course Code: CSDS309	Credits	4
Sem.	VI	Course Title: Operations Research	Hours	75
			Category	A
Course Prerequisites, if any	Basic mathematical and problem solving skills			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs.		
Course Outcomes	● Understand and comprehend the basics of Linear Programing Problem (LPP). ● Learn LPP solving methods and explore duality in LPP. ● Solve assignment problem and its variants. ● Find feasible and optimal solutions for transportation problems. ● Perform critical path analysis and reviewing of a project.			
Unit No.	**Course Content**		**Hours**	
Theory Component				
Unit I	**Introduction** Operation Research – Definition – Characteristics – Techniques – Applications. LPP – Introduction – Applications and components of LPP – Steps in solving LPP.		15	
Unit II	**LPP** Mathematical formulation – Graphical method – Simplex method – Artificial variables – Big-M method - Two-phase method – Degeneracy and unbound solutions – Duality in LPP – Formulation – Relationship between primal and dual problems.		15	
Unit III	**Assignment Model** Mathematical formulations - Hungarian Method – Variants of the Assignment problem.		15	
Unit IV	**Transportation Problem** Mathematical formulation – Finding basic feasible solutions – NWCR, LCM and VAM – Optimal solution – MODI method.		15	
Unit V	**Network Scheduling** Introduction – Basic components – Logical sequencing – Rules of network construction – Concurrent Activities – Critical Path Analysis - Activity Time and Floats – Project Evaluation and Review Technique (PERT) – Three Time Estimates – Critical Path Analysis of PERT network – Probability of completion of Project.		15	
Recommended Learning Resources				
Print Resources	1. Kanti Swarup, P.K. Gupta, Man Mohan, Operations Research, Sultan Chand & Sons, 20th Edition, 2023. 2. Taha H.A., Operations Research: An Introduction, Pearson Education, 10th Edition, 2019.			

Syllabus Design: Dr. G. Krishnapriya, Assistant Professor, PUDoCS

Year	III	Course Code: CSDS310 Course Title : Massive Data Management	Credits	4
Sem	VI		Hours	75
			Category	C
Course Prerequisites, if any	Database management systems			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Understand the Big Data Platform and its Use cases • Apply analytics on Structured, Unstructured Data • Provide an overview of Apache Hadoop • Provide HDFS Concepts and Interfacing with HDFS • Understand Map Reduce system			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Big Data** Evolution of Data Management-Waves of Managing Big data - Big data management architecture- Setting Architectural Foundation-Traditional and Advance analytics - Big data Types - Examine the cloud and big data	9
Unit II	**Big Data Management** Operational database - Non relational database, Key value pair database, Document database, columnar database, Graph database, Spatial database, Polyglot persistence - Map Reduce Fundamentals - origin of Map reduce, Map and Reduce function, map reduce tasks	9
Unit III	**Hadoop Foundation** Hadoop Distributed File system - Hadoop Map Reduce - Big data foundation with Hadoop architecture - managing Resources and application with Hadoop YARN - Interacting with Hadoop Ecosystem	9
Unit IV	**Hadoop Eco System** Application Development in Hadoop- Pig and Pig Latin, Hive, Jaql - Getting data into Hadoop- Basic Copy data, Flume - Other Hadoop Components - Zookeeper, HBase, Oozie, Lucene, Avro	9
Unit V	**Big Data Solutions** Security and Governance for Big data Environments- Security in context with Big data - Data protection options - Data Governance challenges - Importance of Big data to business - Improving Business process with Big data Analysis	9
	Practical Component	
Exercises	1. Install Hadoop and study the HADOOP ecosystem 2. Implement word count program using MapReduce 3. Develop a MapReduce program to calculate the frequency of a given word in a given file	30

	4. Develop a MapReduce program to find the maximum temperature in each year 5. Implement Matrix multiplication using Map-Reduce 6. Implement a MapReduce program that processes a dataset 7. Develop a MapReduce to analyze weather data set and print whether the day is shinny or cool day. 8. Implement Page Rank Algorithm using Map-Reduce 9. Analyse the log files using Pig	
Recommended Learning Resources		
Print Resources	1. Judith Hurwitz, Alan Nugent, Dr.Fern Halper, Macia Kaufman, Big Data for Dummies, John Wiley & Sons, 2013. 2. Chris Eaton, Dirk Deroos, Tom Deutsch, George lapis, Paul Zikopoulos, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGraw-Hill, 2011.	
Syllabus Design Dr. P.Shanthi Bala, Professor, PUDoCS		

Year	III	Course Code: CSDS311 Course Title: Hadoop Ecosystem	Credits	4
Sem	VI		Hours	75
			Category	C
Course Prerequisites, if any	Database Management Systems			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Understand Hadoop ecosystem basics • Build & run simple MapReduce programs • Manage data in HDFS • Apply Hadoop ecosystem tools for data analysis • Understand HiveQL and HBase			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Building Data Analytic Systems with Hadoop - Need for Distributed Analytical System - AI Technologies, Cognitive Computing, Deep Learning and Big Data Analysis, Natural Language Processing and BDAs - SQL and NoSQL Querying - Building BDA systems - Hadoop Ecosystem Implementation - Architecture - Data Visualization and Reporting	9
Unit II	**MapReduce** Introduction - Data Format and A Weather Dataset - Analyzing the Data with Unix Tools and Hadoop - Map and Reduce Functions - Java MapReduce - Scaling Out - Hadoop Streaming and Pipes - Map Reduce working Principle	9
Unit III	**Hadoop Distributed File System** Design and Concepts of HDFS - Blocks, Namenodes, Datanodes - Command-Line Interface –Hadoop File systems - Java Interface - Data Flow - Parallel copying - Hadoop Archives	9
Unit IV	**Advanced MapReduce Development** Configuration and Development Environment - Writing Unit Tests for Map Reduce –Running locally on test data - Running on cluster- Tuning a job - Map Reduce workflow - Map Reduce Types and Formats	9
Unit V	**Hadoop Ecosystem for Data Warehousing and Analysis** Pig Basics and Running Pig Scripts- Pig Latin - User Defined Functions- Data Processing operators - Hive - HiveQL - HBase Basics - HBase versus RDBMS- HBase Use cases	9
	Practical Component	
Exercises	1. Perform weather dataset analysis using Pig script 2. MapReduce program for specific weather dataset analysis with map, reduce, and combiner functions 3. Hadoop MapReduce program for data distribution 4. Hive query for data analysis on weather dataset subset 5. HBase program for employee dataset analysis	30

Recommended Learning Resources	
Print Resources	1. Kerry Koitzsch, Pro Hadoop Data Analytics: Designing and Building Big Data Systems using the Hadoop Ecosystem, Apress, 2018. 2. Tom White, Hadoop: The definitive Guide, O'Reilly Media, 3rd Edition,2012.
Syllabus Design Dr. P.Shanthi Bala, Professor, PUDoCS	

SEMESTER-VII

Year	IV	Course Code :CSDS401 Course Title : Web Engineering	Credits	4
Sem.	VII		Hours	75
			Category	C
Course Prerequisites, if any	Basic understanding of programming concepts.			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs		
Course Outcomes	• Understand the process of web publishing. • Acquire skills developing web pages using HTML. • Acquire skills to style the web pages using CSS. • Acquire skills to build server side web components. • Explore the mobile web development process.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction to World Wide Web** Introduction to web publishing - Web browsers - Web servers - Uniform Resource Locators - Using browser based developer tools.	9
Unit II	**Introduction to HTML and CSS** Structuring a web page with HTML - Basic elements - Lists - Links - Tables - Images - Forms. Using CSS to style a site - CSS for positioning - Integrating Multimedia elements.	9
Unit III	**Introduction to JavaScript** The structure - Operators - Variables - Control structures - Functions - Arrays - Objects - Validation.	9
Unit IV	**Introduction to PHP** Setting up the server - PHP language basics - built-in functions - library functions - using includes - database connectivity - sending email - cookies and sessions-File uploads.	9
Unit V	**Mobile Web** Mobile browsing needs - text on mobile web - design and page layout - links - images and multimedia - CSS for mobile - making use mobile features - Best practices.	9
	Practical Component	
Exercises	1. Build your resume using simple static html. 2. Enrich your resume with CSS. 3. Implement an HTML Form with javascript validation. 4. Build a web application to demonstrate event handling in JavaScript. 5. Add a server side component to the task #3. 6. Build a server side data storage web application. 7. Build a web application to demonstrate session handling. 8. Build a web application to demonstrate cookies handling. 9. Implement mobile web application. 10. Implement file uploads in a web application.	30

Recommended Learning Resource	
Print Resources	1. Laura Lemay, Rafe Coburn, Jennifer Kyrnin, Sams Teach yourself HTML, CSS & Javascript Web Publishing, Pearson Education, 2016
Syllabus Design: Dr.K.S.Kuppusamy, Associate Professor, PUDoCS	

Year	IV	**Course Code: CSDS402**	**Credits**	4
			Hours	75
Sem.	VII	**Course Title: System Modeling and Simulation**	**Category**	C
Course Prerequisites, if any	Basic Knowledge in Statistics			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Understand the fundamentals of modeling and simulation. • Learn about statistical models and input modelling. • Understand the techniques for random number generation. • Perform the simulation of dynamic systems. • Verify simulation models.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Simulation tool - Advantages and disadvantages of Simulation - Areas of application- Systems and system environment - Components of a system - Discrete and continuous systems - Model of a system - Types of Models – DESS - Simulation of queuing systems - General Principles.	9
Unit II	**Statistical Models in Simulation** Review of terminology and concepts - Useful statistical models - Discrete distributions - Continuous distributions - Poisson process - Empirical distributions - General Principles - Characteristics of queuing systems - Queuing notation - Long-run measures of performance of queuing systems - Steady-state behavior of M/G/1 queue - Networks of queues.	9
Unit III	**Random-Number Generation** Properties of random numbers - Generation of pseudo-random numbers - Techniques for generating random numbers - Tests for Random Numbers - Inverse transform technique Acceptance -Rejection technique.	9
Unit IV	**Input Modeling** Data Collection - Identifying the distribution with data - Parameter estimation - Goodness of Fit Tests - Fitting a non-stationary Poisson process - Selecting input models without data – Multivariate & Time - Series input models -Types of simulations with respect to output analysis - Stochastic nature of output data - Measures of performance and their estimation	9
Unit V	**Simulation Models** Measures of performance and their estimation - Output analysis for terminating simulations - Output analysis for steady - state simulations - Verification, Calibration and Validation - Optimization, Model building, verification and	9

	validation - Verification of simulation models - Calibration and validation of models, Optimization via Simulation.	
Practical Component		
Exercises	1. Simulate of Random Number Generation. 2. Implement Chi-square goodness-of-fit test. 3. Implement One-sample Kolmogorov-Smirnov test 4. Implement Test for Standard Normal Distribution 5. Implement Monte-Carlo Simulation. 6. Simulate Single Server Queuing System. 7. Simulate Two-Server Queuing System. 8. Simulate and control a conveyor belt system 9. Implement Two-sample Kolmogorov-Smirnov test.	30
Recommended Learning Resources		
Print Resources	1. Jerry Banks, John S. Carson II, Barry L. Nelson, David M. Nicol, Discrete-Event System Simulation, Pearson Education, 5th Edition, 2013. 2. Lawrence M. Leemis, Stephen K. Park: Discrete – Event Simulation: A First Course, Pearson Education, 2013	
Syllabus Design: Dr. G. Krishnapriya, Assistant Professor, PUDoCS		

Year	IV	**Course Code: CSDS403** **Course Title: Wireless Communication Networks**	**Credits**	4
Sem.	VII		**Hours**	75
			Category	C
Course Prerequisites, if any	Knowledge in computer networks			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practicals): 03 hrs.		
Course Outcomes	• Understand basics of wireless communication networks • Understand the satellite communications concepts and compare various generations of wireless communication • Explore IEEE 802.11 WLAN standard • Explore WAP and its applications • Understand wireless LAN technologies			

Theory Component

Unit No.	Course Content	Hours
Unit I	**Introduction** Wireless Communication Technology- Antennas and Propagation- Antennas, Propagation Modes, Fading in the Mobile Environment. Signal Encoding Techniques- Signal Encoding Criteria, Digital Data-Analog Signals, Analog Data-Analog Signals, Analog Data-Digital Signals.	9
Unit II	**Satellite Communications** Wireless Networking - Satellite Communications- Satellite Parameters and Configurations, Capacity Allocation-Frequency Division, Capacity Allocation-Time Division Cellular Wireless Networks- Principles of Cellular Networks, First-Generation Analog, Second-Generation - TDMA, CDMA, 3G Systems.	9
Unit III	**Wireless LAN Standards** Evolution of IEEE 802.11- Introduction to IEEE 802.11 -General Description- Medium Access Control (MAC) for the IEEE 802.11 - WLANs Physical Layer for IEEE 802.11 -WLANs; Radio Systems -IR Systems Applications.	9
Unit IV	**Mobile IP** Introduction, operation of Mobile IP, Mobile IP terminologies, Wireless Access Protocols: Introduction, Architecture overview, Wireless application environment	9

Unit V	**Wireless LAN Technology** Wireless LAN- application, requirements, Technology: Infrared, spread spectrum, Narrowband microwave (radio), Introduction to Bluetooth Technologies.	9
	Practical Component	
Exercises	1. Study about different Wireless devices like Wi-Fi Dongler, Wireless Access Point, Antenna, Wi-Fi Router. 2. Configure a wireless LAN using CISCO Packet Tracer. 3. Develop a client server application using Wireless LAN. 4. Simulate BlueTooth Communication after pairing in CISCO Packet Tracer.	30
Recommended Learning Resources		
Print Resources	1. William Stallings, Wireless Communications and Networks, Pearson Prentice Hall, 2nd Edition, 2008.	
Syllabus Design: Dr.T.Sivakumar, Assistant Professor, PUDoCS		

Year	IV	Course Code : CSDS404 Course Title : Big Data Analytics	Credits	4
Sem.	VII		Hours	75
			Category	C
Course Prerequisites, if any	Basic understanding of programming			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs		
Course Outcomes	• Understand the Role of Big Data Technologies. • Gain a working knowledge of Big Data Ecosystems • Implement and understand the basics of Hadoop and MapReduce • Develop foundational skills for further exploring Big Data technologies and tools			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction to Big Data** Introduction to Big Data: Types of Digital Data-Characteristics of Data – Evolution of Big Data – Definition of Big Data – Challenges with Big Data – 3Vs of Big Data – Non Definitional traits of Big Data – Business Intelligence vs. Big Data – Data Warehouse and Hadoop environment – Coexistence.	9
Unit II	**Introduction to Big Data Analytics and NoSQL** Big Data Analytics: Classification of analytics – Data Science Terminologies in Big Data – CAP Theorem – BASE Concept. Introduction to NoSQL: Types of Databases – Advantages – NewSQL – SQL vs. NoSQL vs NewSQL.	9
Unit III	**Introduction to Hadoop Ecosystems** Introduction to Hadoop: Features – Advantages – Versions – Overview of Hadoop Ecosystems – Hadoop distributions – Hadoop vs. SQL – RDBMS vs. Hadoop – Hadoop Components – Architecture – HDFS – Map Reduce: Mapper – Reducer – Combiner – Partitioner – Searching – Sorting – Compression.	9
Unit IV	**Introduction to NoSQL Database** No SQL databases: Mongo DB: Introduction – Features – Data types – Mongo DB Query language – CRUD operations – Arrays – Functions: Count – Sort – Limit – Skip – Aggregate – Map Reduce. Cursors – Indexes – Mongo Import – Mongo Export.	9
Unit V	**Introduction to Hive and Pig Architectures** Introduction to Hive – Architecture – data type – File format – HQL – SerDe – User-defined functions.	9
	Practical Component	
Exercises	1. Install Big Data tools (Hadoop, Mongo DB, Hive, and Pig) 2. Explore Hadoop Commands 3. Explore Mongo DB Datasets 4. Implement a simple Hadoop program for word count 5. Implement a simple MapReduce program for finding the average length of words starting with each character. 6. Basic Mongo DB database creation. 7. Implement basic Mongo DB database manipulation commands.	30

Recommended Learning Resources	
Print Resources	1. Seema Acharya, Subhashini Chellappan, Big Data and Analytics, Wiley Publication,2nd Edition, 2019. 2. Tom White, Hadoop: The Definitive Guide, O'Reilly Publications, 2011. 3. Kyle Banker, Mongo DB in Action, Manning Publications Company, 2012. 4. Judith Hurwitz, Alan Nugent, Dr. Fern Halper, Marcia Kaufman, Big Data for Dummies, John Wiley & Sons, Inc., 2013.
Syllabus Design: Dr.K.S.Kuppusamy, Associate Professor, PUDoCS	

Year	IV	Course Code: CSDS405 Course Title: Predictive Analytics	Credits	4
Sem.	VII		Hours	75
			Category	C
Course Prerequisites, if any	Basic problem solving and analyzing skills			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory) : 03 hrs Duration of ESA (Practical) : 03 hrs		
Course Outcomes	• Understand the basis of predictive analytics concepts • Formulate predictive model for a problem • Gain insights into the dataset through various exploratory data analysis techniques • Apply predictive modelling techniques to solve real-world problems • Assess the performance of predictive models			

Unit No.	Course Content	Hours
	Theory Component	
UNIT I	**Overview of Predictive Analytics** Analytics - Predictive Analytics - Business Intelligence - Predictive Analytics vs. Business Intelligence - Similarities between Business Intelligence and Predictive Analytics - Predictive Analytics vs. Statistics - Statistics and Analytics - Predictive Analytics vs. Data Mining – Uses - Challenges	9
UNIT II	**Setting up the Problem** Defining Data for Predictive Modeling - Defining the Target Variable - Defining Measures of Success for Predictive Models - Building Models First - Early Model Deployment - Case Study: Fraud Detection	9
UNIT III	**Data Understanding** Single Variable Summaries - Data Visualization in One Dimension – Histograms - Multiple Variable Summaries - Hidden Value in Variable Interactions: Simpson's Paradox - The Combinatorial Explosion of Interactions - Correlations - Data Visualization, Two or Higher Dimensions	9
UNIT IV	**Predictive Modeling** Decision Trees - The Decision Tree Landscape - Building Decision Trees - Decision Tree Splitting Metrics - Logistic Regression - Neural Networks - K-Nearest Neighbor - Naïve Bayes - Regression Models - Linear Regression	9
UNIT V	**Assessing Predictive Models** Batch Approach to Model Assessment - Percent Correct Classification - Rank-Ordered Approach to Model Assessment - Assessing Regression Models	9
	Practical Component	
Exercises	1. Explore the distribution of target variables and features using histograms, box plots and scatter plots 2. Analyze correlations between features using correlation matrices 3. Split the data into training and testing sets using techniques like cross-validation.	30

	4. Train various machine learning models such as linear regression, logistic regression, decision trees, support vector machines (SVM), k-nearest neighbours (KNN), etc. 5. Evaluate models using appropriate metrics such as accuracy, precision, recall, F1-score, ROC-AUC, etc. 6. Compare the performance of different models using cross-validation or holdout validation.	
Recommended Learning References		
Print Resources	1. Dean Abbott, Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst, Wiley, 1st Edition, 2014 2. Anasse Bari, Mohamed Chaouchi, Tommy Jung, Predictive Analytics for Dummies, Wiley, 2nd Edition, 2016	
Syllabus Design : Dr. M.Nandhini, Professor, PUDoCS		

Year	IV	**Course Code: CSDS406**	**Credits**	4
Sem.	VII	**Course Title: Data Mining**	**Hours**	75
			Category	C
Course Prerequisites, if any	Database Management Systems			
Internal Assessment Marks : 25	End Semester Marks : 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	**Students will be able to** • Gain a comprehensive understanding of data mining concepts • Acquire knowledge in data preprocessing techniques • Gain knowledge in pattern mining • Attain knowledge and skills in classification • Understand various clustering algorithms			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Overview and History - Data Mining - Types of data - Kinds of Patterns - Technologies Used - Applications - Major Issues in Data Mining – Data Objects and Attribute Types – Basic Statistical Descriptions of Data	9
Unit II	**Data Preprocessing & Data Warehouse** Data Preprocessing Overview - Data Cleaning - Data Integration - Data Reduction - Data Transformation- Data Warehouse: Basic Concepts - Data Cube and OLAP - Data Generalization by Attribute-Oriented Induction	9
Unit III	**Pattern Mining** Pattern Mining Concepts - Market Basket Analysis - Frequent Itemsets - Closed Itemsets and Association Rules- Frequent Itemset Mining Methods - Pattern Evaluation Methods	9
Unit IV	**Classification** Fundamentals - Decision Tree Induction - Bayes Classification - Rule Based Classification - Model Evaluation and selection - Techniques to Improve Classification Accuracy	9
Unit V	**Clustering** Cluster Analysis - Partitioning methods - Hierarchical methods - Agglomerative, Divisive hierarchical clustering - DBSCAN - Evaluation	9
	Practical Component	
Exercises	1. Perform preprocessing for the given dataset 2. Program to Integrate two datasets with common attributes 3. Program to transform categorical data into numerical format for analysis 4. Program to create a basic data cube and perform OLAP operations	30

	5. Implement the Apriori algorithm for mining frequent itemsets 6. Implement K-means clustering algorithm 7. Implement K-Medoids algorithm 8. Implement DBSCAN algorithm	
Recommended Learning Resources		
Print Resources	1. Jiawei Hen, Micheline Kambler, Jian Pie, Data Mining Concepts and Techniques, Morgan Kaufman, 4th Edition, 2022 2. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Pearson India Education Services Pvt. Ltd, 2016.	
Syllabus Design: Dr. P.Shanthi Bala, Professor, PUDoCS		

Year	IV	Course Code: CSDS407	Credits	4
Sem.	VII	Course Title : Text and Speech Analytics	Hours	75
			Category	C
Course Prerequisites, if any	Natural Language Processing			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	**Students will be able to** • Understand text and speech processing techniques • Ability to apply computational Linguistics concepts • Understand Markov Models • Understand different methods for extracting relevant features from text and speech data • Analyse the sentence structure and its semantics			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Linguistic Essentials** Part of Speech and Morphology - Phrase Structure - Semantics and Pragmatics - collocations - Frequency - Mean, variance – Hypothesis Testing	9
Unit II	**Statistical Inference & Disambiguation** N-gram models over sparse Data- Bins - statistical Estimators - Combining Estimators - Word sense disambiguation - supervised - dictionary based - unsupervised	9
Unit III	**Makov Models & POS Tagging** Markov Model - HMM - POS Tagging - Markov Model Tagger - HMM Tagger - Transformation based Learning of Tags - Tagging Accuracy - uses of Taggers	9
Unit IV	**Analyzing Sentence structure** Context free Grammar - Parsing - Dependency - Feature - structure Processing - Analyzing Meaning of sentences - propositional Logic - First order Logic - Discourse semantics	9
Unit V	**Automatic Speech Recognition and Text-to-Speech** Automatic Speech Recognition Task, Feature Extraction for ASR: Log Mel Spectrum, Speech Recognition Architecture, CTC, ASR Evaluation: Word Error Rate, TTS, Other Speech Tasks	9

Practical Component		
Exercises	1. Perform pre-processing (tokenization, script validation, stop word removal and stemming) of text 2. Perform Morphological analysis 3. Implement Markov model 4. Implement POS tagging 5. Implement chunking to extract noun phrases 6. Identify semantic relationship between the words from given text (use WordNet Dictionary) 7. Implement reference resolution algorithm 8. Perform Name Entity Recognition(NER) on given text 9. Implement a real life natural language application (use standard datasets)	30

Recommended Learning Resources	
Print Resources	1. Daniel Jurafsky, James H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Pearson Education India, 3rd Edition, 2023. 2. Christopher D. Manning and Hinrich Schuetze, Foundations of Statistical Natural Language Processing, MIT Press, 1999 3. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python, O'Reilly, 1st Edition,2009

Syllabus Design : Dr. P.Shanthi Bala, Professor, PUDoCS

SEMESTER-VIII

Year	IV	**Course Code: CSDS408** **Course Title: Machine Learning**	**Credits**	4
			Hours	75
Sem.	VIII		**Category**	C
Course Prerequisites, if any	Probability and Statistics			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Understand the basic concepts and types of Machine Learning • Prepare the data for ML model, train the model and evaluate the model's performance • Understand the fundamentals of features and feature engineering • Build a ML model with the appropriate supervised algorithm for the data • Build a ML model with the appropriate unsupervised algorithm for the data			

Unit No.	Course Content	Hours
	Theory Component	
UNIT I	**Introduction to Machine Learning** Human Learning - Machine Learning -Types - Supervised learning - Unsupervised Learning - Reinforcement Learning - Applications **Preparing to model** Types of data - structure - quality and remediation - pre-processing.	9
UNIT II	**Modelling and Evaluation** Selecting - Training - Model representation and interpretability - Performance evaluation **Feature Engineering** Introduction -- Transformation - Feature subset selection - Issues in high dimensional data – Feature selection - Key drivers - Measures - process - Approaches	9
UNIT III	**Supervised Learning - Classification** Introduction - Example - Model - Learning steps – Algorithms – k-Nearest neighbor - Decision tree - Random Forest model – Support Vector Machines.	9
UNIT IV	**Supervised Learning - Regression** Introduction - Example - Model - Algorithms - Simple and Multiple linear regression - Assumptions – Problems - Logistic Regression - Maximum Likelihood estimation.	9
UNIT V	**Unsupervised Learning** Introduction - Applications - Clustering - Types - Partitioning methods - Hierarchical clustering - Density-based methods - DBSCAN - Apriori algorithm for association rule learning.	9

Practical Component		
Exercises	1. Develop a Python script that uses a decision tree classifier for prediction 2. Develop a ML model that runs a random forest for classification 3. Create a Python program that uses SVM to classify images from the MNIST dataset. 4. Implement K-Means clustering to segment customers into groups based on their shopping data such as purchase history and customer demographics. 5. Implement a linear regression model. 6. Develop a program to perform multiple linear regression to predict house prices 7. Implement logistic regression to classify emails as spam or not spam.	30
Recommended Learning References		
Print Resources	1. Saikat Dutt, Chandramouli.S, Amit Kumar Das., Machine Learning, Pearson, 1st Edition, 2018. 2. Alpaydin, E., Introduction to Machine Learning, MIT Press, 4th Edition, 2020	
Syllabus Design : Dr. M.Nandhini, Professor, PUDoCS		

Year	IV	Course Code: CSDS409 Course Title: Healthcare Data Analytics	Credits	4
Semester	VIII		Hours	75
			Category	C
Course Prerequisites, if any	Foundations of Data Science			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Explain the different types of healthcare data sources, including Electronic Health Records (EHRs), biomedical images, and clinical text data. • Describe the basic structure and components of EHRs and coding systems used in healthcare. • Discuss the benefits and challenges associated in using healthcare data for analytics. • Describe the role of Natural Language Processing (NLP) in processing clinical text data and extracting meaningful information. • Identify potential challenges and ethical considerations in healthcare data analytics.			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Healthcare Data Sources and Basic Analytics - Advanced Data Analytics for Healthcare - Applications and Practical Systems for Healthcare - Resources for Healthcare Data Analytics.	9
Unit II	**Electronic Health Records** Survey - Components of EHR - Coding Systems - Benefits of EHR - Barriers to Adopting EHR - Challenges of Using EHR Data - Phenotyping Algorithms	9
Unit III	**Biomedical image analysis** Introduction - Biomedical Imaging Modalities - Object Detection Image Segmentation - Image Registration - Feature Extraction	9
Unit IV	**Biomedical signal analysis** Introduction - Types of Biomedical Signals - ECG Signal Analysis - Denoising of Signals – Recent Trends	9
Unit V	**Natural language processing and data mining for clinical text** Introduction - Mining Information from Clinical Text - Challenges for Processing Clinical Reports - Clinical Applications	9
	Practical Component	
Exercises	1. Obtain a sample EHR dataset and explore its structure and contents (patient demographics, diagnoses, medications, and procedures) 2. Create visualizations (e.g., histograms, bar charts) to illustrate the distribution of key variables in the EHR dataset. 3. Cleanse the EHR dataset by addressing common data quality issues such as missing values, outliers and inconsistent formatting. 4. Implement the phenotyping algorithm and apply it to the EHR dataset. Evaluate the algorithm's performance in terms of sensitivity, specificity, and accuracy.	30

	5. Implement an image segmentation algorithm (e.g., U-Net, Mask R-CNN) to segment specific structures or lesions in biomedical images. Train the model using annotated training data and evaluate its performance on a separate validation set. 6. Acquire a biomedical image dataset (e.g., MRI brain scans, X-ray images) and explore its characteristics, including image dimensions, pixel values, and annotation labels . Provide a summary of the dataset properties and visualize sample images along with their corresponding annotations. 7. Apply denoising techniques such as median filtering, wavelet denoising, adaptive filtering on a noisy biomedical signal (e.g., ECG or EEG). Evaluate the effectiveness of each method in improving signal quality.	
Recommended Learning Resources		
Print Resources	1. Chandan K. Reddy, Healthcare Data Analytics , CRC Press, Taylor and Francis Group, 1st Edition, 2015. 2. Philippe J. Giabbanelli , Vijay K. Mago, Advanced Data Analytics in Health: 93 (Smart Innovation, Systems and Technologies), Springer, 1st Edition, 2018.	
Syllabus Design Dr. R.Sunitha, Associate Professor, PUDoCS		

Year	IV	Course Code: CSDS410 Course Title: Business Analytics	Credits	4
Sem.	VIII		Hours	75
			Category	C
Course Prerequisites, if any	Foundations of Data Science			
Internal Assessment Marks : 25	End Semester Marks : 75	Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.		
Course Outcomes	• Understand the various features of Spreadsheets and use them for business analysis • Understand and describe data and its distributions • Describe the various forecasting techniques and perform spreadsheet modelling and analysis. • Make decisions by optimizing business problems • Use Six Sigma for data driven decision making			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Business Analytics:** Introduction - Business Analytics: The Science of Data-Driven Decision Making - Descriptive Analytics - Predictive Analytics - Prescriptive Analytics – Techniques - Big Data Analytics - Web and Social Media Analytics - Analytics Capability Building - Challenges in Data-Driven Decision Making and Future.	9
Unit II	**Descriptive Analytics** Visualization and Exploring data - Descriptive Statistical Measures - Probability Distributions and Data modelling - Sampling and Estimation - Statistical Inference	9
Unit III	**Predictive Analytics** Trendlines and Regression analysis - Forecasting Techniques - Data mining - Spreadsheet modelling and Analysis - Monte Carlo Simulation and Risk Analysis	9
Unit IV	**Prescriptive Analysis** Linear Optimization (LO) - Applications of LO - Integer Optimization - Decision Analysis	9
Unit V	**Six Sigma** Introduction, Origins of Six Sigma, Three-Sigma versus Six-Sigma Process, Cost of Poor Quality, Sigma Score, Industrial Applications of Six Sigma.	9
	Practical Component	
Exercises	Using MS-Excel /Power BI 1. Perform data import/export operations for different file formats. 2. Perform statistical operations – Mean, Median, Mode and Standard deviation, Variance, Skewness, Kurtosis 3. Perform Z-test, T-test & ANOVA	30

	4. Perform data pre-processing operations i) Handling Missing data ii) Normalization 5. Perform bivariate and multivariate analysis on the dataset. 6. Apply and explore various plotting functions on the data set.	
Recommended Learning Resources		
Print Resources	1. U. Dinesh Kumar, Business Analytics: The Science of Data-Driven Decision Making, Wiley, 2nd Edition, 2021. (Unit I,V) 2. R. Evans James, Business Analytics , Pearson, 2nd Edition,2017 (Unit II,III,IV)	
Syllabus Design Dr. R.Sunitha, Associate Professor, PUDoCS		

Year	IV	**Course Code: CSDS411** **Course Title : Social Network Analysis**	**Credits**	4
Sem.	VIII		**Hours**	75
			Category	C
Course Prerequisites, if any	Foundations of Data Science			
Internal Assessment Marks : 25	End Semester Marks : 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Describe the levels of Social networks and network measures. • Outline various network growth models and different rank methods • Illustrate different community structures and link prediction models • Explain different cascade behavior in networks. • Analyse the different applications of Social networks			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Introduction to Social Networks- Applications- Levels of Social Network Analysis – Different Graph Visualization Tools **Network measures** Network Basics, Node Centrality, Assortativity, Transitive and Reciprocity, Similarity, Degeneracy	9
Unit II	**Network growth models** Properties of Real World Networks, Random Network Model, Ring Lattice Model, Watts Strogatz Model, Preferential Attachment Model, Price Model, Local-World Network Growth Model **Link analysis** Applications, Signed Networks, Strong and Weak Ties, Link Analysis and Algorithms, Page Rank, Personalized Page Rank, Divrank, Simrank, PathSim	9
Unit III	**Community structure in networks** Applications, Types of Communities, Community Detection Methods, Disjoint Community Detection, Overlapping Community Detection, Local Community Detection, Community Detection vs Community Search, Evaluation.	9
Unit IV	**Link prediction** Applications, Temporal Changes in a Network, Problem Definition, Evaluating Link Prediction Networks, Heuristic Models, Probabilistic Models, Supervised Random Walk, Information-Theoretic Model	9
Unit V	**Cascade behavior and network effects** Preliminaries, Cascade Model, case study, Probabilistic Cascades, Epidemic Models, Independent Cascade Models, Cascade Prediction. **Application and Case Study -**Recommendation System	9

Practical Component Suggested Open Source tools : Gephie, Cytoscape Suggested Datasets :		
Exercises	1. Implement different types of centrality measures using graph analysis tools 2. Implement the page rank algorithm 3. Program to identify triadic closure, strong and weak ties logics, homophily 4. Program to identify different forms of structural balance and visualization of socio-affiliation network 5. Study and analysis of different community detection algorithms 6. Analysis of network evolution using SNA tools	30
Recommended Learning Resources		
Print Resources	1. Tanmoy Chakraborty, Social Network Analysis, Wiley, 2021. 2. Song Yang, Franziska B. Keller and Lu Zheng, Social Network Analysis: Methods and Examples, Sage Publications, 1st Edition, 2017.	
Syllabus Design Dr. R.Sunitha, Associate Professor, PUDoCS		

Year	IV	**Course Code: CSDS412** **Course Title: Deep Learning**	**Credits**	4
			Hours	75
Sem.	VIII		**Category**	C
Course Prerequisites, if any	Machine Learning			
Internal Assessment Marks: 25	End Semester Marks: 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Understand the basic architecture and workings of neural networks • Understand the fundamentals of deep neural networks • Understand the architectures and workings of deep networks • Build the model for data variants using deep network • Build and train CNN and RNN deep learning architectures			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Foundations of Neural Networks and Deep Learning** Neural Networks - Training Neural Networks - Activation Functions - Loss Functions - Hyperparameters	9
Unit II	**Fundamentals of Deep Networks** Defining Deep Learning- Common Architectural Principles of Deep Networks - Parameters- Layers - Activation Functions - Loss Functions - Optimization Algorithms - Hyperparameters - Building Blocks of Deep Networks - RBMs- Autoencoders- Variational Autoencoders	9
Unit III	**Major Architectures of Deep Networks** Unsupervised Pretrained Networks - Convolutional Neural Networks (CNNs) - Architecture - Input, Convolutional, Pooling, fully connected Layers - Applications - Recurrent Neural Networks (RNN) - Modeling the Time Dimension - 3D Volumetric Input - Architecture - LSTM Networks	9
Unit IV	**Building Deep Networks** Matching Deep Networks to the Right Problem - Modeling CSV Data with Multilayer Perceptron Networks - Modeling Handwritten Images Using CNNs - Modeling Sequence Data using RNN	9
Unit V	**Tuning Deep Networks** Concepts: Matching Input Data and Network Architectures- Relating Model Goal and Output Layers - Working with Layer Count, Parameter Count, and Memory - Feed-Forward Multilayer Neural Networks - Controlling Layer and Parameter Counts - Weight Initialization Strategies- Using Activation Functions - Applying Loss Functions - Understanding Learning Rates - Applying Methods of Optimization - Controlling Epochs and Mini-Batch Size	9

	- Regularization - Max-Norm Regularization - Dropout- Dealing with Overfitting	
	Practical Component	
Exercises	1. Implement a simple perceptron model and train it to perform binary classification on a given dataset. Use the sigmoid activation function and gradient descent for training. 2. Build a multilayer feed-forward neural network from scratch. Train the network using the backpropagation algorithm on a given dataset 3. Implement and train the CNN on the MNIST dataset for handwritten digit classification. 4. Implement RNN for text generation 5. Fine-tune a pre-trained CNN model	30
Recommended Learning References		
Print Resources	1. Josh Patterson and Adam Gibson, Deep Learning – A Practitioner's Approach, O'Reilly Media, 1st Edition, 2017. 2. Nikhil Buduma and Nicholas Locascio, Fundamentals of Deep Learning: Designing Next Generation Machine Intelligence Algorithms, O'Reilly Media, 1st Edition, 2017.	
Syllabus Design : Dr. M.Nandhini, Professor, PUDoCS		

Year	IV	**Course Code: CSDS413** **Course Title : Time Series Analysis**	**Credits**	4
Sem.	VIII		**Hours**	75
			Category	C
Course Prerequisites, if any	Deep Learning			
Internal Assessment Marks: 25	End Semester Marks: 75		Duration of ESA (Theory) : 03 hrs. Duration of ESA (Practical) : 03 hrs.	
Course Outcomes	• Learn concepts related to time series forecasting • Understand the characteristics of time series data • Learn the widely used time series models • Learn time series data visualisation techniques • Understand deep learning architectures used for time series forecasting			

Unit No.	Course Content	Hours
	Theory Component	
Unit I	**Introduction** Time series analysis - forecasting - characteristics - python for time series analysis - deep learning with PyTorch - time series as deep learning problem - Univariate - multivariate - single step vs multistep	9
Unit II	**Preprocessing** Normalisation - trend removal - differencing - sliding window - loss function - training - validating - testing - model optimisation	9
Unit III	**Exploratory Data Analysis For Time Series** Data visualisation - Histograms - Scatterplots - Understanding stationarity - Understanding correlation - 1D,2D,3D visualisations	9
Unit IV	**Autoregressive And Automated Methods** Autoregression - Moving average - Autoregressive moving average - automated machine learning - clustering - decision tree	9
Unit V	**Recurrent Neural Networks** Architecture - importing dataset - Perform training and evaluation of RNN model - Develop GRUs and LSTMs for Time Series Forecasting	9
	Practical Component	
Exercises	Perform time series analysis for 1. Health care data 2. Stock price 3. Financial data 4. Stream data 5. Weather data 6. Sales data	30

Recommended Learning Resources	
Print Resources	1. Francesca Lazzeri, Machine Learning for Time Series Forecasting with Python, Wiley, 2020 2. Ivan Gridin, Time Series Forecasting Using Deep Learning, BPB publications, 2021 3. Aileen Nielsen, Practical Time Series Analysis, O'Reilly, 2019 4. Charu C. Aggarwal, Neural Networks and Deep Learning, Springer International Publishing AG, 2018
Syllabus Design: Dr. V.Uma, Associate Professor, PUDoCS	

Year	IV	**Course Code: CSDS414** **Course Title: Natural Language Processing**	**Credits**	4
			Hours	75
Sem.	VIII		**Category**	C
Course Prerequisites, if any	Basic understanding of linguistics, syntax and semantics			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory) : 03 hrs Duration of ESA (Practical) : 03 hrs		
Course Outcomes	• Understand the theoretical foundations of NLP • Describe N-gram language model • Perform classification tasks using naive Bayes classifiers and logistic regression • Understand vector semantics and embeddings • Implement neural language models for any concerned applications			

Unit No.	Course Content	Hours
	Theory Component	
UNIT I	**Regular Expressions, Text Normalization, Edit Distance** Regular Expressions - Words - Corpora - Word Tokenization - Word Normalization, Lemmatization and Stemming - Sentence Segmentation - Minimum Edit Distance.	9
UNIT II	**N-gram Language Models** N-Grams - Evaluating Language Models – Sampling Sentences from a Language Model - Generalization and Zeros - Smoothing. **Sequence Labelling for Parts of Speech and Named Entities** English Word Classes - Part-of-Speech Tagging - Named Entities and Named Entity Tagging – HMM or Part-of-Speech Tagging - Conditional Random Fields (CRFs).	9
UNIT III	**Naive Bayes, Text Classification and Sentiment** Naive Bayes Classifiers - Training the Naive Bayes Classifier - Optimizing for Sentiment Analysis - Naive Bayes as a Language Model – Evaluation - Test sets and Cross-validation.	9
UNIT IV	**Vector Semantics and Embeddings** Lexical Semantics- Vector Semantics- Words and Vectors - Cosine for measuring similarity – TF-IDF – PMI – Applications of TF_IDF and PMI – Word2Vec	9
UNIT V	**Neural Networks and Neural Language Models** Units - The XOR problem- Feed Forward Neural Networks - Feedforward Networks for NLP: Classification - Training Neural Nets - Feedforward Neural Language Modeling - Training the Neural Language Model. **NLP Applications** Chatbots & Dialogue Systems - Design	9

Practical Component		
Exercises	Implement the following 1. Separate the sentences and tokens of the given text document. 2. Remove all the stop words ('a' , 'the', 'was'…) from the text 3. Perform stemming/convert each token to its root form in the given text and lemmatization on the given text 4. Correct the spelling errors in the text given. 5. Detect if a text is a positive or negative sentiment. 6. Find the dependencies of all the words in the given text. 7. Build a text classifier with available training data using textblob library in python 8. Design a simple chatbot for a domain of your choice	30
	Recommended References	
Print Resources	1. Daniel Jurafsky and James H. Martin, Speech and Language Processing - An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Prentice Hall, 3rd Edition, 2024. 2. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python, O'Reilly Media, 1st Edition, 2009.	
Syllabus Design : Dr. M.Nandhini, Professor, PUDoCS		

MULTIDISCIPLINARY COURSES

Year		**Course Code: COMS101**	**Credits**	3
			Hours	75
Sem.		**Course Title: Introduction to Python Programming**	**Category**	A
Course Prerequisites, if any	Problem-solving skills			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs. Duration of ESA (Practical): 03 hrs.		
Course Outcomes	• Understand Python programming constructs • Learn about different data structures in Python • Write programs using functions • Explore the use of Python modules and packages • Perform Visualization using Python package			

Unit No.	Course Content	Hours
Unit I	**Introduction** **Python Basics:** Working – Identifiers – Comments – Types – Operations – Buit-in, library functions **Strings:** Accessing – Properties – Operations **Control-flow Instructions:** Decision Control – logical operators – conditional expressions Repetition control instruction – break and continue – *pass* Statement	9
Unit II	**Console Input/Output:** Console Input – Console Output – Formatted printing **Lists** Definition – Accessing – Operations – Methods – Varieties – Comprehension **Tuples** Definition – Accessing – Operations – Varieties – Comprehension – Conversion – Iterators and Iterables - zip()	9
Unit III	**Sets** Definition – Accessing – Operations – Functions – Mathematical set operations – Updating set operations **Dictionaries** Definition – Accessing – Operations – Functions – Nested Dictionary	9
Unit IV	**Functions** Definition – Communication – Types – Unpacking – Lambda, Recursive functions **Modules and Packages** Creation and importing	9
Unit V	**Exception handling** Syntax errors – handling exceptions – *try-except* – user-defined exceptions – *else, finally* blocks – Tips **Visualization** - Matplotlib package – Plotting Graphs	9

Recommended Learning Resources	
Print Resources	1. Aditya Kanetkar, Yashavant Kanetkar, Let us Python, BPB Publisher, 6th Edition, 2023

Syllabus Design: Dr. R.Sunitha, Associate Professor, PUDoCS

Year		**Course Code: COMS102**	**Credits**	3
			Hours	75
Sem.		**Course Title: Foundations of Information Technology**	**Category**	A
Course Prerequisites, if any	Basic knowledge of Computers			
Internal Assessment Marks: 25	End Semester Marks: 75	Duration of ESA (Theory): 03 hrs.		
Course Outcomes	• Familiarize the fundamentals of Information Technology. • Understand the management of hardware and software • Describe the basics of networking • Discuss about data management and security aspects of data • Ability to troubleshoot computer systems			
Unit No.	**Course Content**		**Hours**	
Unit I	**Introduction** Overview of IT – Computer Basics – Software fundamentals – Networks & Internet – IT ethics and policies		15	
Unit II	**Hardware and Software Management** Computer Assembly and maintenance - Operating Systems – Software installation and maintenance – Virtualization, Cloud Computing		15	
Unit III	**Networking Essentials** Network Fundamentals - Hardware – Protocols and services – Wireless Networking – Security		15	
Unit IV	**Data Management and Security** Data and fundamentals of Database – Data Backup and recovery – Cyber Security – Encryption and Cryptography		15	
Unit V	**IT Support and Troubleshooting** Help desk and IT support – Troubleshooting methodologies – Diagnostic tools and utilities – Future trends in IT		15	
Recommended Learning Resources				
Print Resources	1. Floyd Fuller, Brian Larson, Computers: Understanding Technology, EMC Paradigm, 4th Edition, 2011. 2. Mike Meyers , CompTIA A+ Certification All-in-One Exam Guide, McGraw-Hill Education, 11th Edition, 2023. 3. Jeffrey S. Beasley, Piyasat Nilkaew, Networking Essentials, Prentice Hall Certification, 3rd Edition, 2012. 4. Charles J. Brooks, Christopher Grow, Philip Craig, and Donald Short , Cybersecurity Essentials, Sybex Publisher, 1st Edition, 2018.			
Syllabus Design: Dr. R.Sunitha, Associate Professor, PUDoCS				