The Curriculum Book

Master of Science

in

Computer Science

(Artificial Intelligence and Data Science)

2 YEAR-PROGRAMME

Choice Based Credit System and NEP-2020
w. e. f. July 2023

DEPARTMENT OF DATA SCIENCE

GURU JAMBHESHWAR UNIVERSITY OF SCIENCE & TECHNOLOGY

HISAR-125001 , HARYANA

SEMESTER I

Sr. No.	Course Types	Course Codes	Nomenclature of the Courses	Hours per week			Credits	Internal	External
				L	T	P			
1.	Core Course (CC)	MDS-CC-101T	Data Structure & Algorithms	3	0	0	3	30	70
2.	Core Course (CC)	MDS-CC-102T	Database Management System	3	0	0	3	30	70
3.	Core Course (CC)	MDS-CC-103T	Fundamentals of Data Science	3	0	0	3	30	70
4.	Core Course (CC)	MDS-CC-104T	Probability and Statistics	3	0	0	3	30	70
5.	Skill Enhancement Course (SEC)	MDS-SEC-101T	Python and R Programming	3	0	0	3	30	70
6.	Core Course (CC)	MDS-CC-101P	Data Structure & Algorithms Lab.	0	0	4	2	50	50
7.	Core Course (CC)	MDS-CC-102P	Database Management System Lab.	0	0	4	2	50	50
8.	Skill Enhancement Course (SEC)	MDS-SEC-101P	Python and R Programming Lab.	0	0	4	2	50	50
Total				15	0	12	21	300	500

SEMESTER II

Sr. No.	Course Types	Course Codes	Nomenclature of the Courses	Hours per week			Credits	Internal	External
				L	T	P			
1.	Core Course (CC)	MDS-CC-201T	Data Analytics	3	0	0	3	30	70
2.	Core Course (CC)	MDS-CC-202T	Data Mining	3	0	0	3	30	70
3.	Core Courses(CC)	MDS-CC-203T	Artificial Intelligence	3	0	0	3	30	70
4.	Core Courses(CC)	MDS-CC-204T	Information Retrieval Systems	3	0	0	3	30	70
5.	Skill Enhancement Course (SEC)	MDS-SEC-201T	Python Tools for Data Science	3	0	0	3	30	70
6.	Core Course (CC)	MDS-CC-201P	Data Analytics Lab.	0	0	4	2	50	50
7.	Core Course (CC)	MDS-CC-202P	Data Mining Lab.	0	0	4	2	50	50
8.	Skill Enhancement Course (SEC)	MDS-SEC-201P	Python Tools for Data Science Lab.	0	0	4	2	50	50
Total				15	0	12	21	300	500

SEMESTER III

Sr. No.	Course Types	Course Codes	Nomenclature of the Courses	Hours per week			Credits	Internal	External
				L	T	P			
1.	Core Course (CC)	MDS-CC-301T	Big Data Analytics	3	0	0	3	30	70
2.	Core Course (CC)	MDS-CC-302T	Machine Learning	3	0	0	3	30	70
3.	Core Courses (CC)	MDS-CC-303T	Data Communication and Networking	3	0	0	3	30	70
4.	Core Courses (CC)	MDS-CC-304T	Operating System	3	0	0	3	30	70
5.	Skill Enhancement Course (SEC)	MDS-SEC-301T	Optional-I (i-iv)	3	0	0	3	30	70
6.	Core Course (CC)	MDS-CC-301P	Big Data Analytics Lab.	0	0	4	2	50	50
7.	Core Course (CC)	MDS-CC-302P	Machine Learning Lab.	0	0	4	2	50	50
8.	Project Work	MDS-PW -301	Project-1	0	0	4	2	100	---
Total				15	0	12	21	350	450

List of Skill Enhancement Course (SEC))- MDS- SEC -301T-Optional -1 (i-iv)

MDS- SEC -301T-(i) Theory of Computations
MDS- SEC -301T-(ii) Cyber Security

MDS- SEC -301T-(iii) Soft Computing

MDS- SEC -301T-(iv) Any MOOC Course with the permission of chairperson from the list approved by department.

SEMESTER IV

Sr. No.	Course Types	Course Codes	Nomenclature of the Courses	Hours per week			Credits	Internal	External
				L	T	P			
1.	Core Course (CC)	MDS-CC-401T	Deep Learning	3	0	0	3	30	70
2.	Core Course (CC)	MDS-CC-402T	Evolutionary and Swarm Intelligence for Optimization	4	0	0	4	30	70
3.	Skill Enhancement Course (SEC))	MDS- SEC -401T	Optional-I (i-iv)	3	0	0	3	30	70
4.	Skill Enhancement Course (SEC))	MDS- SEC -401T	Optional-II (i-iv)	3	0	0	3	30	70
6.	Core Course (CC)	MDS-CC-301P	Deep Learning Lab.	0	0	2	1	50	50
8.	Project Work	MDS-PW-401	Project-II	0	0	12	6	30	70
Total				13	0	14	20	200	400

List of Skill Enhancement Course (SEC))- MDS- SEC -401T-Optional -1 (i-iv)

MDS- SEC -401T-(i) Computer System Architecture and Organization

MDS- SEC -401T-(ii) : Blockchain Technology

MDS- SEC -401T-(iii) Reinforcement Learning

MDS- SEC -401T-(iv) Any MOOC Course with the permission of chairperson from the list approved by department.

List of Skill Enhancement Course (SEC))- MDS- SEC -402T-Optional -2 (i-iv)

MDS- SEC -402T-(i) Pattern Recognition

MDS- SEC -402T-(ii) Cloud Computing

MDS- SEC -402T-(iii) Internet of Things

MDS- SEC -402T-(iv) Any MOOC Course with the permission of chairperson from the list approved by department.

Data Structures and Algorithms

General Course Information

Course Code: MDS-CC-101T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Programming in C

About the Course:

Data Structure and Algorithms is a core and an essential course for every graduate in Computer Science and Engineering. This course introduces data structures like arrays, linked lists, trees and graphs etc. and various operations to be implemented on these data structures for solving real world problems. It includes various sorting and searching algorithms as well. Further, it incorporates complexity analysis of algorithms implemented on various data structures.

Course Content

Unit I

Introduction to data structures and their types, Abstract data types, Linear lists: Arrays and linked lists: memory representations, implementing operations like traversing, searching, inserting and deleting etc. Applications of arrays and linked lists. Representing sets and polynomials using linked lists.

Unit II

Stack and Queue: Static and linked implementations, Operations and Applications. Circular queues, Tress, Binary trees and related terminology, Tree traversals (Recursive), Threaded Binary Trees, Binary Search Trees implementation and operations, Priority queues.

Unit III

Height Balanced or AVL trees and B trees. Graph definitions and related terminology, memory representations and related operations (traversal, insertion, deletion, search), Path Matrix, Warshall's Shortest path algorithm Hashing, Hash tables, hash function and collision resolution.

Unit IV

Sequential and binary search, Sorting algorithms: Bubble sort, Selection sort, Insertion sort, Quick sort, Merge sort, Count sort, Heap sort, Comparison of searching and sorting techniques based on their complexity analysis, Time and space complexity of algorithms: Asymptotic analysis, Big O, Omega, Theta notations.

Text and Reference Books:

1. Aho, A. V., Ullman, J. D., and Hopcroft, J. E., *Data Structures and Algorithms*, Addison-Wesley, 1983.
2. LangsamYedidyah, Augenstein J Moshe, Tenenbaum M Aaron, *Data Structures using C and C++*, 3rd edition, PHI, 2009.
3. Cormen, T. H., Leiserson, C. E., Rivest, R. L. and Stein, C., *Introduction to Algorithms*, MIT Press, 2009.
4. Robert L. Kruse, *Data Structure and Program Design in C*, Pearson Education India, 2007.
5. Weiss, M. A., *Data Structures and Algorithm Analysis in C++*, Addison-Wesley, 2007.
6. Sahni, S., *Data Structures, Algorithms, and Applications in C++*, WCB/McGraw-Hill, 2001.

Database Management System

General Course Information

Course Code: MDS-CC-102T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Exam Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Prerequisite: Knowledge of UNIX/ Windows, programming language and data structures

About the Course:

This course includes a detailed coverage of principles of database design and models. Students learn querying a database using SQL, normalization techniques, transaction processing etc.

Course Content

Unit I

Overview: Overview of File Systems and Database Systems, Characteristics of the Data Base Approach, Database users, Advantages and Disadvantages of a DBMS, Responsibility of Database Administrator.

Data Base Systems Concepts and Architecture: DBMS architecture and various views of Data, Data Independence, Database languages, Data Models: Relational Database Model, Hierarchical Data Model, Network Data Model, Schemas and Instances.

Unit II

E-R Model: Entity Types, Attributes & Keys, Relationships, Roles and Structural Constraints, E-R Diagrams, Reduction of an E-R Diagram to Tables. Relational Model and Query Language: Overview of Relational Database, Key Integrity Constraints, Relational Algebra, Relational Calculus, SQL fundamentals, Basic Operators, Missing information and NULL values, Advanced SQL features

Unit III

Relational Database Design: Overview of normalization, Database Anomalies, Candidate and Super Key, Functional Dependencies, Integrity Constraints, Decomposition, Normal forms: First, Second, Third Normal, Boyce Codd, Normal Form, Multi-valued Functional Dependencies and Fourth Normal Form, Join Dependencies and Fifth Normal Form, Denormalization.

Unit IV

Concurrency Control Techniques: Overview of database Transactions, Transaction states, ACID properties of a Transaction, Transaction Recovery, Concurrency Control, Locking Techniques, Time-stamp ordering, Multi-version Techniques, Deadlock, Recovery Techniques in centralized DBMS.

DDBMS Design: Replication and Fragmentation Techniques.

Text and Reference Books:

1. Elmasri, R., and Navathe, S. B., *Fundamentals of Database Systems*, 3rd Edition, Addison Wesley, 2002.
2. Silberschatz, A., Korth, H. F., and Sudarshan, S., Database *System Concepts*, McGraw Hill, 2011.
3. Pannerselvam R., *Database Management Systems*, 2nd Edition, PHI Learning, 2011.
4. Desai, B. C., *An Introduction to Database System*, Galgotia Publication, 2010.
5. Leon, A., and Leon, M., *Database Management Systems*, Ist Edition, Vikas Publishing, 2009.
6. Mata-Toledo, R., Cushman, P., Sahoo, D., *Database Management Systems*, Schaums' Outline series, TMH, 2007.

Fundamentals of Data Science

General Course Information

Course Code: MDS-CC-103T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Knowledge about Fundamentals of Data Base Management System

About the Course:

This course is designed to introduce students to data science and its practice: how it works and how it can produce insights from social, political, business and economic data.

Course Content

Unit I

Definition – Basic Terminology- Data science Venn diagram- Types of Data- Structured versus Unstructured data- Quantitative versus Qualitative data- The Four Levels of Data- Five steps of Data Science- Data Science Process Overview –Data science classification-Data Science Algorithms- Business Intelligence and Data Science- Components of Data Science, Introduction-Prior Knowledge-Data Preparation-Modeling-Applications-Objectives of Data Exploration-Datasets- Descriptive statistics

Unit II

Data Visualization: Introduction- Types of Data visualization- Technologies for visualization - Various visualization techniques - The Five Cs of Data Visualization, Data Science Methodology- Analytics for Data Science- Data Analytics Examples- Data Analytics Life Cycle- Data Discovery- Data preparation- Model Planning- Model Building- Operationalization.

Unit III

Introduction-Feature Selection: Classifying feature selection methods- Anomaly Detection: Introduction- Distance and Density based outlier detection-Local Outlier Factor-Timeseries Forecasting- Decomposition-Smoothing based methods-Regression based methods-Machine Learning methods.

Unit IV

Introduction to Data Science Tools- SAS- APACHE FLINK –BigML- Excel- Tableau- Matplotlib- TensorFlow- Weka- Applications: Hands-on with Solving Data Problems-Introduction-Collecting and Analyzing Twitter Data-Collecting and Analyzing YouTube Data.

Text and Reference Books:

1. Sanjeev J. Wagh, Manisha S. Bhende, Anuradha D. Thakare, *Fundamentals of Data Science*, 1st Edition, 2022
2. Daimi, Kevin, Ed. Hamid R. Arabnia, *Principles of Data Science*, Springer, 2020.
3. Vijay Kotu, Bala Deshpande, *Data Science: Concepts and Practices*, Morgan Kaufmann Publishers, Second edition, 2019
4. D J Patil, Hilary Mason, Mike Loukides, *Ethics and Data Science*, O' Reilly, 1st edition, 2018
5. Sinan Ozdemir, *Principles of Data Science*, Packt Publishing, December 2016.
6. Jure Leskovek, Anand Rajaraman and Jeffrey *Ullman*, *Mining of Massive Datasets* v2.1, Cambridge University Press, 2014.
7. Cielen, Davy, Arno DB Meysman, Mohamed Ali, *Introducing Data Science: Big Data, Machine Learning, and more, using Python Tools*, Manning Publications Co., 2016

Probability and Statistics

General Course Information

Course Code: MDS-CC-104T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basics of Mathematics and Statistics

About the Course:

Probability and statistics both are the most important concepts for Data Science and Machine Learning. Probability is about predicting the likelihood of future events, while statistics involves the analysis of the frequency of past events. This course will give an additional space for widening the applications of the knowledge in Data Science domain to the various fields of real life.

Course Content

Unit I

Raw Data – Graphical Plots and Charts - Frequency Distribution – Histogram and Frequency Polygons - Relative Frequency Distributions – Cumulative Frequency Distributions – Frequency Curves and Their Types - Measures of Central Tendency: Mean, Median, Mode, Trimmed Mean – Measures of Dispersion: Range, Standard Deviation, Quartile Deviation, Mean and Median Absolute Deviation – Moments - Measures of Skewness and Kurtosis – Notion of Linear Correlation and Linear Regression, Concept of Probability – Axioms of Probability - Conditional Probability – Simple Problems - Independent Events - Bayes' Rule (without proof) and Simple Applications.

Unit II

Discrete and Continuous Random Variables, Probability Distributions for Discrete and Continuous Random Variables – Distribution Functions for Discrete and Continuous Random Variables - Joint Distributions - Independent Random Variables - Probability Distributions of Functions of Random Variables – Marginal and Conditional Distributions – Mathematical Expectation, Notions of Binomial, Poisson Distribution and Normal Distributions – Properties – Relationship Between Binomial and Normal Distributions, Poisson and Normal Distributions – Uniform, Exponential, Gamma Distributions, t, Chi-square and F Distributions - Bivariate Normal Distribution – Simulation: Random Number Generation from Exponential, Gamma and Normal Distributions.

Unit III

Population and Sample - Random Samples – Sampling with and without Replacement, Sampling Distributions, Sampling distributions of Mean, Proportion and Difference of Means, Standard Error. Estimation of Parameters, Properties of Estimators: Unbiasedness, Consistency, Efficiency, Sufficiency. Point and Interval Estimates and Their

Reliability, Confidence Interval Estimates of Population Parameters Based on Normal, t and Chi-square Distributions.

Unit IV

Statistical Decisions, Statistical Hypothesis, Tests of Hypothesis and Significance, One-tail and Two-tail Tests. Parametric Tests: Tests Involving Normal, t, Chi-square and F Distributions - Test for Goodness of Fit, Contingency Tables, Tests for Independence of Attributes, One-way and Two-way Analysis of Variance. Non-parametric Tests: Sign Test, Run Test, Wilcoxon Signed Rank Test, Mann-Whitney U test, Kruskal-Wallis Test.

T**ext and Reference Books:**

1. Montgomery, D. C., and Runger, G. C. *Applied Statistics and Probability for Engineers*, Seventh Edition, John Wiley & Sons, 2018.
2. Bruce, P., Bruce, A., and Gedeck, P. *Practical Statistics for Data Scientists*, Second Edition, O'Reilly Media, 2020.
3. Spiegel, M. R., Schiller, J. J., and Alu Srinivasan, R. *Probability and Statistics*, Fourth Edition, Schaum's Outline Series, McGraw Hill Companies, 2013.

Python and R Programming

General Course Information

Course Code: MDS-SEC-101T Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic knowledge on programming concepts and statistics.

About the Course:

R programming Language and Python are both used extensively for Data Sciences. Both are very useful and open-source languages as well. Python supports a very large community to general-purpose in data science. One of the most basic use for data analysis, primarily because of the fantastic ecosystem of data-centric Python packages. R Programming has a rich ecosystem to use in standard machine learning and data mining techniques. It works in statistical analysis of large datasets, and it offers a number of different options for exploring data and It makes it easier to use probability distributions, apply different statistical tests. The main objective of this course is to provide the basic understanding of both programming language to students to solve real-life problems.

Course Content

Unit I

Introduction to Python, Python Identifiers, Keywords and Indentation, Comments, Python Operators and Expressions, Function: print(), input(), eval(), Data Types: int, float, complex, Variables, Mutable vs Immutable variables, Namespaces, Decision Statements: Boolean Type, Boolean Operators, if statement, else statement, Nested Conditionals Statements, Multi-way Decision Statements (elif statement), Loop Control Statements: While loop, range() Function, For Loop, Nested Loops, Infinite Loop, Break Statement, Continue Statement, Pass Statement, Lists: Operations on List: Slicing, Inbuilt Functions for Lists, List Processing: Searching and Sorting, Dictionaries: Need of Dictionary, Operations on Directories: Creation, Addition, Retrieving Values, Deletion; Tuples, operations on Tuples, Inbuilt Functions for Tuples, Sets and operations on sets.

Unit II

Python Functions, Inbuilt functions, Main function, User Defined functions, Defining and Calling Function, Parameter Passing, Actual and Formal Parameters, Default Parameters, Global and Local Variables, Recursion, Lamda Function, Modules, Importing Own Module, Packages.
Python Object Oriented: Overview of OOP, Classes and objects, accessing attributes, Built-In Class Attributes, Methods, Class and Instance Variables, Polymorphism, Overlapping and Overloading of Operators, Class Inheritance: super (), Method Overriding, Exception Handling, Try-except-else clause, Python Standard Exceptions, User-Defined Exceptions

Unit III

Introduction to R, Installation of R, Installation of R-Studio, Types of Variables, Types of Datatypes: Real, Integer, Complex, Character, Strings, Vectors, Arrays, List, Matrices, Factors, Data Frames, Types of Operators: • Arithmetic, Logical, Relational, Membership, Special Operators, If-else Flow Control, Loops in R (While, For, Break, Next), Nested Loops, Switch-Case

Functions in R, Function declaration with parameters, Function declaration without parameters

Unit IV

R Data Interface: Reading CSV files, Reading XML files, JSON files, Scraping data from the Web, SQL with R, Databases with R

R package: ggplot2 and dplyr

Data Visualization of R: Pie Chart, Bar graph, Line Graph, Scatter plot, Stack Plot, Box-Plot

T**ext and Reference Books:**

1. R. Nageswara Rao, *Core Python Programming*, Dreamtech Press; Second edition (1 January 2018)
2. Allen B. Downey, *Think Python: How to Think Like a Computer Scientist*, 2nd edition, Updated for Python 3, Shroff/O'Reilly Publishers, 2016.
3. Jake Vanderplas, *Python Data Science Handbook: Essential Tools for Working with Data*, 1st Edition, O'Reilly Media, 2016.
4. Norman Matloff, *The Art of R Programming: A Tour of Statistical Software Design*, No Starch Press, First Edition, 2011.
5. Hadley Wickham, Garett Grolemund, *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, O'Reily Publications, First Edition, Feb 2017
6. Reema, Thareja, *Python Programming: Using Problem Solving Approach*, Oxford University Press, June 2017
7. Garrett Grolemund, *Hands-on Programming with R: Write your own functions and simulations*, O'Reilly Publisher, 2014.
8. Saroj Dahiya Ratnoo and Himmat Singh Ratnoo, *Essentials of R for Data Analytics*, Wiley (1 January 2021)

Data Structures and Algorithms Lab.

General Course Information

Course Code: MDS-CC-101P Course Credits: 2 Type: Core Course (CC) Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Programming in C language.

About the Course:

This lab course involves implementation of basic and advance data structures and various operations on these data structures. The objective of the lab course is to train the students to solve the problems related to data structures and choose the appropriate data structure for solving computational problem efficiently.

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Database Management System Lab.

General Course Information

Course Code: MDS-CC-102P Course Credits: 2 Type: Core Course (CC) Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Exposure to programming language, MS Access.

About the Course:

This lab. course on DBMS involves a rigorous training on Oracle programming. It provides a strong formal foundation in database concepts, technology and practice to the students to groom them into well-informed database application developers. The objective of the lab course is to develop proficiency in the execution of commands of the database design and query using Oracle.

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Python and R Programming Lab.

General Course Information

Course Code: MDS-SEC-101P Course Credits: 2 Type: Skill Enhancement Course (SEC) Contact Hours: 4 hours/week Mode: Lab practice and assignments.	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Understanding of programming language.

About the Course:

The major objective of Python programming is to make the students solve real word problem including data science problems efficiently using python library. The understanding and knowledge of R programming helps the students to read the data and its manipulation using R.

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Data Analytics

General Course Information

Course Code: MDS-CC-201T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic programming skills, Probability and Statistics

About the Course:

In this course, the learners will be able to develop expertise in R programming for manipulating, exploring, visualizing, applying descriptive and inferential statistics. In addition, they will learn to implement predictive modelling.

Course Content

Unit I

Data analytics preliminaries: Introduction to data analytics, scales of measurements (Data types) and their implementation in R. Working with vectors, matrices and tabular data (data frames), reading and writing tabular data from and to files (text and CSV). Describing data with statistical summaries (mean, median, mode, variance and standard deviation). Discriminating between sample and population, Quantile-Quantile plot. writing user-defined functions in R.

Unit II

Manipulating tabular data: Sorting, filtering cases, selecting variables, deriving new variables, grouping and summarizing data. working with packages (tidyverse) for data manipulations and transformations.

Exploratory data analysis: random and normally distributed variables, skewed normal distribution, z-score, detecting outliers in data, handling missing values.

Visualizing data through various plots and charts: bar charts, histogram, frequency polygon, density plots, scatter plots, box & whisker plots, heat and contour plots, plotting the above graphs in R, plotting with package ggplot2.

Unit III

Predictive modelling: what is predictive modelling, estimating a function, the trade-off between model accuracy and prediction accuracy and model interpretability, regression versus classification, measuring the quality of fit, The bias and variance trade- off.

Simple and multiple linear regression modelling: estimating the coefficients, assessing the accuracy of the coefficient estimates, assessing the accuracy of the model. Building regression models in R.

Unit IV

Classification Modeling: The process of classification, decision tree, Bayesian, k-nearest neighbor, support vector machine classification models and their implementation in R. evaluating a classification model: confusion matrix, accuracy, sensitivity, specificity, f-measure, kappa statistics, ROC and area under curve. accuracy and interpretability of classification models.

Evaluating the accuracy of a classifier: holdout or random sampling methods, cross-validation, bootstrap methods.

Text and Reference Books:

1. W. N. Venables, D. M. Smith and the R core Team, *An introduction to R, Notes on R: A Programming Environment for Data Analysis and Graphics*, version 3.3.2, 2016.
2. Saroj Dahiya Ratnoo and Himmat Singh Ratnoo, *Essentials of R for Data Analytics*, Wiley, 2021.
3. Hadley Wickham and Garrett Grolemund, *R for Data Science Import, Tidy, Transform and model Data*, O'Reilly, 2017.
4. Paul Teeter, *R Cookbook*, O'Reilly, 2011.
5. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, *An Introduction to Statistical Learning with Applications in R*, Springer, 2013.
6. Han, J., Kamber, M, Pei, J., *Data Mining Concepts and Techniques*, Third edition, Morgan Kaufmann, 2012.

Data Mining

General Course Information

Course Code: MDS-CC-202T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Knowledge of database systems, elementary knowledge of statistics and probability.

About the Course:

Today's era is the era of information. Data is growing exponentially day by day. There is a need to process and analyse the data to extract knowledge from it, so that one can use that knowledge for decision making. This course provides introductory concepts of data mining and data warehousing. The course will be taught with a database as well as machine learning perspectives. The objective of the course is to provide a comprehensive understanding of data mining tasks and evaluation of results obtained out of data mining processes.

Course Content

Unit I

Introduction to Data Mining: Kind of data to be mined, Data Mining Functionalities, Technologies used in Data Mining, Applications of data Mining, Major Issues in Data Mining.
Data Warehouse: Introduction, Data Warehouse and Database Systems, Data Warehouse Architecture, Data Warehouse Models, Data Cube and OLAP, Multidimensional data Model, Concept Hierarchies, OLAP operations
Pattern Mining: Mining Frequent Patterns, Associations and Correlations, Frequent Itemset Mining using Apriori Algorithm, Generating Association Rules from Frequent Itemsets, Pattern Growth Approach for Mining Frequent Itemsets, Pattern evaluation Methods

Unit II

Classification: Introduction, Classification using Decision Tree Induction, Bayesian Classification Methods, Rule Based Classification, Model Evaluation and Selection, Techniques to Improve Classification Accuracy.
Introduction to advanced classifiers: k-Nearest Neighbor, Support Vector Machine, Artificial Neural Network.

Unit III

Cluster Analysis: Introduction, overview of Basic Clustering Methods,
Partitioning Methods: k-mean, k-medoids,

Hierarchical Methods: Agglomerative versus Divisive Hierarchical Clustering, Distance Measures in Algorithmic Methods, Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH), Chameleon: Multiphase Hierarchical Clustering Using Dynamic Modeling, Probabilistic Hierarchical Clustering,
Density-based methods: DBSCAN, OPTICS, DENCLUE,
Grid-based Methods: STING, CLIQUE, **Evaluation of Clustering.**

Unit IV

Outlier Detection: Introduction, types of outliers, challenges of outlier detection.
Outlier detection methods: statistical approaches, proximity-based approaches, clustering based approaches, classification-based approaches, Outlier detection in high dimensional data.

Software/Tools to be learnt: WEKA, RapidMiner, XMiner, MATLAB

T**ext and Reference Books:**

1. Jiawei Han, Micheline Kamber and Jian Pei, *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, Third Edition, July 2011.
2. Alex Berson, Stephen J. Smith, *Data Warehousing, Data Mining & OLAP*, Tata McGraw Hill, 2004.
3. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, *Introduction to Data Mining*, Pearson Education, 2014.
4. K. P. Soman, Shyam Diwakar and V. Ajay, *Insight into Data Mining Theory and Practice*, Easter Economy Edition, Prentice Hall of India, 2009.
5. G. K. Gupta, *Introduction to Data Mining with Case Studies*, Prentice Hall of India, 2006.
6. Daniel T. Larose, *Data Mining Methods and Models*, Wiley, 2006.
7. W. H. Inman, *Building the Data Warehouse*, Wiley India, 2005

Artificial Intelligence

General Course Information

Course Code: MDS-CC-203T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic Knowledge of Algorithms and Probability.

About the Course:

Artificial Intelligence is a core and an essential course for every graduate in Computer Science and Engineering. This course introduces the concepts of Artificial Intelligence and challenges inherent in building intelligent systems. It includes the role of knowledge representation in problem solving and how these are used in making intelligent machine. Further it incorporates the concepts of expert system and its applications.

Course Content

Unit I

Introduction to AI: Introduction, Turing Test, AI problems, State Space Search, production system

Problem Solving Using Search: Blind search techniques - Breadth first search, Depth first search. Heuristic search techniques - Generate and test, Hill Climbing, Best first search, A* Algorithm, AO* Algorithm, The Minimax Search Procedure, Adding Alpha-Beta Cut-offs.

Unit II

Knowledge Representation: Introduction, Knowledge Representation- Representation and Mappings, Symbolic Logic - Propositional logic, Predicate logic- Representing simple facts in logic, Representing Instances and ISA Relationship, Computable functions and Predicates, Unification, Resolution.

Representing Knowledge Using Rules: Procedural versus Declarative Knowledge, Logic Programming, Forward versus Backward Reasoning, Matching, Control Knowledge.

Unit III

Reasoning Under Uncertainty: Introduction to Nonmonotonic Reasoning, Probability and Baye's Theorem, Certainty Factors and Rule-based Systems, Bayesian Networks.

Fuzzy logic system: Introduction, Crisp Set, Fuzzy Sets, Fuzzy Membership Functions, Operations on Fuzzy Sets, Fuzzy Relations.

Unit IV

Planning: Introduction, Components of Planning System, Goal Stack Planning, Nonlinear Planning using Constraint Posting, Hierarchical Planning.

Expert System and Applications: Introduction, Architecture, Rule based Expert Systems, Applications of Expert Systems.

T**ext and Reference Books:**

1. Elaine Rich, Kevin Knight and Shivashankar B Nair, *Artificial intelligence*, McGraw Hill Education. 3rd edition, 2009.
2. Stuart Russel and Peter Norvig, *Artificial intelligence: A modern Approach*, Pearson Education, 3rd edition, 2015.
3. Dan W. Patterson, *Introduction to Artificial Intelligence and Expert System*, Pearson Education.1st edition, 2007.
4. Deepak Khemani, *A first course in Artificial Intelligence*, McGraw Hill Education. 3rd edition, 1st edition, 2013.
5. George F. Luger, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, Pearson Education, 5th edition, 2009.

Information Retrieval Systems

General Course Information:

Course Code: MDS-CC-204T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Data Structures, Data Base Management Systems

About the Course:

This course would enable the students to understand the various aspects of an information retrieval system and its evaluation and to be able to design. The main aim of this course is to give students an understanding about data/file structures that are necessary to design, and implement information retrieval (IR) systems, IR principles to locate relevant information large collections of data, different document clustering algorithms, information retrieval systems for web search tasks etc.

Course content

Unit I

Introduction to Information Retrieval Systems: Definition of Information Retrieval System, Objectives of Information Retrieval Systems, Functional Overview, Relationship to Database Management Systems, Digital Libraries and Data Warehouses. Information Retrieval System Capabilities: Search Capabilities, Browse Capabilities, Miscellaneous Capabilities.

Unit II

Cataloging and Indexing: History and Objectives of Indexing, Indexing Process, Automatic Indexing, Information Extraction. Data Structure: Introduction to Data Structure, Stemming Algorithms, Inverted File Structure, N-Gram Data Structures, PAT Data Structure, Signature File Structure, Hypertext and XML Data Structures, Hidden Markov Models.

Unit III

Automatic Indexing: Classes of Automatic Indexing, Statistical Indexing, Natural Language, Concept Indexing, Hypertext Linkages. Document and Term Clustering: Introduction to Clustering, Thesaurus Generation, Item Clustering, Hierarchy of Clusters.

User Search Techniques: Search Statements and Binding, Similarity Measures and Ranking, Relevance Feedback, Selective Dissemination of Information Search, Weighted Searches of Boolean Systems, Searching the INTERNET and Hypertext.

Unit IV

Information Visualization: Introduction to Information Visualization, Cognition and Perception, Information Visualization Technologies.

Text Search Algorithms: Introduction to Text Search Techniques, Software Text Search Algorithms, Hardware Text Search Systems. Multimedia Information Retrieval: Spoken Language Audio Retrieval, Non-Speech Audio Retrieval, Graph Retrieval, Imagery Retrieval, Video Retrieval.

Text and References Books:

1. Kowalski & Maybury, *Information storage and retrieval systems: theory and implementation* (Vol. 8). Springer Science & Business Media, 2002.
2. Frakes & Baeza-Yates (Eds)., Information *retrieval: data structures and algorithms*. Prentice-Hall, Inc., 1992.
3. Korfhage, *Information Retrieval and Storage*, John Wiley & Sons, 1997
4. Baeza-Yates & Ribeiro-Neto (1999), *Modern information retrieval* (Vol. 463), New York: ACM press, 1999.

Python Tools for Data Science

General Course Information

Course Code: MDS-SEC-201T Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Python Programming Language and Database Management System

About the Course:

Python as a programming language has become very popular in recent times. It has been used in data science, IoT, AI, and other technologies, which has added to its popularity. Python is used as a programming language for data science because it contains costly tools from a mathematical or statistical perspective. It is one of the significant reasons why data scientists around the world use Python. This course focus on study of python tools for data scientists. In this course, powerful python tools -NumPy, Pandas and Matplotlib are to be studied. After studying this course, a student is expected to apply python tool for solving data science problems efficiently.

Course content

Unit I

Datasets: Datasets, Data Preprocessing, Preparing Datasets, Missing Data, Anomalies, and Outliers, Imbalanced Classification.

NumPy: Introduction to NumPy, importing NumPy in Python code, NumPy useful features, NumPy Arrays, working with Loops, Appending Elements to Arrays, Multiplying Lists and Arrays, Lists and Exponents, Arrays and Exponents, Math Operations and Arrays, Arrays and Vector Operations, reshape () Method, Mean and Standard Deviation, NumPy and Matplotlib, Linear Regression, Mean Squared Error (MSE) formula, MSE by Successive Approximation.

Unit II

Pandas: Introduction to Pandas, importing pandas in Python code, Pandas useful features, Pandas Data Frames, Pandas Data Frames and Data Cleaning, Boolean Data Frames, Pandas Data Frames and Random Numbers, Reading CSV Files in Pandas, loc() and iloc() Methods, Converting Categorical Data to Numeric Data, Merging and Splitting Columns in Pandas, Combining Pandas Data Frames, Pandas Data Frames and CSV Files, Managing Columns and rows in Data Frames, Handling Missing Data in Pandas, groupby(), apply() and mapapply() in Pandas, Handling Outliers in Pandas, Pandas Data Frames and Simple Statistics, Pandas Method Chaining, Pandas Profiling.

Unit III

Data Cleaning: Data Cleaning, Data Cleaning in SQL, Replace Multiple Values with a Single Value, Handle Mismatched Attribute Values, Convert Strings to Date Values, Working with Variable Column Counts, Truncating Rows in CSV Files, Converting Numeric Date Formats, Converting Alphabetic Date Formats, Data Cleaning on a Kaggle Dataset.

Unit IV

Data Visualization: Data Visualization, Types of Data Visualization, Matplotlib, Diagonal Lines in Matplotlib, Colored Grid in Matplotlib, Plotting Multiple Lines in Matplotlib, Trigonometric Functions in Matplotlib, Plot Best-Fitting Line in Matplotlib, SkLearn, Pandas, and Iris Dataset (Introduction only), Seaborn and its Features.

Text and Reference Books:

1. Campesato, Oswald. *Python Tools for Data Scientists Pocket Primer*. Stylus Publishing, LLC, 2022.
2. VanderPlas, Jake. *Python data science handbook: Essential tools for working with data*. " O'Reilly Media, Inc.", 2016.
3. Cielen, Davy, and Arno Meysman. *Introducing data science: big data, machine learning, and more, using Python tools*. Simon and Schuster, 2016.
4. Müller, Andreas C., and Sarah Guido. *Introduction to machine learning with Python: a guide for data scientists*. " O'Reilly Media, Inc.", 2016.
5. Boschetti, Alberto, and Luca Massaron. *Python data science essentials.* Packt Publishing Ltd, 2015.
6. Grus, Joel. *Data science from scratch: first principles with python*. O'Reilly Media, 2019.
7. Boschetti, Alberto, and Luca Massaron. *Python data science essentials: A practitioner's guide covering essential data science principles, tools, and techniques*. Packt Publishing Ltd, 2018.
8. Scavetta, Rick J., and Boyan Angelov. *Python and R for the Modern Data Scientist*. " O'Reilly Media, Inc.", 2021.

Data Analytics Lab.

General Course Information

Course Code: MDS-CC-201P Course Credits: 2 Type: Core Course (CC). Course Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Basic programming skills.

About the Course:

The objective of this lab is to enable students to apply advanced data analytics tools for manipulating data, applying statistics, regression and classification.

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Data Mining Lab.

General Course Information

Course Code: MDS-CC-202P Course Credits: 2 Type: Core Course (CC). Course Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Basic programming skills.

About the Course:

The course helps the students to learn how to perform data mining tasks using a data mining toolkit (such as open-source WEKA), understand the data sets and data pre-processing, demonstrate the working of algorithms for data mining tasks such as association rule mining, classification, clustering and regression, and exercise the data mining techniques with varied input values for different parameters.

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Python Tools for Data Science Lab.

General Course Information

Course Code: MDS-SEC-201P Course Credits: 2 Type: Skill Enhancement Course (SEC) Course Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Core Python Programming.

About the Course:

In this course, students will learn how to use Python tools -NumPy and Pandas for data exploration, preparation, and analysis.

Note:
The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Big Data Analytics

General Course Information

Course Code: MDS-CC-301T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Statistics, Data Analytics and Data Mining

About the course:

This course aims to equip the student with emerging field of big data analytics. Students achieve this through developing understanding of Big data analytics techniques and principles in typical real world scenarios. The course teaches students to understand as well as apply data analytics to big data projects.

Course Content

UNIT-I

Introduction: Overviews of Big Data, State of the Practice in Analytics, The Data Scientist, Big Data Analytics in Industry Verticals, Data Analytics Lifecycle Challenges of Conventional Systems.

UNIT-II

Mining Data Streams: Stream Data Model and Architecture, Stream Computing, Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream.

UNIT-III

Components of Hadoop - Analyzing Big data with Hadoop Design of HDFS -- MapReduce - Different Phases, Classic, Components - MapReduce Features - Counters, Sorting, Joins - YARN - Components, Workflow, Scheduling.

UNIT-IV

Introduction to NoSQL - Aggregate Data Models - Key Value Data Model -Columnar Model - Document Data Model - Graph Data Model and its Applications Relationships - Schema-Less Databases - Materialized Views

Text and Reference Books:

1. Michael Berthold, David J. Hand, Intelligent Data Analysis, Springer, 2007.

2. A. Rajaraman, J.D. Ullman, Mining of Massive Datasets, Cambridge University Press, 2012.

3. Bill Franks, Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics, John Wiley & sons, 2012.

4. Glenn J. Myatt, Making Sense of Data, John Wiley & Sons, 2007 5. Pete Warden, Big Data Glossary, O'Reilly, 2011.

Machine Learning

General Course Information

Course Code: MDS-CC-302T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basics of Linear Algebra and Statistics, Basics of Probability Theory, Data Structures and Computer Algorithms.

About the Course:

Machine learning is the study of computer algorithms that improve their performance through experience. Machine learning draws its conceptual foundation from the fields like artificial intelligence, probability and statistics, computational complexity, cognitive science, biology and information theory etc. The course introduces some of the key machine learning algorithms and the theory that form the backbone of these algorithms. The examples of such algorithms are classification algorithms for learning patterns from data, neural network algorithms for pattern recognition etc.

Course Content

Unit I

Introduction to Machine Learning, Well -Posed Learning Problems, Designing a Learning System, Perspectives and Issues in Machine Learning, The concept learning task, Concept learning as search, Finding a maximally specific hypothesis, Version spaces and candidate elimination algorithm, Remarks on version spaces and candidate-eliminations, Inductive bias.

Unit II

Introduction to classification and regression models, Introduction to Artificial Neural Network, Neural Network Representation, Perceptron: Representational Power of Perceptrons, Perceptron Training Rule, Gradient Descent Algorithm and delta learning rule.

Unit III

Multilayer Networks and Backpropagation Algorithm, Convergence and Local Minima, Representational Power of Feedforward Networks, Illustrative ANN examples, Advanced topics in neural networks

Unit IV

Reinforcement Learning: Introduction, the Learning Task, Q Learning, Nondeterministic Rewards and Actions, Temporal Difference Learning, Relationship to dynamic programming.

Text and Reference Books:

1. Tom M. Mitchell, Machine Learning, McGraw-Hill, 1997.
2. Ethem Apaydin, Introduction to Machine Learning , 3rd Edition MIT Press
3. Bishop Christopher, Pattern Recognition and Machine Learning, Springer Verlag, 2006.
4. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction, Springer, 2nd edition, 2009..
5. M. Gopal, Applied Machine Learning, McGraw Hill, 2018

Data Communication and Networking

General Course Information

Course Code: MDS-CC-303T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic knowledge of Digital and Analog Communication

About the Course:

This course has been designed with an aim to provide students with an overview of the concepts and fundamentals of data communication and computer networks. The learner is given an opportunity to grasp various algorithms for routing of data, forwarding data and switching the data from hop to hop. Layered Architecture adds value to the subject contents.

Course Content

Unit - I

Data communication: Components. Network: Uses, Topologies, Network Services, OSI and TCP/IP Reference Models; Network categories: LAN, MAN, WAN; Wireless Transmission Media, Switching Techniques: Circuit Switching, Packet Switching, Message Switching, Networking Devices: Hubs, Repeaters, Bridges, Modems, Switches, Routers, and Gateways.

Unit - II

Data Link Layer-design issues, Framing & Error Handling: Framing Protocols, Error detection and correction mechanisms; Flow Control Protocols: Stop-and-wait, Sliding Window protocols: Go-back N and Selective Repeat, Multiple Access Communication: Random Access-ALOHA, Slotted-ALOHA, CSMA, CSMA-CD, LAN Standards & Ethernet.

Unit - III

Network Layer-Design issues, store and forward packet switching connection less and connection-oriented networks, Routing algorithms: optimality principle, shortest path, flooding, Distance Vector Routing, Count to Infinity Problem, Link State Routing, Hierarchical Routing, Internetworking: IPV4 and IPV6, IP Addressing (Classful Addressing, Private IP Addresses, Classless Addressing, Sub-netting).

Unit - IV

Transport Layer: Transport layer Services: Addressing, Multiplexing, Internet Transport Protocols: UDP& TCP,

Congestion control algorithms, admission control. TCP Segmentation & TCP Connection management. Application Layer: Introduction to DNS, HTTP, SMTP, Electronic Mail, WWW

Text and Reference Books:

1. Andrew S Tanenbaum, Computer Networks, 5th Edition, Pearson publications, 2010.
2. Forouzan, Data Communication and networking ,5th Edition, Tata Mc Graw Hill, 2012.
3. William Stalling, Data & Computer Communication 6th edition, LPE Pearson Education, 2013.
4. Todd Lammle, CCNA Study Guide, 6th Edition, 2013.
5. RFCs and Internet Drafts available from Internet Engineering Task Force.

Operating Systems

General Course Information

Course Code: MDS-CC-304T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Programming in C and knowledge of computer fundamentals.

About the Course:

The objective of this course is to help students become familiar with the fundamental concepts of operating systems and provide them with enough understanding of operating system design.

Course Content

Unit –1

Introductory Concepts: Operating systems functions and characteristics, Computer system organization, Computer system architecture, Operating system structure, Virtual machines, Protection & Security, Operating system services and system calls, Types of Operating systems: Batch operating system, Time-sharing OS, Distributed operating system, Real time systems.NOS, Multiprocessor OS, Mobile OS, RTOS, Cloud OS.

Unit - II

Processes: Process in memory, Process states, PCB, Process scheduling, Inter-process communication, CPU scheduling: Levels of Scheduling, Scheduling criteria, Scheduling algorithms, Multithreading models. Thrashing. File Systems: Types of Files and their access methods, File allocation methods, Directory structure.

Unit –III

Distributed OS- types of distributed operating systems, Network topology, Communication protocols. Issues in Distributed operating systems Deadlocks- Deadlock characterization, Methods for handling deadlocks, Deadlock prevention, Deadlock avoidance, Recovery from deadlock.

Unit– IV

Memory: Basic hardware, Address binding, swapping, logical and physical address space, Contiguous memory allocation, Fragmentation, Paging, TLB, Segmentation, Virtual memory Demand paging, Page replacement algorithms. Case Studies: Comparative study of WINDOW, UNIX & LINUX system.

Text and Reference Books:

1. Advanced Concepts in Operating Systems, by Mukesh Singhal, Niranjan G. Shivaratri, TMH.
2. Operating System Concepts, (6th Edition), by Abraham Silberschatz, Peter Baer Galvin, Greg Gagne.
3. Theory and problem of programming with C, Byron C Gottfried, TMH
4. Teach yourself all about computers by Barry Press and Marcia Press, 2000, IDG Books India.
5. Using Computers and Information by Jack B. Rochester, 1996, Que Education & Training

Theory of Computations

General Course Information

Course Code: MDS-SEC-301T-(i) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: The students are expected to have a strong background in the fundamentals of discrete mathematics like in the areas of symbolic logic, set, induction, number theory, summation, series, combinatory, graph, recursion, basic proof techniques.

About the Course:

Computation theory presents the theoretical aspects of computer science, which lay the foundation for students of Computer Science. The course introduces some fundamental concepts in automata theory and formal languages including grammar, finite automaton, regular expression, formal language, pushdown automaton, and Turing machine.

Course Content

Unit – I

Finite Automata and Regular Expressions: Finite State Systems, Basic Definitions Non-Deterministic finite automata (NDFA), Deterministic finite automata (DFA), Equivalence of DFA and NDFA Finite automata, Introduction to Machines: Concept of basic Machine, Moore and mealy Machines, Equivalence of Moore and Mealy machines. Minimization of finite Automata, Minimization Algorithm.

Unit – II

Regular Sets and Regular Grammars: Regular Expressions, Finite Automata and Regular Expressions, Equivalence of finite automata and Regular Expressions, Regular expression conversion and vice versa, Conversion of NFA to DFA Algebraic Method Using Arden's Theorem. The Pumping Lemma for Regular Sets, Applications of the pumping lemma, Closure properties of regular sets.

Unit - III

Grammars: Definition, Context free and Context sensitive grammar, Ambiguity regular grammar, Reduced forms, Removal of useless Symbols and unit production, Chomsky Normal Form (CNF), Griebach Normal Form (GNF). Pushdown Automata: Introduction to Pushdown Machines, Application of Pushdown Machines.

Unit - IV

Turing Machines: Deterministic and Non-Deterministic Turing Machines, Design of T.M, Halting problem of T.M. Chomsky Hierarchies: Chomsky hierarchies of grammars, Relation between languages of classes.

Text and Reference Books:

1. Hopcroaft & O. D. Ullman, R Mothwani, Introduction to automata theory, language & computations, AW, 2001.
2. K. L. P. Mishra & N. Chandrasekaran, Theory of Computer Sc.(Automata, Languages and computation), PHI, 2000.
3. Peter Linz, Introduction to formal Languages & Automata, Narosa, Publication, 2001.
4. Ramond Greenlaw and H. James Hoover, Fundamentals of the Theory of Computation- Principles and Practice, Harcourt India Pvt. Ltd., 1998.
5. H. R. Lewis & C. H. Papaditriou, Elements of theory of Computation, PHC, 1998. 6. John C. Martin, Introduction to Languages and the Theory of Computation, T.M.H., 2003.

Cyber Security

General Course Information

Course Code: MDS-SEC-301T-(ii) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic knowledge of number theory, complexity theory and basic programming skills for security.

About the Course:

About the Course: The increase in techniques for unauthorised access into systems has led to variety of information and cyber-attacks, To mitigate the exploitation of the vulnerabilities leading to these attacks, it's need to adopt robust security architecture into our premises. In this course students are going to studies various security technologies such as cryptography, digital signatures, key management, program security, database security, cyber laws, Wifi security.

Course Content

Unit - 1

Cyber Security Fundamentals: Network and Security Concepts: Firewalls, Virtualization, DNS, Radio- Frequency Identification, Attacker Techniques and Motivations: Tunnelling Techniques, Fraud Techniques, Threat Infrastructure, Exploitation, Malicious code. Defense and Analysis Techniques.

Unit - II

Ethics in Cyber Security: Privacy, Intellectual property in the cyberspace, Professional ethics, Freedom of speech, Fair user and ethical hacking, Trademarks, Internet fraud, Electronic evidence, Forensic technologies, Digital evidence collections. Tools and methods used in cybercrime: Introduction, Password cracking, Keyloggers and spywares, Virus and worms, Phishing and identity theft, Trojan horses and backdoors, Steganography

Unit – III

Cyber crimes and Cyber security: Cyber crime and legal landscape around the world, Cyber laws, The Indian IT Act, Challenges, Digital signatures and Indian IT Act, Amendments to the Indian IT Act, Cyber crime and punishment, Cost of Cyber crimes and IPR Issues, Web threats for organizations, Social computing and associated challenges for organizations.

Unit – IV

Protecting Critical Infrastructures: Critical Infrastructures: Key Assets, Critical Infrastructure Interdependencies, Internet, Social Media and Cyber Attacks on Critical Infrastructures, Cyber Threat Spectrum- Cyberspace Attacks and Weapons, Framework for improving Critical Infrastructure Cyber security.

Text and Reference Books:

1. James Graham, Richard Howard, “Cyber Security Essentials”, CRC Press, Taylor & Francis Group, ISBN: 978-1-4398-5126-5, 2011.
2. Thomas A. Johnson, “Cyber-Security Protecting Critical Infrastructures from Cyber Attack and Cyber Warfare”, CRC Press, ISBN:978-1-4822-3923-2, 2015.
3. Nina Godhole and Sunit Belapure, Cyber Security, Wiley India, 2011.

Soft Computing

General Course Information

Course Code: MDS-SEC-301T-(iii) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic knowledge of Probability Theory, Set Theory and, Data Structure and Computer Algorithms

About the Course:

We need to learn soft computing techniques to make intelligent machines that possess human like abilities to reason, learn and handle the uncertainty and vagueness often inherent in real world problems. Unlike conventional computing, soft computing techniques are tolerant of imprecision, uncertainty and approximations, and provide low cost, robust and tractable solutions to the complex real-world problems where conventional methods fail to do so. This introductory course on soft computing is going to cover Genetic Algorithms, Artificial Neural Networks and Fuzzy Logic

Course Content

Unit - I

Introduction to Soft Computing and related definitions: Defining soft computing, Differentiating the situations for application of hard and soft computing; Working of a simple Genetic Algorithm: Representation/Encoding Schemes, initializing a GA population, evaluation function, genetic operators, Function optimization using GA. Study of parameters of genetic algorithms and its performance, sampling and selection mechanisms. Scaling of GA population.

Unit - II

Designing Genetic Algorithms for different applications: Different types encoding schemes, role of fitness function, different types of genetic operators, Designing GAs for numerical optimization, knapsack problem and travelling salesperson and other similar problems.

Unit-III

Fuzzy sets: Basic terminology and definitions ,Operations on Fuzzy sets, MF formulations and parameterisation, MFs of one and two dimensions, Derivatives of parameterised MFs, Fuzzy numbers, Extension principle and fuzzy relations, Operations on Fuzzy relations, Linguistic variables, Fuzzy If-Then Rules, Compositional rule of inference.

Unit - IV

Neural networks: Basic terminology and definitions, Model of an artificial neuron, Sigmoid function, Neural

Network Architectures, Rosenblatt's Perceptron, Fixed increment perceptron learning algorithm for a classification problem, Examples of learning of AND/OR gate by perceptron, XOR problem. Back Propagation Neural Networks: Architecture of a backpropagation network, Model for multi-layer perceptron, Back propagation learning, Delta or gradient descent learning rule and effect of learning rate, Back propagation learning algorithm.

Text and Reference Books:

1. David. E. Goldberg, Genetic Algorithms in Search, Optimization and machine learning, Addison Wesley, 1999.
2. Zbigniew Michalewicz, Genetic algorithms + Data Structures = Evolution Programs, Springers-Verlag, 1999.
3. M. Mitchell, An Introduction to Genetic Algorithms, Prentice-Hall, 1998.
4. S. Rajasekaran & G. A. Vijayalakshmi Pai, Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis & Applications, PHI, 2003.
5. S. N. Sivanandam & S. N. Deepa, Principles of Soft Computing, Wiley - India, 2007.
6. J-S. R. Jang, C.-T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing, PHI, 1997.
7. Simon O. Haykin, Neural Networks, A Comprehensive Foundation, PHI, 1994.

Big Data Analytics Lab.

General Course Information

Course Code: MDS-CC-301P Course Credits: 2 Type: Core Course (CC) Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Data Analytics, Big Data Analytics

About the Course:

The course familiarizes students with Hadoop distributions, configuring Hadoop and performing File management tasks. It also includes implementation of storage of big data using MongoDB,MapReduce programs for processing big data for various applications, andProcessing Large datasets using programming tools like PIG & HIVE in Hadoop ecosystem.

Practical Lab based on theory course MDS-CC-301T

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Machine Learning Lab.

General Course Information

Course Code: MDS-CC-302P Course Credits: 2 Type: Core Course (CC) Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Programming in Java, Python, R and Octave/MATLAB.

About the Course:

In this lab course, students learn to solve optimization, supervised, unsupervised and reinforcement learning problems using machine learning tools. Students will use machine learning tools available in Java, R, Python and MATLAB etc. The lab experiments involve downloading datasets and applying machine learning techniques on these datasets. The course has a special focus on interpreting and visualizing results of machine learning algorithms.

Practical Lab based on theory course MDS-CC-302T

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course and the levels of the given course outcomes. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Project Work-I.

General Course Information

Course Code: MDS-PW-301 Course Credits: 2 Type: Project Work (PW) Contact Hours: 4 hours/week Mode: Project Work	**Course Assessment Methods:** **Total Marks: 100** The internal assessment is based on the level of participation in lab. sessions and the timely submission of Project Work /assignments, the quality of solutions designed for the assignments and project work, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The examination is conducted by committee constituted by the Chairperson of the Department

About the Project Work-I:

Students start working on their project work in third semester. Students do the background research for Identifying appropriate problems, methodology and tools for their respective project. At the end of semester, each student is required to prepare a project in the format provided and present it in front of a committee constituted by the Chairperson of the Department. Students can carry out projects in groups of two. In case of group project, the size of the problem should be significant, and members of the group must specify their individual contribution.

Deep Learning

General Course Information

Course Code: MDS-CC-401T Course Credits: 3 Type: Core Course (CC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Linear Algebra, probability and information theory and machine learning

About the Course:

Deep learning has revolutionised the field of machine learning. Deep learning emphasises on learning complex, hierarchical feature representation from raw data. Deep learning algorithms have found tremendous amount of applications in machine learning applications. The course covers fundamental principles of deep learning and elaborates on building and optimizing these highly parameterized models. It involves learning about convolutional neural networks, recurrent and generative adaptive neural network models.

Course Content

Unit-1

Introduction : Learning Algorithms - Capacity, Overfitting and Underfitting, Hyperparameters and Validation Sets, Estimators, Bias and Variance Challenges Motivating Deep Learning.

UNIT-II

Convolutional Networks and Transfer Learning Convolution Operation - Motivation - Pooling - Structured Outputs - Data Types - Popular CNN Architectures - LeNet, AlexNet, VGG - Transfer Learning - DenseNet, PixelNet.

UNIT-III

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 - Challenges of Long-Term Dependencies - Long Short-Term Memory and Gated RNNs Optimization for Long-Term Dependencies.

UNIT-IV

Deep Generative Models :Boltzmann Machines - Restricted Boltzmann Machines - Deep Belief Networks - Deep Boltzmann Machines - Directed Generative Nets - Variational Autoencoders, Generative Adversarial Networks, Generative Moment Matching Networks.

Text Book(s) :

1. lan Goodfellow, Yoshua Bengio and Aaron Courville, "Deep learning", 2016, MIT Press.
2. Dipanjan Sarkar, Raghav Bali, Tamoghna Ghosh, "Hands-On Transfer Learning with Python", 2018, First edition, Packt Publishing.
3. John D. Kellcher, "Deep Learning", 2019, First edition, The MIT Press
4. Charu C. Aggarwal, "Neural Networks and Deep Learning: A Textbook", 2018, First edition, Springer

Evolutionary and Swarm Intelligence for Optimization

General Course Information

Course Code: MDS-CC-402-T Course Credits: 4 Type: Professional Core Contact Hours: 4 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted in open book mode by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the average of the highest marks obtained by a student in any of the two minor examinations will be considered. All the minor examination question papers will be prepared and evaluated by following the Outcome Based Education framework. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basic understanding of computer algorithms, familiarity with basic optimization methods

About the Course:

This course provides the introduction to the concepts, principles and applications of evolutionary and swarm optimization techniques. It covers evolutionary computing, particle swarm optimization, ant colony optimization algorithms. It offers an opportunity to understand and apply these nature inspired algorithms to solve complex optimization problems which otherwise cannot be solved within reasonable time due to their high complexity.

Course content

Unit I

History and need for nature inspired algorithm.
Basics of Genetic Algorithm (GA) and its working: Encoding, Fitness evaluation, Selection, Crossover and Mutation operators. Parameters of GA.
Different Encoding Schemes, Selection and Sampling mechanisms and GA Operators, Designing GAs for solving problems regarding function optimization.

Unit II

Local and Global Convergence, Linear Scaling, Multi-objective optimization, Multi-Objective GAs, Working of NSGA-II. Solving function optimization, knapsack, numeric optimization, routing and scheduling problems using GA, parallel genetic algorithms.

Unit III

Swarm Intelligence, Basic of Particle Swarm Optimization (PSO): Definitions and Concepts of PSO, Working of PSO, swarm size, information links, initialization, Equations of motion, interval confinement, proximity distributions, distribution bias, explosion and maximum velocity, parameters of PSO, Local best and global best PSO, Solving function optimization, knapsack, numeric optimization, routing and scheduling problems using PSO.

Unit IV

Ants' foraging behaviour and optimization, Artificial Ants, Artificial Ants and minimum cost paths, combinatorial optimization, Ant Colony Optimization (ACO) Metaheuristic, Applying ACO, Theoretical considerations on ACO, Convergence Proof, ACO and Model based search, solving travelling sales person and other similar problems using ACO, Ant Systems and its successors, ACO plus local search.

Text and Reference Books:

1. Zbigniew Michalewicz, *Genetic algorithms +Data Structures = Evolution Programs*, Springers-Verlag, 1999.
2. David.E. Goldberg, *Genetic Algorithms in Search, Optimization and machine learning*, Addison Wesley, 1999.
3. Marco Dorigo, Thomas, Stutzle, *Ant Colony Optimization*, MIT Press, 2004.
4. Helio J.C. Barbosa, "*Ant Colony Optimization - Techniques and Applications*", Intech 2013
5. Maurice Clerc, *Particle Swarm Optimization*, ISTE, Ltd., London, UK, 2006.

Computer System Architecture and Organization

General Course Information

Course Code: MDS-SEC-401T-(i) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Digital Electronics and computer systems.

About the Course: Computer Architecture and organization describes the role of instruction set architecture in digital computer, main memory, and input/output devices. It illustrates the simple data path and control design for processors. It helps to understand the different operations and concept of instructions. It would enable the students to learn the basic function and architecture of modern computer systems.

Course Content

Unit - I

Basic Principles: Boolean algebra and Logic gates, Combinational logic blocks (Adders, Subtractors, Multiplexers, Encoders, decoders, de-multiplexers, K-Maps), Sequential logic blocks (Flip-Flops, Registers, Counters); Flynn's classification of computers (SISD, MISD, MIMD);

Unit - II

Computer Organization: Store program control concept, Instruction codes, timing and control, instruction cycle; type of instructions: memory reference, register reference, I/O reference; Basics of Logic Design, accumulator logic,

Unit - III

Control memory; Micro Programmed Control: address sequencing, micro-instruction formats, micro-program sequencer, Implementation of control unit.

Stack Organization, Instruction Formats; addressing modes: register, immediate, direct, indirect, indexed; Operations in the instruction set: Arithmetic and Logical, Data Transfer, Control Flow; Types of interrupts;

Unit - IV

Memory Hierarchy & I/O Techniques: The need for a memory hierarchy (Locality of reference principle, Memory hierarchy in practice: Cache, main memory and secondary memory, Main memory (Semiconductor RAM & ROM organization, memory expansion, Static & dynamic memory types); Cache memory (Associative & direct mapped cache organizations; input-output interface, mode of transfer, DMA (Direct memory transfer).

Text and Reference Books:

1. Mano, M. Morris, Digital Logic and Computer Design, Prentice Hall of India Pvt. Ltd., 1981.
2. M. Morris Mano, Computer System Architecture, Prentice Hall of India Pvt. Ltd., 1993.
3. Milles J. Murdocca, Vincent P. Heuring, Computer Architecture and Organization, An Integrated Approach, John Wiley& Sons Inc., 2007.
4. William Stallings, 10th edition, Computer Organization and Architecture, Prentice Hall, 2016.
5. Heuring, V. P., Jordan, H.F., Computer Systems Design and Architecture, Addison Wesley, 1997.
6. R.P Jain, Modern Digital Electronics, 3rd Edition, Tata McGraw Hill, 2003.

Blockchain Technology

General Course Information

Course Code: MDS-SEC-401T-(ii) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted in open book mode by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the average of the highest marks obtained by a student in any of the two minor examinations will be considered. All the minor examination question papers will be prepared and evaluated by following the Outcome Based Education framework. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basics of Cryptographic Hash Functions

About the Course:

This course provides a broad overview of the essential concepts of blockchain technology by initially exploring the Bitcoin protocol followed by the Ethereum protocol to lay the foundation necessary for developing applications.

Course Contents

Unit I

Overview of Blockchain Technology: Defining Blockchain and Distributed Ledger, Blockchain properties decentralized, transparent, immutable and secure. Blockchain applications. Types of blockchain: Public, private, and consortium based blockchain, when to use, and when not to use blockchain, History of blockchain.

Introduction to computing models and P2P networking: Centralized, Decentralized and Distributed Systems, Decentralization vs distributed, P2P systems, propertied of P2P systems, P2P communication architecture. P2P network applications: File sharing, P2P network for Blockchain

Unit II

Foundational Concepts in Blockchain Data Structure: Cryptographic Hash Functions, Digital Signatures, Public Keys as Identities, Hash Pointers and Hash chain and Merkel tree, Consensus mechanisms
Blockchain Characteristics: Decentralized Identity management, Transactions, incentivizing and mining. Distributed Consensus (PoW), Cryptocurrency as the first blockchain application. Mechanics of Bitcoin, Bitcoin Scripts, Storing and Using Bitcoins, Mining in Bitcoin.

Unit III

Consensus Mechanisms: Proof of storage, proof of stake, proof of deposit, proof of burn, proof of activity. algorithms for adjusting difficulty and retargeting. Limitations of Bitcoin, alternative cryptocurrencies.
Smart Contracts and Ethereum: Purpose and types of smart contracts, Introduction to Ethereum, bitcoin vs Ethereum stack. P2P network in Ethereum, consensus in Ethereum, scripts in Ethereum, Smart contracts (Ethereum Virtual Machine). Developing and executing smart contracts in Ethereum. State and data structure in Ethereum.

Unit IV

Private and Consortium based BlockchainHyperledger: Need for the consortium. Hyperledger stack, Multichainblockchain, Innovation in Hyperledger, smart contracts, and distributed applications in Hyperledger.
Case studies/ Enabling Technologies and applications: Application of blockchain in privacy and security, IoT and smart cities, Business and Industry, Data management, e-Governance.

Text and Reference Books and Links:

1. Imran Bashir, *Mastering Block Chain: Distributed Ledger Technology, Decentralization and Smart Contracts Explained*, Packt Publishing, 2018.
2. Daniel Drescher, *Block Chain Basics*, Apress; 1stedition, 2017.
3. Josh Thompsons, *Block Chain: The Block Chain for Beginners- Guide to Blockchain Technology and Leveraging Blockchain Programming*, Createspace Independent Pub, 2017.
4. Pethuru Raj, Kavita Saini, ChellammalSurianarayanan, *Blockchain Technology and Applications*, CRC Press, 2021.
5. Raj K., *Foundation of Blockchain: The pathway to cryptocurrency and decentralized blockchain application*, 1st ed. Packt Publishing Ltd, 2019.
6. S. Shukla, M. Dhawan, S. Sharma, S. Venkatesan, *Blockchain Technology: Cryptocurrency and Applications*, Oxford University Press, 2019
7. Melanie Swan, *Blockchain: Blueprint for a New Economy*, O'Reilly, 2015.
8. Amit Dua, *Blockchain Technology and Applications: A systematic and Practical approach,* Amazon LLC, 2022

Reinfroncement Learning

General Course Information

Course Code: MDS-SEC-401-T-(iii) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted in open book mode by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the average of the highest marks obtained by a student in any of the two minor examinations will be considered. All the minor examination question papers will be prepared and evaluated by following the Outcome Based Education framework. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Pre-requisites: Probability and linear algebra, python programming, data structures and algorithms, artificial intelligence, machine learning.

About the Course:

Reinforcement learning is a paradigm that aims to model the trial-and-error learning process that is needed in many problem situations where explicit instructive signals are not available. The goal of the course is to introduce the basic foundations of reinforcement learning, model-based learning, temporal difference learning and ensemble learning.

Course Contents

Unit I

Introduction: Reinforcement Learning, Elements of Reinforcement Learning, Limitations and Scope, relationship to dynamic programming.

Multi-armed Bandits: A k-armed Bandit Problem, Action-value Methods, The 10-armed Testbed,Incremental Implementation, Tracking a Nonstationary Problem, Optimistic Initial Values, Upper-Confidence-Bound Action Selection, Gradient Bandit Algorithms.

Unit II

Finite Markov Decision Processes: The Agent–Environment Interface, Goals and Rewards, Reward models (infinite discounted, total, finite horizon, and average), Returns and Episodes, Policies and Value Functions, Optimal Policies and Optimal Value Functions

Dynamic Programming: Policy Evaluation (Prediction), Policy Improvement, Policy Iteration, Value Iteration

Monte Carlo Methods: Monte Carlo Prediction, Monte Carlo Estimation of Action Values, Monte Carlo Control

Unit III

Temporal-Difference Learning: TD Prediction, Advantages of TD Prediction Methods, Optimality of TD (0), Sarsa, Q-learning, Expected Sarsa , Maximization Bias and Double Learning

n-step Bootstrapping: n-step TD Prediction, n-step Sarsa, n-step Off Policy Learning, The n-step Tree Backup Algorithm, A Unifying Algorithm: n-step Q(σ)

Unit IV

Policy Gradient Methods: Policy Approximation and Advantages, Policy Gradient Theorem, Monte Carlo Policy Gradient, Reinforce with Baseline,Actor–Critic Methods, Policy Gradient for Continuing Problems

Applications and case studies: TD-Gammon, Samuel's Checkers Player, Watson's Daily-Double Wagering, Optimizing Memory Control, Mastering the Game of Go, Personalized Web Services, Reinforcement learning in robotics

Text and Reference Books and Links:

1. Sutton and Barto, *Reinforcement Learning: An Introduction*, The MIT Press Cambridge, Massachusetts London, England, 2015.
2. Zhou and Zhi-Hua,*Ensemble Methods: Foundations and Algorithms*, Chapman & Hall/CRC, 2012.
3. Tesauro, *Temporal Difference Learning and TD-Gammon, Communications of the Association for Computing Machinery*, 1995.
4. Dimitri P. Bertsekas, *Reinforcement Learning and Optimal Control*, 1st Edition, Athena Scientific
5.

Pattern Recognition

General Course Information

Course Code: MDS-SEC-402T-(i) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted in open book mode by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the average of the highest marks obtained by a student in any of the two minor examinations will be considered. All the minor examination question papers will be prepared and evaluated by following the Outcome Based Education framework. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Artificial Intelligence, Neural Networks, Formal Language theory, Data mining, data analytics.

About the Course:
The aim of this course to summarize various techniques involved in pattern recognition. It includes supervised, unsupervised and ANN based pattern recognition techniques and their applications.

Course Content

Unit I
Introduction - Basic concepts, Applications, Fundamental problems in pattern Recognition system design, Design concepts and methodologies, The Design Cycle, Learning and Adaptation., Examples of Automatic Pattern recognition systems, Simple pattern recognition model, Decision and Distance Functions - Linear and generalized decision functions, Pattern space and weight space, Geometrical properties, implementations of decision functions, Minimum-distance pattern classifications. Bayes Decision Theory: Bayes Decision Theory, Minimum Error rate Classification.

Unit II

Maximum Likelihood and Bayesian Parameter Estimation: Maximum Likelihood Estimation, Bayesian Estimation, Bayesian Parameter Estimation, Gaussian Case and General Theory. Hidden Markov models; Non Parametric Techniques: Density Estimation, Parzen Windows, K- Nearest Neighbor Estimation, Nearest Neighbour rule, Metrics and Nearest Neighbour Classification, Fuzzy Classification, k-Means Clustering, Self-Organizing Maps.
Non Parametric Decision Making - Introduction, histogram, kernel and window estimation,. Adaptive decision boundaries, adaptive discriminate functions, Minimum squared error Discriminate functions.

Unit III

Linear Discriminant Functions: Linear Discriminant Functions and Decision Surfaces, Generalized Discriminant

Functions, The two-category linearly separable case, Minimizing the perceptron criterion function, relaxation procedures, non- separable behaviour, Minimum Squared- Error procedures. Support vector machines, Algorithm-independent machine learning-Bias and Variance, Bootstrapping-Adaboost Algorithm, Boosting, Bagging

Unit IV

Syntactic Pattern Recognition & Application of Pattern Recognition: Introduction, concepts from formal language theory, formulation of syntactic pattern recognition problem, syntactic pattern description, recognition grammars, automata as pattern recognizers, Application of pattern recognition techniques in bio-metric, facial recognition, IRIS scan, Finger prints, etc..

Text and Reference Books:

7. R. O. Duda, P. E. Hart and D. G. Stork, Pattern classification, John Wiley & Sons, 2002.
8. C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
9. V. N. Vapnik, The Nature of Statistical Learning Theory, Springer, 2000.
10. N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, 2000.

Cloud Computing

General Course Information

Course Code: MDS-SEC-402T-(ii) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basics of Computer Network, Distributed System.

About the Course: The objective of the course is to give students a comprehensive view understanding of the vision and impact of Cloud, cloud Market perspective and Cloud architecture.

Course Content

Unit I

Overview of Cloud Computing- Cloud at a Glance: The Vision of Cloud Computing, Defining a Cloud, Cloud Computing Reference Model, Characteristics and Benefits, Historical Developments: Distributed Systems, Cluster Computing, Grid Computing, Virtualization.

Unit II

Virtualization & Cloud Computing Architecture – Introduction, Characteristics of Virtualized Environments, Taxonomy of Virtualization Techniques: Execution Virtualization, Other Types of Virtualization, Virtualization and Cloud Computing: Pros and Cons of Virtualization, Cloud Architecture: Introduction, Cloud Reference Model Architecture, Infrastructure as a Service, Platform as a Service, Software as a Service, Types of Clouds: Public, Private, Hybrid, Community,

Unit III

Cloud in Industry and Its Applications – Amazon Web Services: Compute Services, Storage Services, Communication Services, Additional Services, Google AppEngine: Architecture and Core Concepts, Application Life-Cycle, Microsoft Azure: Core Concepts, SQL Azure, Windows Azure Platform Appliance, Cloud Applications in various domains.

Unit IV

Security in Cloud – Cloud Information Security Fundamentals, Cloud Security Services, Design Principles, Secure Cloud Software Requirements, Policy Implementation, Cloud Computing Security Challenges, Virtualization Security Management, Cloud Computing Security Architecture.

Text and Reference Books and Links:

1. Rajkumar Buyya, Christian Vecchiola and S. Thamarai Selvi, *Mastering Cloud Computing*, McGraw Hill Publication (India) Private Limited, 2013.
2. Krutz, Vines, *Cloud Security*, Wiley Publication, 2010.
3. Bloor R., Kanfman M., Halper F. Judith Hurwitz, *Cloud Computing for Dummies*, (Wiley India Edition), 2010.
4. John Rittinghouse & James Ransome, *Cloud Computing Implementation Management and Strategy*, CRC Press, 2010.
5. Antohy T Velte , *Cloud Computing : A Practical Approach*, McGraw Hill, 2009.
6. Rajkumar Buyya, James Broberg and Andrez Gosscinski, *Cloud Computing: Principles and Paradigm*, John Wiley and Sons, Inc. 2011.
7. Kai Hwang, Geofferyu C. Fox and Jack J.Dongarra, *Distributed and Cloud Computing*, Elsevier, 2012.

Internet of Things

General Course Information

Course Code: MDS-SEC-402T-(iii) Course Credits: 3 Type: Skill Enhancement Course (SEC) Contact Hours: 3 hours/week Mode: Lectures (L) Examination Duration: 3 hours	**Course Assessment Methods:** **Max. Marks: 100 (Internal: 30; External: 70)** Three minor tests, each of 20 marks, will be conducted. The third minor will be conducted by the Course Coordinator. No date sheet will be issued for the third minor at the level of the Departments. For the purpose of internal assessment, the highest marks obtained by a student in any of the minor examinations will be considered. Class Performance will be measured through percentage of lectures attended (4 marks) Assignments (4 marks) and class performance (2 marks). The end semester examination will be of 70 marks. For the end semester examination, nine questions are to be set by the examiner. All questions carry equal marks. Question number one will be compulsory and based on the entire syllabus. It will contain seven short answers type questions. The remaining eight questions are to be given by setting two questions from each of the four units of the syllabus. A candidate is required to attempt five questions in all, one compulsory and any other four questions by selecting one from each unit.

Pre-requisites: Basics of Computer Network, Distributed System.

About the Course: The objective of the course is to give students a comprehensive view understanding of the vision and impact of IoT, IoT Market perspective and IoT architecture.

Course Content

Unit I

What is the Internet of Things: History of IoT, About IoT, Overview and Motivations, Examples of Applications, Internet of Things Definitions and Frameworks: IoT Definitions, IoT Architecture, General Observations, ITU-T Views, Working Definition, IoT Frameworks, Basic Nodal Capabilities, Basics of Microcontroller, Microprocessor Vs Microcontroller, Types of Sensors, Actuators and their Applications.

Unit II

Identification of IoT Objects and Services, Structural Aspects of the IoT, Environment Characteristics, Traffic Characteristics, Scalability, Interoperability, Security and Privacy, Open Architecture, Key IoT Technologies, Device Intelligence, Communication Capabilities, Mobility Support, Device Power, Sensor Technology, RFID Technology- Introduction, Principle of RFID, Components of an RFID system, Issues, Satellite Technology.

Unit III

IoT Access Technologies: Physical and MAC layers, Topology and Security of IEEE 802.15.4, 802.15.4g, 802.15.4e, 1901.2a, 802.11ah and LoRaWAN, Network Layer: IP versions, Constrained Nodes and Constrained Networks, Optimizing IP for IoT: From 6LoWPAN to 6Lo, Routing over Low Power and Lossy Networks, Application Transport Methods: Supervisory Control and Data Acquisition, Application Layer Protocols: CoAP and MQTT.

Unit IV

Business Models and Business Model Innovation, Value Creation in the Internet of Things, Business Model Scenarios for the Internet of Things. Internet of Things Applications: Smart Metering Advanced Metering Infrastructure, e-Health Body Area Networks, City Automation, Automotive Applications, Home Automation, Smart Cards, Smart Transportation and Smart Shopping.

Text and Reference Books:

1. David Hanes, Gonzalo Salgueiro, Patrick Grossetete, Rob Barton and Jerome Henry, *IoT Fundamentals: Networking Technologies, Protocols and Use Cases for Internet of Things*, Cisco Press, 1st Edition, 2017.
2. Olivier Hersent, David Boswarthick, Omar Elloumi, *The Internet of Things – Key applications and Protocols*, Wiley, 2nd Edition, 2012.
3. Dieter Uckelmann, Mark Harrison, Michahelles, Florian (Eds), *Architecting the Internet of Things*, 1st Edition, Springer, 2011.
4. Michael Margolis, Arduino Cookbook, "*Recipes to Begin, Expand, and Enhance Your Projects*", 2nd Edition, O'Reilly Media, 2011.
5. Arshdeep Bahga, Vijay Madisetti, *Internet of Things – A hands-on approach*, 1st Edition, Universities Press, 2015.

.

Deep Learning Lab.

General Course Information

Course Code: MDS-CC-401P Course Credits: 2 Type: Core Course (CC) Contact Hours: 4 hours/week Mode: Lab practice and assignments	**Course Assessment Methods:** **Total Marks: 100 (internal: 50; external: 50)** The internal and external assessment is based on the level of participation in lab. sessions and the timely submission of lab experiments/assignments, the quality of solutions designed for the assignments, the performance in VIVA VOCE, the quality of lab. file and ethical practices followed. The internal examination is conducted by the course coordinator. The external examination is conducted by external examiner appointed by the Controller of Examination in association with the internal examiner appointed by the Chairperson of the Department.

Pre-requisites: Introduction to deep learning concepts and techniques

About the Course:
This is a lab course for hands on practice of deep learning techniques such as simple neural network, CNN, RNN, Transfer Learning using Keras, TensorFlow and PyTorch. The students will solve and compare various techniques for classification problems.

Note:

The actual experiments/assignments will be designed by the course coordinator. One assignment should be designed to be done in groups of two or three students. The assignments must meet the objective of the course and the levels of the given course outcomes. The list of assignments and schedule of submission will be prepared by the course coordinator at the beginning of the semester.

Project -II

General Project Information

Course Code: MDS-PW-401 Course Credits: 6 Mode: Learning under the guidance of a faculty member.	**Course Assessment Methods (Internal evaluation: 30 marks; External Evaluation marks: 70)** Evaluation is done by the internal examiner (project guide) and external examiner appointed by the Controller of Examination.

About the major project part II:

Students continue working on their project work and they are required to complete it by the end of 4th semester. Students carry out implementation of their respective projects based on the problem identified, methodology and tools suggested. They prepare the final project reports according to the format provided. At the end of 4th semester, each student is required to present his/her project work in front of internal project guide and external examiner appointed by the Controller of Examination.

Guidelines for Preparing Project report (MDS-PW-401)
All the students are required to follow these guidelines for preparing their project report.

Formatting Instructions

The formatting instructions are given in the table below.

Formatting Instructions		
Sr. No.	Item	Formatting
1.	Front Cover	Quality paper suitable for hard binding
2.	No. of pages	Minimum 50 pages excluding the front material
3.	Paper size	A4
4.	Font Type	Times New Roman
5.	Chapter Heading Font	16
6.	Font of Sections and Subsections	14 and 12 in bold style
7.	Numbering style for sections and subsections; Do not use more than three levels.	2., 2.1 and 2.1.1
8.	Normal text size	12
9.	Figures and Tables must be numbered chapter-wise. Table headings on the top of the tables and Figure heading at the bottoms of the figures.	For example, for chapter 2, Figures should be numbered as Fig. 2.1, Fig. 2.2 etc. and Tables as Table 2.1 and Table 2,2 etc.
10.	Page numbering	Place: Centre Bottom Type: Front material in Roman numbers Body of the report: in Arabic numerals. Pagination must start with first page of the first chapter and continue throughout the end of the report.
11.	Margins	Left margin: 3.75 cms (1.5 inch) Right, bottom, top= 2.5 cms (1 inch)
12	References/Bibliography	IEEE format
13	Binding	Hard binding of good quality

Contents of the Project Report

The contents of the report should be organised as described below.

1. The title page as per instructions.
2. Declaration that the student has carried out his work on his own. It is his/her original creation, not plagiarised from any other source and due credit has been given to the source material used in the report through references and citations.
3. Acknowledgement
4. List of figures
5. List of Tables
6. List of Abbreviations
7. Contents

8. Body of the Report

The report must be written in English. The ideas must be organized in a clear and concise fashion. Chapters must be tentatively organized as below.

Chapter 1. Introduction

This includes introduction to relevant area of minor project, problem formulation, objectives of the minor project, and structure of the project report.

Chapter 2. Requirement analysis, solution design framework of the project work and tools used

Chapter 3. Outputs of the project

References/Bibliography

Format of the Title page
The format for the title page of the minor project is given on next page.

TITLE OF THE PROJECT REPORT
(Write in Times New Roman, 16-point size, Bold and Centred and Uppercase font)
Project report submitted to
Guru Jambheshwar University of Science and Technology, Hisar
for the partial award of the degree

(Write in Times New Roman, 12-point size font, Bold, Italics and Centred style after 2 lines gap with 12 font size from the title of the project)
of
(Write in Times New Roman, 12-point size font, Bold, Italics and Centred style after 1 lines gap with 12 fontsize from the text above in three lines)
Master of Science
in
Computer Science
(Artificial Intelligence and Data Science)

(Write in Times New Roman, 14-point size, Bold, Centred style after "***of***" after 1 line gap with 12 font size)
By
(Write in Times New Roman 12-point size, Bold, Italics, and Centred style after the name of the degree with 1 line gap with 12)

Your Name
(Enrolment Number)

(Write in Times New Roman, 14-point size font, Bold, Centred style after 1 line gap with 12 font from "***By***")

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
GURU JAMBHESHWAR UNIVERSITY OF SCIENCE AND TECHNOLOGY, HISAR
Month, Year

(Write in Times New Roman, 14-point size font, Bold, Centred style, after 2 lines gap from logo)