Maharishi DayanandUniversity, Rohtak
SCHEME OF STUDIES AND EXAMINATION

B.Tech (Computer Science and Engineering-Data Science)
4thYear

Semester: 7th&8th
Scheme effective from 2023-24

COURSE CODE ANDDEFINITIONS

Course Code	Definitions
L	Lecture
T	Tutorial
P	Practical
BSC	Basic Science Courses
ESC	EngineeringScienceCourses
HSMC	Humanities and Social Sciences includingManagementcourses
PCC	Professional Core Courses
PEC	Professional Elective Courses
LC	Laboratory Courses
MC	Mandatory Courses
PT	Practical Training
S	Seminar

Maharishi Dayanand University Rohtak
B.Tech. (Computer Science and Engineering- Data Science)
Scheme of Studies/Examination w.e.f. 2023-24
Semester-7

Sr. No	Course Code	Course Title	Hours per week			Total Contact Hrs. per week	Credit	Examination Schedule (Marks)				Duration of Exam (Hrs)
			L	T	P			Internal Assessment	Theory	Practical	Total	
1	PCC-AI-401G	Deep Learning	3	0	0	3	3	25	75		100	3
2	PCC-DS-401G	Advanced Architectures of Computers	3	0	0	3	3	25	75		100	3
3	Refer to Annexure-II	Professional Elective-IV	3	0	0	3	3	25	75		100	3
4	Refer to Annexure-III	Open Elective-I	3	0	0	3	3	25	75	-	100	3
5	LC-DS-441G	Project-II	0	0	4	4	2	50	-	50	100	3
6	LC-AI-441G	Deep LearningLab	0	0	2	2	1	25	-	25	50	3
7	LC-DS-443G	Advanced Programming Lab-III	0	0	2	2	1	25	-	25	50	3
8	LC-DS-445G	Seminar/MOOC	0	0	2	2	1	25	-	25	50	3
9	MC-319G*	Business Basics for Entrepreneurs	0	0	2	2	0	-	-	-	-	-
10	PT-CSE-425G (Common with CSE)	Practical Training -II	-	-	1	-	-	-	-	-	-	-
Total			12	0	13	24	17	225	300	125	650	

*__MC-319G__is a mandatory non –credit course based on Business Basics for Entrepreneurs experts sessions in which the students will be evaluated as per their performance of learning and required toget passing Grade as per below.

PT-CSE-425G PracticalTrainingII:TheevaluationofPracticalTraining-IIwillbebasedonseminar, viva-voce,reportsubmittedbythestudents.

Accordingtoperformance,thestudentswill beawardedgradesA,B,C,F; Astudent whoisawarded 'F'gradeisrequiredtorepeat the above courses (MC-319G & PT-CSE-425G)

Grades : Excellent: A, Good : B, Satisfactory: C, Not Satisfactory: F

Maharishi Dayanand University Rohtak
B.Tech. (Computer Science and Engineering- Data Science)
Scheme of Studies/Examination w.e.f. 2023-24
Semester-8

Sr. No	Course Code	Course Title	Hours per week			Total Contact Hrs. per week	Credit	Examination Schedule (Marks)				Duration of Exam (Hrs)
			L	T	P			Internal Assessment	Theory	Practical	Total	
1	PCC-DS-402G	Software Project Management Essentials	3	0	0	3	3	25	75	-	100	3
2	Refer to Annexure-II	Professional Elective-V	3	0	0	3	3	25	75	-	100	3
3	Refer to Annexure-II	Professional Elective-VI	3	0	0	3	3	25	75	-	100	3
4	Refer to Annexure-II	Professional Elective-VII	3	0	0	3	3	25	75	-	100	3
5	Refer to Annexure-III	Open Elective-II	3	0	0	3	3	25	75	-	100	3
6	LC-DS-442G	Project-III	0	0	4	4	2	50	-	50	100	3
7	LC-DS-444G	Advanced Programming Lab-IV	0	0	2	2	1	25	-	25	50	3
8	MC-320G*	Entrepreneurship Management	0	0	2	2	0	-	-	-	-	-
Total			15	0	8	23	18	200	375	75	650	-

*__MC-320G__is a mandatory non –credit course based on Business Basics for Entrepreneurs experts' sessions in which the students will evaluated as per their performance of learning and required toget passing Grade as per below.

Accordingtoperformance,thestudentswill beawardedgradesA,B,C,F.Astudent whoisawarded 'F'gradeisrequiredtorepeat the above courses (MC-320G)

Grades : Excellent: A, Good : B, Satisfactory: C, Not Satisfactory: F

Annexure-II

Professional Elective Courses

Note-Students will have to choose one each from Professional Elective-IV, Professional Elective- V, Professional Elective- VI & Professional Elective- VII.

	Professional Electives for B.Tech (CSE-Data Science)	
	Course Code	**Course title**
Professional Elective-IV	PEC-DS-403G	Big Data Tools and Techniques
	PEC-DS-405G	Advanced Python Programming
	PCC-AI-403G	Applied Machine Learning
Professional Elective- V	PEC-DS-406G	Image and Video Analytics
	PEC-DS-408G	Predictive Analytics Domains
	PEC-DS-409G	Streaming Data Analytics
	PEC-AI-404G	Social Network Analysis
Professional Elective- VI	PEC-SS-404G	Blockchain and Cryptocurrency Technologies
	PEC-DS-407G	Malware Analysis in Data Science
	PEC-AI-408G	Decision Support and Intelligent System
	PEC-AI-407G	Computer Vision& Pattern Recognitions
	PEC-DS-411G	Database Security and Privacy
Professional Elective- VII	PEC-DS-410G	Big Data Visualization
	PEC-AI-418G	Web Intelligence
	PEC-AI-406G	Recommender Systems
	PEC-DS-412G	Text Mining
	PEC-AI-416G	Natural Language Processing& Applications

Annexure-III

Open Elective Courses

Note-Students will have to choose one each from Open Elective- I & Open Elective- II.

	Professional Electives for BTech (CSE-Data Science)	
	Course Code	**Course title**
Open Elective- I	OEC-DS-431G	Open Source Programming
	OEC-DS-432G	Information Storage and Management
	OEC-AI-432G	Cyber Security and AI
	OEC-AI-434G	R Programming
	OEC-AI-436G	Essentials of Hadoop
Open Elective- II	OEC-AI-431G	Data Visualization and Tableau
	OEC-DS-433G	Virtualization
	OEC-ME-451G	Intelligent Vehicle Technology
	OEC-ME-452G	Hybrid and Electrical Vehicle

Deep Learning					
Course code	PCC-AI-401G				
Category	Professional Core Course				
Course title	Deep Learning				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

COURSE OBJECTIVES:

1. Understand the concepts of Neural Networks and Deep Learning
2. Understand Deep neural network and layered learning approach
3. Study and understand CNN and RNN for deep learning
4. Learn and understand Auto Encoders and its applications
5. Understand concept of transfer learning and its applications with keras

Unit-I

Introduction: Feed forward Neural networks, Gradient descent and the back propagation algorithm, Unit saturation, the vanishing gradient problem, and ways to mitigate it. RelU Heuristics for avoiding bad localminima, Heuristics for faster training, Nestors accelerated gradient descent, Regularization, Dropout.Convolutional Neural Networks: Architectures, convolution / pooling layers, normalization, sequence modeling, applications.

Unit-II

Recurrent NeuralNetworks: LSTM, GRU, Encoder Decoder architectures. Introduction to Simple DNN Platform for Deep Learning ,Deep Learning Software Libraries,Deep Unsupervised Learning: Autoencoders,Variational Auto-encoders, Adversarial Generative Networks, Auto-encoder and DBM Attention andmemory models. Dynamic memory networks.

Unit-III

Applications of Deep Learning to ComputerVision: Image segmentation, object detection, automatic image captioning, Image generation withGenerative adversarial networks, video to text with LSTM models, Attention models for computer visiontasks.
Applications of Deep Learning to NLP: Introduction to NLP and Vector Space Model ofSemantics,
Word Vector Representations: Continuous Skip-Gram Model, Continuous Bag-of-Wordsmodel (CBOW), Glove, Evaluations.

Unit-IV

Applications in word similarity, Analogy reasoning: Named EntityRecognition, Opinion Mining using Recurrent Neural Networks: Parsing and Sentiment Analysis usingRecursive Neural Networks: **Sentence Classification using Convolutional Neural Networks**: DialogueGeneration with LSTMs, Applications of Dynamic Memory Networks in NLP, Factoid Question Answering,similar question detection, Dialogue topic tracking, Neural Summarization.

Suggested References Books:

1. I. Goodfellow, Y. Bengio and A. Courville *Deep Learning(1e),* MIT Press,2016.
2. T. Hastie, R. Tibshirani, and J. Friedman *The Elements of Statistical Learning(2e),* Springer,2013.
3. D. Koller, and N. Friedman *Probabilistic Graphical Models,* MIT Press,2010.
4. S. Haykin, *Neural Networks and Learning Machines,* PHI, 2016.
5. Ng's Notes on *Machine Learning* from CS229
6. JasonBrownlee,"DeepLearningwithPython",ebook,2016.

Course outcomes:

1. Apply basic mathematical concepts in Deep Learning
2. Work with powerful framework for supervised learning
3. Deal with Convolution Neural Networks
4. Analyze various types efficient data encoders
5. Apply various network models in deep learning

Advanced Architectures of Computers					
Course code	PCC-DS-401G				
Category	Professional Core Course				
Course title	Advanced Architectures of Computers				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

COURSE OBJECTIVES:

1. Tomake studentsknowaboutheParallelismconceptsinProgramming.
2. Togivethestudentsanelaborateideaaboutthe differentmemorysystems andbuses.
3. Tointroducetheadvancedprocessorarchitecturestothe students.
4. Tomakethestudentsknowabouttheimportanceofmultiprocessorandmulticomputer.
5. Tostudy aboutdataflowcomputerarchitectures.

Unit-I

Parallel Computer Models: The state of computing, Classification of parallel computers, Multiprocessors and multicomputers, Multivector and SIMD computers, PRAM Vs VLSI.
Program and Network Properties: Conditions of parallelism, Data and resource Dependences, Hardware and software parallelism, Program partitioning and scheduling, Grain Size and latency, Program flow mechanisms, Control flow versus data flow, Data flow Architecture, Demand driven mechanisms, Comparisons of flow mechanisms

Unit-II

Pipelining:Advanced Processor Technology, Superscalar and Vector Processors, Linear pipeline processor, nonlinear pipeline processor, Instruction pipeline Design, Mechanisms for instructionpipelining, Dynamic instruction scheduling, Branch Handling techniques, branch prediction, Arithmetic Pipeline Design, Computer arithmetic principles, Static Arithmetic pipeline, Multifunctional arithmetic pipelines

Unit-III

Arithmetic for Computers: Signed and unsigned Numbers, Addition and Subtraction, Multiplication, Division, Floating Point.
CPU Performance and Its factors, Evaluating performance of CPU.

Unit-IV

Memory hierarchy-cache and shared memory concepts-Cache memory organization-cacheaddressing models, Aliasing problem in cache, cache memory mapping techniques-Sharedmemory organization-

Interleaved memory organization, Lower order interleaving, Higher orderinterleaving. Back plane bus systems-Bus addressing, arbitration and transaction.
Recent Trends in parallel processors, Case study of Crey, PowePC,Sun, Intel, AMD and PARAMparallel systems.

Suggested References Books:

1. Kai Hwang, Motwani,Advanced Computer architecture Parallelism ,scalablity ,Programmablity , Mc Graw Hill,N.Y, 2003
2. Kai Hwang, Advanced computer architecture; TMH. 2000
3. D. A. Patterson and J. L. Hennessey, Computer organization and design, Morgan Kaufmann, 2nd Ed. 2002
4. J.P.Hayes, computer Architecture and organization; MGH. 1998
5. Harvey G.Cragon,‖Memory System and Pipelined processors; Narosa Publication. 1998
6. V.Rajaranam&C.S.R.Murthy, ―Parallel computer; PHI. 2002
7. R.K.Ghose, RajanMoona&Phalguni Gupta, Foundation of Parallel Processing, Narosa Publications, 2003
8. Kai Hwang and Zu, Scalable Parallel Computers Architecture, MGH. 2001
9. Stalling W, ComputerOrganisation& Architecture, PHI. 2000
10. D.Sima, T.Fountain, P.Kasuk, Advanced Computer Architecture-A Design space Approach, Addison Wesley, 1997.
11. M.J Flynn, Computer Architecture, Pipelined and Parallel Processor Design; Narosa Publishing. 1998
12. D.A.Patterson, J.L.Hennessy, Computer Architecture :A quantitative approach; Morgan Kauffmann, feb, 2002.
13. Hwan and Briggs, Computer Architecture and Parallel Processing; MGH. 1999

Course outcomes:

1. UnderstandtheConceptof ParallelProcessinganditsapplications.
2. ImplementtheHardwareforArithmeticOperations.
3. Analyzethe performanceof differentscalarComputers.
4. DevelopthePipeliningConceptforagivensetof Instructions.
5. Distinguishtheperformanceofpipeliningandnon-pipeliningenvironmentin aprocessor.

Project-II					
Coursecode	LC-DS-441G				
Category	Laboratory Courses				
Course title	Project-II				
Schemeand Credits	L	T	P	Credits	
	0	0	4	2	
Class work	50 Marks				
Exam	50Marks				
Total	100Marks				
Durationof Exam	03Hours				

Course Objectives:

7. To prepare the student to gain major design and or research experience as applicable to the profession
8. Apply knowledge and skills acquired through earlier course work in the chosen project.
9. Make conversant with the codes, standards, application software and equipment
10. Carry out the projects within multiple design constraints
11. Incorporate multidisciplinary components
12. Acquire the skills of comprehensive report writing

Students will be assigned projects(Applications/Research based) individually or in a group of not more than 3 students depending on the efforts required for completion of project in the subject(s)/area/ skills delivered in this semester using current tools/technology(ies) .

The project will have 4 stages:
(*Marks for internal evaluation are given in brackets)
1. Synopsis submission
(10marks),
2. 1st mid-term progress evaluation(Literature Survey in case of research project) (10 marks)
3.2nd mid-term progress evaluation (Paper Publishing/acceptance in a reputed Journal or Conferenceacceptance/ Presenting) (10 marks)
4. Final submission evaluation (20 marks).

The external examiner will evaluate the project on the basis of idea/quality of project, implementation of the project, project report and/or publication and viva.

Course Outcomes:

Design a system / process or gain research insight into a defined problem as would be encountered in engineering practice taking into consideration its impacton global, economic, environmental and social context.

Deep Learning Lab					
Course code	LC-AI-443G				
Category	Laboratory Course				
Course title	Deep Learning Lab				
Scheme and Credits	L	T	P	Credits	
	0	0	2	1	
Class work	25 Marks				
Exam	25 Marks				
Total	50 Marks				
Duration of Exam	03 Hours				

NOTE:

3. **Lab programs/activities can be designed and developed by the subject faculty using Python, Python Library/suitable Open Source tools/ software.**
4. **A min 15 Lab activities will be carried out from the offered coursecontentsofDeep Learning in the semester.**

In this course, various experiments will be performed, covering various Deep Learning techniques.Experiments covering various deep learning techniques for computer vision, NLP, word vector representations,word similarity, Analogy reasoning, sentence classifications, questions answering, handwritten digits recognitions, sentiment analysis using Keras etc.

Advanced Programming Lab-III					
Course code	LC-DS-443G				
Category	Laboratory Course				
Course title	Advanced Programming Lab-III				
Scheme and Credits	L	T	P	Credits	
	0	0	2	1	
Branches (B. Tech.)	Computer Science and Engineering				
Class work	25 Marks				
Exam	25 Marks				
Total	50 Marks				
Duration of Exam	03 Hours				

NOTE:

A min 15Lab activities relatedto the coursecontentsofProfessional Electives (PE-IV of 7th Sem) can be designed and developed by the subject faculty using suitable Python or any available Open Source tools/ software.

Seminar/MOOC				
Course code	LC-DS-445G			
Category	Laboratory Course			
Course title	Seminar/MOOCs			
Scheme and Credits	L	T	P	Credits
	-	-	2	1
Class work*	25 Marks			
Exam	25 Marks			
Total	50 Marks			
Duration of Exam	03 Hours			

Course Objectives:

The purpose of learning this course is to:

1. Identify an area of interest within the program or a related one (multidisciplinary), carry out a literature survey on it, gain understanding and present the same before an audience. OR
2. Identify a MOOC of interest within the program or a related one (multidisciplinary) available at the MOOCs Platform, registered themselves and earned the certification.

Students have to choose either a seminar or MOOC enrollment.

1. For seminar, Students have to identify a topic/area of interest within the program or a related one (multidisciplinary), carry out a literature survey on it, gain understanding, publish a literature survey paper in a reputed journal / conference/proceeding and present the same work before an audience of the department. Students may do it either individually or in a group of not more than 2 students depending on the efforts required for completion of seminar.
2. In case of MOOC, one(individually)has to Identify a MOOCCourse of interest within the program or a related one (multidisciplinary) available at the MOOCs Platform, registered themselves,gain understanding and earned the certification by qualify the evaluation process.

The Evaluation of this course will have 4 stages:
(*Marks for internal evaluation are given in brackets)

1. Synopsis submission of Seminar/ MOOC identification and registration (5 marks)
2. 1st mid-term progress evaluation (Literature Survey/MOOCs Learning Stage) (5 marks)
3. 2nd mid-term progress evaluation- (5 marks)
   1. Paper Publishing/acceptance or Conference acceptance/ Presenting OR
   2. MOOCs Learning Last Stage
4. Final Presentation evaluation (10 marks).

The external examiner will evaluate the Seminar/MOOC on the basis of idea/quality of literature survey, paper publish/presentation of the topic/MOOC certification or subsequent stage, seminar/MOOC course report and viva.

Course Outcomes:

At the end of this course, learners will be able to:

1. Carry out a self-study of an area of interest and communicate the same to others with clarity

Practical Training-II					
Course code	PT-CSE-425G				
Category	Laboratory Course				
Course title	Practical Training-II				
Scheme and Credits	L	T	P	Credits	Remarks- Common with CSE
	0	0	1	0	
Class work	-				
Exam	-				
Total	-				
Duration of Exam	-				

The evaluation of Practical Training-I will be based on seminar, viva-voce, report submitted by the students. According to performance, the students are awarded grades A, B, C, F. A student who is awarded 'F' grade is required to repeat Practical Training.

Grades :

Excellent: A

Good : B

Satisfactory: C

Not Satisfactory: F

Software Project Management Essentials					
Course code	PCC-DS-402G				
Category	Professional Core Course				
Course title	Software Project Management Essentials				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

COURSE OBJECTIVES:

1. To understand the importance of software project management and identify main stages and stakeholders of a software project
2. To explain the purpose of a project's planning documents and construct the scope statement and the work breakdown structure
3. To portray how the software can assist in project management and articulate what is involved in quality assurance, planning and control on projects
4. To demonstrate RUP, Microsoft project 2010/Above & open source software project management tools

Unit-I

Introduction to Project Management: Importance of software project management - Stages of Project - The Stakeholder of Project – ProjectManagement Framework - Software Tools for Project Management – Microsoft Project 2010/Above – Softwareprojects versus other types of project – Contract management and technical project management

Unit-II

Project Planning : Integration Management: Project Plan Development - Plan Execution Scope Management: Methods forSelecting Projects - Project Charter - Scope Statement - WBS. Stepwise Project Planning: Main Steps inProject Planning Use of Software to Assist in Project Planning Activities

Project Scheduling : Time Management: Importance of Project Schedules - Schedules and Activities - Sequencing and Scheduling Activity Project Network Diagrams: Network Planning Models - Duration Estimating and Schedule Development - Critical Path Analysis - Program Evaluation and Review Technique (PERT) Use of Software to Assist in Project Scheduling Activities - Software Metrics for Project Management: Metrics Sets for Project Management

Unit-III

Software Risk Management : Perspectives of Risk Management - Risk Definition – Risk Categories – Risk Assessment: Approaches,techniques and good practices – Risk Identification / Analysis / Prioritization – Risk Control (Planning /Resolution / Monitoring) – Risk Retention – Risk Transfer - Failure Mode and Effects Analysis (FMEA) –Operational Risks – Supply Chain Risk Management.
Project Cost Management: Importance and Principles of Project Cost Management - Resource Planning - Cost Estimating - Cost Budgeting - Cost Control - Use of Software to assist in Cost Management
Software Quality Management : Project Quality: Stages of Software Quality Management - Quality Planning - Quality Assurance -Quality Control – Quality Standards – Tools for Quality control

Unit-IV

People Management : Leadership styles – Developing Leadership skills – Leadership assessment – Motivating People – Organizational strategy – Management – Team building – Delegation – Art of Interviewing People – Team Management – Rewarding - Client Relationship Management - Organizational behavior: a background, Selecting the right person for the job –Instruction in the best methods– The Oldham-Hackman job characteristics model. Recent Trends and tools.

Suggested References Books:

1. Information Technology Project Management, Kathy Schwalbe, Seven Edition 2013
2. Software Project Management in Practice, Pankaj Jalote, Pearson, 2015.
3. Murali Chemuturi, Thomas M. Cagley, ―Mastering Software Project Management: Best Practices, Tools and Techniques, J. Ross Publishing, 2010
4. Bole Hughes and Mike Cotterell, "Software Project Management", Tata McGraw Hill, Third Edition, 2002
5. Microsoft Project 2010 Bible,ElaineMarmel

Course outcomes:

At the end of course student should be able to -

1. Actively participate or successfully manage a software development project by applying project management concepts
2. Demonstrate knowledge of project management terms and techniques
3. Analyze the Steps involved in analyzing the Software projects and concepts to meet the estimation of the software Projects.
4. Work on Microsoft project, IBM RUP & open source software project management tools.
5. Estimate the organizing team based on industry exposure.

Project-III					
Coursecode	LC-DS-442G				
Category	Laboratory Courses				
Course title	Project-III				
Schemeand Credits	L	T	P	Credits	
	0	0	4	2	
Class work	50 Marks				
Exam	50Marks				
Total	100Marks				
Durationof Exam	03Hours				

Course Objectives:

1. To prepare the student to gain major design and or research experience as applicable to the profession
2. Apply knowledge and skills acquired through earlier course work in the chosen project.
3. Make conversant with the codes, standards, application software and equipment
4. Carry out the projects within multiple design constraints
5. Incorporate multidisciplinary components
6. Acquire the skills of comprehensive report writing

Students will be assigned projects(Applications/Research based) individually or in a group of not more than 3 students depending on the efforts required for completion of project in the subject(s)/area/ skills delivered in this semester using current tools/technology(ies) .

The project will have 4 stages:
(*Marks for internal evaluation are given in brackets)
1. Synopsis submission (10marks)
2. 1st mid-term progress evaluation (Literature Survey in case of research project) (10 marks)
3. 2nd mid-term progress evaluation (Paper Publishing/acceptance in a reputed Journal or Conference acceptance/ Presenting) (10 marks)
4. Final submission evaluation (20 marks)

The external examiner will evaluate the project on the basis of idea/quality of project, implementation of the project, project report and/or publication and viva.

Course Outcomes:

Design a system / process or gain research insight into a defined problem as would be encountered in engineering practice taking into consideration its impacton global, economic, environmental and social context.

Advanced Programming Lab-IV					
Course code	LC-DS-444G				
Category	Laboratory Course				
Course title	Advanced Programming Lab-IV				
Scheme and Credits	L	T	P	Credits	
	0	0	2	1	
Class work	25 Marks				
Exam	25 Marks				
Total	50 Marks				
Duration of Exam	03 Hours				

NOTE:

1. **Lab programs/activities can be designed and developed by the subject faculty using Python or any suitable OpenSource tools/ software.**
2. **Min 25 Lab activities will be carried out from the offered coursecontentsofProfessional Electives in the semester.**
3. **Cases studies may be given to the students based on above courses.**

Professional Electives

Big Data Tools & Techniques					
Course code	PEC-DS-403G				
Category	Professional Elective Course				
Course title	Big Data Tools & Techniques				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Gain knowledge about the various tools and techniques used in big data analytics
2. Learn the fundamentals of Hadoop and the related technologies
3. Understand the basics of development of applications using MapReduce, HDFS, YARN
4. Learn the basics of Pig, Hive and Sqoop
5. Learn the basics of Apache Spark, Flink and understand the importance of NoSQL databases
6. Learn about Enterprise Data Science and data visualization tools

Detailed Course

Unit-I

Overview of Big Data Analytics, Introduction to data analytics and big data, Big data mining, Technical elements of the Big Dataplatform , Analytics Toolkit, Components of theanalytics toolkit, Distributed and Parallel Computing for Big Data, Cloud computing and Big Data, In-Memory Computing Technology for Big DataFundamentals of Hadoop, Hadoop Ecosystem, The core modules of Hadoop , Introduction to Hadoop MapReduce, Introduction to Hadoop YARN

Unit-II

MapReduce - Analyzing data with Unix tools andHadoop, Scaling Out – Data Flow, CombinerFunctions, Hadoop Streaming, HDFS, Hadoop filesystems, Java Interface to Hadoop, YARN, Job Scheduling, Hadoop I/O ,Data Integrity , Compression , Serialization, File based Data Structures, Developing a MapReduce Application

Unit-III

Setting up a Hadoop cluster, Cluster specification and setup, Hadoop configuration, YARN configuration, Introduction to Pig, Installing and running pig, Basics of Pig Latin, Introduction to Hive
Installing and running Hive , Introduction to HiveQL, Introduction to Zookeeper, Installing and running Zookeeper, The Zookeeper Service, Flume Architecture, Introduction to Sqoop; Introducing Oozie, Apache Spark, Limitations of Hadoop and overcoming the limitations, Core components and architecture ofSpark; Introduction to Apache Flink, Installing Flink, Batch analytics using Flink;

Unit-IV

Big Data Mining with NoSQL, Why NoSQL?,NoSQL databases; Introduction to HBase, Introduction to MongoDB, Cassandra.
Enterprise Data Science- Overview, Data Science Solutions in the enterprise, Visualizing Big Data, Enterprise data science – MachineLearning and AI, Enterprise Infrastructure solutions, Using Python and R for visualization, Big Data Visualization Tools, Data Visualization with Tableau, Case Studies: Hadoop, Spark, NoSQL

References Books:

1. Tom White, Hadoop: The Definitive Guide, 3rd Edition,O'Reilly,2012.
2. Sridhar Alla,Big Data Analytics with Hadoop3,Packt,2018.
3. NatarajDasgupta,Practical Big Data Analytics,Packt,2018.
4. DTEditorialServices,Big Data:Black Book,2016.

Course Outcomes:

After successfully completing the course the student should be able to

1. Use the var ious tools and techniques in big data analytics
2. Apply Hadoop and related technologies to big data analytics
3. Apply MapReduce, HDFS and YARN develop big data applications
4. Develop applications using Pig, Hive and Sqoop
5. Apply Apache Spark and Flink to applications and understand the importance of NoSQL databases
6. Understand the applications of Enterprise Data Science and data visualization tools

Advanced Python Programming				
Course code	PEC-DS-405G			
Category	Professional Elective Course			
Course title	Advanced Python Programming			
Scheme and Credits	L	T	P	Credits
	3	0	0	3
Class work	25 Marks			
Exam	75 Marks			
Total	100 Marks			
Duration of Exam	03 Hours			

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1.To be able to apply advanced python programming concepts for industry standard problems.
2.To perform advanced Data Preprocessing tasks like Data Merging and Mugging
3.To be able to develop powerful Web-Apps using Python

Detailed Course

Unit-I

Data Structures: Problem solving using Python Data Structures : LIST, DICT, TUPLES and SET-Functions and Exceptions – Lamda Functions and Parallel processing – MAPS – Filtering - Itertools – Generators.

Classes & Objects: Classes as User Defined Data Type ,Objects as Instances of Classes, Creating Class and Objects, Creating Objects By Passing Values, Variables & Methods in a Class Data , Abstraction, Data Hiding, Encapsulation, Modularity, Inheritance, Polymorphism

Unit-II

Python Multithreading: Python Multithreading and Multiprocessing Multithreading and multiprocessing Basics – Threading module and example – Python multithreading - Multithreaded Priority Queue.

Data Processing: Handling CSV, Excel and JSON data - Creating NumPy arrays, Indexing and slicing in NumPy, Downloading and parsing data, Creating multidimensional arrays, NumPy Data types, Array Attribute, Indexing and Slicing, Creating array views copies, Manipulating array shapes I/O – MATPLOT LIB

Unit-III

Data Science Perspectives: Using multilevel series, Series and Data Frames, Grouping, aggregating, Merge Data Frames, Generate summary tables, Group data into logical pieces, Manipulate dates, Creating metrics for analysis.

Data Handling Techniques: Data wrangling, Merging and joining, - Loan Prediction Problem, Data Mugging using Pandas

Unit-IV

Web Applications: Web Applications With Python – Django / Flask / Web2Py – Database Programming – NoSQL databases - Embedded Application using IOT Devices - Building a Predictive Model for IOT and Web programming; Recent Trends and Contemporary issues.

References Books:

1. Doug Farrell, The Well Grounded Python Developer; Manning Publications, 2021
2. Paul Barry, Head-First Python, O-Reilly Media, 2016
3. Zed A Shaw, Learn Python the Hard Way - A Very Simple Introduction to the Terrifyingly Beautiful World of Computers and Code, Addison Wesley Press, 2013
4. Eric Mathews, Python Crash Course, Second Edition, No Starch Press, 2019
5. Michael Kennedy, Talk Python: Building Data-Driven Web Apps with Flask and SQLAlchemy, Manning Publications, 2020

Course Outcomes:

After successfully completing the course the student should be able to

1. Understand the nuances of Data Structures
2. Derive an understanding of a classes and objects and their potential
3. Gain knowledge of multithreading concepts and implementing the same
4. Appreciate the difference between different data processing techniques
5. Learn to apply Python features for Data Science
6. Get an insight into Metrics Analysis
7. Develop web-apps and build models for IoT

Applied Machine Learning					
Course code	PCC-AI-403G				
Category	Professional Core Course				
Course title	Applied Machine Learning				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Analyze the text data using Machine Learning
2. Analyze the audio data using Machine Learning
3. Analyze Time series and Sequential data using Machine Learning
4. Analyze the Image Content using Machine Learning
5. Visualize the data

Unit-I

Text Feature Engineering-Introduction, Cleaning text data, Preprocessing data using, tokenization , tagging and categorising words, Sequential tagging, Backoff tagging, Creating features from text data-
Stemming,Lemmatising, Bagging using random forests, Implementing bag of words, Testing prepared dataAnalyze the results, Building a text classifier, Analyzing the sentiment of a sentence, Implement the sentiment analysis ofa sentence, Identifying patterns in text using topicmodeling, Implement identifying patterns in text using topic modeling, Case study- Twitter Data ;
Speech Recognition- Introduction, Reading audio data, Plotting audio data, Transforming audio signals into the frequency domain, Apply Fourier transform signal and plot, Generating audio signals with customparameters, Generate the time axis,

Unit-II

Speech Recognition- Synthesizing music, Construct the audio sample -amplitudeand frequency, synthesizer function, Extracting frequency domain features, MFCC and filter bank features, Building Hidden Markov Models, HMM training and prediction, Building a speech recognizer, MFCC features, Case study.
Dissecting Time Series and SequentialData – Introduction, Transforming data into the time seriesformat, Pandas and Numpy to convert Time Series data, Plotting time series data, Slicing time series data, Operating on time series data, Plotting sliced time series data, Operating on time series data, Extracting statistics from time seriesdata, Correlation coefficients, Plotting and understanding correlations, Building Hidden Markov Models forsequential data, Prepare the Time Series data, Train Gaussian HMM, Visualizing the model, Building Conditional Random Fields forsequential text data, CRF Model, Analyzing stock market data using, Hidden Markov Models, Train the HMM and visualize

Unit-III

Image Content Analysis, Computer Vision, Operating on images using OpenCV-Python, Learn to extract and load the image, Detecting edges, Histogram equalization, Sobel filter, Laplacian edge detector, Canny edge detector, Histogram equalization, Visualize gray scale image, Detecting corners, Understand the output corner detection image, Detecting SIFT feature points, SIFT feature detection, Visualize the feature detected image, Building a Star feature detector, Detect features using the Star featuredetector, Visualize keypoints on the input image, Creating features using visual codebookand vector quantization, Method to quantize the data points

Unit-IV

Biometric Face Recognition - Face detection from the image and video, Capturing and processing video from awebcam, Resizing and Scaling, Building a face detector using Haarcascades , determine the location of a face in thevideo frames captured from the webcam, Face detector on the grayscale image, Building eye and nose detectors, Face cascade classifier, Visualize eye and nose detector , Performing Principal ComponentsAnalysis, PCA in face recognition systems, Convert the dataset from a five-dimensional set to a two-dimensionalset , Kernel Principal Components Analysis , Plot Kernel PCA-transformed data, Performing blind source separation, Independent Components Analysis, Perform Kernel PCA, Plot the PCA-transformed data

References Books:

1. PrateekJoshiandco,Python:RealWorldMachineLearning,PacktPublishing,2016
2. SebastianRaschka,PythonMachineLearning,PacktPublishing,2013.
3. RichertCoelho,BuildingMachineLearningSystemswithPython,PacktPublishing,2016
4. MichaelBowles,MachineLearninginPython,Wiley&Sons,2015

Course Outcomes:

After successfully completing the course the student should be able to

1. Identifying patterns in text using topic modeling
2. Building a speech recognizer
3. Extracting statistics from time series data,
4. Building Conditional Random Fields for sequential text data
5. Building an object recognizer

Blockchain and Cryptocurrency Technologies				
Course code	PEC-DS-404G			
Category	Professional Elective Course			
Course title	Blockchain and Cryptocurrency Technologies			
Scheme and Credits	L	T	P	Credits
	3	0	0	3
Class work	25 Marks			
Exam	75 Marks			
Total	100 Marks			
Duration of Exam	03 Hours			

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

1. To understand the mechanism of Blockchain and Cryptocurrency.
2. To understand the functionality of current implementation of blockchain technology.
3. To understand the required cryptographic background.
4. To explore the applications of Blockchain to cryptocurrencies and understanding limitations of current Blockchain.
5. An exposure towards recent research.

UNIT-I

Introduction to Cryptography and Cryptocurrencies: Cryptographic Hash Functions, Hash Pointers and Data Structures, Digital Signatures, PublicKeys as Identities, A Simple Cryptocurrency.

How Blockchain Achieves and How to Store and Use: Decentralization-Centralization vs. Decentralization-Distributed consensus, Consensus with- out identity using a blockchain, Incentives and proof of work. Simple Local Storage, Hot and Cold Storage, Splitting and Sharing Keys, Online Wallets and Exchanges, Payment Services, Transaction Fees, Currency Exchange Markets.

UNIT-II

Mechanics of Bitcoin: Bitcoin transactions, Bitcoin Scripts, Applications of Bitcoin scripts, Bitcoin blocks, The Bit- coin network, Limitations and improvements.

Bitcoin Mining: The task of Bitcoin miners, Mining Hardware, Energy consumption and ecology, Mining pools, Mining incentives and strategies

UNIT-III

Bitcoin and Anonymity: Anonymity Basics, How to De-anonymize Bitcoin, Mixing, Decentralized Mixing, Zerocoin andZerocash.

Community, Politics, and Regulation: Consensus in Bitcoin, Bitcoin Core Software, Stakeholders: Who‟s in Charge, Roots of BitcoinGovernments Notice on Bitcoin, Anti Money Laundering Regulation, New York‟s Bit License Proposal. Bitcoin as a Platform: Bitcoin as an Append only Log, Bitcoins as

Smart Property, Secure Multi Party Lotteries in Bitcoin, Bitcoin as Public Randomness, Source-Prediction Markets, and Real World Data Feeds.

UNIT-IV

Altcoins and the CryptocurrencyEcosystem: Altcoins: History and Motivation, A Few Altcoins in Detail, Relationship Between Bitcoin and Altcoins, Merge Mining-Atomic Crosschain Swaps-6 BitcoinBacked Altcoins, Side Chains,Ethereum and Smart Contracts.Recent Trends and applications.

Reference Books:

1. Narayanan, A., Bonneau, J., Felten, E., Miller, A., and Goldfeder, S. (2016). Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press.
2. Antonopoulos, A. M. (2014). Mastering Bitcoin: unlocking digital cryptocurrencies. OReilly Media, Inc.".
3. Franco, P. (2014). Understanding Bitcoin: Cryptography, engineering and economics. John Wiley and Sons.

Course Outcomes:

After successfully completing the course the student should be able to

1. To Understand and apply the fundamentals of Cryptography in Cryptocurrency
2. To gain knowledge about various operations associated with the life cycle of Blockchain and Cryptocurrency
3. To deal with the methods for verification and validation of Bitcoin transactions
4. To demonstrate the general ecosystem of several Cryptocurrency
5. To educate the principles, practices and policies associated Bitcoin business

Image and Video Analytics				
Course code	PEC-DS-406G			
Category	Professional Elective Course			
Course title	Image and Video Analytics			
Scheme and Credits	L	T	P	Credits
	3	0	0	3
Class work	25 Marks			
Exam	75 Marks			
Total	100 Marks			
Duration of Exam	03 Hours			

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

1. To impart knowledge on the concepts of computer vision and intelligent video processing with analytics to process massive video feed and derive insights with machine intelligence
2. To explore and demonstrate real time video analytics in solving practical problems of commercial and scientific interests.

Unit-I

Introduction Digital Image Processing: Characteristics of Digital Image, basic relationship between pixels, fundamental operations on image, image sampling and quantization, image transformations models;

Unit-II

Basic Techniques of image processing Fundamentals of spatial filtering: spatial correlation and convolution, smoothing blurring, sharpening, basics of filtering in the frequency domain: smoothing, blurring, sharpening, histograms and basic statistical models of image;

Unit-III

Transformations andSegmentations: Colour models and Transformations, image and video Segmentation, image and videodemonising, image and Video enhancement- Image and Video compression;

Unit-IV

Detection andClassification Object detection and recognition in image and video, texture models Image and videoclassification models object tracking in Video; **Applications and Case studies Industrial**-Transportation& travel, remote sensing, video Analytics: IoT Video Analytics Architectures.

References:

1. R.C Gonzalez and R.E Woods, *Digital Image Processing*, Pearson Education, 4th edition, 2018.
2. N.M. Tekalp, *Digital Video Processing*, (1e), Pearson, 2017
3. A.K. Jain, *Fundamentals of Digital Image Processing*, PHI, New Delhi, 1995
4. Rick Szelisk, *Computer Vision: Algorithms and Applications*, Springer 2011.

5. C. Shan, F. Porikli, T. Xiang and S. Gong, *Video Analytics for Business Intelligence*,(1e), Springer, 2012

Course Outcomes:

After successfully completing the course the student should be able to

1. Understand basic image and video processing concepts
2. Explore both the theoretical and practical aspects of intelligent perception and understanding of images.
3. Apply principles and techniques of video processing in applications related to intelligent and automated visual system design and analysis.
4. Analyze techniques for action representation and recognition
5. Develop algorithms that can perform high-level visual recognition tasks on real-world images and videos.

Malware Analysis in Data Science					
Course code	PEC-DS-407G				
Category	Professional Elective Course				
Course title	Malware Analysis in Data Science				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. To understand and analyse malware using static and dynamic analysis
2. To observe malware behaviour
3. To build and analyse Malware Networks
4. To identify adversary groups through shared code analysis
5. To catch vulnerabilities by building your own machine learning detector
6. To measure malware detector accuracy
7. To identify malware campaigns, trends, and relationships through data visualization

Unit-I

Static and Dynamic Malware Analysis: Basic Static Malware Analysis: Static Analysis Definition - Microsoft Windows PE format – Dissecting PE format using PE file – Examining Malware images – Strings – Factors that Limit StaticAnalysis ;

Introduction to Dynamic Analysis: Why use Dynamic Analysis – Dynamic analysis for data science – Basic tools for dynamic analysis – Limitation of basic dynamic analysis

Unit-II

Identifying Attacks: Identifying Attack Campaigns using Malware Networks: Bipartite Networks – Building and Visualizing Malware Networks – Building a shared image relationship network

Shared Code Analysis: Samples comparisons by extracting features – Jaccard Index to quantify similarity – Evaluate Malware Shared Code estimation methods – Building a Similarity Graph – Persistent Malware Similarity Search System

Unit-III

Malware Detectors and Evaluation: Machine Learning Based Malware Detectors **:** Steps for building detector – Understanding FeatureSpaces and Decision Boundaries – Overfitting and Underfitting – Major

Types of Machine LearningAlgorithms: Logistic Regression – K-Nearest Neighbors – Decision Trees – Random Forest - Toy Decision Tree based Detector – Real World Learning Detectors with sklearn – Industrial Strength Detector;

Evaluating Malware Detection System: Four possible Detection Outcomes – Considering base rates in evaluation- Evaluating the Detector‘s performance;

Unit-IV

Visualizing Malware Trends: Understanding our Malware Dataset – Using matplotlib to visualize data – Using seaborn to visualize Data; Deep Learning Basics - Building a Neural Network Malware Detector with Keras; Contemporary issues

Reference Books:

1. Malware Data Science – Attack Detection and Attribution , Joshua Saxe and Hillary Sanders, No Starch Press, 2018
2. Machine Learning and Security: Protecting Systems with Data and Algorithms, Clarence Chio, David Freeman, 1st Edition, O‘Reilly Media, Feb 2018.
3. Mastering Malware Analysis: The complete malware analyst's guide to combating malicious software, APT, cybercrime, and IoT attacks, Alexey Kleymenov, Amr Thabet, 1st Edition,Packt publishing, 2019.
4. Practical Malware Analysis, Michael Sikorski, Andrew Honig, No Starch Press, 2012

Course Outcomes:

After successfully completing the course the student should be able to

1. Analyse malware behaviour and identify its adversary groups
2. Build your own machine learning detector system to catch vulnerabilities and to measure its accuracy
3. Visualize malware threat data to reveal attack campaigns and trends

Predictive AnalyticsDomains					
Course code	PCC-DS-408G				
Category	Professional Elective Course				
Course title	Predictive Analytics Domains				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. It introduces theoretical foundations, algorithms, methodologies for analysing data in various domains such Retail, Finance, Risk and Healthcare.

Unit-I

Retail Analytics: Understanding Customer: Profiling and Segmentation, Modelling Churn. Modelling LifetimeValue, Modelling Risk, Market Basket Analysis.

Risk Analytics : Risk Management and Operational Hedging: An Overview, Supply Chain Risk Management, A Bayesian Framework for Supply Chain Risk Management, Credit Scoring and Bankruptcy Prediction

Unit-II

Financial Data Analytics: Financial News analytics: Framework, techniques, and metrics, News events impact market sentiment, Relating news analytics to stock returns

Financial Time Series Analytics: Financial Time Series and Their Characteristics, Common Financial Time Series models,Autoregressive models, Markov chain models, Time series models with leading indicators, Long termforecasting

Unit-III

Health care Analytics: Introduction to Healthcare Data Analytics, Electronic Health Records, Privacy-Preserving Data Publishing Methods in Healthcare, Clinical Decision Support Systems

Healthcare Data Analytics : Natural Language Processing and Data Mining for Clinical Text: Core NLP Components, Information Extraction and Named Entity Recognition, Social Media Analytics for Healthcare: Tracking of Infectious Disease Outbreaks, Readmission risk Prediction

Unit-IV

Genomic Data Analytics : Microarray Data, Microarray Data Analysis, Genomic Data Analysis for Personalized Medicine,Patient Survival Prediction from Gene Expression Data, Genome Sequence Analysis; Current trends and Contemporary issues

References Books:

1. Chris Chapman, Elea McDonnell Feit "R for Marketing Research and Analytics", Springer, 2015.
2. Olivia Parr Rud "Data Mining Cookbook: Modeling Data for Marketing, Risk, and Customer Relationship Management", Wiley, 2001.
3. Chandan K. Reddy, Charu C. Aggarwal "Healthcare Data Analytics", CRC Press, 2015.
4. Rene Carmona "Statistical Analysis of Financial Data in R", Springer, 2014.
5. James B. Ayers "Handbook of Supply Chain Management" Auerbach Publications, 2006.
6. PanosKouvelis, Ling xiu Dong, OnurBoyabatli, Rong Li "The Handbook of Integrated Risk Management in Global Supply Chains", Wiley, 2012.

Course Outcomes:

After successfully completing the course the student should be able to

1. Recognize challenges in dealing with data sets in domains such as finance, risk and healthcare.
2. Identify real-world applications of machine learning in domains such as finance, risk and healthcare.
3. Identify and apply appropriate algorithms for analyzing the data for variety of problems in finance, risk and healthcare.
4. Make choices for a model for new machine learning tasks based on reasoned argument

Streaming Data Analytics					
Course code	PEC-DS-409G				
Category	Professional Elective Course				
Course title	Streaming Data Analytics				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. It introduces theoretical foundations, algorithms, methodologies, and applications ofstreaming data and also provide practical knowledge for handling and analyzing streaming data.

Unit-I

Introduction : Characteristics of the data streams, Challenges in mining data streams Requirements andprinciples for real time processing, Concept drift Incremental learning..

Data Streams : Basic Streaming Methods, Counting the Number of Occurrence of the Elements in a Stream,Counting the Number of Distinct Values in a Stream, Bounds of Random Variables, PoissonProcesses, Maintaining Simple Statistics from Data Streams, Sliding Windows, Data Synopsis, Change Detection: Tracking Drifting Concepts, Monitoring the Learning Process

Unit-II

Decision Trees : Very Fast Decision Tree Algorithm (VFDT), The Base Algorithm, Analysis of the VFDT Algorithm,Extensions to the Basic Algorithm: Processing Continuous Attributes, Functional Tree Leaves,Concept Drift.

Clustering from Data Streams : Clustering Examples: Basic Concepts, Partitioning Clustering - The Leader Algorithm, Single Pass k-Means, Micro Clustering, Clustering Variables: A Hierarchical Approach

Unit-III

Frequent Pattern Mining: Mining Frequent Item sets from Data Streams- Landmark Windows, Mining Recent Frequent Itemsets, Frequent Item sets at Multiple Time Granularities, Sequence Pattern Mining-Reservoir Sampling for Sequential Pattern Mining over data streams

Evaluating Streaming Algorithms: Evaluation Issues, Design of Evaluation Experiments, Evaluation Metrics, Error Estimators using a Single Algorithm and a Single Dataset, Comparative Assessment, The 0-1 loss function, Evaluation Methodology in Non-Stationary Environments, The Page-Hinkley Algorithm.

Unit-IV

Complex Event Processing: Introduction to Complex Event Processing, Features of CEP, Need for CEP, CEP Architectural Layers, Scaling CEP, Events, Timing and Causality, Event Patterns, Rules and Constraint, STRAW- EPL, Complex Events and Event Hierarchies; Current trends and Contemporary issues

References Books:

1. Joao Gama, "Knowledge Discovery from Data Streams", CRC Press, 2010.
2. David Luckham, "The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems", Addison Wesley, 2002.
3. Charu C. Aggarwal, "Data Streams: Models And Algorithms", Kluwer Academic Publishers, 2007

Course Outcomes:

After successfully completing the course the student should be able to

1. Recognize the characteristics of data streams that make it useful to solve real-world problems.
2. Identify and apply appropriate algorithms for analyzing the data streams for variety of problems.
3. Implement different algorithms for analyzing the data streams
4. Identify the metrics and procedures to evaluate a model

Big Data Visualization					
Course code	PEC-CSE-410G				
Category	Professional Elective Course				
Course title	Big Data Visualization				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Understand the key techniques used in visualization which includes data models, graphical perception and techniques specifically for visual encoding and interaction
2. Obtain an exposure to common data domains and thecorresponding analysis tasks which includes multivariate data and text
3. Get hands-on experience in building and evaluating visualization systems
4. Gain knowledge in data visualization aides
5. Understand the significance of data by placing it in a visual context
6. Utilize the knowledge by reading and discussingresearch papers from the visualization literature

Detailed Course

Unit-I

Introduction to Big Data Visualization, Challenges of Big Data Visualization , Categorization, Visualization Philosophies, Approaches to Big Data Visualization, Quality of Visualization, Infographics versus Data Visualization, Exploration versus Explanation, Informative versus Persuasive versusVisual Art, Ingredients of Successful Visualizations, Choose Appropriate Visual EncodingsNaturalOrdering,DistinctValues, Redundant Encoding, Defaults versus Innovative Formats,Readers' Context,Compatibility with Reality,Patterns and Consistency, Selecting Structure, Position: Layout and Axes, The Meaning of Placement and Proximity, Patterns of Organization-Specific GraphsLayouts, and Axis Styles, Appropriate Use of Circles and Circular Layouts

Unit-II

Definitions and explanations ofvisualization categories, Exploring RIn big data, Example with Patient Medical History, Digging in with R, No looping , Comparisons and Contrasts, Tendencies, Dispersion, Data quality categorized, Data Manager, Data Manager and big data, Example-Reformatting-A little Setup, Adding Script Code, Executing the scene, Status and relevance, Naming the nodes, Consistency ,Reliability ,Appropriateness , Accessibility and Other Output nodes

Unit-III

An Introduction to Visualization tools ,Visualization tools and big data, Example 1 – Sales transactions Adding more context, Wrangling the data, *Trifacta* Script panel, A visualization dashboard, Experimenting with the data and build thevisualization, Data *pane_core* details, Constructing Dashboards, Saving and Presenting the work, Visualization re-coloring, resizing, adding orchanging labels, Filters and Measure Names, Example-Promotion Spend Effect on Sales , Sales and spend, Sales v Spend and Spend as % of SalesTrend , Tables and indicators

Unit-IV

Introduction to D3, D3 and big data, Basic Examples, Getting started with D3, D3 visualization sample templates, Big data visualization using D3, Displaying Results Using D3, Create a summary file for visualization, Visualization using HTML document, Data visualization showing the stackedview, Visual transitions, Multiple donuts, Another twist on bar chart visualizationswith examples, D3 Stacked Area via Nest template, Adopting the sample, Visualization changes format

Case Studies: 1: Color considerations with a dark background, 2: Leveraging animation in thevisuals youpresent, 3: Logic in order, 4: Strategies for avoiding thespaghettigraph, 5: Alternatives to pies

Recent Trends and Contemporary issues.

References Books:

1. Big Data Visualization,James D. Miller,Copyright © 2017 Packt Publishing
2. Designing Data Visualizations,by Noah Iliinsky and Julie Steele© 2011
3. Storytelling with data - a data visualization guide for business professionals by colenussbaumerknaflic,Wiley publications
4. Tableau Your Data! by Daniel G. Murray and the InterWorks BI Team, Wiley publications

Course Outcomes:

After successfully completing the course the student should be able to

1. Design and exploring the result with data visualizations
2. Conducting exploratory data analysis using visualization techniques and tools.
3. Visual presentations of data for effective communication.
4. Designing and evaluating color palettes for visualization based on principles of perception.
5. Using the knowledge of perception and cognition to evaluate visualization design alternatives
6. Identifying opportunities for the application of data visualization in various domains.

Web Intelligence					
Course code	PEC-AI-418G				
Category	Professional Elective Course				
Course title	Web Intelligence				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Understand the topics of Web Intelligence
2. Study models of information retrieval, semantic webs, search engines, and web mining.
3. Gain knowledge on the algorithmic aspect of Web Intelligent systems
4. Acquire knowledge on Data mining techniques
5. Understand the impact of Social Network Design for Web Intelligence
6. Gain Knowledge on different approaches required for studying the impact of social network for Web Intelligence

Detailed Course

Unit-I

Introduction to Web Intelligence, What is Web Intelligence?, Benefits of Intelligent Web:Whatapplications can benefit from web intelligence, Wisdom Web, Ingredients of Intelligent Web, Topics of Web Intelligence, How can I build intelligence in my own application?, Examples of intelligent web applications, Fallacies of Intelligent applications, Related Technologies

Unit-II

Information Retrieval-Introduction, Document Representation, Retrieval Models, Evaluation of Retrieval Performance, Semantic Web-Introduction, The Layered-Language Model, Metadata and Ontologies, Ontology Languages for the Web, Tool Environment for the Ontology, RDFferret-Full Text Search and RDFQuerying, OntoShare-Community support, Onto Edit-Ontology Development, OntoView-Change Management forOntologies, Sesame-Repositories for Ontologies andData, CORPORIUM- Information Extraction

Unit-III

Data Mining Techniques-Classification, Clustering andAssociation, Associations, Web Usage Mining-Web-Log processing, Web Usage Mining -Analyzing Web Logs, Applications of Web Usage Mining, Clustering of Web Users, Applications of Web Usage Mining-Classification Modeling of Web Users,

Applications of Web Usage Mining-Association Mining of Web Usages, Sequence-Pattern Analysis of Web Logs;
Web Content Mining- Introduction, Web Crawlers, Search Engines, Personalization of Web Content, Multimedia Information Retrieval, Web Structure Mining- Modeling WebTopology, PageRank Algorithm, Hyperlink-Induced Topic Search (HITS), Random Walks on the Web;

Unit-IV

Social Network Design for WebIntelligence: Introduction, Web Intelligence , Overview of Social Intelligence Design: The Travelling Conversation Model, A Broadcast-Based Approach, AConversational Agent-Based Approach, Smart Environment based approach, Psychological Evaluation, TechnicalIssues, Groups and Communities, Issues ofSocial Intelligence Design, Applications ofSocial Intelligence Design, Case Study-Putting it all together : anintelligent news portal, Applying Web Intelligence for Business Intelligence

References Books:

1. Akerkar, R. &Lingras, P. (2008). Building an Intelligent Web: Theory and practice. Jones and Bartlett Publishers, Sudbury, Massachusetts. ISBN-13: 978-0-7637-4137-2
2. Marmanis&Babenko: Algorithms of the Intelligent Web, Manning Publications, 2009, ISBN:978-1933988665
3. Witten, Ian H. & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edition, Morgan Kaufman. ISBN 0120884070, 9780120884070
4. Bing Liu: Web Data Mining, Springer, 2nd ed. 2011 (view online or download fromSpringerlink)
5. Manning, Raghavan and Schuetze: Introduction to Information Retrieval, CambridgeUniversityPress,2008(bookavailableonline)
6. N. Zhong, J.M. Liu, Y.Y. Yao, Web Intelligence (Springer, 2003)

Course Outcomes:

After successfully completing the course the student should be able to

1. Acquire the knowledge on topics and benefits of Web Intelligence
2. Acquire the ability to build models of information retrieval, semantic webs, search engines, and web mining.
3. Understand the basic ideas of Multimedia Information Retrieval
4. Acquire knowledge to use web crawlers and fetch relevant information
5. Acquire knowledge to refine the social network design approached used for developing intelligent web
6. Apply the knowledge of different web intelligence based algorithms in practical applications

Recommender System					
Course code	PEC-AI-406G				
Category	Professional Elective Course				
Course title	Recommender System				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. To provide the students a foundation of recommender systems concepts
2. To expose the students to a variety of recommender systems algorithms
3. To provide the students a knowledge on the different evaluation methods of recommendersystems
4. To provide the students an ability to classify the different recommender systems solutions
5. To build up the capability to develop a recommender systems solution

Unit-I

Introduction: Recommender system functions, Linear Algebra notation: Matrix addition, Multiplication,transposition, and inverses; covariance matrices, Understanding ratings, Applications of recommendationsystems, Issues with recommender system.

Collaborative Filtering: User-basednearest neighbor recommendation, Item-based nearest neighbor recommendation, Model based and preprocessingbased approaches, Attacks on collaborative recommender systems.

Unit-II

Content-basedrecommendation: High level architecture of content-based systems, Advantages and drawbacksof content-based filtering, Item profiles, discovering features of documents, obtaining item features fromtags, representing item profiles, Methods for learning user profiles, Similarity based retrieval, Classificationalgorithms.

Knowledge based recommendation: Knowledge representation and reasoning, Constraintbased recommenders, Case based recommenders.

Hybrid approaches: Opportunities for hybridization,Monolithic hybridization design: Feature combination, Feature augmentation, Parallelized hybridizationdesign: Weighted, Switching, Mixed, Pipelined hybridization design: Cascade Meta-level, Limitations ofhybridization strategies.

Unit-III

Evaluating Recommender System: Introduction, General properties ofevaluation research, Evaluation designs, Evaluation on historical datasets, Error metrics, Decision-Supportmetrics, User-Centred metrics.
Context-Sensitive RecommenderSystems - The Multidimensional Approach -Contextual Pre-filtering: A Reduction-Based Approach - Post-Filtering Methods - Contextual Modeling
Time and Location sensitive Recommender System: Temporal Collaborative Filtering - Discrete Temporal Models - Location-Aware Recommender Systems
Social and Trust Centric Recommender System: Multidimensional Models for Social Context - Network-Centric and Trust-Centric Methods – UseInteraction in Social Recommenders

Unit-IV

Attack Resistant Recommender System: Understanding the Trade-Offs in Attack Models - Types of Attacks -Detecting Attacks on Recommender Systems - Strategies for Robust Recommender Design
Advanced Topics: Learning to Rank - Group Recommender Systems - Multi-Criteria Recommender - Active Learning in Recommender Systems - Privacy in Recommender Systems - Application Domains - Portal Content Personalization - Google News Personalization - Computational Advertising versus Recommender Systems; Recent trends and Contemporary Issues.

References Books:

1. C.C. Aggarwal , *Recommender Systems: The Textbook* (1e), Springer, 2016.
2. N Manouselis, H. Drachsler, K. Verbert and E. Duval., *Recommender Systems for Learning (1e),*Springer 2013.
3. F. Ricci, L. Rokach, D. Shapira and B.P. Kantor, *Recommender Systems Handbook (1e),*Springer ,2011.
4. K. Falk, Practical recommender systems. Shelter Island, NY: Manning Publications Company, 2019.(ISBN : 9781617292705)
5. D. Jannach, Recommender systems. New York: Cambridge University Press, 2011. (ISBN: 9780521493369)
6. D. Agarwal and B. Chung-Chen, Statistical methods for recommender systems. New York,NY: Cambridge University Press, 2016.(ISBN: 9781107036079)
7. R. Banik, Hands-On Recommendation Systems with Python. Birmingham, United Kingdom:PACKT Publishing Limited, 2018.(ISBN: 9781788993753)
8. S. Berkovsky, I.Cantador and D. Tikk, Collaborative Recommendations: Algorithms, Practical Challenges and Applications. UK: World Scientific Publishing Co, 2019.(ISBN:9789813275348)

Course Outcomes:

After successfully completing the course the student should be able to

1. Characterize different types of Recommender Systems, map a given real world problem to
2. appropriate model, understand and identify the stages and issues in the deployment of the system
3. Apply principles and techniques of recommender systems in applications related to recommender systems design and analysis
4. Analyze and evaluate various recommender algorithms
5. Implement appropriate recommender system for real world applications

Decision Support and Intelligent System					
Course code	PEC-AI-408G				
Category	Professional Elective Course				
Course title	Decision Support and Intelligent System				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Familiarize with Business Intelligence, Analytics and Decision Support
2. Understand the technologies for Decision making
3. Familiarize with predictive modeling techniques
4. Familiarize with sentiment analysis techniques
5. Understand about Multi-criteria Decision-making systems
6. Familiarize with Automated decision systems

Detailed Course

Unit-I

Information Systems Support forDecision Making, An Early Framework for ComputerizedDecision Support, The Concept of Decision SupportSystems, A Framework for Business Intelligence, Business Analytics Overview, Brief Introduction to Big Data Analytics, Clickstream Analysis, Metrics, Practical Solutions, Competitive Intelligence Analysis

Unit-II

Decision Making: Introduction and Definitions , Phases of the Decision Making Process, The Intelligence Phase, Design Phase, Choice Phase,Implementation Phase, Decision Support Systems: Capabilities, Classification, Components, modeling, Structure of mathematical models fordecision support; Decision making under certainty, Uncertainty and Risk, Decision modeling with spreadsheets, Mathematical programming optimization; Decision analysis-introduction , Decision tables, Decision Trees, Multi-criteria decision making, Pairwise comparisons

Unit-III

Basic Concepts of Neural Networks , Developing Neural NetworkBased Systems, Illuminating the Black Box of ANN withSensitivity, Support Vector Machines, A ProcessBased Approach to the Use of SVM, Nearest Neighbor Method for Prediction; Sentiment Analysis- Overview, Applications,Process, analytics, Speech Analytics.

Unit-IV

Automated Decision Systems, The Artificial Intelligence field ,Basic concepts of Expert Systems, Applications of Expert Systems, Structure of Expert Systems, Knowledge Engineering, Development of Expert Systems, Location based Analytics, Cloud Computing, Business Intelligence

References Books:

1. Ramesh Sharda, DursunDelen, EfraimTurban, J.E.Aronson,Ting-Peng Liang, David King, "Business Intelligence and Analytics: System for Decision Support", 10thEdition, Pearson GlobalEdition, 2013.

Course Outcomes:

After successfully completing the course the student should be able to

1. Gain knowledge on Business Intelligence, Analytics and Decision Support
2. Understand the technologies for Decision making
3. Apply predictive modeling techniques
4. Apply sentiment analysis techniques
5. Gain knowledge on Multi-criteria Decision-making systems
6. Gain knowledge on Automated decision systems

Computer Vision and Pattern Recognition					
Course code	PEC-AI-407G				
Category	Professional Elective Course				
Course title	Computer Vision and Pattern Recognition				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Recognize and describe both the theoretical and practical aspects of computing with images. Connect issues from Computer Vision to Human Vision
2. Describe the foundation of image formation and image analysis. Understand the basics of 2D and 3D Computer Vision.
3. Become familiar with the major technical approaches involved in computer vision. Describe various methods used for registration, alignment, and matching in images.
4. Get an exposure to advanced concepts leading to object and scene categorization from images.
5. Build computer vision applications.

Detailed Course

Unit-I

Introduction: Computer Vision-Definition and Overview, Image Formation,Geometric primitives, Lighting,Reflectance, Sampling and aliasing, Human Color Perception and Inference fromcolor, Image Transformation, 2D,3D Transformation, 3D to 2D Projection.

Image Processing: Point operation, Pixel transforms Color transforms Linear and Non-Linear Filtering,Transform Filtering Techniques, Interpolation and multi resolution.

Unit-II

Feature Detection and Matching: Feature detection, descriptor, matching, Feature tracking, Edge detection, Lines Detection, Edge linking, Successive approximationHough transform,Vanishing points, ObjectRecognition, Principal Component Analysis, SHIFT and HOG Feature.

Unit-III

Shape Detection andSegmentation: Active contours, Snakes, Dynamic snakes and CONDENSATION

Scissors, Level Sets, Region Split and merge, Graph cut and Energy based methods,2D and 3D feature-based alignment, Pose estimation, Medical Image Segmentation.

Unit-IV

Motion Estimation: Triangulation, Two-frame structure from motion, Projective reconstruction, Self-calibration, Perspective and projective factorization, Bundle adjustment, Exploiting sparsity, Constrained structure and motion, Hierarchical motion estimation, Fourier-based alignment, Incremental refinement, Parametric Motion, Spline based motion, Optical Flow,Kalman Filtering, application of motion estimation to video stabilization, Medical Image Registration.

References Books:

1. R. Szeliski, *Computer Vision: Algorithms and Applications*, (2e), Springer International Publishing, 2021.
2. D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, (2e), PHI learning 2012.
3. C. M. Bishop, *Pattern Recognition and Machine Learning*, (1e), Springer, 2011.
4. RichardSzeliski,"ComputerVision:AlgorithmsandApplications",Springer,2010.
5. Forsyth/Ponce,"ComputerVision:AModernApproach",PearsonEducationIndia;2ed(2015)
6. S.Nagabhushana,"ComputerVisionandImageProcessing",NewAgeInternationalPvtLtd; 1st ed(2005)
7. Rafael C. GonzaLez"'Digital Image Processing",Pearson Education; Fourth edition (2018)

Course Outcomes:

After successfully completing the course the student should be able to

1. Provide an introduction to computer vision including fundamentals of image formation
2. Provide a clear view of image formation
3. Provide a clear view of image processing Provide knowledge about Computational photography
4. Provide knowledge about Image rendering

Database Security and Privacy					
Course code	PEC-DS-411G				
Category	Professional Elective Course				
Course title	Database Security and Privacy				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Understand the fundamentals of security relates to information
2. How security is maintained in information systems
3. Understand the concept of security models in database
4. Implementation of virtual private database
5. Learn the procedures of database auditing
6. Implementation of data mining algorithms for PPDM

Detailed Course

Unit-I

Security Architecture: Introduction, Information Systems, Database Management Systems, Information Security Architecture- Database Security Asset Types and value, Security Methods, Operating System Security Fundamentals:Introduction, Operating System Overview, Security Environment, Security Components, Authentication Methods, User Administration, Password Policies, Vulnerabilities, Email Security, Internet security

Unit-II

Administration of Users-Introduction, Authentication, Creating Users, SQL Server User, Removing, Modifying Users Default users, Remote Users, Database Links, Linked Servers, Remote Servers, Practices for Administrators andManagers-Profiles, Password Policies, Privileges andRoles: Introduction , Defining and Using Profiles, Designing and Implementing Password, Policies, Best Practices Granting and Revoking User Privileges, Creating, Assigning and Revoking User Roles, Best practices

Unit-III

Database Application Security Models: Introduction-Types of Users -Security Models, Application Types-Application Security Models Data Encryption, Virtual Private Databases: Introduction-Overview of VPD, Implementation of VPD using Views, Application Context in Oracle, Implementing Oracle VPD-

Implementing Oracle VPD, Viewing VPD Policies, VPD using views, Application contexts using Data Dictionary Policy manager implementation

Unit-IV

Auditing Database Activities-introduction, Oracle Database Activities, Creating DLL Triggers with Oracle, Auditing Database Activities with Oracle, Auditing Server Activity with SQL Server2000, Auditing Server Activity with SQL Server, Auditing Server Activity with Oracle, Security and Auditing, Casestudy: projest security and auditing

Data Mining Techniques:Privacy Preserving Data Mining Algorithm, General Survey-Data Mining TechniquesRandomization Method, Privacy Preserving Data MiningTechniques: Introduction, Randomization Methods, Group Based Anonymization, Distributed Privacy Preserving Data Mining, Curse of Dimensionality, Application of Privacy Preserving DataMining, Casestudyon PPDM.

References Books:

1. HassanA.Afyouni,"DatabaseSecurityandAuditing",ThirdEdition,CengageLearning,2009.
2. RonBenNatan,"ImplementingDatabaseSecurityandAuditing",ElsevierDigitalPress,2005
3. Charu C. Aggarwal, Philip S Yu, "Privacy Preserving Data Mining": Models and Algorithms, Kluwer
   Academic Publishers, 2008

Course Outcomes:

After successfully completing the course the student should be able to

1. Acquire the knowledge of information system and information security
2. Able to manage the security of information system as well as database
3. Able to design and develop the security model in database
4. Able to implement VPD in various database
5. Able to audit the database activities, users, security
6. Apply the security mechanism in PPDM using various algorithms

Text Mining Essentials					
Course code	PEC-DS-412G				
Category	Professional Elective Course				
Course title	Text Mining				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Understand the fundamentals of text mining
2. Utilize text for prediction techniques
3. Understand the relevance between information retrieval and text mining
4. Understand the goals of information extraction
5. Analyze different case studies related to text mining

Detailed Course

Unit-I

Overview: Overview of text mining Special about Text Mining, Structured Data, Unstructured Data, Is text different from numbers, Types of Problem can be solved, Document Classification, Informational Retrieval,
Prediction and Evaluation, From Textual Information to NumericalVectors, Collecting Documents, Document Standardization, Tokenization Lemmatization, Inflectional Stemming, Stemming to a Root, Vector Generation for Prediction, Multiword Features
Text Extraction: Introduction, Rapid automatic keyword extraction: candidate keywords, keyword scores, adjoining keywords, extracted keywords, benchmark evaluation: precision and recall, efficiency, stop list generation, Evaluation on new articles.

Unit-II

Clustering: Multilingual document clustering: Multilingual LSA,Tucker1 method, PARAFAC2 method, LSA with term alignments, LMSA, LMSA with term alignments.
Classification: Content-based spam email classification using machine-learning algorithms, utilizingnonnegative matrix factorization for email classification problems, Constrained clustering with k-means type
algorithms;

Unit-III

Anomaly and trend detection: Text Visualization techniques such as tag clouds, authorshipand change tracking, Data Exploration, and the search for noval patterns, sentiment tracking, visualanalytics and Future Lens, scenario discovery, adaptive threshold setting for novelty mining.

Text streams

Text streams: Introduction, Text streams, Feature extraction and data reduction, Event detection, Trend detection, Event, and trend descriptions, embedding semantics in LDA

Unit-IV

Topic models: Introduction, vectorspace modelling, latent semantic analysis, probabilistic latent semantic analysis, Latent Dirichlet allocation,embedding external semantics from Wikipedia, data-driven semantic embedding.

Case study: Search Engine, E-mail filtering, Market Intelligence from the web, Generating Model cases for Help deskApplication, Extracting Named Entities from Documents, Mining Social Media, Customized Newspapers, Lightweight DocumentMatching for Digital Libraries, Assigning topics to news articles

References Books:

1. M. W. Berry and J. Kogan , *Text Mining Applications and Theory*, Wiley publications, 2010
2. C. C. Aggarwal, and C.X. Zhai, *Mining text data*. Springer Science & Business Media, 2012.
3. G. Miner, et al. *Practical text mining and statistical analysis for non-structured text dataapplications*. Academic Press, 2012.
4. N. Srivastava and M. Sahami, *Text mining: Classification, clustering, and applications,* Chapman and Hall/CRC, 2009.
5. Sholom M. Weiss, Nitin Indurkhya, Tong Zhang.,Fundamentals of Predictive Text Mining

Course Outcomes:

After successfully completing the course the student should be able to

1. Acquire knowledge on fundamentals of text mining
2. Perform prediction from text and evaluate it
3. Perform document matching
4. Identify patterns and entities from text
5. Understand how text mining is implemented

Natural Language Processing& Applications					
Course code	PEC-AI-416G				
Category	Professional Elective Course				
Course title	Natural Language Processing				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

COURSE OBJECTIVES:

1. Teach students the leading trends and systems in natural language processing.
2. Make them understand the concepts of morphology, syntax, semantics and pragmatics of the language
   and that they are able to give the appropriate examples that will illustrate the above mentioned concepts.
3. Teach them to recognize the significance of pragmatics for natural language understanding.
4. Enable students to be capable to describe the application based on natural language processing and to
5. show the points of syntactic, semantic and pragmatic processing.
   To conceive basics of knowledge representation, inference, and relations to the artificial intelligence.
6. To understand natural language processing and to learn how to apply basic algorithms in this field

Unit-I

Introduction to Natural Language Processing, Steps – Morphology – Syntax – Semantics, Morphological Analysis (Morphological Parsing), Stemming – Lemmatization , Parts of Speech Tagging , Approaches on NLP Tasks (Rule-based, Statistical, Machine Learning), N-grams, Multiword Expressions, Collocations (Association Measures, Coefficients and Context Measures), Vector Representation of Words, Language Modeling

Unit-II

Syntax Parsing, Dependency Parsing, Semantics, Semantic Parsing , Word Sense Disambiguation, Lexical Disambiguation, Structural Disambiguation, Word, Context andSentence-level Semantics, Pronoun Resolution, Semantic Representation of text, Introduction to Semantic , Relations, Semantic Relations, Semantic Role Labeling, Semantic Frames, Ontology and Semantics, Semantic Network andKnowledge Graph.

Unit-III

Intent Detection and Classification, Paraphrase Extraction, Discourse ,Coreference Resolution, Text Coherence, DiscourseStructure Coherence, Discourse Planning. Corpora and Lexicon: characteristics of Gold Standard Corpora like Treebank, Wordnet, Sentiwordnetetc

Information Extraction and itsapproaches, Information Retrieval, Semantic Search, Summarization, Extractive Vs Abstractive, Summarization, Information Fusion, Single and Multi-document, Summarization – Question Answering, Introduction to Chatbot Applications, Retrieval based-Conversation basedNLU and NLG, Machine Translation.

Unit-IV

Introduction to ProbabilisticApproaches, Statistical Approaches to NLP Tasks, Sequence Labeling, Problems - Similarity Measures, Word Embeddings , CBOW , Skip-gram, Sentence Embeddings, Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM).

Application: Semantic Text Similarity,Sentence or Document Classification, Machine translation, Text Summarization.

Suggested References Books:

1. Daniel Jurafsky and Prentice Hall James H Martin, Speech and Language Processing: An introduction on Natural Language Processing, Computational Linguistics and Speech Recognition‖, Prentice Hall, 2nd , Edition,2018.
2. C.Manning and H.Schutze, ―Foundations of Statistical Natural Language Processing‖, MIT, Press. Cambridge, MA:,1999
3. James Allen, Bejamin/cummings, NaturalLanguageUnderstanding,2nd edition,1995
4. Yoav Goldberg, Neural Network Methods for Natural Language Processing.
5. http://mccormickml.com/2106/04/19/word2vec-tutorial-the-skip-gram-model/
6. https://nlp.stanford.edu/pubs/glove.pdf

Course outcomes:

1. Understand approaches to syntax and semantics in NLP.
2. Understand approaches to discourse, generation, dialogue and summarization within NLP.
3. Understand current methods for statistical approaches to machine translation.
4. Understand machine learning techniques used in NLP, including the probabilistic context-free grammars andunsupervised methods, as applied within NLP
5. Understand the knowledge of various levels of analysis involved in NLP
6. Gain knowledge in automated Natural Language Generation and Machine Translation

Social Network Analysis					
Course code	PEC-AI-404G				
Category	Professional Elective Course				
Course title	Social Network Analysis				
Scheme and Credits	L	T	P	Credits	
	2	0	0	2	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

COURSE OBJECTIVES:

1. Understand the concepts of Social Web
2. Understand Network features Visualizing approach
3. Study and understand Link prediction
4. Learn and understand various analysis algorithms
5. Understand the concept of social influence and actions in marketing

Unit-I

Introduction to Social Web: Nodes, Edges and Network measures, Describing Nodes and Edges, Describing Networks, Layouts;

Unit-II

Visualizing Network features: The role of Tie Strength, Measuring TieStrength, Tie Strength and Network Structure, Tie Strength and Network Propagation, Link Prediction,Entity Resolution;
Link Prediction: Case Study Friend Recommendation, Introduction to CommunityDiscovery, Communities in Context, Quality Functions;

Unit-III

Algorithms: The Kernighan-Lin algorithm,Agglomerative Algorithms, Spectral Algorithms, Multi-level Graph Partitioning, Markov Clustering, OtherApproaches;

Unit-IV

Introduction to Social Influence: Influence Related Statistics, Social Similarity andInfluence, Homophile, Existential Test for Social Influence, Influence and Actions, Influence and Interaction,Influence Maximization in Viral Marketing

Suggested References Books:

1. S.P. Borgatti ,M.G. Everett , J.C. Johnson, *Analyzing Social Networks (2e*) SAGE Publications Ltd,2018
2. J. Goldbeck, *"Analyzing the Social Web",* Morgan Kaufmann Publications, 2013.
3. C. C. Aggarwal, "*Social Network Data Analytics*", Springer Publications, 2011.
4. J. Scott, *"Social Network Analysis", (3e),* SAGE Publications Limited, 2013.
5. S. Kumar, F. Morstatter and H. Liu, "*Twitter Data Analytics*", Springer Publications, 2013

Course outcomes:

1. Learn basic concepts in Social web
2. Work with Network features Visualizing approaches
3. Deal with Link predictions and recommendation systems
4. Analyze various types efficient network analysis algorithms
5. Learn social influence and related statistics in influence maximization

Open Electives

Open Source Programming					
Course code	OEC-DS-431G				
Category	OpenElective Course				
Course title	Open Source Programming				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. To comprehend and analyze the basic concepts of web frameworks
2. To describe how different frameworks work and to choose the framework depending on the application.
3. To demonstrates the uses of different web frameworks.

Detailed Course

Unit-I

Django Framework: Introduction and Installation – MVT Structure – Creating a project and app in Django – Django, Forms – creation of forms – render forms - form fields – form fields widgets – formsets – Django Templates– Template filters – Template Tags – Variables – Operators – for loop- If-Django Templates – Templateinheritance

Django Model: Django Views – Function based views – Class based generic views – Models – ORM – Basic App Model –Intermediate fields - Uploading Images – Render Model – Build-in and custom field validations – HandlingAjax Request – Django Admin interface

Unit-II

Ruby on Rails Framework : Ruby of Rails introduction – Installation – MVC architecture - IDE – Rails scripts - Directory structure- Database setup – Active records - RVM – Bundler - Rails Migration – controllers –routes– views – layouts - scaffolding – sessions – file upload – filters - Ajax

ExpressJS: Introduction – installation – Node JS Environment Setup – Routing – HTTP Methods – URLBuilding – Middleware – Templating – Different template Engines– Static Files – Form Data

Unit-III

ExpressJS& Database: Database– Mongo DB – Mongoose – Cookies,sessions – Authentication – RESTFUL APIs – Scaffolding – Error Handling – File upload

Angular JS: Introduction – Environment setup – First application – Data binding & Directives – Expressions – Controllers – Scopes – Events – Services – Filters - Modules

Unit-IV

Angular JS – Routing: HTML DOM -Forms – Validation – Routing – Includes – AJAX – Views – Dependency Injection- Custom Directives – Single Page applications

Recent Trends and contemporary issues

References Books:

1. AidasBendoraitis, Jake Kronika, Django 3 Web Development Cookbook: Actionable solutions to common, Packt Publishing; 4th edition, 2020.
2. Michael Hartl, Ruby on Rails Tutorial, Addison-Wesley Professional; 6th edition, 2020.
3. Adam Freeman, Pro Angular 9: Build Powerful and Dynamic Web Apps, Apress, 4th Edition,2020.
4. Ethan Brown, Web Development with Node and Express, 2e: Leveraging the JavaScript Stack, O′Reilly; 2nd edition, 2019.
5. Lopatin, Ben, Django Standalone Apps, Apress, 1st Edition, 2020.
6. Simon D. Holmes and Clive Harbe, Getting MEAN with Mongo, Express, Angular, and Node, Second Edition, Manning Publications,2017.

Course Outcomes:

After successfully completing the course the student should be able to

1. Use Django framework to create basic website.
2. Use Ruby on Rails framework to quickly develop websites.
3. Use Express framework along with Node JS to render webpages effectively
4. Use Mongo DB along with Express to display dynamic web content
5. Use Angular JS to extend an enhance HTML pages
6. Implementing web-based solution effectively using different web frameworks.

Data Visualization& Tableau					
Course code	OEC-AI-431G				
Category	OpenElective Course				
Course title	Data Visualization& Tableau				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Course objectives:

The basic objective is to understand the data analysis & visualize your data & method not just a tool-oriented Analyst.

Unit-I

Overview of Data analysis, Introduction to Data visualization,Working with statistical formulas - Logical and financial functions,Data Validation& data models, Power Map for visualize data, Power BI-Business Intelligence, Data Analysis using statistical methods, Dashboard designing.

Unit-II

Heat Map, Tree Map, Smart Chart, Azure Machine learning, Column Chart, Line Chart , Pie, Bar, Area ,Scatter Chart, Data Series ,Axes ,Chart Sheet ,Trendline, Error Bars, Sparklines, Combination Chart, Gauge, Thermometer Chart ,Gantt Chart ,Pareto Chartetc,Frequency, Distribution, PivotChart, Slicers, Tables: Structured References, Table Styles, What-If Analysis: Data Tables, GoalSeek, Quadratic Equation Transportation Problem, Maximum Flow Problem, Sensitivity Analysis, Histogram, Descriptive,Statistics,Anova ,F-Test,t-Test ,Moving,Average ,Exponential Smoothing, Correlationmodel, Regressionmodel; SQL- Overview, data types, operators, database query operations

Unit-III

What is Tableau? What does the Tableau product suite comprise of? How Does Tableau Work?Tableau Architecture, What is My Tableau Repository?, Connecting to Data & Introduction to data source concepts, Understanding the Tableau workspace, Dimensions and Measures, Data Types & Default Properties, Tour of Shelves & Marks Card, Using Show Me!, Building basic views, Saving and Sharing your work-overview, Demo related to above.

Unit-IV

Tableau Date Aggregations and Date parts,Cross tab & Tabular charts, Totals & Subtotals, Bar Charts & Stacked Bars, Trend lines, Reference Lines, Forecasting, Filters, Context filters, Line Graphs with Date & Without Date, Tree maps, Scatter Plots, Individual Axes, Blended Axes, Dual Axes & Combination chart, Edit axis, Parts of Views, Sorting.Demo related to above.

Books Recommended

- Big Data Analytics Beyond Hadoop: Real-Time Applications with Storm, Spark, and More Hadoop Alternatives, 1e Pearson Education India; 1 edition (2015)
- Big Data Fundamentals: Concepts Drivers: Concepts, Drivers and Techniques, Erl/Khattak/Buhler, Pearson Education India; First edition (2016)
- Ebook: Data Visualization Tools (Innovation Trends Series), BBVA Innovation Center, Kindle Edition

Course Outcomes:

On completion of this course, the students will be able to understand the data analysis & visualize your data & method not just a tool-oriented Analyst.

1. Understand data fundamental, analyse the data methodology, techniques, powerful dashboards, Power BI & Visualization power of data along with a strong focus on case studies to ensure hands on learning.
2. Learn the powerful data visualization tool like Advanced version of Excel, Power Map, Power BI, Business Intelligence software, Tableau software & other open source tools etc to present your analysis.

Information Storage and Management					
Course code	OEC-DS-432G				
Category	OpenElective Course				
Course title	Information Storage and Management				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Understand the components of storage infrastructure.
2. Gain knowledge to evaluate storage architectures including storagesubsystems
3. Understand the business continuity, backup and recovery methods.
4. Acquire knowledge on information security framework
5. Introduce the working principle of storage infrastructure with monitoring principles
6. Understand the structure of cloud computing and its techniques

Detailed Course

Unit-I

Introduction to Information Storage Management, Evolution of Storage Architecture, Data Centre Infrastructure, Virtualization and Cloud Computing, Key challenges in managing information, Data Center Environment: Application, Database Management System (DBMS), Host : Connectivity, Storage Disk Drive Components,Disk Drive Performance, Intelligent Storage System and its Components, Storage Provisioning, Types of Intelligent Storage Systems, Creation of Virtual storage machine , Navigation of storage system .

Unit-II

Virtualization and Cloud Computing : FiberChannel: Overview , SAN and its Evolution , Components of FC SAN, FCConnectivity,FC Architecture, IPSAN-iSCSI components, iSCSI Protocol StackiSCSI Names; NAS: General Purpose Servers versusNAS Devices, Benefits of NAS- File Systems andNetwork File Sharing, Components of NAS, I/O Operations, Implementations, File Sharing Protocols; Object Based Storage Devices, Content Addressed Storage, Configuration and Tracing of FC scan and iSCSI scan

Unit-III

Business Continuity And Back UpRecovery :Business Continuity: Information Availability, BC-Terminology, Planning life cycle, Failure Analysis, Business Impact Analysis, Technology Solutions; Backup and Archive: Backup Purpose , Backup Considerations, Backup Granularity,

Recoveryconsiderations, Backup Methods, Backup Architecture, Backup and Restore Operations, Backup Topologies, Backup in NAS Environments, Backup Targets, Data Deduplication forBackup , Backup in Virtualized Environments, Sharing Files between host and VirtualMachines, Usage of Backup techniques.

Unit-IV

Storage Security And Management : Information Security Framework , Risk Triad, Storage SecurityDomains, Security Implementations in StorageNetworking, Securing Storage Infrastructure inVirtualized and Cloud Environments, RSA and VMware Security Products, Monitoring the Storage Infrastructure, Monitoring Parameters,Components Monitored, Monitoringexamples, Storage Infrastructure Management Activities, Storage Infrastructure ManagementChallenges, Storage ManagementExamples, Storage Allocation to a New Server/Host;

Cloud Computing:Cloud Enabling Technologies, Characteristics, Benefits, Service Models, Deployment models;Cloud Infrastructure Mechanism: Logical Network Perimeter, Virtual Server, Storage Device, Usage Monitor, Resource Replication, Ready Made environment, Challenges, Container, Adoption Considerations, Usage of Cloud services with open sourcecloud tools (like Eucalyptus, Openstack,Open Nebula and others)

References Books:

1. EMC Corporation, "Information Storage and Management",2nd edition Wiley India, ISBN13:978-1118094839
2. Thomas Erl, "Cloud Computing: Concepts, Technology & Architecture", Prentice Hall, 2013, ISBN:9780133387568
3. UifTroppen Rainer Wolfgang Muller,"Storage Networks Explained", India, Wiley, 2010, ISBN13: 978-0470741436

Course Outcomes:

After successfully completing the course the student should be able to

1. Acquire the knowledge on the components of storage infrastructure
2. Acquire the ability to evaluate storage architectures including storagesubsystems
3. Understand the business continuity, backup and recovery methods.
4. Appreciate the concepts of storage security and information security applied to virtual machine
5. Apply the knowledge for storage infrastructure
6. Acquire the knowledge on structure of cloud computing and its techniques

Virtualization					
Course code	OEC-DS-433G				
Category	OpenElective Course				
Course title	Virtualization				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. To identify and select suitable hypervisor for a cloud environment.
2. To acquire the knowledge of various virtualization techniques and tools.
3. To understand the process of data center automation and secure virtualized environment.

Unit-I

Introduction : Characteristics Virtualization definition – virtual machine basics – benefits – need for virtualization – limitations −traditional vs. contemporary virtualization process – virtual machines – taxonomy – challenges.
Hypervisors : Introduction to Hypervisors – Type 1 Hypervisors – Type 2 Hypervisors – comparing hypervisors – virtualization considerations for cloud providers.

Unit-II

Hardware Virtualization: Full virtualization - para virtualization - server virtualization - OS level virtualization - emulation – binary translation techniques – managing storage for virtual machines

Types of Virtualization: Application virtualization - desktop virtualization - network virtualization - storage virtualization - comparing virtualization approaches.

Unit-III

Virtualization Management: Management life cycle - managing heterogeneous virtualization environment – customized and modifying virtual machines – virtual machine monitoring – management tools.
Automation: Benefits of data center automation – virtualization for autonomic service provisioning – software defined data center - backup - disaster recovery.

Unit-IV

Security: Mapping Design (Models) to Code – Testing - Usability – Deployment – Configuration Management – Maintenance; Current trends and Contemporary issues

References Books:

1. Nelson Ruest, Danielle Ruest, Virtualization, A beginners guide, 2009, MGH.
2. Nadeau,TimCerng, Je Buller, Chuck Enstall, Richard Ruiz, Mastering MicrosoftVirtualization, Wiley Publication, 2010.
3. William Von Hagen, Professional Xen Virtualization, Wiley Publication, 2008.
4. Matthew Portney, Virtualization Essentials, John Wiley & Sons, 2012.
5. Dave Shackleford, Virtualization security, protecting virtualized environment, John Wiley, 2012.

Course Outcomes:

After successfully completing the course the student should be able to

1. Illustrate the process of virtualization.
2. Create and configure the hypervisors in cloud.
3. Apply the virtualization concepts in server and manage the storage capacity.
4. Analyze, identify and select suitable type of virtualization.
5. Use the management tools for managing the virtualized cloud infrastructure.
6. Apply suitable automation and security methods on data centre

Cyber Security and AI					
Course code	OEC-AI-432G				
Category	Open Elective Course				
Course title	Cyber Security and AI				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. To learn the need of AI for Cyber Security
2. To learn the detection of DDOS using AI techniques
3. To learn the intrusion detection using Neural Networks
4. To learn the various applications of AI to detect cyber attacks

Unit-I

Fundamentals of Cyber Security: Identity, authentication, confidentiality, privacy, anonymity, availability,and integrity, exploring cryptographic algorithms together with major attacks (using a break-understandand-fixapproach), Exploring high-level security protocols;
Fundamentals of AI for Security: deep learning fundamentals from a security perspective., case studies;

Unit-II

Web Application Security: Injection, Broken authentication, Sensitive data exposure, XML External Entities (XXE), Broken access control, Securitymisconfiguration, Cross-Site Scripting (XSS), Insecure deserialization, Using components with knownvulnerabilities, Insufficient logging, and monitoring.

Unit-III

Secure Web: making websites secure using AItechniques for injection using regular expressions and identifying patterns and matching with existingscores. Case studies;
Deep learning applications: Pattern detection and model behavior for anomalousbehavior, Advanced Malware Detection Case studies;

Unit-IV

Secure AI Development:foundations of securesoftware design, secure programming, and security testing. The section requires a basic understanding ofApplication Programming Interface (API) and example APIs of companies referred to are: Darktrace,Vectra and Cylance;
Impact of AI on Cyber Security: Threat hunting in memory, file system and networkdata, analysis of malicious programs; Contemporary issues in Artificial Intelligence for Cyber security.

Reference Books:

1. A. Parisi, Hands-On Artificial Intelligence for Cybersecurity: Implement smart AI systems for preventingcyber-attacks and detecting threats and network anomalies, (1e) Packt Publishing, 2019
2. S. Halder, Hands-On Machine Learning for Cybersecurity: Safeguard your system by making your machines intelligent using the Python ecosystem, (1e), Packt Publishing, 2018
3. L.F. Sikos (Ed.), AI in Cybersecurity, Springer International Publishing, 2019
4. E. Tsukerman, Machine Learning for Cybersecurity Cookbook, Packt Publishing, 2019

Course Outcomes:

After successfully completing the course the student should be able to

1. Understand the cyber threats, attacks and vulnerabilities and its defensive mechanism
2. Understand and implement various AI techniques to detect cyber attacks
3. The recent challenges in AI related to cyber security and able to develop new security solutions to the real time applications

R Programming					
Course code	OEC-AI-434G				
Category	OpenElective Course				
Course title	R Programming				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course: The objective of this course are:

1. Understand what R is and what it can be used for
2. Why would you choose R over another tool
3. Troubleshoot software installs (keep your fingers crossed)
4. Gain familiarity with using R from within the RStudio IDE
5. Get to know the basic syntax of R functions
6. Be able to install and load a package into your R library

Details Syllabus

Unit-I

INTRODUCTION: Getting **R,** R Version, 32-bit versus 64-bit, The **R** Environment, Command Line Interface, RStudio, Revolution Analytics RPE. **R** Packages: Installing Packages, Loading Packages, Building a Package. R Basics: Basic Math, Variables, Data Types, Vectors, Calling Functions, Function, Documentation, Missing Data, Advanced Data Structures: data frames, Lists, Matrices, Arrays.

Unit-II

R DATA : Reading Data into **R:** Reading CSVs, Excel Data, Reading from Databases, Data from OtherStatistical Tools, R Binary Files, Data Included with R, Extract Data from Web Sites, Statistical Graphics: Base Graphics, ggplot2.

R FUNCTIONS & STATEMENTS: Writing **R** Functions: Hello, World!, Function Arguments, Return Values, do.call, Control Statements: if and else, switch, ifelse, Compound Tests, Loops: for Loops, while Loops, Controlling Loops.

Unit-III

DATA MANIPULATION: Group Manipulation: Apply Family, aggregate, plyr, data.table, Data Reshaping: cbind and rbind, Joins, reshape2, Manipulating Strings: paste, sprint, Extracting Text, Regular.

R STATISTICS & LINEAR MODELING: Probability Distributions: Normal Distribution, Binomial Distribution, Poisson, Basic Statistics: Summary Statistics, Correlation and Covariance, T-Tests 200, ANOVA, Linear Models: Simple Linear Regression, Multiple Regression, Generalized Linear Models:

Logistic Regression, Poisson, Model Diagnostics: Residuals, Comparing Models, Cross-Validation, Bootstrap, StepwiseVariable Selection

Unit-IV

NON-LINEAR MODELING: Nonlinear Models: Nonlinear Least Squares, Splines, Generalized Additive Models, DecisionTrees, Random Forests, Clustering: K-means, PAM, Hierarchical Clustering

Course Outcomes:

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

1. Familiarize themselves with R and the RStudio IDE
2. Understand and use R functions
3. Install and load a package into your R library
4. Get insight into the capabilities of the language as a productivity tool for data manipulation and statistical analyses.

REFERENCES:

- Jared P. Lander, R for Everyone: Advanced Analytics and Graphics, Pearson Edu. Inc.
- Christian Heumann, Michael Schomaker and Shalabh, Introduction to Statistics and Data Analysis - With Exercises, Solutions and Applications in R , Springer, 2016
- Pierre Lafaye de Micheaux, Rémy Drouilhet, Benoit Liquet, The R Software Fundamentals of Programming and Statistical Analysis, Springer 2013
- Alain F. Zuur, Elena N. Ieno, Erik H.W.G. Meesters, A Beginner's Guide to R (Use R) Springer 2009

Essentials of Hadoop					
Course code	OEC-AI-436G				
Category	OpenElective Course				
Course title	Essentials of Hadoop				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

Provide the skills needed for building computer system for various applications in a career in Computer Science field.

1) Explain the characteristics of Big Data
2) Describe the basics of Hadoop and HDFS architecture
3) List the features and processes of MapReduce
4) Describe the basics of Pig

Unit-I

What is Big Data and where it is produced? Rise of Big Data, Compare Hadoop vs traditional systems, Limitations and Solutions of existing Data Analytics Architecture, Attributes of Big Data, Types of data, other technologies vs Big Data.

Hadoop Architecture and HDFS - What is Hadoop? Hadoop History, Distributing Processing System, Core Components of Hadoop, HDFS Architecture, Hadoop Master – Slave Architecture, Daemon types - Learn Name node, Data node, Secondary Name node.

Unit-II

Hadoop Clusters and the Hadoop Ecosystem- What is Hadoop Cluster? Pseudo Distributed mode, Type of clusters, Hadoop Ecosystem, Pig, Hive, Oozie, Flume, SQOOP. Hadoop MapReduce Framework - Overview of MapReduce Framework, MapReduce Architecture, Learn about Job tracker and Task tracker, Use cases of MapReduce, Anatomy of MapReduce Program.

Unit-III

MapReduce programs in Java- Basic MapReduce API Concepts, Writing MapReduce Driver, Mappers, and Reducers in Java, Speeding up Hadoop Development by Using Eclipse, Unit Testing MapReduce Programs, and Demo on word count example.

Hive and HiveQL- What is Hive?, Hive vs MapReduce, Hive DDL – Create/Show/Drop Tables, Internal and External Tables, Hive DML – Load Files & Insert Data, Hive Architecture & Components, Difference between Hive and RDBMS, Partitions in Hive.

Unit-IV

PIG vs MapReduce, PIG Architecture & Data types, Shell and Utility components, PIG Latin Relational Operators, PIG Latin: File Loaders and UDF, Programming structure in UDF, PIG Jars Import, limitations of PIG. Apache SQOOP, Flume

- Why and what is SQOOP? SQOOP Architecture, Benefits of SQOOP, Importing Data Using SQOOP, Apache Flume Introduction, Flume Model and Goals, Features of Flume, Flume Use Case.

HBase- What is HBase? HBase Architecture, HBase Components, Storage Model of HBase, HBase vs RDBMS, Introduction to Mongo DB, CRUD, Advantages of MongoDB over RDBMS, Use case.

Oozie and Zookeeper

Topics - Oozie – Simple/Complex Flow, Oozie Workflow, Oozie Components, Demo on Oozie Workflow in XML, What is Zookeeper? Features of Zookeeper, Zookeeper Data Model

Reference Books:

1. Hadoop – The Definitive Guide by Tom White, 4th EditionORelly, 2015
2. Expert Hadoop Administration: Managing, Tuning, and Securing Spark, YARN, and HDFS by Alapati Sam R.,2017
3. Big Data and Hadoop- Learn by Exampleby Mayank Bhushan,BPB Pub,2018
4. Big Data and Hadoop by V. K. Jain, Khana Pub., 2017

Course Outcomes:

1. Understanding of Big Data problems with easy to understand examples.
2. History and advent of Hadoop right from when Hadoop wasn't even named Hadoop.
3. What is Hadoop Magic which makes it so unique and powerful.
4. Understanding the difference between Data science and data engineering, which is one of the big confusions in selecting a carrier or understanding a job role.
5. And most importantly, demystifying Hadoop vendors like Cloudera, MapR and Hortonworks by understanding about them.

Intelligent Vehicle Technology					
Course code	OEC-ME-451G				
Category	OpenElective Course				
Course title	Intelligent Vehicle Technology				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Acquire knowledge of about Intelligent vision system
2. Know the architecture of Intelligent transportation system
3. Impart the techniques of adaptive control
4. Know the architecture for autonomous vehicles
5. Study the autonomous vehicle cases

Detailed Course

Unit-I

Introduction to Intelligent Vision System: Vision Based Driver Assistance System –Vehicle optical Sensor , Laser Radar, Non Contact ground velocity detectingSensor, Road Surface Recognition Sensor, Vehicle Sensors for Electronic TollCollection System, Components of a Vision Sensor System; Driver Assistance on Highways –LaneRecognition, TrafficSign Recognition; Driver Assistance in Urban Traffic-StereoVision , Shapebase analysis, Pedestrian Recognition

Unit-II

Vehicle Information System and Intelligent Transportation: Intelligent Transportation System (ITS) – Vision for ITS Communications, Multimedia communication in a car, Current ITS Communication Systems andServices, Vehicle to Vehicle CommunicationSystems, Road to Vehicle CommunicationSystems, Inter Vehicle Communication, Intra Vehicle Communication, VANETS-Devices, Optical Technologies, Millimeter Wave technologies

Unit-III

Adaptive Control Techniques for Intelligent Vehicles: Automatic Control Of Highway Traffic AndMoving Vehicles, Adaptive Control Of Highway Traffic AndMoving Vehicles, Adaptive Control Overview, Gain Scheduling, Model Reference Adaptive Control, Self-Tuning Adaptive Control System Model, System Identification Basics,Recursive Parameter Estimation, Estimator Initialization, Design Of Self-Tuning Controllers, Generalized Minimum Variance (GMV)Control, Pole Placement Control

Model Predictive Control Overview andExamples.

Decisional Architectures for Autonomous Vehicles: Control Architectures, Motion Autonomy, Deliberative Architectures, Reactive Architectures, Hybrid Architecture Overview and Examples,

Unit-IV

Decisional Architectures for Autonomous Vehicles: Overview Of Sharp Architecture,Models Of Vehicles, Concepts Of Sensor Based Maneuver, Reactive Trajectory Following,Parallel Parking , Platooning, Main Approaches To Trajectory Planning,Non-Homonymic Path Planning.

Autonomous Vehicle and Case Studies:DARPA Challenge Case Study, ARGO Prototype Vehicle, The Gold System, The inverse Perspective Mapping, Lane Detection, Obstacle Detection, Vehicle Detection, Software systems architecture, Computational Performances, ARGO Prototype vehicle Hardware, Functionalities- ARGO Prototype vehicle, Data acquisition System, Processing System, Control System, Overview Pedestrian Detection

References Books:

1. LjuboVlacic, Michel Parent and Fumio Harashima, "Intelligent Vehicle Technologies", Butterworth -Heinemann publications, Oxford, 2001-ISBN 0 7506 5093 1
2. Ronald K Jurgen, "Automotive Electronics Handbook ", Automotive Electronics Series, SAE, USA, 1998.
3. NicuBizon,Lucian D Ascalescu And NaserMahdavitAbatabaei "Autonomous Vehicles

Course Outcomes:

After successfully completing the course the student should be able to

1. Understand the intelligent vision system used in automobiles
2. Understand the architecture of intelligent transportation system
3. Understand adaptive control techniques of an autonomous vehicle
4. Understand about the successful autonomous vehicle projects
5. Know the case studies of Autonomous vehicle

Hybrid and Electrical Vehicle					
Course code	OEC-ME-452G				
Category	Open Elective Course				
Course title	Hybrid and Electrical Vehicle				
Scheme and Credits	L	T	P	Credits	
	3	0	0	3	
Class work	25 Marks				
Exam	75 Marks				
Total	100 Marks				
Duration of Exam	03 Hours				

Note: Examiner will set nine questions in total. Question one will be compulsory. Question one will have 6 parts of 2.5 marks each from all units and remaining eight questions of 15 marks each to be set by taking two questions from each unit. The students have to attempt five questions in total, first being compulsory and selecting one from each unit.

Objectives of the course:

The objective of this course are:

1. Provide an insight into how electric vehicle operate
2. Demonstrate the functional requirements of Battery management system in detail.
3. Demonstrate how Electric and Hybrid Vehicle vary as per design requirements.
4. Perform the detailed analysis on the drives and driveline.
5. Selection of the appropriate drive and driveline system for the different cases.

Unit-I

Electric Vehicle Propulsion and Energy Sources: Basic concepts and problems concerning the electrification in Mobility, Functional components in an electric andhybrid vehicle, Vehicle Mechanics – Kinetics, Vehicle Mechanics – Dynamics &Roadway Fundamentals, Propulsion System Design - Force VelocityCharacteristics,Calculation Of Tractive Power And EnergyRequired, Electric Vehicle Power Source – BatteryCapacity, Battery Construction and Types State of Charge and Discharge , Calculation of Specific Energy and SpecificPower &Ragone Plot RelationshipBattery Modeling - Run Time BatteryModel, First Principle Model, Battery Management System- SOC Measurement, Battery Cell Balancing.Traction Batteries - Nickel Metal Hydride Battery, Li-Ion, Li-Polymer Battery.

Unit-II

Electric Vehicle Powerplant And Drives : Basic concepts of electric vehicle powerplant, Power and Torque plot, Construction of Induction Machines,Operating cycle and application in traction, Construction of Permanent MagnetMachines, Construction of Switch ReluctanceMachines, Role of Power Electronic ConvertersDC/DCConverters, Description of Buck Boost Converter Isolated DC/DC Converter, Functional Requirements and Operatinglimits, Two Quadrant Chopper Switching Modes, AC Drives- PWM , Current Control Method, Role of Switch Reluctance Machine Drives, Voltage Control, Current Control

Unit-III

Hybrid and Electric Drivetrains : Functional requirements of Hybrid Vehicle, Operational difference between the FullyElectric, Hybrid and Mild Hybrid, Topological Phenomena and SocialImportance of e-mobility, Role of modern drivetrain and theconversion efficiency and powerconsumption, Description of Hybrid Traction, Description of Electric Traction. Topological Optimization for HybridTraction Topological Optimization for ElectricTraction, Power Flow Control & Energy Efficiency Analysis,

Configuration and Control of DC MotorDrives, Induction Motor Drive. Permanent Magnet Motor Drives, SwitchReluctance Motor, Drives, Drive System Efficiency.

Unit-IV

Electric and Hybrid Vehicle Design: Design perspectives of Hybrid vehicle, Power plant energy distribution, Matching the Electric Machine and theInternal Combustion Engine, Parameter optimization – IC Engine, Position and Types of arrangements, Parameter optimization – Motor Position and Types of arrangements, Sizing of Propulsion Motor , Power Electronics & Drive System
Selection of Energy Storage Technology Topological Optimization, Communications & Supporting Subsystem, Energy Management Strategies in Hybrid, Vehicles- Classification,Comparison, Implementation

Electric And Hybrid Vehicles Case Studies: Parallel Hybrid, Series Hybrid -ChargeSustaining, Parallel Hybrid, Series Hybrid –ChargeDepleting, Hybrid Vehicle Case Study –Toyota Prius, Honda Insight, Chevrolet Volt; 42 V System for Traction Applications, Lightly Hybridized Vehicles and Low
Voltage System; Electric Vehicle Case Study - GM EV1,Nissan Leaf, Mitsubishi Miev;Hybrid Electric Heavy-Duty Vehicles, FuelCell Heavy DutyVehicles

Reference Books:

1. Iqbal Husain, "Eclectic and Hybrid vehicles Design Fundamentals" , CRC Press, second edition 2013, ISBN 9781439811757
2. James Larminie, John Lowry, "Electric vehicle technology Explained" second Edition, Wiley 2012, ISBN-13: 9781119942733
3. Ali Emadi, "Hand book of Automotive Power Elect ronics and Motor Drives", CRC Press 2005, ISBN 9780824723613.
4. Ali Emadi, Mehrdad Ehsani, John M. Muller, "Vehicular Electric Power Systems" Marcel Dekker, Inc., 2004

Course Outcomes:

After successfully completing the course the student should be able to

1. Learn the basic concepts of electric vehicle technology and electric vehicles.
2. Develop and analyze hybrid and electric drive trains.
3. Interpret various vehicle power sources in hybrid vehicle technology
4. Analyze data to determine appropriate design calculations of hybrid system under study.
5. Apply the concepts in sizing the electric motors