VISVESVARAYA TECHNOLOGICAL UNIVERSITY BELAGAVI

Scheme of Teaching and Examinations and Syllabus

M.Tech in Data Science (SDS)

(Effective from Academic year 2021-22)

VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI
Scheme of Teaching and Examinations (with effect from 2021-22)

M.Tech in Data Science (SDS)

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

I SEMESTER

Sl. No.	Course	CourseCode	Course Title	Teaching Hours / Week			Examination				Credits
				Theory	Practicals / Seminar	Skill Development Activity	urations in Hours	CIE Marks	SEE Marks	Total Marks	
1	PCC	20SDS11	Mathematical Foundations of Computer Science	03	--	02	03	40	60	100	4
2	PCC	20SDS12	Data Science	03	--	02	03	40	60	100	4
3	PCC	20SDS13	Advanced Database Management System	03	--	02	03	40	60	100	4
4	PCC	20SDS14	Predictive Analytics	03	--	02	03	40	60	100	4
5	PCC	20SDS15	Data acquisition and Productization	03	--	02	03	40	60	100	4
6	PCC	20SDSL16	Data Science & Analytics lab	--	04	--	03	40	60	100	2
7	PCC	20RMI17	Research Methodology and IPR	01	--	02	03	40	60	100	2
TOTAL				**16**	**04**	**12**	**21**	**280**	**420**	**700**	**24**

Note: PCC: Profession Core, PEC

Note: PCC: Profession Core

Skill development activities:

Students and course instructor/s to involve either individually or in groups to interact together to enhance the learning and application skills. The students should interact with industry (small, medium and large), understand their problems or foresee what can be undertaken for study in the form of research/ testing / projects, and for creative and innovative methods to solve the identified problem. The students shall Gain confidence in Modeling of systems and algorithms.

Work on different software/s (tools) to Simulate, analyze and authenticate the output to interpret and conclude. Operate the simulated system under changed parameter conditions to study the system with respect to thermal study, transient and steady state operations, etc. Handle advanced instruments to enhance technical talent. Involve in case studies and field visits/ field work. Accustom with the use of standards/codes etc., to narrow the gap between academia and industry.

All activities should enhance student's abilities to employment and/or self-employment opportunities, managementskills, Statistical analysis, fiscal expertise, etc.

Internship: All the students have to undergo mandatory internship of 6 weeks during the vacation of I and II semesters and /or II and III semesters. A University examination shall be conducted during III semester

and the prescribed credit shall be counted for the same semester. Internship shall be considered as a head of passing and shallbe considered for the award of degree. Those, who do not take-up/complete the internship shall be declared as fail in internship course and have to complete the same during the subsequent University examination after satisfying the internship requirements.
Note: (i) Four credit courses are designed for 50 hours Teaching – Learning process. (ii) Three credit courses are designed for 40 hours Teaching – Learning process.

VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI
Scheme of Teaching and Examinations (with effect from 2021-22)

M.Tech., in Data Science (SDS)

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

II SEMESTER

SL. No.	Course	Course Code	Course Title	Teaching Hours / Week			Examination				Credits
				Theory	Practicals / Seminar	Skill Development	Durations in Hours	CIE Marks	SEE Marks	Total Marks	
1	PCC	20SDS21	Advanced Machine Learning	03	--	02	03	40	60	100	4
2	PCC	20SDS22	Information Retrieval	03	--	02	03	40	60	100	4
3	PCC	20SDS23	Data Visualization	03	--	02	03	40	60	100	4
4	PEC	20SDS24X	Professional elective 1	04	--	--	03	40	60	100	4
5	PEC	20SDS25X	Professional elective 2	04	--	--	03	40	60	100	4
6	PCC	20SDSL26	Visualization & Data Science and Mini Project Lab	--	04	--	03	40	60	100	2
7	PCC	20SDS27	Technical Seminar	--	02	--	--	100	--	100	2
			TOTAL	**17**	**06**	**06**	**18**	**340**	**360**	**700**	**24**

Note: PCC: Profession Core, PEC: Professional Elective Course

Professional Elective-1		**Professional Elective-2**	
Course Code 20SDS24X	**Course Title**	**Course Code 20SDS25X**	**Course Title**
20SDS241	Managing Big Data	20SDS251	Data Security and Privacy
20SDS242	Natural Language Processing	20SDS252	Time Series Analysis and Forecasting
20SDS243	Soft and Evolutionary Computing	20SDS253	Applied Social Network Analysis
20SDS244	Pattern Recognition	20SDS254	Image and Video Analytics

Note:

1. **Technical Seminar:** CIE marks shall be awarded by a committee comprising of HoD as Chairman, Guide/co- guide, if any, and a senior faculty of the department. Participation in the seminar by all

postgraduate students of the program shall be mandatory. The CIE marks awarded for Technical Seminar, shall be based on the evaluation of Seminar Report, Presentation skill and performance in Question-and-Answer session in the ratio 50:25:25.

2. **Internship:** All the students shall have to undergo mandatory internship of 6 weeks during the vacation of I and II semesters and /or II and III semesters. A University examination shall be conducted during III semester and the prescribed internship credit shall be counted in the same semester. Internship shall be considered as a head of passing and shall be considered for the award of degree. Those, who do not take-up/complete the internship shall be declared as fail in internship course and have to complete the same during the subsequent University examination after satisfying the internship requirements.

VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI
Scheme of Teaching and Examinations (with effect from 2021-22)

M.Tech in Data Science (SDS)

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

III SEMESTER

SL. No.	Course	Course Code	Course Title	Teaching Hours / Week			Examination				Credits
				Theory	Practical / Seminar	Skill Development Activity	Duration in Hours	CIE Marks	SEE Marks	Total Marks	
1	PCC	20SDS31	Deep Learning	03	--	02	03	40	60	100	4
2	PEC	20SDS32X	Professional elective 3	04	--	--	03	40	60	100	3
3	PEC	20SDS33X	Professional elective 4	04	--	--	03	40	60	100	3
4	Project	20SDS34	Project work phase - 1	--	02	--	--	100	--	100	2
5	PCC	20SDS35	Mini-Project	--	02	--	--	100	--	100	2
6	Internship	20SDSI36	Internship	(Completed during the intervening vacation of I and II semesters and / or II and III semesters)			03	40	60	100	6
			TOTAL	**11**	**04**	**02**	**12**	**360**	**240**	**600**	**20**

Note: PCC: Profession Core, PEC: Professional Elective Course

Professional Elective-3		**Professional Elective-4**	
Course Code 20SDS32X	**Course Title**	**Course Code 20SDS33X**	**Course Title**
20SDS321	Business Analytics	20SDS331	Blockchain Technology
20SDS322	Cyber security and Cyber Law	20SDS332	Supply Chain Management
20SDS323	Cloud Computing for Data Analysis	20SDS333	Spatial Data Analysis
20SDS324	Data Stream Mining	20SDS334	Health Care Data Analytics

Note:

1. **Project Work Phase-1:** Students in consultation with the guide/co-guide if any, shall pursue literature survey and complete the preliminary requirements of selected Project work. Each student shall prepare relevant introductory project document and present a seminar.CIE marks shall be awarded by a committee comprising of HoD as Chairman, Guide/co-guide if any, and a senior faculty of the department. The CIE marks awarded for project work phase -1, shall be based on the evaluation of Project Report, Project Presentation skill and Question and Answer session in the ratio 50:25:25. SEE (University examination) shall be as per the University norms.
2. **Internship:** Those, who have not pursued /completed the internship shall be declared as fail in internship course and have to complete the same during subsequent University examinations after satisfying the internship requirements. Internship SEE (University examination) shall be as per the University norms.

VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI

Scheme of Teaching and Examinations (with effect from 2021-22)

M.Tech in Data Science (SDS)

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

IV SEMESTER

SL. No.	Course	CourseCode	Course Title	Teaching Hours / Week			Examination				Credits
				Theory	Practical / Seminar	Skill Development Activity	Duration in Hours	CIE Marks	SEE Marks	Total Marks	
1	Project	20SDS41	Project work phase 2	--	04	03	03	40	60	100	20
TOTAL				**--**	**04**	**03**	**03**	**40**	**60**	**100**	**20**

Note:

Project Work Phase-2:

CIE marks shall be awarded by a committee comprising of HoD as Chairman, Guide/co-guide, if any, and a Senior faculty of the department. The CIE marks awarded for project work phase -2, shall be based on the evaluation of Project Report subjected to plagiarism check, Project Presentation skill and Question and Answer session in the ratio 50:25:25.

SEE shall be at the end of IV semester. Project work evaluation and Viva-Voce examination (SEE), after satisfying the plagiarism check, shall be as per the University norms.

M.TECH IN DATA SCIENCE (SDS) Choice Based Credit System (CBCS) and Outcome Based Education (OBE) SEMESTER – I			
MATHEMATICAL FOUNDATION OF COMPUTER SCIENCE			
Course Code	**20SDS11**, 20LNI11, 20SCS11, 20SCE11, 20SFC11, 20SCN11, 20SSE11, 20SIT11, 20SAM11, 20SIS11	CIE Marks	40
Teaching Hours/Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03
Module-1			
Vector Spaces: Vector spaces; subspaces Linearly independent and dependent vectors Basis and dimension; coordinate vectors-Illustrative examples. Linear transformations, Representation of transformations by matrices; (RBT Levels: **L1 & L2)** (Textbook:1)			
Module-2			
Orthogonality and least squares: Inner product, orthogonal sets, orthogonal projections, orthogonal bases. Gram-Schmidt orthogonalization process. QR factorizations of a matrices, least square problems, applications to linear models (least square lines and least square fitting of other curves). (RBT Levels: **L2 & L3)** (Textbook:1)			
Module-3			
Symmetric and Quadratic Forms: Diagonalization, Quadratic forms, Constrained Optimization, The Singular value decomposition. Applications to image processing and statistics, Principal Component Analysis (RBT Levels: **L2 & L3)** (Textbook:1)			
Module-4			
Statistical Inference: Introduction to multivariate statistical models: Correlation and Regression analysis, Curve fitting (Linear and Non-linear) (RBT Levels: **L2 & L3)** (Textbook:3)			
Module-5			
Probability Theory: Random variable (discrete and continuous), Probability mass function (pmf), Probability density function (pdf), Mathematical expectation, Sampling theory: testing of hypothesis by ***t***-test, 3^2- test. (RBT Levels: **L1 & L2)** (Textbook:3)			

Course Outcomes:

On completion of this course, students are able to:

1. Understand the numerical methods to solve and find the roots of the equations.
2. Apply the technique of singular value decomposition for data compression, least square approximation in solving inconsistent linear systems
3. Understand vector spaces and related topics arising in magnification and rotation of images.
4. Utilize the statistical tools in multi variable distributions.
5. Use probability formulations for new predictions with discrete and continuous RV's.

Question Paper Pattern:

- The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.
- The question paper will have ten full questions carrying equal marks.
- Each full question consisting of 20 marks
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbooks:

Sl. No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Linear Algebra and its Applications	David C. Lay, Steven R. Lay and J. J. McDonald	Pearson Education Ltd	5th Edition 2015.
2	Numerical methods for Scientific and Engineering Computation	M K Jain, S.R.K Iyengar, R K. Jain	New Age International	6th Edition 2014
3	Probability, Statistics and Random Process	T. Veerarajan	Tata Mc-Graw Hill Co	3rd Edition 2016
Reference Books				
1	Optimization: Theory & Applications Techniques	Rao. S.S	Wiley Eastern Ltd New Delhi.	
2	Signals, Systems, and Inference	Alan V. Oppenheim and George C.Verghese	Spring	2010
3	Foundation Mathematics for Computer Science	John Vince	Springer International	

M.TECH IN DATA SCIENCE (SDS)

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

SEMESTER – I

DATA SCIENCE

Course Code	**20SDS12,** 20SCS21, 20SAM14, 20SIS22	CIE Marks	40
Teaching Hours/Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

Introduction: What is Data Science? Big Data and Data Science hype – and getting past the hype, Why now? – Datafication, Current landscape of perspectives, Skill sets. Needed Statistical Inference: Populations and samples, Statistical modelling, probability distributions, fitting a model, - Introduction to R.

Text Book 1: Chapter 1

Module-2

Exploratory Data Analysis and the Data Science Process: Basic tools (plots, graphs and summary statistics) of EDA, Philosophy of EDA, The Data Science Process, Case Study: Real Direct (online realestate firm). Three Basic Machine Learning Algorithms: Linear Regression, k-Nearest Neighbours (k- NN), k-means.

Text Book 1: Chapter 2

Module-3

One More Machine Learning Algorithm and Usage in Applications: Motivating application: Filtering Spam, Why Linear Regression and k-NN are poor choices for Filtering Spam, Naive Bayes and why it works for Filtering Spam, Data Wrangling: APIs and other tools for scrapping the Web.

Text Book 1: Chapter 7

Module-4

Feature Generation and Feature Selection (Extracting Meaning from Data): Motivating application: user (customer) retention. Feature Generation (brainstorming, role of domain expertise, and place for imagination), Feature Selection algorithms. Filters; Wrappers; Decision Trees; Random Forests. Recommendation Systems: Building a User-Facing Data Product, Algorithmic ingredients of a Recommendation Engine, Dimensionality Reduction, Singular Value Decomposition, Principal Component Analysis, Exercise: build your own recommendation system.

Text Book 2: Chapter 6

Module-5

Mining Social-Network Graphs: Social networks as graphs, Clustering of graphs, Direct discovery of

communities in graphs, Partitioning of graphs, Neighbourhood properties in graphs, Data Visualization:Basic principles, ideas and tools for data visualization. Data Science and Ethical Issues, Discussions on privacy, security, ethics, Next-generation data scientists.

Text Book 1: Chapter 8

Course outcomes:

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

- Define data science and its fundamentals
- Demonstrate the process in data science
- Explain machine learning algorithms necessary for data sciences
- Illustrate the process of feature selection and analysis of data analysis algorithms
- Visualize the data and follow of ethics

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year
1	Doing Data Science	Cathy O'Neil and Rachel Schutt	O'Reilly Media, Inc	2013
2	Mining of Massive Datasets.	Anand Rajaraman and Jeffrey D. Ullman	Cambridge University Press	2010

Reference Books

1	Machine Learning: A Probabilistic Perspective	Kevin P. Murphy	MIT Press	2013
2	Data Mining: Concepts andTechniques	Jiawei Han, Micheline Kamber and Jian Pei	Elsevier	2011.
3	Practical Statistics for Data Scientists	Peter Bruce and Andrew Bruce	O'Reilly Media, Inc	2017

M.TECH IN DATA SCIENCE (SDS)

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

SEMESTER - I

ADVANCED DATABASE MANAGEMENT SYSTEM

Course Code	**20SDS13,** 20SCS13, 20SCE252, 20SIT14, 20SSE15	CIE Marks	40
Teaching Hours / Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

Review of Relational Data Model and Relational Database Constraints:

Relational model concepts; Relational model constraints and relational database schemas; Update operations, anomalies, dealing with constraint violations, Types and violations.

Object and Object-Relational Databases:

Overview of Object Database Concepts, Object Database Extensions to SQL, The ODMG Object Model and the Object Definition Language ODL, Object Database Conceptual Design, The Object Query Language OQL, Overview of the C++ Language Binding in the ODMG Standard.

Text Book 1: Chapter 12

Module-2

Disk Storage, Basic File Structures, Hashing, and Modern Storage Architectures:

Introduction, Secondary Storage Devices, Buffering of Blocks, Placing File Records on Disk Operations on Files, Files of Unordered Records (Heap Files), Files of Ordered Records (Sorted Files), Hashing Techniques, Other Primary File Organizations, Parallelizing Disk Access Using RAID Technology, Modern Storage Architectures.

Distributed Database Concepts:

Distributed Database Concepts, Data Fragmentation, Replication, and Allocation Techniques for Distributed Database Design, Overview of Concurrency Control and Recovery in Distributed Databases, Overview of Transaction Management in Distributed Databases, Query Processing and Optimization in Distributed Databases, Types of Distributed Database Systems , Distributed Database Architectures, Distributed Catalogue Management.

Text Book 1: Chapter 16. 17, 23

Module-3

NOSQL Databases and Big Data Storage Systems:

Introduction to NOSQL Systems, The CAP Theorem, Document-Based NOSQL Systems and MongoDB,NOSQL Key-Value Stores, Column-Based or Wide Column NOSQL Systems, NOSQL Graph Databasesand Neo4j.

Big Data Technologies Based on MapReduce and Hadoop:

What Is Big Data? Introduction to MapReduce and Hadoop, Hadoop Distributed File System (HDFS), MapReduce: Additional Details Hadoop v2 alias YARN, General Discussion

Text Book 1: Chapter 24, 25

Module-4

Enhanced Data Models: Introduction to Active, Temporal, Spatial, Multimedia, and Deductive Databases: Active Database Concepts and Triggers, Temporal Database Concepts, Spatial Database Concepts, Multimedia Database Concepts, Introduction to Deductive Databases.

Introduction to Information Retrieval and Web Search: Information Retrieval (IR) Concepts, Retrieval Models, Types of Queries in IR Systems, Text pre-processing, Inverted Indexing, Evaluation Measures of Search relevance, web Search and Analysis. Trends in Information Retrieval.

Text Book 1: Chapter 26, 27

Module-5

Data Mining Concepts: Overview of Data Mining Technology, Association Rules, Classification, Clustering, Approaches to Other Data Mining Problems, Applications of Data Mining, Commercial Data Mining Tools.

Overview of Data Warehousing and OLAP: Introduction, Definitions, and Terminology, Characteristics of Data Warehouses, Data Modeling for Data Warehouses, building a Data Warehouse, Typical Functionality of a Data Warehouse, Data Warehouse versus Views, Difficulties of Implementing Data.

Text Book 1: Chapter 28, 29

Course outcomes:

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

- Select the appropriate high-performance database like parallel and distributed database
- Infer and represent the real-world data using object-oriented database
- Interpret rule set in the database to implement data warehousing of mining
- Discover and design database for recent applications database for better interoperability

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Books

Sl. No	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year

1	Fundamentals of Database Systems	Elmasri and Navathe	Pearson	2013
Reference Books				
1	Database System Concepts	Abraham Silberschatz, Henry F. Korth, S. Sudarshan	McGraw Hill	6th Edition, 2010
2	Database Management Systems	Raghu Ramakrishnan and Johannes Gehrke	McGraw-Hill	3rd Edition, 2013.

M.TECH IN DATA SCIENCE (SDS) **Choice Based Credit System (CBCS) and Outcome Based Education (OBE)** **SEMESTER – I**			
PREDICTIVE ANALYTICS			
Course Code	**20SDS14**	CIE Marks	40
Teaching Hours/Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03
Module-1			
Overview of Supervised Learning: Introduction, Variable Types and Terminology, Two Simple Approaches to Prediction: **Linear Methods for Regression and Classification:** Introduction, Linear regression models and least squares, , Subset selection , Shrinkage Methods, A Comparison of the Selection and Shrinkage Methods, Linear Discriminant Analysis, Logistic regression. **Text Book 1: Chapters 2.1 – 2.3, 3.1 – 3.4, 3.6, 4.1, 4.3 – 4.4**			
Module-2			
Model Assessment and Selection: Bias, Variance, and model complexity, The Bias-variance Decomposition, Optimism of the training error rate, Estimate of In-sample prediction error, The Effective number of parameters, Bayesian approach and BIC, Cross- validation, Boot strap methods, Conditional or Expected Test Error. **Text Book 1: Chapters 7.1 – 7.7, 7.10 – 7.12**			
Module-3			
Additive Models, Trees, and Related Methods: Generalized additive models, Tree-Based Methods, **Boosting and Additive Trees:** Boosting Methods, Exponential Loss and AdaBoost, Example: Spam Data, Numerical Optimization via Gradient Boosting , Illustrations (California Housing , New Zealand Fish, Demographic Data)			

Text Book 1: Chapters 9.1 – 9.2, 10.4, 10.8, 10.10, 10.13

Module-4

Neural Networks: Introduction, Fitting Neural Networks, Some Issues in Training Neural Networks

Support Vector Machines: Introduction, The Support Vector Classifier, Support Vector Machines and Kernels

Unsupervised Learning and Random forests: Association rules, Cluster analysis, Details of Random Forests, Random forests and analysis.

Text Book 1: Chapters 11.1, 11.3 – 11.5, 12.1 – 12.3, 14.1 – 14.3, 15.1 – 15.4

Module-5

Assessing Performance of a classification Algorithm (t-test, McNemar's test, Paired t-test, F-test), Analysis of Variance, Creating data for analytics through designed experiments.

Text Book 2: Chapter 19

Course outcomes:

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

- Apply Regression and classification models to solve real world problems
- Identify and analyze different analytical models
- Identify and apply Additive models to different data science related problems
- Apply Supervised and Unsupervised learning techniques
- Choose appropriate assessment evaluation criterion for different analytical methods

Question paper pattern:

- The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.
- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year
1	The Elements of Statistical Learning-Data Mining, Inference, and Prediction	Trevor Hastie, Robert Tibshirani, Jerome Friedman	Springer	2009

2	Introduction to Machine Learning	E. Alpaydin	PHI	2010
Reference Books				
1	Pattern Recognition and Machine Learning	Christopher M. Bishop	Springer	2007
2	All of statistics	L.Wasserman	Springer	2004
3	An Introduction to statistical learning with applications in R	G. James, D. Witten, T. Hastie, R. Tibshirani	Springer	2017

M.TECH IN DATA SCIENCE (SDS) **Choice Based Credit System (CBCS)and Outcome Based Education (OBE)** **SEMESTER–I**			
DATA ACQUISITION AND PRODUCTIZATION			
Course Code	**20SDS15**	CIE Marks	40
Teaching Hours / Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03
Module-1			
Introduction to Data Warehouse- OLTP and OLAP concepts, Introduction to Data Mining- Data Objects and Attribute Types, Basic Statistical Descriptions of Data, Exploratory Data analysis, Measuring Data Similarity and Dissimilarity, Graphical representation of data. **Text Book 1: Chapter 4** **Text Book 2: Chapter 1**			
Module-2			
Introduction to Data Acquisition, Applications, Process, Data Extraction, Data Cleaning and Annotation, Data Integration, Data Reduction, Data Transformation, Data Discretization and Concept Hierarchy Generation. **Text Book 1: Chapter 3**			
Module-3			
Visualization-Introduction, Terminology, Basic Charts and Plots- Multivariate Data Visualization, Data Visualization Techniques, Pixel-Oriented Visualization Techniques, Geometric Projection Visualization			

Techniques, Icon-Based Visualization Techniques, Hierarchical Visualization Techniques, Visualizing Complex Data and Relations, Data Visualization Tools, Rank Analysis Tools, Trend Analysis Tools, Multivariate Analysis Tools, Distribution Analysis Tools, Correlation Analysis Tools, Geographical Analysis Tools.

Text Book 1: Chapter 2

Module-4

IoT Overview, IoT Design methodology, Semantic Web Infrastructure, Intelligence Applications, Programming Framework for IoT, Distributed Data Analysis for IoT, Security and Privacy in IoT, Applied IoT, Cloud Based Smart Facilities Management.

Text Book 3: Chapter 2, 5

Module-5

Virtualization on Embedded Boards IoT, Stream Processing in IoT, Internet of Vehicles and Applications, Case study on Data Acquisition using Dashboards, Android and iOS apps.

Text Book 3: Chapter 8, 9

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

- Apply of data pre-processing, extraction, cleaning, annotation, integration on data.
- Apply the suitable visualization techniques to output analytical results.
- Explore on applications using Internet of things.

Question paper pattern: The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/Textbooks

Sl. No.	Title of the book	Name of the Author/s	PublisherName	Edition and year
1	Data mining: concepts and techniques	Jiawei Han, Jian Pei Micheline Kamber	Elsevier	3rd Edition, 2011.
2	Data Mining: Introductory and Advanced Topics	Margaret H. Dunham	Pearson Education	2012
3	Internet of Things -A hands-on approach	Arshdeep Bahga, Vijay Madisetti	Universities Press	2015

Reference Books

1	Intel Galileo and Intel Galileo Gen 2: API Features and Arduino Projects for Linux Programmers	Manoel Carlos Ramon	Apress	2014
2	Learning Qlikview Data Visualization	Karl Pover	Packt	2013
3	Internet of Things: Principles and Paradigms	Rajkumar Buyya, Amir Vahid Dastjerdi	Elsevier	2016

M.TECH IN Data SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – I

DATA SCIENCE AND ANALYTICS LAB

Course Code	**20SDSL16**	CIE Marks	40
Teaching Hours/Week (L:P:S)	0:4:0	SEE Marks	60
Credits	02	Exam Hours	03

The purpose of this laboratory is to get you acquainted with Python/R and use them in implementing Data Science and Algorithms.

- **Iris Data set**

Iris is a particularly famous toy dataset (i.e. a dataset with a small number of rows and columns, mostly used for initial small-scale tests and proofs of concept). This specific dataset contains information about the Iris, a genus that includes 260-300 species of plants. The Iris dataset contains measurements for 150 Iris flowers, each belonging to one of three species: Virginica, Versicolor and Setose. (50 flowers for each of the three species). Each of the 150 flowers contained in the Iris dataset is represented by 5 values:

- Sepal length, in cm
- Sepal width, in cm
- petal length, in cm
- petal width, in cm

Iris species, one of: iris-setose, iris-versicolor, iris-virginica. Each row of the dataset represents a distinct flower (as such, the dataset will have 150 rows). Each row then contains 5 values (4 measurements and a species label). The dataset is described in more detail on the UCI Machine Learning Repository website. The dataset can either be downloaded directly from there (iris.data file), or from a terminal, using the wget tool. The following command downloads the dataset from the original URL and stores it in a file named iris.csv.

$ wget "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data" -O iris.csv

- **Citybik.es**

Citybik.es is a website that offers an Application Programming Interface (or API, for short) for the usage of bike-sharing services throughout the world. Among the others, data for one of Turin's bike sharing system is available. The information available is at a "station" granularity. This means that all the data available regards the bike stations: some of the useful information available is the station name, its position (in terms of latitude and longitude), the number of available bikes and the number of free docks. The data is offered in near real-time (i.e. it is updated every 15-30 minutes).

The API endpoint to request the data about for the Bike service is the following: http://api.citybik.es/v2/networks/to-bike. This dataset is in the JSON (JavaScript Object Notation) format.

- **MNIST**

The MNIST dataset is another particularly famous dataset as CSV file. It contains several thousands of hand-written digits (0 to 9). Each hand-written digit is contained in a 28 × 28 8-bit grayscale image. This means that each digit has 784 (282) pixels, and each pixel has a value that ranges from 0 (black) to 255 (white). The dataset can be downloaded from the following URL: https://raw.githubusercontent.com/dbdmg/data-science-lab/master/datasets/mnist_test.csv. Each row of the MNIST datasets represents a digit. For the sake of simplicity, this dataset contains only a small fraction (10,000 digits out of 70,000) of the real MNIST dataset, which is known as the MNIST test set. For each digit, 785 values are available.

Exercises

1. Iris dataset

Load the Iris dataset as a list of lists (each of the 150 lists should have 5 elements). Compute and print the mean and the standard deviation for each of the 4 measurement columns (i.e. sepal length and width, petal length and width). Compute and print the mean and the standard deviation for each of the 4 measurement columns, separately for each of the three Iris species (Versicolor, Virginica and Setose). Which measurement would you consider "best", if you were to guess the Iris species based only on those four values?

2. Citybik.es dataset

Load the Citybik.es dataset as a Python dictionary. Use of the json module. Count and print the number of active stations (a station is active if its extra.status field is "online"). Count and print the total number of bikes available (field free_bikes) and the number of free docks (field empty_slots) throughout all stations. Given the coordinates (latitude, longitude) of a point (e.g. 45.074512, 7.694419), identify the closest bike station to it that has available bikes. For computing the distance among two points (given their coordinates), you can use the function distance_coords() defined in the code snippet below (which is an implementation of the great-circle distance):

from math import cos, acos, sin defdistance_coords(lat1, lng1, lat2, lng2):

"""Compute the distance among two points.""" deg2rad = lambda x: x * 3.141592 / 180

lat1, lng1, lat2, lng2 = map(deg2rad, [lat1, lng1, lat2, lng2])

R = 6378100 # Radius of the Earth, in meters

return R * acos(sin(lat1) * sin(lat2) + cos(lat1) * cos(lat2) * cos(lng1 - lng2))

3. MNIST dataset

Load the MNIST dataset. Create a function that, given a position $1 \leq k \leq 10,000$, prints the kthdigit of the dataset (i.e. thekthrow of the csv file) as a grid of 28×28 characters. More specifically, you should map each range of pixel values to the following characters:

[0, 64) → " "

[64, 128) → "."

[128, 192) → "*"

[192, 256) → "#"

Compute the Euclidean distance between each pair of the 784-dimensional vectors of the digits at the following positions: 26th, 30th, 32nd, 35th. Based on the distances computed in the previous step and knowing that the digits listed are 7, 0, 1, 1, can you assign the correct label to each of the digits?

4. Tips dataset

Read the dataset "Tips.csv" as a dataframe "Data". Extract the columns in the following sequence - Time, TotalBill, Tips. Plot a histogram for the variable 'TotalBill' to check which range has the highest frequency. Draw bar chart for the variable "Day". Identify the category with the maximum count. Demonstrate the data distributions using box, scatter plot, histogram, and bar chart on iris dataset. Demonstrate the correlation plot on iris dataset and perform exploratory visualization giving an overview of relationships among data with covariance analysis.

5. Split the Iris dataset into two the datasets - IrisTest_TrainData.csv, IrisTest_TestData.csv. Read them as two separate data frames named Train_Data and Test_Data respectively.

Answer the following questions:

a) How many missing values are there in Train_Data?
b) What is the proportion of Setosa types in the Test_Data?
c) What is the accuracy score of the K-Nearest Neighbor model (model_1) with 2/3 neighbors using Train_Data and Test_Data?
d) Identify the list of indices of misclassified samples from the 'model_1'.
e) Build a logistic regression model (model_2) keeping the modelling steps constant. Find the accuracy of the model_2

6. Implementation of the problem solving strategies: either using Forward Chaining or Backward Chaining.

7. For the given dataset mtcars.csv (www.kaggle.com/ruiromanini/mtcars), plot a histogram to check the frequency distribution of the variable „mpg" (Miles per gallon)

8. Train a regularized logistic regression classifier on the iris dataset (https://archive.ics.uci.edu/ml/machine-

learning-databases/iris/ or the inbuilt iris dataset) using sklearn.Train the model with the following hyperparameter C = 1e4 and report the best classification accuracy.

9. Train an SVM classifier on the iris dataset using sklearn. Try different kernels and the associated hyperparameters. Train model with the following set of hyperparameters RBF-kernel, gamma=0.5, onevs-rest classifier, no-feature-normalization. Also try C=0.01,1,10C=0.01,1,10. For the above set of hyperparameters, find the best classification accuracy along with total number of support vectors on the test data.

10. Consider the dataset spiral.txt (https://bit.ly/2Lm75Ly). The first two columns in the dataset corresponds to the co-ordinates of each data point. The third column corresponds to the actual cluster label. Compute the rand index for the following methods:

K – means Clustering

Single – link Hierarchical Clustering

Complete link hierarchical clustering. Also visualize the dataset and which algorithm will be able to recover the true clusters.

Course outcomes:

At the end of this course the students will be able to:

- Demonstrate proficiency with statistical analysis of data.
- Illustrate the ability to build and assess data-based models.
- Optimize the data using Classifiers.
- Apply clustering algorithms and logistic regressions on data sets.
- Apply kernel techniques on datasets.

Conduction of Practical Examination:

- All laboratory experiments (nos) are to be included for practical examination.

Evaluation:

- Students are allowed to pick one experiment from the list.
- Strictly follow the instructions as printed on the cover page of answer script for breakup of marks
- Change of experiment is allowed only once and marks allotted to the procedure part to be made Zero.

M.TECH IN DATA SCIENCE (SDS)

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

SEMESTER – I

RESEARCH METHODOLOGY AND IPR

(Professional Core Course) and (Common to all M.Tech Programmes)

Course Code	**20RMI17**	CIE Marks	40
Teaching Hours/Week (L:P:SDA)	1:0:2	SEE Marks	60
Credits	02	Exam Hours	03

Course objectives:

- To give an overview of the research methodology and explain the technique of defining a research problem
- To explain the functions of the literature review in research.
- To explain carrying out a literature search, its review, developing theoretical and conceptual frameworks and writing a review.
- To explain various research designs and their characteristics.
- To explain the details of sampling designs, and also different methods of data collections.
- To explain the art of interpretation and the art of writing research reports.
- To explain various forms of the intellectual property, its relevance and business impact in the changing global business environment.
- To discuss leading International Instruments concerning Intellectual Property Rights.

Module-1

Research Methodology: Introduction, Meaning of Research, Objectives of Research, Motivation in Research, Types of Research, Research Approaches, Significance of Research, Research Methods versus Methodology, Research and Scientific Method, Importance of Knowing How Research is Done, Research Process, Criteria of Good Research, and Problems Encountered by Researchers in India.

Text Book 1: Chapter 1

Module-2

Defining the Research Problem: Research Problem, Selecting the Problem, Necessity of Defining the Problem, Technique Involved in Defining a Problem, An Illustration.

Reviewing the literature: Place of the literature review in research, Bringing clarity and focus to your research problem, Improving research methodology, Broadening knowledge base in research area, Enabling contextual findings, How to review the literature, searching the existing literature, reviewing the selected literature, Developing a theoretical framework, Developing a conceptual framework, Writing about the literature reviewed.

Text Book 1: Chapter 2

Text Book 2: Chapter 3

Module-3

Research Design: Meaning of Research Design, Need for Research Design, Features of a Good Design, Important Concepts Relating to Research Design, Different Research Designs, Basic Principles of Experimental Designs, Important Experimental Designs.

Design of Sample Surveys: Introduction, Sample Design, Sampling and Non-sampling Errors, Sample Survey versus Census Survey, Types of Sampling Designs.

Text Book 1: Chapter 3, 4

Module-4

Data Collection: Experimental and Surveys, Collection of Primary Data, Collection of Secondary Data, Selection of Appropriate Method for Data Collection, Case Study Method. **Interpretation and Report Writing:** Meaning of Interpretation, Technique of Interpretation, Precaution in Interpretation, Significance of Report Writing, Different Steps in Writing Report, Layout. **Interpretation and Report Writing (continued):** of the Research Report, Types of Reports, OralPresentation, Mechanics of Writing a Research Report, Precautions for Writing Research Reports.

Text Book 1: Chapter 6, 14

Module-5

Intellectual Property: The Concept, Intellectual Property System in India, Development of TRIPS Complied Regime in India, Patents Act, 1970, Trade Mark Act, 1999,The Designs Act, 2000, The Geographical Indications of Goods (Registration and Protection) Act1999, Copyright Act,1957,The Protection of Plant Varieties and Farmers' Rights Act, 2001,The Semi-Conductor Integrated Circuits Layout Design Act, 2000, Trade Secrets, Utility Models, IPR and Biodiversity, The Convention on Biological Diversity (CBD) 1992, Competing Rationales for Protection of IPRs, Leading International Instruments Concerning IPR, World Intellectual Property Organisation (WIPO),WIPO and WTO, Paris Convention for the Protection of Industrial Property, National Treatment, Right of Priority, Common Rules, Patents, Marks, Industrial Designs, Trade Names, Indications of Source, Unfair Competition, Patent Cooperation Treaty (PCT), Advantages of PCT Filing, Berne Convention for the Protection of Literary and Artistic Works, Basic Principles, Duration of Protection, Trade Related Aspects of Intellectual Property Rights(TRIPS) Agreement, Covered under TRIPS Agreement, Features of the Agreement, Protection of Intellectual Property under TRIPS, Copyright and Related Rights, Trademarks, Geographical indications, Industrial Designs, Patents, Patentable Subject Matter, Rights Conferred, Exceptions, Term of protection, Conditions on Patent Applicants, Process Patents, Other Use without Authorization of the Right Holder Layout-Designs of Integrated Circuits, Protection of Undisclosed Information, Enforcement of Intellectual Property Rights, UNSECO.

Course outcomes:

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

- Discuss research methodology and the technique of defining a research problem
- Explain the functions of the literature review in research, carrying out a literature search, developing theoretical and conceptual frameworks and writing a review.
- Explain various research designs and their characteristics.
- Explain the art of interpretation and the art of writing research reports

Question paper pattern:

Examination will be conducted for 100 marks with question paper containing 10 fullquestions, each of 20 marks.

- Each full question can have a maximum of 4 sub questions.
- There will be 2 full questions from each module covering all the topics of the module
- Students will have to answer 5 full questions, selecting one full question from eachmodule
- The total marks will be proportionally reduced to 60 marks as SEE marks is 60

Text Books

Sl. No.	Title of the book	Name of the Author/s	PublisherName	Edition and year
1	Research Methodology: Methods and Techniques,	C.R. Kothari, Gaurav Garg	New Age International	4th Edition, 2018
2	Research Methodology a step-by-step guide for beginners. (For the topic Reviewing the literature under module 2),	Ranjit Kumar,	Sage Publications	3rd Edition, 2011
3	Study Material (For the topic Intellectual Property under module 5)	Professional Programme Intellectual Property Rights, Law and Practice,	The Institute of Company Secretaries of India, Statutory Body Under an Act of Parliament,	2013

Reference Books

1	Research Methods: the concise knowledgebase	Trochim	Atomic Dog Publishing	2005
2	Conducting Research Literature Reviews: From the Internet to Paper.	Fink A	Sage Publications	2009

M.TECH IN DATA SCIENCE (SDS) Choice Based Credit System (CBCS) and Outcome Based Education (OBE) SEMESTER – II			
ADVANCED MACHINE LEARNING			
Course Code	**20SDS21**	CIE Marks	40
Teaching Hours/Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

Ensemble Methods: Rationale for ensemble method, methods for constructing an Ensemble classifier, Bias-Variance decomposition, Bagging, Boosting, Random forests, Empirical comparison among Ensemble methods, **Class Imbalance Problem, Multiclass Problem**

Anomaly Detection: preliminaries, Statistical Approaches, Proximity-Based Outlier Detection, Density-Based Outlier Detection, Cluster-Based Techniques

Text book 2: Chapter 5.5 – 5.7 and 10.1 – 10.5

Module-2

Learning Sets of Rules: Introduction, Sequential Covering Algorithms, Learning Rule Sets, Learning First Order Rules, Learning Sets of First Order Rules, Induction as Inverted Deduction, Inverting Resolution.

Text book 1: Chapter 10.1 – 10.7

Module-3

Analytical learning: Introduction, Learning with Perfect domain theories, Remarks on Explanation based learning; Explanation based learning of search control knowledge.

Combining Inductive and Analytical learning: Motivation, Inductive-Analytical approaches to learning, Using prior knowledge to initialize the hypothesis, Using prior knowledge to initialize the hypothesis, Using prior knowledge to alter the search objective, Using prior knowledge to augment search.

Text book 1: Chapter 11.1 – 11.4 and 12.1 – 12.5

Module-4

Reinforcement Learning: Introduction, The Learning Task, Q Learning, Nondeterministic Rewards and Actions, Temporal Difference learning, Generalization from Examples, Relationship to Dynamic programming.

Text book 1: Chapter 13.1 – 13.7

Module-5

Transfer Learning: Introduction, Transfer in inductive learning: Inductive transfer, Bayesian transfer, Hierarchical transfer, Transfer with Missing Data or Class Labels.

Transfer in reinforcement learning: Starting-Point Methods, Imitation Methods, Hierarchical Methods, Alteration Methods, New RL Algorithms**.**

Avoiding negative transfer: Rejecting Bad Information, Choosing a Source Task, Modeling Task Similarity;

Automatically mapping tasks: Equalizing Task Representations, Trying Multiple Mappings, Mapping by Analogy; The future of transfer learning.

Course outcomes:

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

- Apply Ensemble techniques and identify anomalies in the data
- Apply Learning Sets of Rules to solve different problems
- Make use of Inductive and Analytical learning based methods to solve different problems
- Identify and apply a sequence of decisions and then tries to improve its next step to get the best outcome.
- Transfer, or cross-utilize the knowledge to solve machine learning problems.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.
The question paper will have ten full questions carrying equal marks.
Each full question is for 20 marks.
There will be two full questions (with a maximum of four sub questions) from each module.
Each full question will have sub question covering all the topics under a module.
The students will have to answer five full questions, selecting one full question from eachmodule.

Text Books

Sl. No.	Title of the book	Name of the Author/s	PublisherName	Edition and year
1	Machine Learning	Tom M. Mitchell,	Mc Graw Hill	2013
2	Introduction to Data Mining	Pang-Ning Tan, Michael Steinbach, Vipin Kumar	Pearson	1st Edition 2014
3	A Comprehensive Hands on Guide to Transfer Learning with Real World Applications in Deep Learning • (https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a) • https://ruder.io/transfer-learning/			

Reference Books

1	Reinforcement Learning	Marco Wiering Martijn van Otterlo (Eds.)	Springer	2012

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

INFORMATION RETRIEVAL

Course Code	**20SDS22,** 20SSE243	CIE Marks	40
Teaching Hours/Week (L:T:P)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

Introduction: Motivation, Basic concepts, Past, present, and future, The Retrieval Process.

Modeling: Introduction, A taxonomy of information retrieval models, Retrieval: Adhoc and filtering, A formal characterization of IR models, Classic information retrieval, Alternative set theoretic models, Alternative algebraic models, Alternative probabilistic models, Structured text retrieval models, Models for browsing.

Text Book 1: Chapter 1, 2

Module -2

Retrieval Evaluation: Introduction, Retrieval performance Evaluation, Reference collections.

Query Languages: Introduction, keyword-based querying, Pattern matching, Structural queries, Query protocols.
Query Operations: Introduction, User relevance feedback, Automatic local analysis, Automatic global analysis.

Text Book 1: Chapter 3, 4, 5

Module – 3

Text and Multimedia Languages and Properties: Introduction, Metadata, Text, Markup languages, Multimedia.
Text Operations: Introduction, Document preprocessing, Document clustering, Text compression, Comparing text compression techniques.

Text Book 1: Chapter 6, 7

Module-4

User Interfaces and Visualization: Introduction, Human-Computer interaction, The information access process, Starting pints, Query specification, Context, Using relevance judgments, Interface support for the search process.
Searching the Web: Introduction, Challenges, Characterizing the web, Search engines, Browsing, Meta searchers, Finding the needle in the haystack, Searching using hyperlinks.

Text Book 1: Chapter 10, 13

Module-5

Indexing and Searching: Introduction; Inverted Files; Other indices for text; Boolean queries; Sequential searching; Pattern matching; Structural queries; Compression.

Parallel and Distributed IR: Introduction, Parallel IR, Distributed IR.

Text Book 1: Chapter 8, 9

Course outcomes:

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

- Upon completion of the course, the students will be able to
- Build an Information Retrieval system using the available tools
- Identify and design the various components of an Information Retrieval system
- Apply machine learning techniques to text classification and clustering which is used for efficient Information Retrieval
- Analyze the Web content structure
- Design an efficient search engine

Question paper pattern:

- The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.
- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl.No.	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Modern Information Retrieval	Ricardo Baeza-Yates	Pearson	1999

Reference Books

Sl. No.	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Information Retrieval Algorithms and Heuristics	David A. Grossman, Ophir Frieder	Springer	2nd Edition 2004

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

DATA VISUALIZATION

Course Code	**20SDS23**	CIE Marks	40
Teaching Hours/Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

What Is Visualization?, History of Visualization, Relationship between Visualization and Other Fields, The Visualization Process, Types of Data, Structure within and between Records, Data Preprocessing, Perception in Visualization, Metrics, The Visualization Process in Detail, Semiology of Graphical Symbols, The Eight Visual Variables, Taxonomies.

Text Book 1: Chapter 1.1 – 1.4, 2.1 – 2.3, 3.4 – 3.5, 4.1 – 4.3, 4.5

Module -2

Visualization Techniques for Spatial Data, Visualization Techniques for Geospatial Data, Visualization Techniques for Multivariate Data

Text Book 1: Chapter 5.1 – 5.5, 6.1 – 6.5, 7.1 – 7.4

Module – 3

Visualization Techniques for Time-Oriented Data, Visualization Techniques for Trees, Graphs, and Networks, Text and Document Visualization

Text Book 1: Chapter 8.1 – 8.3, 9.1 – 9.6

Module-4

Interaction Concepts: Interaction Operators, Interaction Operands and Spaces, A Unified Framework,

Interaction Techniques: Screen Space, Object Space (3D Surfaces), Data Space (Multivariate Data Values), Attribute Space (Properties of Graphical Entities), Data Structure Space (Components of Data Organization), Visualization Structure Space (Components of the Data Visualization), Animating Transformations,

Designing Effective Visualizations: Steps in Designing Visualizations, Problems in Designing Effective Visualizations.

Text Book 1: Chapter 10.1 – 10.3, 11.1 – 11.7, 12.1 – 12.2

Module-5

Comparing and Evaluating Visualization Techniques, Visualization Systems, Research Directions in Visualization

Text Book 1: Chapter 13.1 – 13.7, 14.1 – 14.5, 15.1 – 15.6

Course outcomes:

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

- Explain the complete life cycle of BI/Analytical development
- Illustrate technology and processes associated with Business Intelligence framework
- Demonstrate a business scenario, identify the metrics, indicators and make recommendations toachieve the business goal.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of theAuthor/s	Publisher Name	Edition andyear
1	Interactive Data Visualization: Foundations, Techniques, and Applications	Matthew O. Ward, Georges Grinstein, Daniel Keim	CRC Press	2015

Reference Books

1	The Visual Display of Quantitative Information	Edward Tufte	Graphics Press	2001

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

MANAGING BIG DATA

Course Code	**20SDS241**, 20SSE322, 20SIT31, 20LNI251, 20SEE21, 20SFC331	CIE Marks	40
Teaching Hours/Week (L:P:S)	4 : 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

Meet Hadoop: Data!, Data Storage and Analysis, Querying All Your Data, Beyond Batch, Comparison with Other Systems: Relational Database Management Systems, Grid Computing, Volunteer Computing Hadoop Fundamentals **MapReduce:** A Weather Dataset: Data Format, Analyzing the Data with Unix Tools, Analyzing the Data with Hadoop:

Map and Reduce, Java MapReduce, Scaling Out: Data Flow, Combiner Functions, Running a Distributed MapReduce Job, Hadoop Streaming **The Hadoop Distributed File system** The Design of HDFS, HDFS Concepts: Blocks, Namenodes and Datanodes, HDFS Federation, HDFS High-Availability, The Command-Line Interface, Basic Filesystem Operations, Hadoop Filesystems Interfaces, The Java Interface, Reading Data from a Hadoop URL, Reading Data Using the FileSystem API, Writing Data, Directories, Querying the Filesystem, Deleting Data, Data Flow: Anatomy of a File Read, Anatomy of a File Write. **Text Book 1: Chapter 1, 2, and 3**
Module 2
YARN Anatomy of a YARN Application Run: Resource Requests, Application Lifespan, Building YARN Applications, YARN Compared to MapReduce, Scheduling in YARN: The FIFO Scheduler, The Capacity Scheduler, The Fair Scheduler, Delay Scheduling, Dominant Resource Fairness. **Hadoop I/O** Data Integrity, Data Integrity in HDFS, Local FileSystem, Checksum File System, Compression, Codecs, Compression and Input Splits, Using Compression in MapReduce, Serialization, The Writable Interface, Writable Classes, Implementing a Custom Writable, Serialization Frameworks, File-Based Data Structures: SequenceFile **Text Book 1: Chapter 4**
Module 3
Developing a MapReduce Application The Configuration API, Combining Resources, Variable Expansion, Setting Up the Development Environment, Managing Configuration, Generic Options Parser, Tool, and Tool Runner, Writing a Unit Test with MRUnit: Mapper, Reducer, Running Locally on Test Data, Running a Job in a Local Job Runner, Testing the Driver, Running on a Cluster, Packaging a Job, Launching a Job, The MapReduce Web UI, Retrieving the Results, Debugging a Job, Hadoop Logs, Tuning a Job, Profiling Tasks, MapReduce Workflows: Decomposing a Problem into MapReduce Jobs, JobControl, Apache Oozie **How MapReduce Works** Anatomy of a MapReduce Job Run, Job Submission, Job Initialization, Task Assignment, Task Execution, Progress and Status Updates, Job Completion, Failures: Task Failure, Application Master Failure, Node Manager Failure, Resource Manager Failure, Shuffle and Sort: The Map Side, The Reduce Side, Configuration Tuning, Task Execution: The Task Execution Environment, Speculative Execution, Output Committers. **Text Book 1: Chapter 5, 6**
Module 4
MapReduce Types and Formats: MapReduce Types, Input Formats: Input Splits and Records, Text Input, Binary Input, Multiple Inputs, Database Input (and Output) Output Formats: Text Output, Binary Output, Multiple Outputs, Lazy Output, Database Output, **Flume** Installing Flume, An Example, Transactions and Reliability, Batching, The HDFS Sink, Partitioning and Interceptors, File Formats, Fan Out, Delivery Guarantees, Replicating and Multiplexing Selectors, Distribution: Agent Tiers, Delivery Guarantees, Sink Groups, Integrating Flume with Applications, Component Catalog **Text Book 1: Chapter 7 and Reference Book 2**
Module 5

Pig Installing and Running Pig, Execution Types, Running Pig Programs, Grunt, Pig Latin Editors, An Example: Generating Examples, Comparison with Databases, Pig Latin: Structure, Statements, Expressions, Types, Schemas, Functions, Data Processing Operators: Loading and Storing Data, Filtering Data, Grouping and Joining Data, Sorting Data, Combining and Splitting Data.

Spark An Example: Spark Applications, Jobs, Stages and Tasks, A Java Example, A Python Example, Resilient Distributed Datasets: Creation, Transformations and Actions, Persistence, Serialization, Shared Variables, Broadcast Variables, Accumulators, Anatomy of a Spark Job Run, Job Submission, DAG Construction, Task Scheduling, Task Execution, Executors and Cluster Managers: Spark on YARN

Text Book 1: Chapter 11 and Reference Book 1

Course outcomes:

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

- Understand managing big data using Hadoop and SPARK technologies
- Explain HDFS and MapReduce concepts
- Install, configure, and run Hadoop and HDFS.
- Perform map-reduce analytics using Hadoop and related tools
- Explain SPARK concepts

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Hadoop: The Definitive Guide	Tom White	O'Reilley	3rd Edition, 2012
Reference Books				
1	SPARK: The Definitive Guide	Bill Chambers MateiZaharia	O'Reilley	2018
2	Apache Flume: Distributed Log Collection for Hadoop	D'Souza and Steve Hoffman	O'Reilley	2014

M.TECH IN DATA ACIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

NATURAL LANGUAGE PROCESSING

Course Code	**20SDS242,** 20SCS242, 20SCE243, 20SAM23	CIE Marks	40
Teaching Hours / Week (L:P:S)	4 : 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

OVERVIEW AND LANGUAGE MODELING: Overview: Origins and challenges of NLP-Language and Grammar-Processing Indian Languages- NLP Applications-Information Retrieval. Language Modeling: Various Grammar- based Language Models-Statistical Language Model.

Text Book 1: Chapter 1, 2

Module 2

WORD LEVEL AND SYNTACTIC ANALYSIS: Word Level Analysis: Regular Expressions-Finite-State Automata-Morphological Parsing-Spelling Error Detection and correction-Words and Word classes- Part-of Speech Tagging. Syntactic Analysis: Context-free Grammar-Constituency- Parsing- Probabilistic Parsing.

Text Book 1: Chapter 3, 2

Module 3

Extracting Relations from Text: From Word Sequences to Dependency Paths: Introduction, Subsequence Kernels for Relation Extraction, A Dependency-Path Kernel for Relation Extraction and Experimental Evaluation. Mining Diagnostic Text Reports by Learning to Annotate Knowledge Roles: Introduction, Domain Knowledge and Knowledge Roles, Frame Semantics and Semantic Role Labeling, Learning to Annotate Cases with Knowledge Roles and Evaluations. A Case Study in Natural Language Based Web Search: InFact System Overview, The GlobalSecurity.org Experience.

Module 4

Evaluating Self-Explanations in iSTART: Word Matching, Latent Semantic Analysis, and Topic Models: Introduction, iSTART: Feedback Systems, iSTART: Evaluation of Feedback Systems, Textual Signatures: Identifying Text-Types Using Latent Semantic Analysis to Measure the Cohesion of Text Structures: Introduction, Cohesion, Coh-Metrix, Approaches to Analysing Texts, Latent Semantic Analysis, Predictions, Results of Experiments. Automatic Document Separation: A Combination of Probabilistic Classification and Finite-State Sequence Modeling: Introduction, Related Work, Data Preparation, Document Separation as a Sequence Mapping Problem, Results. Evolving Explanatory Novel Patterns for Semantically based Text Mining: Related Work, A Semantically Guided Model for Effective Text mining.

Module 5

Information Retrieval and Lexical Resources: Information Retrieval: Design features of Information Retrieval

Systems-Classical, non classical, and Alternative Models of Information Retrieval – valuation Lexical Resources: World Net-Frame Net- Stemmers-POS Tagger- Research Corpora.

Course outcomes:

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

- Analyse the natural language text.
- Generate the natural language.
- Demonstrate Text mining.
- Apply information retrieval techniques.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl. No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Natural Language Processing and Information Retrieval	Tanveer Siddiqui, U.S. Tiwary	Oxford University Press	2008
2	Natural Language Processing and Text Mining	Anne Kao and Stephen R. Potee	Springer	2007

Reference Books

1	Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition	Daniel Jurafsky and James H Martin	PHI	2nd Edition, 2008
2	Natural Language Understanding	James Allen	Benjamin / Cummings publishing company	2nd Edition, 1995
3	Information Storage and	Gerald J. Kowalski	Kluwer academic	2000.

	Retrieval systems	and Mark T Maybury		
4	Natural Language Processing with Python	Steven Bird, Ewan Klein, Edward Loper	O'Reilly	2009
5	Foundations of Statistical Natural Language Processing	Christopher D. Manning and Hinrich Schutze	MIT Press	1999

M.TECH IN DATA ACIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

SOFT AND EVOLUTIONARY COMPUTING

Course Code	**20SDS243,** 20SCS323, 20SSE31, 20SAM22	CIE Marks	40
Teaching Hours/Week (L:P:S)	4 : 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

Introduction to Soft computing: Neural networks, Fuzzy logic, Genetic algorithms, Hybrid systems and its applications. Introduction to classical sets and fuzzy sets: Classical relations and fuzzy relations, Membershipfunctions.

Text Book 1: Chapter 1, 7, 8.1 – 8.4, 9

Module 2

Defuzzification, Fuzzy Arithmetic and Fuzzy Measures, Fuzzy Rule Base and Approximate Reasoning, Fuzzy Decision Making, Architecture and Operation of FLC System and applications.

Text Book 1: Chapter 10, 11, 12, 13and 14.3 – 14.4

Module 3

Genetic algorithms: Introduction, Basic operations, Traditional algorithms, Simple GA General Genetic Algorithm, Operators in Genetic Algorithm, Stopping Condition for Genetic Algorithm Flow, Constraints in Genetic Algorithm, Problem Solving Using Genetic Algorithm, The schema theorem, Genetic programming, applications.

Text Book 1: Chapter 15.1 – 15.13, 15.16

Module 4

Swarm Intelligence System: Introduction, background of SI, Ant colony system Working of ant colony optimization, ant colony for TSP.

Text Book 2: Chapter 8

Module 5

Unit commitment problem, particle Swarm Intelligence system Artificial bee colony system, Cuckoo search system.

Course outcomes:

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

- Implement machine learning through neural networks.
- Design Genetic Algorithm to solve the optimization problem.
- Develop a Fuzzy expert system.
- Model Neuro Fuzzy system for clustering and classification

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Principles of Soft computing	Shivanandam, Deepa S. N	Wiley	2011
2	Soft Computing with MATLAB Programming	N. P. Padhy, S. P. Simon	Oxford	2015

Reference Books

1	Neuro-fuzzy and soft computing	J.S.R. Jang, C.T. Sun, E. Mizutani	PHI	2012
2	Soft Computing: Fundamentals, Techniques and Applications	Saroj Kaushik, SunitaTiwari	Mc Graw Hill	2018

M.TECH IN DATA ACIENCE (SDS)

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

SEMESTER – II

PATTERN RECOGNITION

Course Code	**20SDS244,** 20SCS244, 20SCE242, 20SAM253	CIE Marks	40
Teaching Hours/Week (L:P:S)	4 : 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

Introduction: Definition of PR, Applications, Datasets for PR, Different paradigms for PR, Introduction to probability, events, random variables, Joint distributions and densities, moments. Estimation minimum risk estimators, problems

Text Book 1: Chapter 1

Text Book 2: Chapter 1

Module 2

Representation: Data structures for PR, Representation of clusters, proximity measures, size of patterns, Abstraction of Data set, Feature extraction, Feature selection, Evaluation of Classifiers and Clustering

Text Book 1: Chapter 2

Module 3

Nearest Neighbour based classifiers & Bayes classifier: Nearest Neighbour Algorithm, Variants of NN Algorithm, Use of NN for Transaction Databases, Efficient Algorithms, Data Reduction, Prototype Selection, Bayes theorem, Minimum Error Rate Classifier, Estimation of Probabilities, Comparison with NNC, Naive Bayes classifier, Bayesian belief network.

Text Book 1: Chapter 3 and 4

Module 4

Hidden Markov models: Markov Models for Classification, Hidden Markov Models and Classification Using HMMS.

Decision Trees: Introduction, Decision Trees for Pattern Recognition, Construction of Decision Trees, Splitting at the Nodes, Over fitting & Pruning, Example of Decision Tree Induction.

Text Book 1: Chapter 5 and 6

Module 5

Clustering: Hierarchical (Agglomerative, single/complete/average linkage, wards, Partitional (Forgy's, k-means, Isodata), Clustering Large Data Sets, examples, An application: Handwritten Digit recognition

Text Book 1: Chapter 9

Course outcomes:

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

- Explain pattern recognition principals
- Develop algorithms for Pattern Recognition.
- Develop and analyse decision tress.
- Design the nearest neighbour classifier.
- Apply Decision tree and clustering techniques to various applications

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Pattern Recognition	V Susheela Devi, M Narsimha Murthy	Universities Press	2011
2	Pattern Recognition and Image Analysis	Earl Gose, Richard Johnsonbaugh, Steve Jost	PHI	1996
Reference Books				
1	Pattern Classification	Duda R. O., P.E. Hart, and D. G. Stork	Wiley	2000.

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER –II

DATA SECURITY AND PRIVACY

Course Code	**20SDS251**	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

Classical Encryption Techniques Symmetric Cipher Model, Cryptography, Cryptanalysis and Brute-Force Attack, Substitution Techniques, Caesar Cipher, Monoalphabetic Cipher, Playfair Cipher, Hill Cipher, Poly alphabetic Cipher, One Time Pad.

Block Ciphers and the data encryption standard: Traditional block Cipher structure, stream Ciphers and block Ciphers, Motivation for the feistel Cipher structure, the feistel Cipher, The data encryption standard, DES encryption, DES decryption, A DES example, results, the avalanche effect, the strength of DES, the use of 56-Bit Keys, the nature of the DES algorithm, timing attacks, Block cipher design principles, number of rounds, design of function F, key schedule algorithm.

Text Book1: Chapter 3, Chapter 4

Module-2

Public-Key Cryptography and RSA: Principles of public-key cryptosystems. Public- key cryptosystems. Applications for public-key cryptosystems, requirements for public key cryptosystems. Public-key cryptanalysis. The RSA algorithm, description of the algorithm, computational aspects, the security of RSA.

Other Public-Key Cryptosystems: Diffiehellman key exchange, The algorithm, key exchange protocols, man in the middle attack, Elgamal Cryptographic systems, Elliptic curve arithmetic, abelian groups, elliptic curves over real numbers, elliptic curves over Zp, elliptic curves overGF(2m), Elliptic curve cryptography, Analog of Diffie-hellman key exchange, Elliptic curve encryption/ decryption, security of Elliptic curve cryptography, Pseudorandom number generation based on a asymmetric cipher

Text book 1: Chapter 9, Chapter 10

Module-3

Key Management and Distribution: Symmetric key distribution using Symmetric encryption, A key distribution scenario, Hierarchical key control, session key lifetime, a transparent key control scheme, Decentralized key control, controlling key usage, Symmetric key distribution using asymmetric encryption, simple secret key distribution, secret key distribution with confidentiality and authentication, A hybrid scheme, distribution of public keys, public announcement of public keys, publicly available directory, public key authority, public keys certificates, X-509 certificates. Certificates, X-509 version 3, Public Key infrastructure.

Text Book 1: Chapter 14

Module-4

An Introduction to privacy preserving data mining: Privacy-Preserving Data Mining Algorithms, The Randomization Method, Group Based Anonymization.

Text Book 2: Chapter 1 -1.1, 1.2 , Chapter 2 - 2.2, 2.3

Module-5

Distributed Privacy-Preserving Data Mining, Privacy-Preservation of Application Results, Limitations of Privacy: The Curse of Dimensionality, Applications of Privacy- Preserving Data Mining

Text Book 2: Chapter 2 - 2.4, 2.5, 2.6, 2.7

Course outcomes:

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

- Analyze the vulnerabilities in any computing system and hence be able to design a security solution.
- Identify the security issues in the network and resolve it.
- Evaluate security mechanisms using rigorous approaches, including theoretical.
- Describe importance of data privacy, limitations and applications

Question paper pattern: The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year
1	Cryptography and Network Security,	William Stallings	Pearson	7th Edition 2017
2	Privacy Preserving Data Mining: Models and Algorithms	Charu C. Aggarwal, Philip S Yu,	Kluwer Academic	2008
Reference Books				
1	Cryptography and Network Security	Atul Kahate	McGraw Hill	2019
2	Cryptography and Information Security,	Pachghare V K	PHI	2019

M.TECH IN DATA ACIENCE (SDS)

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

SEMESTER – II

TIME SERIES ANALYSIS AND FORECASTING

Course Code	20SDS252	CIE Marks	40
Teaching Hours/Week (L:P:S)	4: 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

An Introduction to Forecasting: Forecasting and Data. Forecasting Methods. Errors in Forecasting. Choosing a Forecasting Technique. An Overview of Quantitative Forecasting Techniques. REGRESSION ANALYSIS: The Simple Linear Regression Model. The Least Squares Point Estimates. Point Estimates and Point Predictions. Model Assumptions and the Standard Error. Testing the Significance of the Slope and y Intercept. Confidence and Prediction Intervals. Simple Coefficients of Determination and Correlation. An F Test for the Model.

Text book 1: Chapter 1 & 2 (Part-I)

Module 2

Multiple Linear Regressions: The Linear Regression Model. The Least Squares Estimates, and Point Estimation and Prediction. The Mean Square Error and the Standard Error. Model Utility: R2, Adjusted R2, and the Overall F Test. Model Building and Residual Analysis: Model Building and the Effects of Multicollinearity. Residual Analysis in Simple Regression. Residual Analysis in Multiple Regressions. Diagnostics for Detecting Outlying and Influential Observations.

Text book 1: Chapter 3, 4 & 5 (Part-II)

Module 3

Time Series Regression: Modeling Trend by Using Polynomial Functions. Detecting Autocorrelation. Types of Seasonal Variation. Modeling Seasonal Variation by Using Dummy Variables and Trigonometric Functions. Growth Curves. Handling First-Order Autocorrelation. Decomposition Methods: Multiplicative Decomposition. Additive Decomposition. The X-12-ARIMA Seasonal Adjustment Method. Exercises. Exponential Smoothing: Simple Exponential Smoothing. Tracking Signals. Holt's Trend Corrected Exponential Smoothing. Holt-Winters Methods. Damped Trends and Other Exponential.

Text book 1: Chapter 6, 7 & 8 (Part- III)

Module 4

Non-seasonal Box-Jenkins Modeling and Their Tentative Identification: Stationary and Non-stationary Time Series. The Sample Autocorrelation and Partial Autocorrelation Functions: The SAC and SPAC. An Introduction to Non-seasonal Modeling and Forecasting. Tentative Identification of Non-seasonal Box-Jenkins Models. Estimation, Diagnostic Checking, and Forecasting for Non-seasonal Box-Jenkins Models: Estimation. Diagnostic Checking. Forecasting. A Case Study. Box-Jenkins Implementation of Exponential Smoothing.

Text book 1: Chapter 9 & 10 (Part- IV)

Module 5

Box-Jenkins Seasonal Modeling: Transforming a Seasonal Time Series into a Stationary Time Series. Examples of Seasonal Modeling and Forecasting. Box-Jenkins Error Term Models in Time Series Regression. Advanced Box-Jenkins Modeling: The General Seasonal Model and Guidelines for Tentative Identification. Intervention Models. A Procedure for Building a Transfer Function Model Causality in time series: Granger causality. Hypothesis testing on rational expectations. Hypothesis testing on market efficiency.

Text book 1: Chapter 11 & 12 (Part- IV)

Course outcomes:

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

- Describe the fundamental advantage and necessity of forecasting in various situations.
- Identify how to choose an appropriate forecasting method in a particular environment.
- Apply various forecasting methods, which include obtaining the relevant data and carrying out the necessary computation using suitable statistical software.
- Improve forecast with better statistical models based on statistical analysis.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and Year
1	Forecasting, Time Series, and Regression	Bruce L. Bowerman, Richard O'Connell, Anne Koehler	Cengage Learning	2004

Reference Books

1	The Econometric Modelling of Financial Time Series	Terence C. Mills, Raphael N. Markellos	Cambridge University Press,	2008
2	Time Series Models	Andrew C. Harvey	MIT	1993
3	Introduction to Time Series and Forecasting. (http://home.iitj.ac.in/~parmod/ document/introduction %20time%20series.pdf)	P. J. Brockwell, R. A. Davis	Springer	1996

4	Time series analysis : with applications in R	Jonathan D Cryer, Kung-sik Chan	Springer	2008
5	Applied Econometric Time Series	W Enders	Wiley	1995

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

APPLIED SOCIAL NETWORK ANALYSIS

Course Code	20SDS253	CIE Marks	40
Teaching Hours/Week (L:P:S)	4: 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

What is a Network?- Basic Network Concepts, Adjacency Matrices, Graphs, and Notation, Nodes and Links, Good Will Hunting Problem, Formal and Informal Networks, summary. **Centrality measures**- What is "Centrality" and Why do we Study It?, calculating Nodal Centrality Measures, Directed Networks and Centrality Measures, Location in the Network. Graph Level Measures- Density , Diameter , Centralization , Average Centralities, Network Topology

Textbook 1: Chapter 1, 2, 3

Module 2

Social Theory: **Social Links-** Individual Actors, Social Exchange Theory, Social Forces, Graph Structure, Agent Optimization Strategies in Networks, Hierarchy of Social Link Motivation, Summary. **Subgroup Analysis:** Subgroups, Organizational Theory, Random Groups, Heuristics for Subgroup Identification, Analysis Methods, Summary. **Diffusion and Influence:** Applications for Social Diffusion, Strain Theory, Social Context, Group Impacts on Diffusion, Network Structure and Diffusion, Group Influence Strategies and Bases of Power.

Textbook 1: Chapter 4, 5, 6

Module 3

Meta-Networks and Relational Algebra: Modes of Data, Source, Target, Direction, Multimode Networks, Bridging a Meta-Network, Strength of Ties. **Sources of Data:** Network Sampling, Measuring Links, Data Quality, Additional Ethnographic Data Collection Methods, Anonymity Issues.

Textbook 1: Chapter 7, 8

Module 4

Information Networks and the World Wide Web: The Structure of the web, Link Analysis and Web Search, Sponsored Search Markets.

Textbook 2: Chapter 13, 14, 15

Module 5

Network Dynamics: Structural Models, The Small-World Phenomenon, Epidemics .

Textbook 2: Chapter 19, 20, 21

Course outcomes:

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

- Define notation and terminology used in network science.
- Demonstrate, summarize and compare networks.
- Explain basic principles behind network analysis algorithms.
- Analyzing real world network

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and Year
1	Social Network Analysis with Applications	Ian McCulloh, Helen Armstrong and Anthony Johnson	Wiley	2013
2	Networks, Crowds, and Markets: Reasoning About a Highly Connected World	David Easley and JohnKleinberg	Cambridge University Press	2010
Reference Books				
1	Social and Economic Networks	Matthew O. Jackson	Princeton University Press	2008

M.TECH IN DATA ACIENCE (SDS)

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

SEMESTER – II

IMAGE AND VIDEO ANALYTICS

Course Code	20SDS254	CIE Marks	40
Teaching Hours/Week (L:P:S)	4: 0 : 0	SEE Marks	60
Credits	04	Exam Hours	03

Module 1

Digital image representation- Visual Perception- Sampling and Quantization- Basic Relations between Pixels- Mathematical Tools Used in Digital Image Processing: Fundamental Operations –Vector and Matrix Operations- Image Transforms (DFT, DCT, DWT, Hadamard).

Text Book 1: Chapter 1, 2.1 – 2.6

Module 2

Fundamentals of spatial filtering: spatial correlation and convolution-smoothing blurring- sharpening- edge detection - Basics of filtering in the frequency domain: smoothing-blurring- sharpening--Histograms and basic statistical models of image.

Text Book 1: Chapter 3.1 – 3.6

Module 3

Detection of Objects of Interest, Tracking of Objects of Interest in a Sequence of Images, Tracking Objects of Interest Through a Camera Network.

Text Book 2: Chapter 7, 8, 9

Module 4

Biometric Techniques Applied to Video Surveillance, Vehicle Recognition in Video Surveillance, Activity Recognition.

Text Book 2: Chapter 10, 11, 12

Module 5

Unsupervised Methods for Activity Analysis and Detection of Abnormal Events, Analysis of Crowded Scenes in Video, Detection of Visual Context, Example of an Operational Evaluation Platform: PPSL

Text Book 2: Chapter 13, 15, 16, 17

Course outcomes:

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

- Apply preprocessing techniques for better understanding of images
- Apply spatial filtering to images for the preprocessing purposes

- Apply color transformations to image and video data
- Make use of detection and classification methods for image and video data
- Apply and analyze the techniques through case studies

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Text Book / Text Books

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Digital Image Processing	R.C. Gonzalez and R.E. Woods	Pearson	2009
2	Intelligent Video Surveillance Systems	Jean-Yves Dufour	Wiley	2013
Reference Books				
1	Computer Vision: Algorithms and Applications	Rick Szelisk	Springer	2011
2	Nonparametric and Semi parametric Models	Härdle W, Müller M, Sperlich S, Werwatz, A.	Springer	2004

M.TECH IN Data SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER – II

VISUALIZATION & DATA SCIENCE AND MINI PROJECT LAB

Subject Code	20SDSL26	**CIE Marks**	40
Number of Contact Hours/Week (L:P:S)	0:4:0	**SEE Marks**	60
Credits	02	**Exam Hours**	3

Course Learning Objectives: This course will enable students to:

- Make use of Data sets in implementing the data visualization techniques
- Implement the data visualization techniques
- Integrate machine learning libraries and mathematical and statistical tools that are suitable for the Data Science applications under consideration.

Descriptions (if any):

PART A:

- Write the programs using Python/R/equivalent Programming Language.
- Execute the programs in either Visual Studio Code or PyCharm Community/equivalent Edition.
- Use appropriate libraries as needed to execute the programs.
- Data sets can be downloaded from standard repositories (https://archive.ics.uci.edu/ml/datasets.html) or constructed by the students

PART B:

- Data Science applications is to be developed with suitable Graphical User Interface.

Installation procedure of the required software must be demonstrated, carried out in groups and documented in the journal.

PART A:

1	Load the specified dataset a) Using numpy and compute mean, median, variance and Standard deviation and illustrate Indexing, Slicing, Splitting, Iterating, Filtering, Sorting, Combining, and Reshaping b) Using pandas and compute mean, median, variance and Standard deviation and illustrate Indexing, Slicing, Iterating, Filtering, Sorting and Reshaping
2	For a given set of training data examples stored in a .CSV file, implement and demonstrate the (Note: Import Matplotlib) a) Visualizing through a Line Plot b) Creating a Bar Plot c) Creating a Stacked Bar Plot to Visualize a specified parameter

	d) Comparing specific parameters using Stacked Area Chart
3	For a given set of training data examples stored in a .CSV file, implement and demonstrate the (Note: Import Matplotlib)
4	a) Histogram and a Box Plot to Visualize the given parameter b) Scatter Plot to Visualize Correlation c) Scatter Plot with Marginal Histograms d) Plotting Multiple Images in a Grid
5	Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs
6	Write a program to parse HTML using Beutifulsoup/equivalent Library support

PART B

Mini Project

- Use Java, C#, PHP, Python, or any other similar front-end tool. Developed mini projects must be demonstrated on desktop/laptop as a stand-alone or web based application
- Installation procedure of the required software must be demonstrated, carried out in groups and documented in the journal.
- Indicative areas include: health care, education, agriculture, banking, library, agent based systems, registration systems, industry, reservation systems, facility management, super market etc., Similar to but not limited to:

  Handwritten Digit Recognition

  Prediction of Cardiac Arrhythmia type using Clustering and Regression Approach

  Hybrid Regression Technique for House Prices Prediction

  An Iris Recognition Algorithm for Identity Authentication

  An Approach to Maintain Attendance using Image Processing Techniques

  Unconstrained Face Recognition

  Vehicle Number Plate Detection System

  Detection of Fake News

  Stock Prediction using Linear Regression

  Prediction of Weather Report

  Analyzing Bike Sharing Trends

  Sentiment Analysis for Movie Reviews

  Analyzing and Recommendations of Music Trends

  Forecasting Stock and Commodity Prices

  Diabetes Prediction Speech Recognition

Spam Detection using neural Networks in Python Combining satellite imagery and to predict poverty
Conduct of Practical Examination:
• Experiment distribution ✓ For laboratories having only one part: Students are allowed to pick one experiment from the lot with equal opportunity. ✓ For laboratories having PART A and PART B: Students are allowed to pick one experiment from PART A and one experiment from PART B, with equal opportunity. • Change of experiment is allowed only once and marks allotted for procedure to be made zero of the changed part only.
• Marks Distribution (Subjected to change in accordance with university regulations) a) For laboratories having only one part – Procedure + Execution + Viva-Voce: 15+70+15 = 100 Marks b) For laboratories having PART A and PART B i. Part A – Procedure + Execution + Viva = 6 + 28 + 6 = 40 Marks ii. Part B – Procedure + Execution + Viva = 9 + 42 + 9 = 60 Marks

TECHNICAL SEMINAR			
Course Code	20SDS27	**CIE Marks**	100
Number of contact Hours/week (L:P:SDA)	0:2:0	**SEE Marks**	--
Credits	02	**Exam Hours**	--

Course objectives:

The objective of the seminar is to inculcate self-learning, face audience confidently, enhance communication skill, involve in group discussion and present and exchange ideas.

Each student, under the guidance of a Faculty, is required to

- Choose, preferably through peer reviewed journals, a recent topic of his/her interest relevant to the Course of Specialization.
- Carryout literature survey, organize the Course topics in a systematic order.
- Prepare the report with own sentences.
- Type the matter to acquaint with the use of Micro-soft equation and drawing tools or any such facilities.
- Present the seminar topic orally and/or through power point slides.
- Answer the queries and involve in debate/discussion.
- Submit two copies of the typed report with a list of references.

The participants shall take part in discussion to foster friendly and stimulating environment in which the students are motivated to reach high standards and become self-confident.

The CIE marks for the seminar shall be awarded (based on the relevance of the topic, presentation skill, participation in the question and answer session and quality of report) by the committee constituted for the purpose by the Head of the Department. The committee shall consist of three faculties from the department with the senior most acting as the Chairperson.

Marks distribution for CIE of the course 21XXX27 seminar:

Seminar Report: 30 marks

Presentation skill: 50 marks

Question and Answer: 20 marks

M.TECH IN DATA SCIENCE (SDS)

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

SEMESTER – III

DEEP LEARNING

Course Code	**20SDS31,** 20SCS31, 20SAM31, 20SIS334	CIE Marks	40
Teaching Hours/Week (L:P:S)	3:0:2	SEE Marks	60
Credits	04	Exam Hours	03

Module-1

Machine Learning Basics: Learning Algorithms, Capacity, Overfitting and Underfitting, Hyperparameters and Validation Sets, Estimator, Bias and Variance, Maximum Likelihood Estimation, Bayesian Statistics, Supervised Learning Algorithms, Unsupervised Learning Algorithms, Stochastic Gradient Decent, Building a Machine Learning Algorithm, Challenges Motivating Deep Learning.

Text Book 1: Chapter 5

Module-2

Deep Feedforward Networks: Gradient-Based Learning, Hidden Units, Architecture Design, Back- Propagation.

Regularization for Deep Learning: Parameter Norm Penalties, Norm Penalties as Constrained Optimization, Regularization and Under-Constrained Problems, Dataset Augmentation, Noise Robustness, Semi- Supervised Learning, Multi-Task Learning, Early Stopping, Parameter Tying and Parameter Sharing, Sparse Representations, Bagging, Dropout.

Text Book 1: Chapter 6.2 – 6.5 and 7.1 – 7.11

Module-3

Optimization for Training Deep Models: How Learning Differs from Pure Optimization, Challenges in Neural Network Optimization, Basic Algorithms. Parameter Initialization Strategies, Algorithms with Adaptive Learning Rates.

Convolutional Networks: The Convolution Operation, Motivation, Pooling, Convolution and Pooling as an Infinitely Strong Prior, Variants of the Basic Convolution Function, Structured Outputs, Data Types, Efficient Convolution Algorithms, Random or Unsupervised Features.

Text Book 1: Chapter 8.1 – 8.5 and 9.1 – 9.9

Module-4

Sequence Modelling: Recurrent and Recursive Nets: Unfolding Computational Graphs, Recurrent Neural Networks, Bidirectional RNNs, Encoder-Decoder Sequence-to-Sequence Architectures, Deep Recurrent Networks, Recursive Neural Networks. Long short-term memory

Text Book 1: Chapter 10.1 – 10.6 and 10.10

Module-5

Practical Methodology: Performance Metrics, Default Baseline Models, Determining Whether to Gather More Data, Selecting Hyperparameters, Debugging Strategies, Example: Multi-Digit NumberRecognition. **Applications:** Vision, NLP, Speech.

Text Book 1: Chapter 11 and 12

Course outcomes:

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

- Identify the deep learning algorithms which are more appropriate for various types of learningtasks in various domains.
- Implement deep learning algorithms and solve real-world problems.
- Execute performance metrics of Deep Learning Techniques.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from eachmodule.

Textbook/ Textbooks

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Deep Learning	Ian GoodfellowYoshuaBengio, and Aaron Courville	MIT Press https://www.deeplearningbook.org/	2016
Reference Books				
1	Neural Networks: A systematic Introduction	Raul Rojas	Springer	1996
2	Pattern Recognition and Machine Learning	Chirstopher Bishop	Springer	2006

M.TECH IN DATA SCIENCE (SDS)

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

SEMESTER – III

BUSINESS ANALYTICS

Course Code	**20SDS321**	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03

Module-1

Development Steps, BI Definitions, BI Decision Support Initiatives, Development Approaches, Parallel Development Tracks, BI Project Team Structure, Business Justification, Business Divers, Business Analysis Issues, Cost – Benefit Analysis, Risk Assessment, Business Case Assessment Activities, Roles Involved in These Activities, Risks of Not Performing Step, Hardware, Middleware, DBMS Platform, Non-Technical Infrastructure Evaluation.

Textbook 1: Chapter: 0, 1, 2

Module -2

Managing The BI Project, Defining and Planning The BI Project, Project Planning Activities, Roles and Risks Involved In These Activities, General Business Requirement, Project Specific Requirements,Interviewing Process.

Textbook 1: Chapter: 3, 4

Module – 3

Differences in Database Design Philosophies, Logical Database Design, Physical Database Design, Activities, Roles And Risks Involved In These Activities, Incremental Rollout, Security Management, Database Backup And Recovery.

Textbook 1: Chapter: 8

Module-4

Descriptive Analytics: Visualizing and Exploring data, Descriptive Statistical Measures, Sampling and Estimation, Statistical Inference

Textbook 2: Chapter: 3, 4, 6, 7

Module-5

Predictive Analytics: Trendlines and Regression Analysis, Forecasting Techniques, Spreadsheet Modeling and Analysis, Basic concepts on Prescriptive Analytics

Textbook 2: Chapter: 8, 9, 11, and Part 4

Course outcomes:

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

- Explain the complete life cycle of BI/Analytical development
- Illustrate technology and processes associated with Business Intelligence framework
- Demonstrate a business scenario, identify the metrics, indicators and make recommendations to achieve the business goal.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.

The students will have to answer five full questions, selecting one full question from eachmodule.

Textbook/ Textbooks

Sl No	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year
1	Business Intelligence Roadmap:The Complete Project Lifecycle for Decision Support Applications	Larissa T Moss and ShakuAtre	Addison Wesley Information Technology Series	2003.
2	Business Analytics	James Evans	Pearson	2nd Edition 2016
Reference Books				
1	Business Intelligence: The Savvy Manager's Guide	David Loshin	Morgan Kaufmann	2003
2	Delivering Business Intelligence with Microsoft SQL Server 2005	Brian Larson	McGraw Hill	2006
3	Foundations of SQL Server 2008 Business Intelligence	Lynn Langit	Apress	2011
4	Fundamentals of Business Analytics	R N Prasad, Seema Acharya	Wiley India	2011.

M.TECH DATA SCIENCE (SDS) Choice Based Credit System (CBCS) and Outcome Based Education (OBE) SEMESTER – III			
CYBER SECURITY AND CYBER LAW			
Course Code	**20SDS322,**20SCR15, 20LNI242, 20SCE244, 20SIT244, 20SAM244	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03

Module-1

Introduction to Cybercrime: Cybercrime: Definition and Origins of the Word, Cybercrime and Information Security, Who are Cybercriminals?, Classifications of Cybercrimes, Cybercrime: The Legal Perspectives, Cybercrimes: An Indian Perspective, Cybercrime and the Indian ITA 2000, A Global Perspective on Cybercrimes, Cybercrime Era: Survival Mantra for the Netizens. Cyberoffenses: How Criminals Plan Them: How Criminals Plan the Attacks, Social Engineering, Cyberstalking, Cybercafe and Cybercrimes, Botnets: The Fuel for Cybercrime, Attack Vector, Cloud Computing.

Textbook 1: Chapter 1, 2

Module -2

Cybercrime: Mobile and Wireless Devices: Introduction, Proliferation of Mobile and Wireless Devices, Trends in Mobility, Credit Card Frauds in Mobile and Wireless Computing Era, Security Challenges Posed by Mobile Devices, Registry Settings for Mobile Devices, Authentication Service Security, Attacks on Mobile/Cell Phones, Mobile Devices: Security Implications for organizations, Organizational Measures for Handling Mobile, Organizational Security Policies and Measures in Mobile Computing Era, Laptops.

Textbook 1: Chapter 3

Module – 3

Tools and Methods Used in Cybercrime: Introduction, Proxy Servers and Anonymizers, Phishing, Password Cracking, Keyloggers and Spywares, Virus and Worms, Trojan Horses and Backdoors, Steganography, DoS and DDoS Attacks, SQL Injection, Buffer Overflow, Attacks on Wireless Networks. Phishing and Identity Theft: Introduction, Phishing, Identity Theft (ID Theft).

Textbook 1: Chapter 4, 5

Module-4

Understanding Computer Forensics: Introduction, Historical Background of Cyber forensics, Digital Forensics Science, The Need for Computer Forensics, Cyber forensics and Digital Evidence, Forensics Analysis of E-Mail, Digital Forensics Life Cycle, Chain of Custody Concept, Network Forensics, Approaching a Computer Forensics Investigation, Setting up a Computer Forensics Laboratory: Understanding the Requirements, Computer Forensics and Steganography, Relevance of the OSI 7 Layer Model to Computer Forensics, Forensics and Social Networking Sites: The Security/Privacy Threats, Computer Forensics from Compliance Perspective,

Challenges in Computer Forensics, Special Tools and Techniques, Forensics Auditing, Anti forensics.

Textbook 1: Chapter 7

Module-5

Introduction to Security Policies and Cyber Laws: Need for An Information Security Policy, Information Security Standards – Iso, Introducing Various Security Policies and Their Review Process, Introduction to Indian Cyber Law, Objective and Scope of the it Act, 2000, Intellectual Property Issues, Overview of Intellectual - Property - Related Legislation in India, Patent, Copyright, Law Related to Semiconductor Layout and Design, Software License.

Textbook 2: Chapter 4

Course outcomes:

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

- Define cyber security, cyber law and their roles
- Demonstrate cyber security cybercrime and forensics.
- Infer legal issues in cybercrime,
- Demonstrate tools and methods used in cybercrime and security.
- Illustrate evidence collection and legal challenges

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Cyber Security: Understanding Cyber Crimes, Computer Forensics And Legal Perspectives	Nina Godbole and Sunit Belapure	Wiley India Pvt Ltd	2013
2	Introduction to information security and cyber laws	Surya Prakash Tripathi, Ritendra Goyal, PraveenKumar Shukla	Dreamtech Press	2015

Reference Books

1	Cybersecurity: Managing Systems, Conducting Testing, and Investigating Intrusions	Thomas J. Mowbray	Wiley	2013
2	Cyber Security Essentials	James Graham, Ryan Olson, and Rick Howard	Auerbach Publications	2010

M.TECH IN DATA SCIENCE (SDS) **Choice Based Credit System (CBCS) and Outcome Based Education (OBE)** **SEMESTER –III**			
CLOUD COMPUTING FOR DATA ANALYSIS			
Course Code	**20SDS323,** 20SCS243, 20LNI15, 20SCE14, 20SIT22, 20SSE251, 20SCN31, 20SIS12	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03
Module-1			
Introduction, Cloud Infrastructure: Cloud computing, Cloud computing delivery models and services, Ethical issues, Cloud vulnerabilities, Cloud computing at Amazon, Cloud computing the Google perspective, Microsoft Windows Azure and online services, Open-source software platforms for private clouds, Cloud storage diversity and vendor lock-in, Energy use and ecological impact, Service level agreements, User experience and software licensing. Exercises and problems. **Text Book 1: Chapter 1: 1.3 – 1.6, Chapter 3: 3.1 – 3.5, 3.7 – 3.8, 3.10 – 3.11, 3.14**			
Module 2			
Cloud Computing: Application Paradigms.: Challenges of cloud computing, Architectural styles of cloud computing, Workflows: Coordination of multiple activities, Coordination based on a state machine model: The Zookeeper, The Map Reduce programming model, A case study: The Gre The Web application, Cloud for science and engineering, High-performance computing on a cloud, Cloud computing for Biology research, Social computing, digital content and cloud computing. **Text Book 1: Chapter 4**			
Module 3			
Cloud Resource Virtualization: Virtualization, Layering and virtualization, Virtual machine monitors, Virtual Machines, Performance and Security Isolation, Full virtualization and paravirtualization, Hardware support for virtualization, Case Study: Xen a VMM based paravirtualization, Optimization of network virtualization, vBlades, Performance comparison of virtual machines, The dark side of virtualization, Exercises and problems. **Text Book 1: Chapter 5.1 – 5.12, 5.16**			

Module 4

Cloud Resource Management and Scheduling: Policies and mechanisms for resource management, Application of control theory to task scheduling on a cloud, Stability of a two-level resource allocation architecture, Feedback control based on dynamic thresholds, Coordination of specialized autonomic performance managers, A utility-based model for cloud-based Web services, Resourcing bundling: Combinatorial auctions for cloud resources, Scheduling algorithms for computing clouds, Fair queuing, Start-time fair queuing, Borrowed virtual time, Cloud scheduling subject to deadlines, Scheduling MapReduce applications subject to deadlines, Resource management and dynamic scaling, Exercises and problems.

Text Book 1: Chapter 6

Module 5

Cloud Security, Cloud Application Development: Cloud security risks, Security: The top concern for cloud users, Privacy and privacy impact assessment, Trust, Operating system security, Virtual machine Security, Security of virtualization, Security risks posed by shared images, Security risks posed by a management OS, A trusted virtual machine monitor, Amazon web services: EC2 instances, Connecting clients to cloud instances through firewalls, Security rules for application and transport layer protocols in EC2, How to launch an EC2 Linux instance and connect to it, How to use S3 in java, Cloud-based simulation of a distributed trust algorithm, A trust management service, A cloud service for adaptive data streaming, Cloud based optimal FPGA synthesis. Exercises and problems.

Text Book 1: Chapter 9.1 –9.9, 9.11, and 11.1 – 11.5, 11.10 – 11.14

Course outcomes:

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

- Compare the strengths and limitations of cloud computing
- Identify the architecture, infrastructure and delivery models of cloud computing
- Apply suitable virtualization concept.
- Choose the appropriate cloud player
- Address the core issues of cloud computing such as security, privacy and interoperability
- Design Cloud Services
- Set a private cloud

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each

module.				
Textbook/ Textbooks				
Sl No	**Title of the book**	**Name of theAuthor/s**	**Publisher Name**	**Edition and year**
1	Cloud Computing Theory and Practice	Dan C Marinescu	Elsevier	2013.
Reference Books				
1	Computing Principles and Paradigms	Rajkumar Buyya, James Broberg, Andrzej Goscinski	Willey	2014
2	Cloud Computing Implementation, Management and Security	John W Rittinghouse, James F Ransome	CRC Press	2013

M.TECH IN DATA SCIENCE (SDS) **Choice Based Credit System (CBCS) and Outcome Based Education (OBE)** **SEMESTER –III**			
DATA STREAM MINING			
Course Code	**20SDS324**	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03
Module-1			
Introduction to Data Streams: Data stream models, basic streaming methods, data synopsis, sampling, histograms, Wavelets, Discrete Fourier Transform **Text Book 1: Chapter 2.1 – 2.2**			
Module 2			
Clustering from Data Streams: Basic concepts, Leader Algorithm, partitioning clustering, hierarchical clustering, grid clustering **Text Book 1: Chapter 6.1 – 6.2**			
Module 3			
Frequent Pattern Mining from Data Streams: Search space, landmark windows, mining recent frequent item sets, sequence pattern mining, reservoir sampling for sequential pattern mining. **Text Book 1: Chapter 7**			
Module 4			
Decision Trees from Data Streams: Introduction, The very fast decision tree Algorithm, Extensions to the Basic			

Algorithm. Novelty Detection in Data Streams: introduction, Learning and Novelty, Novelty detection as a one class classification problem.

Text Book 1: Chapter 8.1 – 8.3, 9.1 – 9.3

Module 5

Time Series Data Streams: Introduction to Time Series analysis, Time Series prediction, Similarity between time series, Symbolic Approximation-SAX. Ubiquitous Data Mining: Intoduction to Ubiquitous Data Mining, Distributed Data Stream Monitoring, Distributed Clustering.

Text Book 1: Chapter 11.1 – 11.4, 12.1 – 12.3

Course outcomes:

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

- Understand the concept of data stream models
- Apply the basic clustering techniques for data streams.
- Understand the difference between FPM in data streams and traditional DM techniques
- Identify the classification techniques for data streams
- Understand the change detection concept in data stream mining

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.

The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of theAuthor/s	Publisher Name	Edition and Year
1	Knowledge Discovery from Data Streams	Joao Gama	CRC	1st Edition 2010
Reference Books				
1	Data Streams: Models and Algorithms	Charu C Aggarwal	Springer	2007
2	Introduction to Data Mining	Pang-Ning Tan, Michael Steinbach, and Vipin Kumar	Pearson.	2011
3	Stream Data Mining: Algorithms and Their	Leszek Rutkowski, Maciej Jaworski, Piotr Duda	Springer	1st Edition 2020

	Probabilistic Properties			

M.TECH IN DATA SCIENCE (SDS) Choice Based Credit System (CBCS) and Outcome Based Education (OBE) SEMESTER – III			
BLOCK CHAIN TECHNOLOGY			
Course Code	**20SDS331**, 20SCR244, 20SCN15, 20SCS23, 20SAM254	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03
Module-1			
Blockchain 101: Distributed systems, History of Blockchain, Introduction to Blockchain, Types of Blockchain, CAP theorem and Blockchain, Benefits and limitations of Blockchain. **Text Book 1: Chapter 1**			
Module-2			
Decentralization and Cryptography: Decentralization using Blockchain, Methods of decentralization, Routes to decentralization, Decentralized organizations. Cryptography and Technical Foundations: Cryptographic primitives, Asymmetric cryptography, Public and private keys. **Text Book 1: Chapter 2.1 – 2.2, 2.5, 3.1 – 3.3**			
Module-3			
Bitcoin and Alternative Coins A: Bitcoin, Transactions, Blockchain, Bitcoin payments B: Alternative Coins, Theoretical foundations, Bitcoin limitations, Namecoin, Litecoin, Primecoin, Zcash. **Text Book 1: Chapter 4, 5**			
Module-4			
Smart Contracts and Ethereum 101: Smart Contracts: Definition, Ricardian contracts. Ethereum 101: Introduction, Ethereum Blockchain, Elements of the Ethereum Blockchain, Precompiled contracts. **Text Book 1: Chapter 6, 7.1 – 7.3**			
Module-5			
Alternative Blockchains: Blockchains Blockchain-Outside of Currencies: Internet of Things, Government, Health, Finance, Media. **Text Book 1: Chapter 10, 11**			

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

- Understand the types, benefits and limitation of Blockchain.
- Explore the Blockchain decentralization and cryptography concepts.
- Enumerate the Bitcoin features and its alternative options.
- Describe and deploy the smart contracts
- Summarize the Blockchain features outside of currencies.

Question paper pattern: The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl. No.	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year
1	Mastering Blockchain Distributed ledgers, decentralization and smart contracts explained.	Imran Bashir	Packet Publishing Ltd	Second Edition, 2017

Reference Books

Sl. No.	Title of the book	Name of theAuthor/s	Publisher Name	Edition and year
1	Bitcoin and Cryptocurrency Technologies	Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, Steven Goldfeder	Princeton University	2016
2	Blockchain Basics: A Non-Technical Introduction in 25 Steps	Daniel Drescher	Apress	1st Edition, 2017
3	Mastering Bitcoin: Unlocking Digital Cryptocurrencies	Andreas M. Antonopoulos	O'Reilly Media,	1st Edition, 2014

M.TECH IN DATA SCIENCE (SDS) Choice Based Credit System (CBCS) and Outcome Based Education (OBE) SEMESTER - III			
SUPPLY CHAIN MANAGEMENT			
Course Code	**20SDS332,** 20SSE324, 20SIT321	CIE Marks	40
Teaching Hours/Week (L:T:P)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03
Module-1			
Introduction to Supply Chain Management: Supply chain – objectives – importance – decision phases – process view – competitive and supply chain strategies – achieving strategic fit – supply chain drivers – obstacles – framework – facilities – inventory – transportation – information – sourcing – pricing. **Text Book 1: Chapter 1, 2, 3**			
Module -2			
Designing the supply chain network: Designing the distribution network – role of distribution – factors influencing distribution – design options – e-business and its impact – distribution networks in practice – network design in the supply chain – role of network – factors affecting the network design decisions – modeling for supply chain. **Text Book 1: Chapter 4, 5**			
Module – 3			
Designing and Planning Transportation Networks.: Role of transportation - modes and their performance - transportation infrastructure and policies - design options and their trade-offs - Tailored transportation. **Text Book 1: Chapter 14**			
Module-4			
Sourcing and Pricing: Sourcing – In-house or Outsource – 3rd and 4th PLs – supplier scoring and assessment, selection – design collaboration – procurement process – sourcing planning and analysis. Pricing and revenue management for multiple customers, perishable products, seasonal demand, bulk andspot contracts. **Text Book 1: Chapter 15, 16**			
Module-5			
Information Technology in the supply chain: IT Framework – customer relationship management – internal supply chain management – supplier relationship management –transaction management – futureof IT. **Text Book 1: Chapter 17**			
Course outcomes: At the end of the course the student will be able to:			

- Discuss SCM Models,
- Formulate of QRM, CPFR.
- Implement various Inventory Models and third-party logistics.

Question paper pattern:

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl.No.	Title of the book	Name of theauthor/s	Publisher Name	Edition and year
1	Supply Chain Management – Strategy, Planning and Operation	Sunil Chopra and Peter Meindl	Pearson	6th Edition, 2015

Reference Books

1	Supply Chain Logistics Management	Donald J Bowersox, Dand J Closs, M Bixby Coluper	McGraw Hill	5th Edition,2020
2	Principles of Supply Chain Management A Balanced Approach	G. Leong, K C Tan, Joel Wisner	South-Western College	2015
3	Designing and Managing the Supply Chain - *Concepts, Strategies, and Case Studies*	David Simchi-Levi, Philip Kaminsky, Edith Simchi-Levi, Ravi Shankar	McGraw Hill	2008
4	The management of Business Logistics – A supply Chain Perspective	Coyle, Bardi, Longley	South-Western College	2002
5	Supply Chain Management	Janat Shah	Pearson	2009

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER - III

SPATIAL DATA ANALYSIS

Course Code	**20SDS333**	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03

Module-1

The context for spatial data analysis: Spatial data analysis - scientific and policy context, Spatial data analysis in science, Place and space in specific areas of scientific explanation. The nature of spatial data: The spatial data matrix - conceptualization and representation issues, The spatial data matrix: its form, The spatial data matrix: its quality, Quantifying spatial dependence.

Text Book 1: Chapter 1.1 – 1.2, 2.1 – 2.4

Module-2

Spatial data: obtaining data and quality issues: Obtaining spatial data through sampling - Sources of spatial data, Spatial sampling. Data quality: implications for spatial data analysis - Errors in data and spatial data analysis, Data resolution and spatial data analysis, Data consistency and spatial data analysis, Data completeness and spatial data analysis.

Text Book 1: Chapter: 3.1 – 3.2, Chapter 4

Module-3

The exploratory analysis of spatial data: Exploratory spatial data analysis: conceptual models - EDA and ESDA, Conceptual models of spatial variation. Exploratory spatial data analysis: visualization methods - Data visualization and exploratory data analysis, Visualizing spatial data. Exploratory spatial data analysis: numerical methods - Smoothing methods, The exploratory identification of global map properties: overall clustering.

Text Book 1: Chapter 5, 6, 7.1, 7.2

Module-4

Hypothesis testing and spatial autocorrelation: Hypothesis testing in the presence of spatial dependence - Spatial autocorrelation and testing the mean of a spatial data set, Spatial autocorrelation and tests of bivariate association.

Modelling spatial data: Models for the statistical analysis of spatial data - Descriptive models, Explanatory models

Text Book 1: Chapter 8, 9

Module-5

Modelling spatial data: Statistical modelling of spatial variation: descriptive modelling - Models for representing spatial variation, Some general problems in modelling spatial variation, Hierarchical Bayesian models. Statistical modelling of spatial variation: explanatory modelling - Methodologies for spatial data modelling, Some applications of linear modelling of spatial data.

Text Book 1: Chapter 10, 11

Course outcomes:

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

- Analyze techniques and models for spatial data.
- Apply analytics for implications of spatial data.
- Analyze the conceptual models and visualization methods for spatial data.

Question paper pattern:

- The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.
- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks

Sl No	Title of the book	Name of the Author/s	Publisher Name	Edition and year
1	Spatial Data Analysis: Theory and Practice	Robert Haining	Cambridge University Press	2003
Reference Books				
1	Distance and space. A Geographical Perspective.	Anthony C Gatrell.	Oxford. University Press	1983
2	Models of spatial Processes	Arthur Getis, Barry Boots	Cambridge University Press,	2008

M.TECH IN DATA SCIENCE (SDS)
Choice Based Credit System (CBCS) and Outcome Based Education (OBE)
SEMESTER –III

HEALTH CARE DATA ANALYTICS

Course Code	**20SDS334**	CIE Marks	40
Teaching Hours/Week (L:P:S)	4:0:0	SEE Marks	60
Credits	03	Exam Hours	03

Module-1

An Introduction to Healthcare Data Analytics, Electronic Health Records-A survey: Components of HER, Coding Systems, Benefits of HER, Barrier to Adopting HER Challenges, Phenotyping Algorithms

Chapter 1.1 – 1.6, 2.3 – 2.6, 2.8

Module-2

Biomedical Image Analysis, Mining of Sensor Data in Healthcare, Biomedical Signal Analysis.

Chapter 3, 4, 5

Module-3

Natural Language Processing and Data Mining for Clinical Text, Mining the Biomedical.

Chapter 7, 8

Module-4

Advanced Data Analytics for Healthcare: Review of Clinical Prediction Models, Temporal Data Mining for Healthcare Data, Visual Analytics for Healthcare, Privacy, Preserving Data Publishing Methods in Healthcare.

Chapter 10, 11, 12, 15

Module-5

Applications and Practical Systems for Healthcare: Data Analytics for Pervasive Health- Fraud Detection in Healthcare- Data Analytics for Pharmaceutical Discoveries- Clinical Decision Support Systems- Computer-Assisted Medical Image Analysis Systems.

Chapter 16, 17, 18, 19, 20

Course outcomes:

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

- Analyze health care data using appropriate analytical techniques.
- Apply analytics for decision making in healthcare services.
- Apply data mining to integrate health data from multiple sources and develop efficient clinical decision support systems.

Question paper pattern: The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 60.

- The question paper will have ten full questions carrying equal marks.
- Each full question is for 20 marks.
- There will be two full questions (with a maximum of four sub questions) from each module.
- Each full question will have sub question covering all the topics under a module.
- The students will have to answer five full questions, selecting one full question from each module.

Textbook/ Textbooks				
1	Healthcare data analytics	Chandan K. Reddy and Charu C Aggarwal,	Taylor & Francis	1st Edition, 2015
Reference Books				
1	Healthcare Analytics: From Data to Knowledge to Healthcare Improvement	Hui Yang and Eva K. Lee	Wiley	2016

PROJECT WORK PHASE – 1			
Course Code	20SDS34	CIE Marks	100
Number of contact Hours/Week	2	SEE Marks	--
Credits	02	Exam Hours	--

Course objectives:

- Support independent learning.
- Guide to select and utilize adequate information from varied resources maintaining ethics.
- Guide to organize the work in the appropriate manner and present information (acknowledging thesources) clearly.
- Develop interactive, communication, organisation, time management, and presentation skills.
- Impart flexibility and adaptability.
- Inspire independent and team working.
- Expand intellectual capacity, credibility, judgement, intuition.
- Adhere to punctuality, setting and meeting deadlines.
- Instil responsibilities to oneself and others.
- Train students to present the topic of project work in a seminar without any fear, face audience confidently, enhance communication skill, involve in group discussion to present and exchange ideas.

Project Phase-1 Students in consultation with the guide/s shall carry out literature survey/ visit industries to finalize the topic of the Project. Subsequently, the students shall collect the material required for the selected project, prepare synopsis and narrate the methodology to carry out the project work.

Seminar: Each student, under the guidance of a Faculty, is required to

- Present the seminar on the selected project orally and/or through power point slides.
- Answer the queries and involve in debate/discussion.
- Submit two copies of the typed report with a list of references.

The participants shall take part in discussion to foster friendly and stimulating environment in which the students are motivated to reach high standards and become self-confident.

Course outcomes:

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

- Demonstrate a sound technical knowledge of their selected project topic.
- Undertake problem identification, formulation, and solution.
- Design engineering solutions to complex problems utilising a systems approach.
- Communicate with engineers and the community at large in written an oral forms.
- Demonstrate the knowledge, skills and attitudes of a professional engineer.

Continuous Internal Evaluation

CIE marks for the project report (50 marks), seminar (30 marks) and question and answer (20 marks) shall be awarded (based on the quality of report and presentation skill, participation in the question and answer session by the student) by the committee constituted for the purpose by the Head of the Department. The committee shall consist of three faculty from the department with the senior most acting as the Chairperson.

MINI PROJECT			
Course Code	20SDS35	CIE Marks	40
Number of contact Hours/Week	2	SEE Marks	60
Credits	02	Exam Hours/Batch	03

Course objectives:

- To support independent learning and innovative attitude.
- To guide to select and utilize adequate information from varied resources upholding ethics.
- To guide to organize the work in the appropriate manner and present information (acknowledging thesources) clearly.
- To develop interactive, communication, organisation, time management, and presentation skills.
- To impart flexibility and adaptability.
- To inspire independent and team working.
- To expand intellectual capacity, credibility, judgement, intuition.
- To adhere to punctuality, setting and meeting deadlines.
- To instil responsibilities to oneself and others.
- To train students to present the topic of project work in a seminar without any fear, face audienceconfidently, enhance communication skill, involve in group discussion to present and exchange ideas.

Mini-Project: Each student of the project batch shall involve in carrying out the project work jointly in constant consultation with internal guide, co-guide, and external guide and prepare the project report as per the norms avoiding plagiarism.

Course outcomes:

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

- Present the mini-project and be able to defend it.
- Make links across different areas of knowledge and to generate, develop and evaluate ideas andinformation so as to apply these skills to the project task.
- Habituated to critical thinking and use problem solving skills.
- Communicate effectively and to present ideas clearly and coherently in both the written and oralforms.
- Work in a team to achieve common goal.

- Learn on their own, reflect on their learning and take appropriate actions to improve it.

CIE procedure for Mini - Project:

The CIE marks awarded for Mini - Project, shall be based on the evaluation of Mini - Project Report, Project Presentation skill and Question and Answer session in the ratio 50:25:25.The marks awarded for Mini - Projectreport shall be the same for all the batch mates.

Semester End Examination

SEE marks for the mini-project shall be awarded based on the evaluation of Mini-Project Report, Presentation skill and Question and Answer session in the ratio 50:25:25 by the examiners appointed by the University.

INTERNSHIP / PROFESSIONAL PRACTICE			
Course Code	20SDSI36	CIE Marks	40
Number of contact Hours/Week	2	SEE Marks	60
Credits	06	Exam Hours	03

Course objectives:

Internship/Professional practice provide students the opportunity of hands-on experience that include personal training, time and stress management, interactive skills, presentations, budgeting, marketing, liability and risk management, paperwork, equipment ordering, maintenance, responding to emergencies etc. The objective are further,

To put theory into practice.

To expand thinking and broaden the knowledge and skills acquired through course work in the field.
To relate to, interact with, and learn from current professionals in the field.

To gain a greater understanding of the duties and responsibilities of a professional.
To understand and adhere to professional standards in the field.

To gain insight to professional communication including meetings, memos, reading, writing, public speaking, research, client interaction, input of ideas, and confidentiality.

To identify personal strengths and weaknesses.

To develop the initiative and motivation to be a self-starter and work independently.

Internship/Professional practice: Students under the guidance of internal guide/s and external guide shall take part in all the activities regularly to acquire as much knowledge as possible without causing any inconvenience at the place of internship.

Seminar: Each student, is required to

- Present the seminar on the internship orally and/or through power point slides.
- Answer the queries and involve in debate/discussion.
- Submit the report duly certified by the external guide.
- The participants shall take part in discussion to foster friendly and stimulating environment in which

the students are motivated to reach high standards and become self-confident.
Course outcomes: At the end of the course the student will be able to: • Gain practical experience within industry in which the internship is done. • Acquire knowledge of the industry in which the internship is done. • Apply knowledge and skills learned to classroom work. • Develop a greater understanding about career options while more clearly defining personal career goals. • Experience the activities and functions of professionals. • Develop and refine oral and written communication skills. • Identify areas for future knowledge and skill development. • Expand intellectual capacity, credibility, judgment, intuition. • Acquire the knowledge of administration, marketing, finance and economics.
Continuous Internal Evaluation CIE marks for the Internship/Professional practice report (20 marks), seminar (10 marks) and question and answer session (10 marks) shall be awarded (based on the quality of report and presentation skill, participation in the question and answer session by the student) by the committee constituted for the purpose by the Head of the Department. The committee shall consist of three faculty from the department with the senior most acting as the Chairperson.
Semester End Examination SEE marks for the internship report (30 marks), seminar (20 marks) and question and answer session (10 marks) shall be awarded (based on the quality of report and presentation skill, participation in the question and answer session) by the examiners appointed by the University.

PROJECT WORK PHASE -2			
Course Code	20SCS41	CIE Marks	40
Number of contact Hours/Week	4	SEE Marks	60
Credits	20	Exam Hours	03

Course objectives:

- To support independent learning.
- To guide to select and utilize adequate information from varied resources maintaining ethics.
- To guide to organize the work in the appropriate manner and present information (acknowledging thesources) clearly.
- To develop interactive, communication, organisation, time management, and presentation skills.
- To impart flexibility and adaptability.
- To inspire independent and team working.
- To expand intellectual capacity, credibility, judgement, intuition.
- To adhere to punctuality, setting and meeting deadlines.
- To instil responsibilities to oneself and others.
- To train students to present the topic of project work in a seminar without any fear, face audienceconfidently, enhance communication skill, involve in group discussion to present and exchange ideas.

Project Work Phase - II: Each student of the project batch shall involve in carrying out the project work jointly in constant consultation with internal guide, co-guide, and external guide and prepare the project report as per the norms avoiding plagiarism.

Course outcomes:

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

- Present the project and be able to defend it.
- Make links across different areas of knowledge and to generate, develop and evaluate ideas andinformation so as to apply these skills to the project task.
- Habituated to critical thinking and use problem solving skills
- Communicate effectively and to present ideas clearly and coherently in both the written and oralforms.
- Work in a team to achieve common goal.
- Learn on their own, reflect on their learning and take appropriate actions to improve it.

Continuous Internal Evaluation:

Project Report: 20 marks. The basis for awarding the marks shall be the involvement of the student in the project and in the preparation of project report. To be awarded by the internal guide in consultation with external guide if any.

Project Presentation: 10 marks.

The Project Presentation marks of the Project Work Phase -II shall be awarded by the committee constituted for the purpose by the Head of the Department. The committee shall consist of three faculty from the department with the senior most acting as the Chairperson.

Question and Answer: 10 marks.

The student shall be evaluated based on the ability in the Question and Answer session for 10 marks.

Semester End Examination

SEE marks for the project report (30 marks), seminar (20 marks) and question and answer session (10 marks) shall be awarded (based on the quality of report and presentation skill, participation in the question and answer session) by the examiners appointed by the University.