



**Nagarjuna College of Engineering &
Technology, Bengaluru**

An Autonomous Institute, Affiliated to VTU Belagavi

2021 Batch

Scheme & Syllabus
of
VI Sem

As per the NEP 2020 Guidelines,
Choice-Based Credit System
&
Outcome-Based Education

CSE (Data Science)

Academic Year 2023-2024

Vision

To prepare the next generation practitioners and researcher for data centric world by bringing together interdisciplinary faculty across the globe.

Mission

M1: To provide Skill Based Education to master the students in problem solving and analytical skills to enhance their niche expertise in the field Data Science

M2: To educate the students with latest technologies to update their knowledge in the field of Data Science

M3: To enable students to experience the Content Based Learning with premier quality data science education, research and industrial collaboration

M4: To enable students to become leaders in the Industry and Academia Nationally as well as internationally

M5: To guide students in research on Data Science, with the aim of having an ethical impact on society by tackling societal grand challenges

PROGRAM OUTCOMES (POs): Graduates of the Computer Science and Engineering – Data Science Program will be able to achieve the following

POs:

PO1: Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and Computer Science and Engineering principles to the solution of complex problems in Computer Science and Engineering.

PO2: Problem Analysis: Identify, formulate, research literature, and analyses complex Computer Science and Engineering problems reaching substantiated conclusions using first principles of mathematics and engineering sciences.

PO3: Design/Development of Solutions: Design solutions for complex Computer Science and Engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4: Conduct investigations of Complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions related to Computer Science and Engineering problems.

PO5: Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex Computer Science and Engineering activities with an understanding of the limitations.

PO6: The Engineer and Society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional Computer Science and Engineering practice.

PO7: Environment and Sustainability: Understand the impact of the professional Computer Science and Engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8: Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the Computer Science and Engineering practice.

PO9: Individual and Team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10: Communication: Communicate effectively on complex Computer Science and Engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: Project Management and Finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage Computer Science and Engineering projects and in multidisciplinary environments.

PO12: Life Long Learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Specific Outcome (PSO)

PSO1: Ability to analyse complex computing issues and apply the principles to achieve related solution.

PSO2: Ability to design, implement and evaluate computing based solutions to meet range of computing requirements based in the data science.

PSO3: Ability to effectively communicate within diverse work group related to professional framework.

Program Educational Objectives (PEOs)

PEO 1: To make students competent for higher studies and employable, to meet industrial requirements.

PEO 2: To develop students having core competence in science, mathematics and fundamentals of Data Science to address ever changing industrial requirements globally.

PEO 3: To create academically conducive environment to learn engineering skills in the domains such as Data Analytics, Data Modelling, Data Visualization and Allied Technologies.

PEO 4: To enrich students with professional ethics, leadership qualities, and entrepreneurial skills.

PEO 5: An ability to engage in lifelong learning for effective adaptation to technological developments.

NAGARJUNA COLLEGE OF ENGINEERING & TECHNOLOGY, BENGALURU

B.E. in CSE (Data Science)

Scheme of Teaching and Examinations 2021

**Outcome-Based Education (OBE) and Choice Based Credit System
(CBCS)**

(Effective from the academic year 2021 - 22)

VI SEMESTER

Sl. No	Course and Course Code	Course Title	Teaching Department (TD) and Question Paper Setting Board (PSB)	Teaching Hours /Week				Examination				Credits
				Theory Lecture	Tutorial	Practical/ Drawing	Self-Study	Duration in hours	CIE Marks	SEE Marks	Total Marks	
				L	T	P	S					
1	PCC 21CDT61	Cloud Computing	Any CS Board Department	3	0	0	0	03	50	50	100	3
2	IPCC 21CDI62	Statistical Inference for Data Science		3	0	2	0	03	50	50	100	4
3	PCC 21CDT63	Big Data Analytics		3	0	0	0	03	50	50	100	3
4	PEC 21XX64x	Professional Elective Course-I		3	0	0	0	03	50	50	100	3
5	OEC 21XX65x	Open Elective Course-I	Concerned Department	3	0	0	0	03	50	50	100	3
6	PCCL 21CDL66	Big Data Analytics Lab	Any CS Board Department	0	0	2	0	03	50	50	100	1
7	MP 21CDP67	Mini Project		Two contact hours /week for interaction between the faculty and students.				--	100	--	100	2
8	INT 21INT68	Innovation/Entrepreneurship /Societal Internship	Completed during the intervening period of IV and V semesters.				--	100	--	100	3	
Total									500	300	800	22

Professional Elective - I

21CDT641	Digital Image Analytics	21CDT643	Agile Methodologies
21CDT642	Advanced Algorithms	21CDT644	Business Economics

Open Electives – I offered by the Department to other Department students

21CDT651	Introduction to JAVA Programming	21CDT654	Introduction to Database Management System
21CDT652	Introduction to Cyber Security	21CDT655	Introduction to Data Science
21CDT653	Artificial Intelligence & Expert System		

Note: HSMC: Humanity and Social Science & Management Courses, **IPCC:** Integrated Professional Core Course, **PCC:** Professional Core Course, **PEC:**

Professional Elective Courses, **OEC**–Open Elective Course, **MP** –Mini Project, **INT** –Internship.

L –Lecture, **T** – Tutorial, **P** - Practical / Drawing, **S** – Self Study Component, **CIE:** Continuous Internal Evaluation, **SEE:** Semester End Examination.

Integrated Professional Core Course (IPCC): Refers to Professional Theory Core Course Integrated with Practical of the same course. Credit for IPCC can be 04 and its Teaching – Learning hours (L : T : P) can be considered as (3 : 0 : 2) or (2 : 2 : 2). The theory part of the IPCC shall be evaluated both by CIE and SEE. The practical part shall be evaluated by CIE only and there shall be no SEE. For more details, the regulation governing the Degree of Bachelor of Engineering /Technology (BE/B.Tech) 2021-22 may be referred.

Professional Elective Courses (PEC):

A professional elective (PEC) course is intended to enhance the depth and breadth of educational experience in the Engineering and Technology curriculum. Multidisciplinary courses that are added supplement the latest trend and advanced technology in the selected stream of engineering. Each group will provide an option to select one course out of five courses. The minimum students' strength for offering professional electives is 10. However, this conditional shall not be applicable to cases where the admission to the programme is less than 10.

Mini-project work: Mini Project is a laboratory-oriented course which will provide a platform to students to enhance their practical knowledge and skills by the development of small systems/applications.

Based on the ability/abilities of the student/s and recommendations of the mentor, a single discipline or a multidisciplinary Mini- project can be assigned to an individual student or to a group having not more than 4 students.

CIE procedure for Mini-project:

(i) Single discipline: The CIE marks shall be awarded by a committee consisting of the Head of the concerned Department and two faculty members of the Department, one of them being the Guide. The CIE marks awarded for the Mini-project work shall be based on the evaluation of project report, project presentation skill, and question and answer session in the ratio of 50:25:25. The marks awarded for the project report shall be the same for all the batch mates.

(ii) Interdisciplinary: Continuous Internal Evaluation shall be group-wise at the college level with the participation of all the guides of the project. The CIE marks awarded for the Mini-project, shall be based on the evaluation of project report, project presentation skill, and question and answer session in the ratio 50:25:25. The marks awarded for the project report shall be the same for all the batch mates.

No SEE component for Mini-Project.

CLOUD COMPUTING			
Course Code	21CDT61	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Prerequisite: Basic Network Concepts			
Course Objectives:			
<ul style="list-style-type: none"> • Discuss the concepts, characteristics, delivery models and benefits of cloud computing. • Explore the key technical, organizational and compliance challenges of cloud computing. • Grasp the concepts of virtualization efficiently. • Explore the security issues that arise from cloud computing architectures intended for delivering Cloud based enterprise IT services. 			
Teaching-Learning Process (General Instructions)			
<p>These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.</p> <ol style="list-style-type: none"> 1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes. 2. Use of Video/Animation to explain functioning of various concepts. 3. Encourage collaborative (Group Learning) Learning in the class. 4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking. 5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it. 6. Introduce Topics in manifold representations. 7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them. 8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding. 			
Module – I			
<p>Introduction to Cloud Computing : Cloud Computing at a Glance, Defining a Cloud, Cloud deployment models, Cloud Computing Reference Model, Characteristics and Benefits, Challenges Ahead, Building Cloud Computing Environments- Application Development ,Infrastructure and System Development. Computing Platforms and Technologies-Google App Engine, Microsoft Azure, Hadoop, Force.com and Salesforce.com, Manjra soft Aneka. 08 Hours</p>			
Module – II			
<p>Virtualization- Characteristics of virtualized environments, Taxonomy of virtualization techniques- Execution virtualization, Hardware-level virtualization, Virtualization and cloud computing Cloud Computing Architecture: Introduction, Cloud Reference Model, Architecture, Infrastructure / Hardware as a Service, Platform as a Service, Software as a Service, Types of Clouds, Public Clouds, Private Clouds, Hybrid Clouds, Community Clouds. 08 Hours</p>			
Module – III			
<p>Aneka Cloud Application Platform- Framework overview- Fabric services, Foundation services, Application services. Building Aneka clouds- Infrastructure organization, Logical organization, Private cloud deployment mode, Public cloud deployment mode, Hybrid cloud deployment mode. 08 Hours</p>			
Module – IV			
<p>Concurrent Computing- Introducing parallelism for single-machine computation, Programming applications with threads- Thread APIs, POSIX Threads, Threading support in java and .NET.</p>			

Techniques for parallel computation with threads- Domain decomposition, Functional decomposition
thread programming model- Aneka thread vs. common threads, Interface compatibility, Thread life cycle, Thread synchronization, Thread priorities, Type serialization. **08 Hours**

Module – V

Cloud Applications- Scientific applications: Healthcare: ECG analysis in the cloud, Biology: protein structure prediction, Biology: gene expression data analysis for cancer diagnosis, Geoscience: satellite image processing

Business and consumer applications- CRM and ERP, Social networking. Media applications

08 Hours

Teaching-Learning Process

Chalk and Talk, Power point presentation, flip teaching, YouTube videos

COURSE OUTCOMES:

After successful completion of this course, the students will be able to:

CO1: Compare the strengths and limitations of cloud computing

CO2: Identify the architecture, infrastructure and delivery models of cloud computing

CO3: Demonstrate the working of VM and VMM on any cloud platforms(public/private), and run a software service on that.

CO4: Identify the known threats, risks, vulnerabilities and privacy issues associated with Cloud based IT services.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Suggested Learning Resources:

Text Books:

1. Rajkumar Buyya, Amir Vahid Dastjerdi, "Internet of Things: Principles and Paradigms", Elsevier, 2016.
2. Internet of things in the cloud -A Middleware Perspective, Honbo Zhou, CRC Press, Taylor & Francis Group
3. R. Chandrasekaran, "Essentials of Cloud computing", 2nd Edition, Chapman and Hall/CRC, 2015.
4. Amita Kapoor, "Hands on Artificial intelligence for IoT", 1st Edition, Packt Publishing, 2019.
5. David Etter, "IoT Security: Practical Guide Book", Create Space Independent Publishing Platform, 2016.

Reference Books:

1. John Soldatos, "Building Blocks for IoT Analytics", River Publishers, 2016.
2. John E. Rossman, "The Amazon way on IoT", Volume 2, John E. Rossman publication, 2016.

E-Books:

1. [http://index-of.co.uk/Cloud-Computingbooks/Essentials%20of%20cloud%20computing%20\(2015\).pdf](http://index-of.co.uk/Cloud-Computingbooks/Essentials%20of%20cloud%20computing%20(2015).pdf)
2. <https://www.iottechexpo.com/2018/11/iot/the-iot-analytics-lifecycle-from-generating-data-topredicting-the-future-losant/>

MOOC:

1. <https://www.coursera.org/learn/cloud-iot-platform>
2. <https://www.udemy.com/course/iothacking1/>

STATISTICAL INFERENCE FOR DATA SCIENCE

Course Code	21CDI62 (IC)	CIE Marks	50
Teaching Hours/Week (L: T: P: S) (3:0:2:0)	Credits (3:0:1:0)	SEE Marks	50
Total Hours of Pedagogy	40 hours Theory + 13 Lab slots	Total Marks	100
Credits	04	ExamHours	03

Prerequisite: Python for Data Science/ R for Data Science

Course Objectives:

This course will enable students to:

- This course aims at giving the foundation knowledge of Probability and Statistical Inference.
- Statistics provides tools and methods to find structure and to give deeper data insights.
- Be able to restate an investigative question in terms of a statistical model or algorithm.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teacher can use to accelerate the attainment of the various course outcomes.

- Lecture method (L) does not mean only traditional lecture method, but a different type of teaching methods may be adopted to develop the outcomes.
- Show Video/animation films to explain the different concepts of Linear Algebra & Signal Processing.
- Encourage collaborative (Group) Learning in the class.
- Ask at least three HOTS (Higher Order Thinking) questions in the class, which promotes critical thinking.
- Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop thinking skills such as the ability to evaluate, generalize, and analyze information rather than simply recall it.
- Topics will be introduced in a multiple representation.
- Show the different ways to solve the same problem and encourage the students to come up with their own creative ways to solve them.
- Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.
- Adopt Flipped class technique by sharing the materials / Sample Videos prior to the class and have discussions on the topic in the succeeding classes.
- Give Programming Assignments.

Syllabus

Module – I

EXPLORATORY ANALYSIS: Elements of Structured Data, Rectangular Data, Measure of Central tendency-Mean, Median and Robust Estimates, Example, Estimates of Variability-Standard Deviation, Exploring the Data Distribution-Frequency Tables and Histograms Correlation.

08 Hours

Module – II

Random Sampling and Sample Bias, Selection Bias, The Bootstrap, Confidence Intervals, Normal distribution, Binomial distribution, Poisson distribution

08 Hours

Module – III																																							
HYPOTHESIS: A/B Testing, Hypothesis Tests- null, one-way, two-way, Statistical Significance and p-Values:z-test, p-Value, Type 1 & 2 errors,t-tests, multiple testing, degrees of freedom, ANOVA, Chi-Square Tests, Power and Sample Size.																																							
08 Hours																																							
Module – IV																																							
REGRESSION AND PREDICTION: Simple Linear Regression, Multiple Linear Regression, Confidence and Prediction Intervals, Multicollinearity, Polynomial and Spline Regression.																																							
08 Hours																																							
Module – V																																							
Naive Bayes, Discriminant Analysis, Logistic Regression, Evaluating Classification Models.																																							
08 Hours																																							
Teaching-Learning Process for all modules	Chalk and Talk, Power point presentation, flip teaching,YouTube videos																																						
PRACTICAL COMPONENTS																																							
Sl. No	Experiments																																						
1	Rainfall prediction data set – draw correlation between the features																																						
2	Find the outliers in the Housing Price dataset																																						
3	For a given dataset, display a chosen feature using different mean values																																						
4	Display the confidence interval of a chosen feature based on a sample																																						
5	Perform t-test on a feature in a dataset																																						
6	Create Boxplots for different groups of a feature																																						
7	Create a Linear Regression model for a dataset and display the error measures																																						
8	Chose a dataset with categorical data and apply linear regression model																																						
9	Apply Naive Bayes algorithm on a dataset and estimate the accuracy																																						
10	Apply Logistic Regression algorithm on a dataset and estimate the accuracy																																						
Course Outcomes:																																							
On completion of this course, students will be able to:																																							
CO 1: Perform exploratory analysis on the datasets																																							
CO 2: Understand the various distribution and sampling																																							
CO 3: Perform Hypothesis Testing on datasets																																							
CO 4: Apply statistical inference for Regression																																							
CO 5: Apply statistical inference for Classification																																							
Assessment Details (both CIE and SEE)																																							
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th colspan="2" style="text-align: center;">Component</th> <th colspan="2" style="text-align: center;">Weightage (%)</th> </tr> </thead> <tbody> <tr> <td rowspan="3" style="text-align: center;">CIE's</td> <td>CIE 1 5th week</td> <td style="text-align: center;">20</td> <td rowspan="3" style="text-align: center;">60</td> <td rowspan="3" style="text-align: center;">Average of 3 tests for 20 marks</td> </tr> <tr> <td>CIE 2 10th week</td> <td style="text-align: center;">20</td> </tr> <tr> <td>CIE 3 15th week</td> <td style="text-align: center;">20</td> </tr> <tr> <td style="text-align: center;">AAT's</td> <td>AAT-1 10th week</td> <td colspan="2"></td> <td style="text-align: center;">10</td> </tr> <tr> <td></td> <td>Lab Test</td> <td style="text-align: center;">30</td> <td colspan="2" style="text-align: center;">Reduced to 10</td> </tr> <tr> <td></td> <td>Lab Record</td> <td style="text-align: center;">20</td> <td colspan="2" style="text-align: center;">10</td> </tr> <tr> <td colspan="5" style="text-align: center;">Continuous Internal Evaluation Total Marks : 100. Reduced to 50 Marks</td> </tr> <tr> <td colspan="5" style="text-align: center;">Semester End Examination (SEE) Total Marks :100. Reduced to 50 Marks</td> </tr> </tbody> </table>		Component		Weightage (%)		CIE's	CIE 1 5 th week	20	60	Average of 3 tests for 20 marks	CIE 2 10 th week	20	CIE 3 15 th week	20	AAT's	AAT-1 10 th week			10		Lab Test	30	Reduced to 10			Lab Record	20	10		Continuous Internal Evaluation Total Marks : 100. Reduced to 50 Marks					Semester End Examination (SEE) Total Marks :100. Reduced to 50 Marks				
Component		Weightage (%)																																					
CIE's	CIE 1 5 th week	20	60	Average of 3 tests for 20 marks																																			
	CIE 2 10 th week	20																																					
	CIE 3 15 th week	20																																					
AAT's	AAT-1 10 th week			10																																			
	Lab Test	30	Reduced to 10																																				
	Lab Record	20	10																																				
Continuous Internal Evaluation Total Marks : 100. Reduced to 50 Marks																																							
Semester End Examination (SEE) Total Marks :100. Reduced to 50 Marks																																							

Text Books:

1. Peter Bruce, Andrew Bruce, and Peter Gedeck. Practical statistics for data scientists: 50 essential concepts. " O'Reilly Media, Inc.", 2ND Edition,2020.

Reference Books:

1. Dodge, Yadolah, ed. Statistical data analysis and inference. Elsevier, 2014.
2. Ismay, Chester, and Albert Y. Kim. Statistical Inference via Data Science: A Modern Dive into R and the Tidyverse. CRC Press, 2019.

E-Books:

1. <https://leanpub.com/LittleInferenceBook>

MOOC:

1. <https://www.coursera.org/learn/statistical-inference>
2. <https://www.datacamp.com/community/open-courses/statistical-inference-and-data-analysis>
3. <https://www.udemy.com/course/python-for-statistical-analysis/>

BIG DATA ANALYTICS			
Course Code	21CDT63	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Prerequisite: Good knowledge skill on Database and Data Structures			
<p>Course Objectives: This course will enable students to</p> <ul style="list-style-type: none"> • Understand Big Data and its fundamentals, Designing Data Architecture, data sources, quality, pre-processing and storing Big Data Platform and its Use cases. • Explore Hadoop, HDFS and its ecosystem. • Understand Map Reduce Jobs. • Exposure to Data Analytics with R. 			
<p>Teaching-Learning Process (General Instructions) These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.</p> <ol style="list-style-type: none"> 1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes. 2. Use of Video/Animation to explain functioning of various concepts. 3. Encourage collaborative (Group Learning) Learning in the class. 4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking. 5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it. 6. Introduce Topics in manifold representations. 7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them. 8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding. 			
Module - I			
<p>Introduction to Big Data: Big Data definition, characteristics, types, classification, handling techniques, Designing Data Architecture, Data Sources, Quality, Pre-Processing and Storing, Data Storage and Analysis, Big Data Analytics Applications and Case Studies. 08 Hours</p>			
Module - II			
<p>Introduction to Hadoop : Introduction to Hadoop Distributed File System Basics to Hadoop , Ecosystem, Hadoop Distributed File System, Hadoop Components. 08 Hours</p>			
Module - III			
<p>Hadoop Eco System Pig : Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators. Hive : Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions. Hbase : HBasics, Concepts, Clients, Example, Hbase Versus RDBMS. 08 Hours</p>			
Module - IV			
<p>Map Reduce: Hadoop Map Reduce Framework, anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features. 08 Hours</p>			

Module - V

Data Analytics with R
Machine Learning: Introduction, Supervised Learning, Unsupervised Learning, Collaborative Filtering.
Big Data Analytics with BigR

08 Hours

Teaching-Learning Process for all modules

Chalk and board, Active Learning, PPT Based presentation, Video

Course Outcomes

On completion of this course, the students are able to:

CO1: Master the concepts of Big Data, Data Architecture, data sources, quality, pre-processing and storing. Identify Big Data and its Business Implications.

CO2: List the components of Hadoop and Develop Big Data Solutions using Hadoop Eco System

CO3: Manage Job Execution in Hadoop Environment.

CO4: Apply Machine Learning Techniques using R.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Text Books:

- 1) Raj Kamal and Preeti Saxena, "Big Data Analytics Introduction to Hadoop, Spark, and Machine-Learning", McGraw Hill Education, 2019 ISBN: 9789353164966, 9353164966.
- 2) Tom White "Hadoop: The Definitive Guide" Fourth Edition, O'reily Media, 2015.
- 3) Seema Acharya, Subhasini Chellappan, First Edition "Big Data Analytics".Mc Graw Hill,2018

Reference Books:

- 1) Boris Lublin sky, Kevin T. Smith, Alexey Yakubovich, —Professional Hadoop Solutions",1st Edition, Wrox Press, 2014ISBN-13: 978-8126551071
- 2) Eric Sammer, —Hadoop Operations: A Guide for Developers and Administrators",1st Edition, O'Reilly Media, 2012.ISBN-13: 978-9350239261
- 3) Douglas Eadline,"Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem", 1stEdition, Pearson Education, 2016. ISBN-13: 978-9332570351
- 4) Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
- 5) Jay Liebowitz, "Big Data and Business Analytics" Auerbach Publications, CRC press (2013)
- 6) Tom Plunkett, Mark Hornick, "Using R to Unlock the Value of Big Data: Big Data Analytics with Oracle R Enterprise and Oracle R Connector for Hadoop", McGraw-Hill/Osborne Media (2013), Oracle press.

E-Resources:

- 1) https://www.tutorialspoint.com/big_data_tutorials.htm
- 2) <https://nptel.ac.in/courses/106/104/106104189/>

Understand the Big Data Platform and its Use cases • Provide an overview of Apache Hadoop • Provide HDFS Concepts and Interfacing with HDFS • Understand Map Reduce Jobs • Provide hands on Hadoop Eco System • Apply analytics on Structured, Unstructured Data. • Exposure to Data Analytics with R.

DIGITAL IMAGE ANALYTICS			
Course Code	21CDT641	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Prerequisite: Digital Signal Processing			
Course Learning Objectives <ul style="list-style-type: none"> • To learn and understand the fundamentals of digital image processing • To learn various image Transforms, Image Enhancement Techniques, Image restoration Techniques and methods • To learn image compression and Segmentation used in digital image processing • Students will examine how to perceive, communicate and create compelling digital images with an emphasis on the experimental and conceptual 			
Teaching-Learning Process (General Instructions) <p>These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.</p> <ol style="list-style-type: none"> 1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes. 2. Use of Video/Animation to explain functioning of various concepts. 3. Encourage collaborative (Group Learning) Learning in the class. 4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking. 5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it. 6. Introduce Topics in manifold representations. 7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them. 8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding. 			
Module – I			
DIGITAL IMAGE FUNDAMENTALS: Introduction – Fundamental steps in Image Processing Systems – Image Acquisition – Sampling and Quantization – Pixel Relationships – Mathematical Tools Used in Digital Image Processing. Some Basic Intensity Transformation Functions: Image Negatives, Log Transformations, Power-Law Transformations - Histogram Processing. Color Fundamentals - Fundamentals of Spatial Filtering - Smoothing Spatial Filters - Sharpening Spatial Filters.			
			08 Hours
Module – II			
MORPHOLOGICAL IMAGE PROCESSING: Morphological Image Processing: Fundamentals - Erosion and Dilation - Opening and Closing – Hit or Miss Transform - Some Basic Morphological Algorithms – Morphological Reconstruction – Grayscale Morphology.			
			08 Hours
Module – III			
IMAGE SEGMENTATION: Introduction - Point, Line, and Edge Detection – Thresholding: Foundation, Basic Global thresholding, Optimum Global Thresholding using Otsu's Method, Multiple Thresholds, Variable Thresholding – Segmentation by Region Growing and by Region Splitting and Merging – Image Segmentation: Active Contours: Snakes and Level Sets.			
			08 Hours

Module – IV	
FEATURE EXTRACTION: Background - Representation – Boundary Preprocessing – Boundary Feature Descriptors: Some Basic Boundary Descriptors, Shape Numbers, Fourier Descriptors, Statistical Moments - Regional Feature Descriptors: Some Basic Descriptors, Topological and Texture Descriptors, Moment Invariants – Principal Components as Feature Descriptors – Whole-image Features Object – Scale-Invariant Feature Transform (SIFT).	
08 Hours	
Module – V	
IMAGE PATTERN CLASSIFICATION: Background -Patterns and Pattern Classes – Pattern Classification by Prototype Matching: Minimum-Distance Classifier, Using Correlation for 2-D prototype matching, Matching SIFT Features, Matching Structural Prototypes - Optimum (Bayes) Statistical Classifiers - Neural Networks and Deep Learning: Background - The Perceptron - Multilayer Feed forward Neural Networks - Deep Convolutional Neural Networks.	
08 Hours	
Teaching-Learning Process for all modules	Chalk and Talk, Power point presentation, flip teaching, YouTube videos
Self Study	
Sl. No	Topics
1	Practical Component: Use Python/ MATLAB Apply various intensity transformations functions.
2	Computing and plotting image histograms and use standard image processing toolbox Spatial filters.
3	Implementation of Morphological operations.
4	Implementation of Optimum Global Thresholding using Otsu’s Method.
5	Implementation of Boundary Feature Descriptors
<p>COURSE OUTCOMES:</p> <p>After successful completion of this course, the students will be able to:</p> <p>CO1: Infer the basics and fundamentals of digital image processing and Apply the various techniques for intensity transformations functions. Implement Color image Smoothing and Sharpening.</p> <p>CO2: Illustrate Morphological operation and Apply Some Basic Morphological Algorithms.</p> <p>CO3: Apply image segmentation techniques such as Optimum Global Thresholding using Otsu’s Method, Active Contours: Snakes and Level Sets for various real-time applications.</p> <p>CO4: Analysis various Feature Extraction methods and Implement for various real-time applications.</p> <p>CO5: Apply and Analysis various Image Pattern Classification methods such as Minimum-Distance Classification, Optimum (Bayes) Statistical Classification, and Deep Convolutional Neural Network.</p>	

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Text Books:

1. Rafael C Gonzalez, Richard E Woods, "Digital Image Processing", 4th Edition, Pearson, 2018.

Reference Books:

1. Kenneth R. Castleman, Digital Image Processing Pearson, 2006.
2. Anil K. Jain, "Fundamentals of Digital Image Processing", Person Education, 2003.

E-Books:

1. https://www.academia.edu/19746149/Digital_Image_Processing_3rd_Edition_Instructors_Manual_Rafael_C._Gonzalez
2. https://www.academia.edu/18324189/Digital_image_processing_using_matlab_gonzalez
3. <https://pdfs.semanticscholar.org/15bd/427a1a5f9bc57a7f67fb1b1fc85c5bb39f46.pdf>

MOOC:

1. <https://www.coursera.org/learn/digital>
2. <https://www.udemy.com/topic/digital-image-processing/>
3. <https://www.edx.org/course/image-processing-and-analysis-for-life-scientists>

ADVANCED ALGORITHMS

Course Code	21CDT642	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Prerequisites:

1. A course on “Computer Programming & Data Structures”
2. A course on “Advanced Data Structures & Algorithms

Course Objectives:

This course will enable students to:

1. Introduces the recurrence relations for analyzing the algorithms
2. Introduces the graphs and their traversals.
3. Describes major algorithmic techniques (divide-and-conquer, greedy, dynamic programming, Brute Force, Transform and Conquer approaches) and mention problems for which each technique is appropriate;
4. Describes how to evaluate and compare different algorithms using worst-case, average-case and best-case analysis.
5. Introduces string matching algorithms

Teaching-Learning Process (General Instructions)

These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students’ Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module – I

Introduction: Role of Algorithms in computing, Order Notation, Recurrences, Probabilistic Analysis and Randomized Algorithms. Sorting and Order Statistics: Heap sort, Quick sort and Sorting in Linear Time.
Advanced Design and Analysis Techniques: Dynamic Programming- Matrix chain Multiplication, Longest common Subsequence and optimal binary Search trees.

08 Hours

Module – II

Greedy Algorithms - Huffman Codes, Activity Selection Problem. Amortized Analysis.
Graph Algorithms: Topological Sorting, Minimum Spanning trees, Single Source Shortest Paths, Maximum Flow algorithms.

08 Hours

Module – III

Sorting Networks: Comparison Networks, Zero-one principle, bitonic Sorting Networks, Merging Network, Sorting Network.

Matrix Operations- Strassen's Matrix Multiplication, Inverting matrices, Solving system of linear Equations

08 Hours

Module – IV

String Matching: Naive String Matching, Rabin-Karp algorithm, matching with finite Automata, Knuth-Morris - Pratt algorithm.

08 Hours

Module – V

NP-Completeness and Approximation Algorithms: Polynomial time, polynomial time verification, NP-Completeness and reducibility, NP-Complete problems. Approximation Algorithms- Vertex cover Problem, Travelling Sales person problem

08 Hours

Teaching-Learning Process for all modules

Chalk & board, Problem based learning, MOOC, Demonstration

Course Outcomes

1. Analyze algorithm complexity using recurrence relations, asymptotic analysis, aggregate analysis, worst case analysis.
2. Prove correctness of algorithms.
3. Ability to analyze the performance of algorithms
4. Ability to choose appropriate data structures and algorithm design methods for a specified application
5. Ability to understand how the choice of data structures and the algorithm design methods impact the performance of programs.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Text Books:

1. Introduction to Algorithms, Thomas H., Cormen, Charles E. Leiserson, Ronal L. Rivest, Clifford Stein, 3rd Edition, PHI, 2010.
2. Algorithm Design, John Kleinberg and Eva Tardos, Pearson Publication 2006

REFERENCE BOOKS:

1. Fundamentals of Computer Algorithms, Ellis Horowitz, Satraj Sahni and Rajasekharam, Galgotia publications pvt. Ltd.
2. Design and Analysis Algorithms - Parag Himanshu Dave, Himanshu Bhalchandra Dave Publisher: Pearson
3. Algorithm Design: Foundations, Analysis and Internet examples, M.T. Goodrich and R. Tomassia, John Wiley and sons.
4. Data structures and Algorithm Analysis in C++, Allen Weiss, Second edition, Pearson education

Online Resources

1. <https://www.coursera.org/courses?query=advanced%20algorithms>
2. <https://ocw.mit.edu/courses/6-854j-advanced-algorithms-fall-2005/pages/lecture-notes/>

AGILE METHODOLOGIES			
Course Code	21CDT643	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Pre-requisite: Knowledge of software engineering fundamentals.			
Course Objectives: This course will enable students to, <ul style="list-style-type: none"> • Identify the limitations of traditional (waterfall) software development methodologies. Interpret how to prioritize user stories while estimating. • Explain the concept and methodology behind Planning Poker as a technique used in Agile for estimating effort or complexity in user stories. • Explain the core principles and values underpinning Agile methodologies, emphasizing flexibility, collaboration, iterative development, and responding to change. • Explain the principles, roles, ceremonies, and artifacts of the Scrum framework in Agile software development methodologies. • Explain the Agile software development lifecycle and its iterative nature, emphasizing the integration of testing throughout the development process. 			
Teaching-Learning Process (General Instructions) These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes. <ol style="list-style-type: none"> 1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes. 2. Use of Video/Animation to explain functioning of various concepts. 3. Encourage collaborative (Group Learning) Learning in the class. 4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking. 5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it. 6. Introduce Topics in manifold representations. 7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them. 8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding. 			
Module - I			
Introduction: Need of Agile software development, agile context– Manifesto, Principles, Methods, Values, Roles, Artifacts, Stakeholders, and challenges. Business benefits of software agility. Project Planning: Recognizing the structure of an agile team– Programmers, Managers, Customers. <p style="text-align: right;">08 Hours</p>			
Module - II			
Project planning: User stories– Definition, Characteristics and content. Estimation – Planning poker, Prioritizing, and selecting user stories with the customer, projecting team velocity for releases and iterations. <p style="text-align: right;">08 Hours</p>			
Module - III			
Project Design: Fundamentals, Design principles–Single responsibility, Open-closed, Liskov			

substitution, Dependency-inversion, Interface-segregation.

Extreme Programming- Core principles, values and practices. Kanban, Feature-driven development
08 Hours

Module - IV

Design Methodologies: Need of scrum, Scrum practices –Working of scrum, Project velocity, Burn down chart, Sprint backlog, Sprint planning and retrospective, Daily scrum, Scrum roles– Product Owner, Scrum Master, Scrum Team.
08 Hours

Module - V

Testing: The Agile lifecycle and its impact on testing, Test driven development– Acceptance tests and verifying stories, writing a user acceptance test, Developing effective test suites, Continuous integration, Code refactoring. Risk based testing, Regression tests, Test automation.
08 Hours

Teaching-Learning Process for all modules Chalk and board, Active Learning, PPT Based presentation, Video

Course Outcomes:

On completion of this course, the students will be able to,

- Understanding the Limitations of Traditional Methodologies. Prioritization of User Stories and Estimation.
- **Understanding the Concept of Planning Poker. Application and Implementation of Planning Poker.**
- **Understanding Agile Principles, Values. Emphasizing Flexibility in Agile Methodologies**
- **Understanding roles and responsibilities of Scrum Principles.**
- **Emphasizing Iterative and Incremental Development, Integration of Testing Across Development Phases**

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Text Books:

1. Ken Schawber, Mike Beedle, “Agile Software Development with Scrum”, International Edition, Pearson
2. Robert C. Martin, “Agile Software Development, Principles, Patterns and Practices”, First International Edition, Prentice Hall.
3. Pedro M. Santos, Marco Consolaro, and Alessandro Di Gioia, “Agile Technical Practices Distilled: A learning journey in technical practices and principles of software design”, First edition, Packt Publisher.

REFERENCES

1. Lisa Crispin, Janet Gregory, “Agile Testing: A Practical Guide for Testers and Agile Teams”, International edition, Addison Wesley
2. Alistair Cockburn, “Agile Software Development: The Cooperative Game”, 2nd Edition, Addison-Wesley

Online Courses and Video lectures

1. “Agile Software Development”, <https://www.edx.org/course/agile-software-development>
Accessed on August 27, 2021.
2. “Agile Software Development”, <https://www.coursera.org/learn/agile-software-development>
Accessed on August 27, 2021

BUSINESS ECONOMICS			
Course Code	21CDT644	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Pre-requisite:			
Course objectives:			
<ul style="list-style-type: none"> • Analyze the global business environment. • Analyze the local business environment • Use critical thinking skills in business situations • Apply an ethical understanding and perspective to business situations. 			
Teaching-Learning Process (General Instructions)			
<p>These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.</p> <ol style="list-style-type: none"> 1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes. 2. Use of Video/Animation to explain functioning of various concepts. 3. Encourage collaborative (Group Learning) Learning in the class. 4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking. 5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it. 6. Introduce Topics in manifold representations. 7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them. 8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding. 			
Module – I			
INTRODUCTION TO ECONOMICS: Introduction to Economics- Flow in an economy, Law of supply and demand, Concept of Engineering Economics – Engineering efficiency, Economic efficiency, Scope of engineering economics			
08 Hours			
Module – II			
COST ANALYSIS: Types of Cost, Element of costs, Marginal cost, Marginal Revenue, Sunk cost, Opportunity cost, Breakeven analysis, Economies of Scale Cost Classification			
08 Hours			
Module – III			
CONSUMER'S AND PRODUCER'S BEHAVIOUR: Consumer Behavior: Law of Diminishing Marginal utility – Equimarginal Utility – Consumer's Equilibrium - Indifference Curve – Production: Law of Variable Proportion – Laws of Returns to Scale– Producer's equilibrium – Economies of Scale Cost Classification.			
08 Hours			

Module – IV

BUDGET: Process of budgeting in India –classification of budgets trends – evaluation systems – types of deficits – fiscal policy – indicators — taxation – centre, state and local – public debt and management.
08 Hours

Module – V

FINANCE: Basics of finance and financial environment – instruments of financial markets – financial intermediation – investment banking and brokerage services – securities – types of securities – market for securities – how and where traded – initial public offering (IPO) – secondary markets – trading on exchanges and trading with margins.

08 Hours

Teaching-Learning Process for all modules

Chalk and board, Active Learning, PPT Based presentation, Video

Course Outcomes:

On completion of this course, the students will be able to,

CO1: Describe the introduction of economics.

CO2: Demonstrate about cost analysis.

CO3: Build knowledge about consumer's and producer's behavior.

CO4: Explain about budget.

CO5: Educate about financial services.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Text Books:

1. S.Shankaran, Business Economics - Margham Publications.
2. H.L. Ahuja, Business Economics – Micro & Macro - Sultan Chand & Sons – New Delhi – 55.

References Books:

1. S.A.Ross, R.W.Westerfield, J.Jaffe and Roberts: Corporate Finance, McGraw-Hill.
2. Joseph E Stiglitz: Economics of the Public Sector.

MOOC:

1. [https://sites.google.com/site/readbookpdf7734/pdf-download-business-economics-by--- mark-taylor-read-online](https://sites.google.com/site/readbookpdf7734/pdf-download-business-economics-by---mark-taylor-read-online)
2. <https://bookboon.com/en/economics-ebooks>

INTRODUCTION TO JAVA PROGRAMMING

Course Code	21CDT651	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Pre-requisite: OOPS concepts, C++.

Course Objectives:

This course will enable students to:

- Learn the basic concepts of object-oriented programming.
- Understand the basics of JAVA Programming using classes and objects.
- Gain the knowledge of Inheritance and packages.
- Expose to the concepts of exceptions that occur while programming in JAVA.
- Acquire the knowledge of multi-threaded programming in JAVA.

Teaching-Learning Process (General Instructions)

These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module – I

Introduction to Object Oriented Concepts: Procedure–Oriented Programming system, Object Oriented Programming System, Comparison of Object Oriented Language with C.

Introduction to Java: Java's magic, The Byte code, Java Development Kit (JDK), Java Buzzwords, Object- oriented programming, IO Streams, Data types, variables and arrays, reference variables, Operators, Control Statements. Simple Java programs.

08 Hours

Module – II

Classes: Classes fundamentals, Declaring objects, this keyword, garbage collection. **Methods:** Method Prototyping, Member functions and data members, Constructors, Objects and methods, Method Overloading, Objects and arrays, Access modifiers, Setters and getters, Nested classes, Console I/O.

08 Hours

Module – III

Inheritance: Inheritance basics, using super, creating multi-level hierarchy, method overriding, using Abstract classes, using final. **Packages:** Packages: Access Protection, Importing Packages. **08 Hours**

Module – IV

Interfaces, Exceptions, Applets: Interfaces, Exception handling fundamentals, exception types, uncaught exceptions, using try and catch, using multiple catch clauses, nested try statements, throw, throws, finally, Exception handling in Java, Applets, Types of Applets, Applet basics and class, Applet Architecture.

08 Hours

Module – V

Event Handling and Multi-Threaded Programming : Two event handling mechanisms, The delegation event model, Event classes, Sources of events, Event listener interfaces, Using the delegation event model, Adapter classes, Inner classes. Multi-Threaded Programming: What are threads? How to make the classes threadable, Extending threads, Implementing runnable, Synchronization, Changing state of the thread, Bounded buffer problems, read-write problem.

08 Hours

Teaching-Learning Process for all modules

Chalk and board, Case studies

Course Outcomes

On completion of this course, the students will be able to:

- Explain the difference between Procedure and Object Oriented Programming.
- Develop basic JAVA programs.
- Apply Inheritance properties and packages in solving real world problems.
- Use exception handling methods efficiently.
- Demonstrate the programs by using multithreaded concepts.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Text Books:

1. Herbert Schildt, "Java The Complete Reference", 7th Edition, Tata McGraw 2013, ISBN-13: 978-0072263855, (Chapters 1-11).

Reference Books:

1. Herbert Schildt, "The Complete Reference C++", 4th Edition, Tata McGraw Hill, 2013, ISBN- 13: 978-0072226805.
2. E Balagurusamy, "Programming with Java-A primer", 2nd Edition, Tata McGraw Hill companies, 2009, ISBN-13: 978-9351343202.

E-Resources:

1. www.geeksforgeeks.org/java/
2. www.tutorialspoint.com/java/index.htm

INTRODUCTION TO CYBER SECURITY			
Course Code	21CDT652	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03
Pre-requisite: Networking, Cryptography, Hacking.			
Course Learning Objectives			
CLO 1. To familiarize cybercrime terminologies and ACTs			
CLO 2. Understanding cybercrime in mobiles and wireless devices along with the tools for Cybercrime and prevention			
CLO 3. Understand the motive and causes for cybercrime, cybercriminals, and investigators			
CLO 4. Understanding criminal case and evidence, detection standing criminal case and evidence.			
Teaching-Learning Process (General Instructions)			
These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.			
<ol style="list-style-type: none"> 1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes. 2. Use of Video/Animation to explain functioning of various concepts. 3. Encourage collaborative (Group Learning) Learning in the class. 4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking. 5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it. 6. Introduce Topics in manifold representations. 7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them. 8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding. 			
Module – I			
Introduction to Cybercrime:			
Cybercrime: Definition and Origins of the Word, Cybercrime and Information Security, Who are Cybercriminals? Classifications of Cybercrimes,			
Cybercrime: The Legal Perspectives,			
Cybercrimes: An Indian Perspective, Cybercrime and the Indian ITA 2000.			
Textbook1:Ch1 (1.1 to 1.8).			08 Hours
Module – II			
Cyber offenses:			
How Criminals Plan Them: Introduction, How Criminals Plan the Attacks, Social Engineering, Cyber stalking, Cyber cafe and Cybercrimes.			
Botnets: The Fuel for Cybercrime, Attack Vector			
Textbook1: Ch2 (2.1 to 2.7).			08 Hours
Module – III			
Tools and Methods Used in Cybercrime: Introduction, Proxy Servers and Anonymizers, Phishing, Password Cracking, Key loggers and Spywares, Virus and Worms, Trojan Horses and Backdoors, Steganography, DoS and DDoS Attacks, Attacks on Wireless Networks.			
Textbook1: Ch4 (4.1 to 4.9, 4.12).			08 Hours

Module – IV

Understanding the people on the scene: Introduction, understanding cyber criminals, understanding cyber victims, understanding cyber investigators.

The Computer Investigation process: investigating computer crime.

Understanding Cybercrime Prevention: Understanding Network Security Concepts, Understanding Basic Cryptography Concepts, Making the Most of Hardware and Software Security

Textbook 2:Ch3,Ch 4, Ch 7.

08 Hours

Module – V

Cybercrime Detection Techniques: Security Auditing and Log Firewall Logs, Reports, Alarms, and Alerts, Commercial Intrusion Detection Systems, Understanding E-Mail Headers Tracing a Domain Name or IP Address.

Collecting and preserving digital Evidence: Introduction, understanding the role of evidence in a criminal case, collecting digital evidence, preserving digital evidence, recovering digital evidence, documenting evidence.

TextBook 2:Ch 9, Ch 10.

08 Hours

Teaching-Learning Process for all modules

Chalk and board, Case studies

Course Outcomes:

At the end of the course the student will be able to:

CO 1. Describe the cyber crime terminologies

CO 2. Analyze cybercrime in mobiles and wireless devices along with the tools for Cybercrime and prevention

CO 3. Analyze the motive and causes for cybercrime, cybercriminals, and investigators

CO 4. Apply the methods for understanding criminal case and evidence, detection standing criminal case and evidence.

Assessment Details (both CIE and SEE)

	Component	Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Textbooks

1. SunitBelapure and Nina Godbole, "Cyber Security: Understanding Cyber Crimes, Computer Forensics And Legal Perspectives", Wiley India Pvt Ltd, ISBN: 978-81- 265-21791, 2013

2. Debra Little John Shinder and Michael Cross, "Scene of the cybercrime", 2nd edition, Syngress publishing Inc, Elsevier Inc, 2008

Reference Books:

1. Robert M Slade, "Software Forensics", Tata McGraw Hill, New Delhi, 2005.

2. Bernadette H Schell, Clemens Martin, "Cybercrime", ABC – CLIO Inc, California, 2004.

3. Nelson Phillips and EnfingerSteuart, "Computer Forensics and Investigations", Cengage Learning, New Delhi, 2009.

4. Kevin Mandia, Chris Prosis, Matt Pepe, "Incident Response and Computer Forensics", Tata McGraw -Hill, New Delhi, 2006.

Weblinks and Video Lectures (e-Resources):

1. <https://www.youtube.com/watch?v=czDzUP1HclQ>
2. <https://www.youtube.com/watch?v=qS4ViqnjkC8>
3. https://www.trendmicro.com/en_nz/ciso/21/h/cybercrime-today-and-the-future.html

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Real world problem solving: Demonstration of projects related to Cyber security.

ARTIFICIAL INTELLIGENCE & EXPERT SYSTEM

Course Code	21CDT653	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Pre-requisite: Statistics, Mathematical Modeling

Course Objectives

- To have insight into the fundamentals of Artificial Intelligence (AI) and Robotics that includes the various peculiar search strategies for AI, Programming the Robots and Controlling Autonomous Robots etc.
- To have proficiency in developing the techniques to solve real world problems unconventionally with optimality.

Teaching-Learning Process (General Instructions)

These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module – I

Introduction to AI: The AI Problems, Under lying assumptions, AI technique, Level of the model, Criteria for success (1.1 to 1.5 from Rich and Knight)

Problems: Problem spaces and search Problem as a state space search, Production systems, Problem characteristics, Production system characteristics, Issues in the design of search problems, additional problems (2.1 to 2.6 from Rich and Knight)

08 Hours

Module – II

Search and control Strategies: Introduction, Generate and Test, Hill Climbing, Simulated annealing (3.1, 3.2 from Rich and Knight)

Expert systems Architectures : Introduction, Rule-Based System Architectures, Nonproduction System Architectures, Dealing with Uncertainty, Knowledge Acquisition and Validation (15.1 to 15.6 from Dan W. Patterson),

08 Hours

Module – III

Introduction to Robotics: The Seven Criteria of Defining a Robot, Robot Categories, Sensors, Actuators, End Effectors, Controllers, Scenario, Giving the robot instructions. (Chapter 1 from Cameron Hughes)

<p>Robot Vocabularies and RSVP: Additional Effort, Actions, The Autonomous Robot's ROLL Model, RSVP (Robot Scenario Visual Planning): Mapping the Scenario, Pseudo code and Flow charting RSVP, State charts for Robots and Objects. (Chapter 2 and 3 from Cameron Hughes)</p>																															
08 Hours																															
Module – IV																															
<p>Actual Capabilities of Robot: The Reality Check for the Micro controller, Sensor Reality Check, Determine Your Robot's Sensor, Limitations, Actuators End-Effectors Reality Check. (Chapter 4 from Cameron Hughes)</p>																															
<p>Sensors: Types of Sensors, Sensor Interfacing with Micro controllers, Attributes of Sensors, Sensor Calibration. (Chapter 5 from Cameron Hughes)</p>																															
08 Hours																															
Module – V																															
<p>Sensors: Types of Sensors, Sensor Interfacing with Microcontrollers, Attributes of Sensors, Sensor Calibration. (Chapter 5 from Cameron Hughes)</p>																															
08 Hours																															
Teaching-Learning Process for all modules	<i>L₁–Remembering, L₂–Understanding, L₃–Applying, L₄–Analysing, L₅–Evaluating.</i>																														
<p>Course outcomes: At the end of the course the student will be able to: CO1: Explain the fundamentals of artificial intelligence, robotics and systems. CO2: Identify knowledge associated and represent it by onto logical engineering to plan a strategy to solve given problem. CO3: Apply the suitable algorithms to solve AI problems CO4: Solve problem using problem decomposition and planning CO5: Design smart system using different in formed search / uninformed search or heuristic approaches</p>																															
Assessment Details (both CIE and SEE)																															
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th colspan="2" style="text-align: center;">Component</th> <th colspan="2" style="text-align: center;">Weightage (%)</th> </tr> </thead> <tbody> <tr> <td rowspan="3" style="text-align: center; vertical-align: middle;">CIE's</td> <td>CIE 1 5th week</td> <td style="text-align: center;">20</td> <td rowspan="3" style="text-align: center; vertical-align: middle;">60</td> </tr> <tr> <td>CIE 2 10th week</td> <td style="text-align: center;">20</td> </tr> <tr> <td>CIE 3 15th week</td> <td style="text-align: center;">20</td> </tr> <tr> <td rowspan="3" style="text-align: center; vertical-align: middle;">AAT's</td> <td>AAT-1 10th week</td> <td colspan="2" style="text-align: center;">10</td> </tr> <tr> <td>AAT-2</td> <td colspan="2" style="text-align: center;">10</td> </tr> <tr> <td>AAT-3</td> <td colspan="2" style="text-align: center;">20</td> </tr> <tr> <td colspan="4" style="text-align: center;">Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks</td> </tr> <tr> <td colspan="4" style="text-align: center;">Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks</td> </tr> </tbody> </table>		Component		Weightage (%)		CIE's	CIE 1 5 th week	20	60	CIE 2 10 th week	20	CIE 3 15 th week	20	AAT's	AAT-1 10 th week	10		AAT-2	10		AAT-3	20		Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks				Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			
Component		Weightage (%)																													
CIE's	CIE 1 5 th week	20	60																												
	CIE 2 10 th week	20																													
	CIE 3 15 th week	20																													
AAT's	AAT-1 10 th week	10																													
	AAT-2	10																													
	AAT-3	20																													
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks																															
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks																															
Textbooks																															
<ol style="list-style-type: none"> Elaine Rich, Kevin Knight and Shiva Shankar B .Nair, Artificial Intelligence, TMH Education (P) Ltd, New Delhi, 3rd Edition, 2010. Dan W. Patterson, Introduction to Artificial Intelligence and Expert Systems, Prentice Hall of India, Private Ltd, New Delhi, 1st Edition, 2015. Cameron Hughes Tracey Hughes, Robot Programming: A Guide to Controlling Autonomous Robots, Pearson Education, 1st Edition, 2016. 																															
Reference Books																															
<ol style="list-style-type: none"> Stuart Russell and Peter Norvig, Artificial Intelligence: A modern approach, Pearson Education, India, 3rd Edition, 2016. Saroj Kaushik, Artificial Intelligence, Cengage Learning India, 1st Edition, 2011. Robin R. Murphy, Introduction to AI Robotics, MIT Press, 1st Edition, 2000. Saha S. K. Introduction to Robotics, TMH Publications, 1st Edition, 2008 																															

Weblinks and Video Lectures:

1. <https://nptel.ac.in/courses/106105077>
2. <https://nptel.ac.in/courses/106106126>
3. <https://aima.cs.berkeley.edu>
4. [https://ai.berkeley.edu/project_overview.html\(forPractical's\)](https://ai.berkeley.edu/project_overview.html(forPractical's))

INTRODUCTION TO DATABASE MANAGEMENT SYSTEMS

Course Code	21CDT654	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Pre-requisite: Data Structures.

Course Learning Objectives

- CLO 1. Understand the basic concepts and the applications of database systems.
- CLO 2. Understand the relational database design principles.
- CLO 3. Master the basics of SQL and construct queries using SQL.
- CLO 4. Familiar with the basic issues of transaction processing and concurrency control.

Teaching-Learning Process (General Instructions)

These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Module – I

Introduction to Databases: Introduction, Characteristics of database approach, Advantages of using the DBMS approach, History of database applications.

Overview of Database Languages and Architectures: Data Models, Schemas, and Instances. Three schema architecture and data independence, database languages, and interfaces, The Database System environment.

Conceptual Data Modelling using Entities and Relationships: Entity types, Entity sets, attributes, roles, and structural constraints, Weak entity types, ER diagrams, Examples

Textbook 1: Ch 1.1 to 1.8, 2.1 to 2.6, 3.1 to 3.7

08 Hours

Module – II

Relational Model: Relational Model Concepts, Relational Model Constraints and relational database schemas, Update operations, transactions, and dealing with constraint violations.

Relational Algebra: Relational algebra: introduction, Selection and projection, set operations, renaming, Joins, Division, syntax, semantics. Operators, grouping and ungrouping, relational

comparison. Examples of Queries in relational algebra.

Mapping Conceptual Design into a Logical Design: Relational Database Design using ER-to-Relational mapping.

Textbook 1: ch5.1 to 5.3, 8.1 to 8.5, 9.1;

08 Hours

Module – III

SQL: SQL data definition and data types, specifying constraints in SQL, retrieval queries in SQL, INSERT, DELETE, and UPDATE statements in SQL, Additional features of SQL.

Advances Queries: More complex SQL retrieval queries, Specifying constraints as assertions and action triggers, Views in SQL, Schema change statements in SQL. Database

Textbook 1: Ch 6.1 to 6.5, 7.1 to 7.4; Textbook 2: 6.1 to 6.6;

08 Hours

Module – IV

Normalization: Database Design Theory – Introduction to Normalization using Functional and Multivalued Dependencies: Informal design guidelines for relation schema, Functional Dependencies, Normal Forms based on Primary Keys, Second and Third Normal Forms, Boyce-Codd Normal Form, Multivalued Dependency and Fourth Normal Form, Join Dependencies and Fifth Normal Form. Examples on normal forms.

Textbook 1: Ch 14.1 to -14.7, 15.1 to 15.6

08 Hours

Module – V

Transaction management and Concurrency –Control Transaction management: ACID properties, serializability and concurrency control, Lock based concurrency control (2PL, Deadlocks), Time stamping methods, optimistic methods, database recovery management.

Textbook 1: Ch 20.1 to 20.6, 21.1 to 21.7;

08 Hours

Teaching-Learning Process

Chalk& board, Problem based learning, Demonstration, MOOC

Course Outcomes

At the end of the course the student will be able to:

- CO 1. Identify, analyze and define database objects, enforce integrity constraints on a database using RDBMS
- CO 2. Use Structured Query Language (SQL) for database manipulation.
- CO 3. Design and build simple database systems
- CO 4. Develop application to interact with databases.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Textbooks

1. Fundamentals of Database Systems, RamezElmasri and Shamkant B. Navathe, 7th Edition, 2017, Pearson.
2. Database management systems, Ramakrishnan, and Gehrke, 3rd Edition, 2014, McGraw Hill

Weblinks and Video Lectures (e-Resources):

1. <https://www.youtube.com/watch?v=3EJlovevfcA>
2. <https://www.youtube.com/watch?v=9TwMRs3qTcU>
3. <https://www.youtube.com/watch?v=ZW10Xow304I>
4. <https://www.youtube.com/watch?v=4YilEjkNPrQ>
5. <https://www.youtube.com/watch?v=CZTkgMoqVss>
6. <https://www.youtube.com/watch?v=HI4NZB1XR9c>
7. https://www.youtube.com/watch?v=EGEwkad_1IA
8. <https://www.youtube.com/watch?v=t5hsV9IC1rU>

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Real world problem solving: Developing and demonstration of models / projects based on DBMS application

INTRODUCTION TO DATA SCIENCE

Course Code	21CDT655	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:0:0	SEE Marks	50
Total Hours of Pedagogy	40	Total Marks	100
Credits	03	Exam Hours	03

Pre-requisite

- Basic knowledge of statistics
- Familiarity with programming (preferably Python)
- Basic understanding of mathematics

Course Learning Objectives

This course will enable students to:

1. Provide a strong foundation for data science and application areas related to it.
2. Learn the process of working with data on large scale.
3. Explore the concepts of Data Processing.
4. Learn basic concepts of Machine Learning.
5. Prepare students for advanced courses in Data Science

Teaching-Learning Process (General Instructions)

These are sample Strategies; which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

Syllabus

Module – I

Introduction to Data Science: Importance of data Science-Need for Data Science, What is Data Science? Data Science Process, prerequisites for data science, Components of Data Science, Tools and Skills needed.

Statistics: Data Types, Variable Types, Statistics, Sampling Techniques.

[Text Book1:1(1.1,1.2,1.3,1.5,1.6,1.7),2(2.1,2.2,2.3,2.4,2.5)].

08 Hours

Module – II

Probability: Probability Theory, Probability types, Probability Distribution Functions, Bayes Theorem.
Data Modeling and Analytics: Data Science Methodology-Analytics for data science, Example of Data Analytics, Data Analytics Life Cycle-Data Discovery, Data Preparation, Model Planning, Model Building, Communicate Results, Operationalization.

[Text Book1:2(2.6,2.7,2.8,2.9),4(4.1,4.2,4.3)]

08 Hours

Module – III

Machine learning –Designing a Learning System, Perspective and Issues in Machine Learning, Supervised learning, Unsupervised learning, Semi- supervised learning, Reinforcement Learning, Role of Machine Learning in Data Science, Data Science vs Machine Learning.

[Text Book2:1.2(1.2.1,1.2.2,1.2.3,1.2.4,1.2.5),

<https://www.geeksforgeeks.org/data-science-vs-machine-learning>

<https://www.zucisystems.com/blog/what-is-the-role-of-machine-learning-in-data-science>

08 Hours

Module – IV

Databases for Data Science – SQL-for Data Science, Basic Statistics with SQL, Data Wrangling, Filtering, Joins, Aggregation, Advanced NoSQL for Data Science, Document Databases for Datascience, Wide Column Databases for Datascience, Graph Databases for Data Science.

TextBook1:3.1(3.1.1,3.1.2,3.1.3), 3.2(3.2.1,3.2.2,3.2.3,3.2.4).

08 Hours

Module – V

Data Analytics and Text Mining: What is Text Mining?, Process of Text Mining, Difference between Text Mining and Data Mining, Major Text Mining Areas, Text Analytics, Text Analytics Steps, Basic Text Analytics Steps.

Introduction to NLP: Introduction, Major Components of NLP, Stages of NLP, Statistical Processing of NLP, Applications of NLP.

TextBook1:6.1(6.1.1,6.1.1.1,6.1.1.2,6.1.1.3),6.2(6.2.1,6.2.2),6.3(6.3.1,6.3.2,6.3.3,6.3.4)

08 Hours

Teaching-Learning Process for all modules

Chalk and board, Active Learning, PPT Based presentation, Video

Course Outcomes

At the end of the course the student will be able to :

- 1: Apply the fundamental concepts of data science
- 2: Evaluate the data analysis techniques for applications handling large data and Demonstrate the data science process.
- 3: Analyze the concept of machine learning used in the data science process.
- 4: Demonstrate and present the inference using various tools.
- 5: Analyze to think through the ethics surrounding privacy, data sharing.

Assessment Details (both CIE and SEE)

Component		Weightage (%)	
CIE's	CIE 1 5 th week	20	60
	CIE 2 10 th week	20	
	CIE 3 15 th week	20	
AAT's	AAT-1 10 th week	10	
	AAT-2	10	
	AAT-3	20	
Continuous Internal Evaluation Total Marks: 100. Reduced to 50 Marks			
Semester End Examination (SEE) Total Marks: 100. Reduced to 50 Marks			

Suggested Learning Resources:

Text Books:

1. Fundamentals of Data Science, Sanjeev J. Wagh, Manisha S. Bhende, and Anuradha D. Thakare, First edition published 2022 by CRC Press.
2. Machine Learning, [Tom Mitchell](#), McGraw Hill, 1997.

E-Resources

1. <https://www.geeksforgeeks.org/data-science-vs-machine-learning>
2. <https://www.zucisystems.com/blog/what-is-the-role-of-machine-learning-in-data-science>

BIGDATA ANALYTICS LAB			
Course Code	21CDL66	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	0:0:2:0	SEE Marks	50
Total Hours of Pedagogy	24	Total Marks	100
Credits	1	Exam Hours	03
Prerequisites: Good knowledge skill on Database and Data Structures			
Course Objectives: This course will enable students to			
<ul style="list-style-type: none"> • Explore Hadoop, HDFS and its ecosystem. • Explore Hadoop tools like PIG ,HIVE and PySpark • Understand Map Reduce Jobs • Understand how to build ,analyze charts using Tableau. 			
Sl. No.	Experiments		
1	Installation of Hadoop Framework, it's components and study the HADOOP ecosystem.		
2	Hive: Introduction Creation of Database and Table, Hive Partition, Hive Built in Function and Operators, Hive View and Index.		
3	Pig: Pig Latin Basic Pig Shell, Pig Data Types, Creating a Pig Data Model, Reading and Storing Data, Pig Operations.		
4	Write a program to implement Twitter Sentiment Analysis System where we populate real-time sentiments for crisis management, service adjusting and target marketing using PYSpark.		
5	Write a program to implement word count program using MapReduce.		
6	Implementing Page Rank Algorithm Using Map-Reduce.		
7	Visualization: Connect to data, Build Charts and Analyze Data, Create Dashboard, Create Stories using Tableau.		
Course Outcomes On completion of this course, the students are able to:			
<ol style="list-style-type: none"> 1. Configure Hadoop and perform File Management Tasks 2. Apply MapReduce programs to real time issues like word count, weather dataset and sales of a company 3. Critically analyze huge data set using Hadoop distributed file systems and MapReduce. 4. Apply different data processing tools like Pig and Hive,Spark. 			
Text Books:			
1. Tom White, "Hadoop: The Definitive Guide" Fourth Edition, O'reilly Media, 2015.			
Reference Books:			
1. Glenn J. Myatt, Making Sense of Data , John Wiley & Sons, 2007 Pete Warden, Big Data Glossary, O'Reilly, 2011.			
2. Michael Berthold, David J.Hand, Intelligent Data Analysis, Spingers, 2007.			
3. Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Paul Zikopoulos, Uderstanding Big Data : Analytics for Enterprise Class Hadoop and Streaming Data, McGrawHill Publishing, 2012.			
4. Anand Rajaraman and Jeffrey David UIlman, Mining of Massive Datasets Cambridge University Press, 2012.			
5. Boris Lublin sky, Kevin T. Smith, Alexey Yakubovich, —Professional Hadoop Solutions",1st Edition, Wrox Press, 2014ISBN-13: 978-8126551071			

6. Eric Sammer, —Hadoop Operations: A Guide for Developers and Administrators",1st Edition, O'Reilly Media, 2012.ISBN-13: 978-9350239261

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each course. The student has to secure not less than 35% (18 Marks out of 50) in the semester-end examination (SEE).

Continuous Internal Evaluation (CIE):

CIE marks for the practical course is **50 Marks**.

The split-up of CIE marks for record/ journal and test are in the ratio **60:40**.

□ Each experiment to be evaluated for conduction with an observation sheet and record write-up. Rubrics for the evaluation of the journal/write-up for hardware/software experiments designed by the faculty who is handling the laboratory session and is made known to students at the beginning of the practical session.

□ Record should contain all the specified experiments in the syllabus and each experiment write-up will be evaluated for 10 marks.

□ Total marks scored by the students are scaled down to 30 marks (60% of maximum marks).

□ Weightage to be given for neatness and submission of record/write-up on time.

□ Department shall conduct 02 tests for 100 marks, the first test shall be conducted after the 8th week of the semester and the second test shall be conducted after the 14th week of the semester.

□ In each test, test write-up, conduction of experiment, acceptable result, and procedural knowledge will carry a weightage of 60% and the rest 40% for viva-voce.

□ The suitable rubrics can be designed to evaluate each student's performance and learning ability.

Rubrics suggested in Annexure-II of Regulation book

□ The average of 02 tests is scaled down to 20 marks (40% of the maximum marks).

□ **The Sum of scaled-down marks scored in the report write-up/journal and average marks of two tests is the total CIE marks scored by the student.**

Semester End Evaluation (SEE):

□ SEE marks for the practical course is 50 Marks.

□ SEE shall be conducted jointly by the two examiners of the same institute, examiners are appointed by the University

□ All laboratory experiments are to be included for practical examination.

□ (Rubrics) Breakup of marks and the instructions printed on the cover page of the answer script to be strictly adhered to by the examiners. **OR** based on the course requirement evaluation rubrics shall be decided jointly by examiners.

□ Students can pick one question (experiment) from the questions lot prepared by the internal /external examiners jointly.

□ Evaluation of test write-up/ conduction procedure and result/viva will be conducted jointly by examiners.

□ General rubrics suggested for SEE are mentioned here, writeup-20%, Conduction procedure and result in -60%, Viva-voce 20% of maximum marks. SEE for practical shall be evaluated for 100 marks and scored marks shall be scaled down to 50 marks (however, based on course type, rubrics shall be decided by the examiners)

□ Students can pick one experiment from the questions lot of PART A with an equal choice to all the students in a batch. For PART B, the project group (Maximum of 4 students per batch) should demonstrate the mini-project.

□ Weightage of marks for PART A is 60% and for PART B is 40%. General rubrics suggested to be followed for part A and part B.

□ Change of experiment is allowed only once and Marks allotted to the procedure part to be made zero (Not allowed for Part B).

□ The duration of SEE is 03 hours

Rubrics suggested in Annexure-II of Regulation book

E-Resources:

1. https://www.tutorialspoint.com/big_data_tutorials.html
2. <https://nptel.ac.in/courses/106/104/106104189/>