Course Structure
Master of Technology Computer Sc & Engineering specialization in Data Science

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Semester I

Sr. No.	Course Code	Course Name	C/E	L	T	P	Hr	Cr
1	22M11CI111	Advanced Data Structures	C	3	0	0	3	3
2	22M11MA111	Mathematical Foundations for Data Science	C	3	0	0	3	3
3	22M11CI112	Introduction to Data Science	C	3	0	0	3	3
4		DE-I	E	3	0	0	3	3
5		DE-II	E	3	0	0	3	3
6		DE-III	E	3	0	0	3	3
7	22M17CI171	Advanced Data Structures Lab	C	0	0	1	2	1
8	22M17CI172	Data Science Lab	C	0	0	2	4	2
9	22M17MA171	Mathematical Foundations for Data Science Lab	C	0	0	2	4	2
		Total Credits						23

Semester II

Sr. No.	Course Code	Course Name	C/E	L	T	P	Hr	Cr
	22M11CI211	Soft computing	C	3	0	0	3	3
	22M11CI212	Deep Learning Techniques	C	3	0	0	3	3
	22M11CI213	Big Data Analytics	C	3	0	0	3	3
		DE-IV	E	3	0	0	3	3
		DE-V	E	3	0	0	3	3
		DE-VI	E	3	0	0	3	3
	22M17CI271	Soft Computing Lab	C	0	0	1	2	1
	22M17CI272	Deep Learning Techniques Lab	C	0	0	2	4	2
	22M17CI273	Big Data using Hadoop Lab	C	0	0	2	4	2
		Total Credits						23

SemesterIII

Sr. No.	Course Code	Course Name	C/E	L	T	P	Cr
1	22M19CI391	Seminar	C	0	0	4	2
2	22M19CI392	Project, Part I	C	0	0	24	12
		Total Credit					14

Semester IV

Sr. No.	Course Code	Course Name	C/E	L	T	P	Cr
1	22M19CI491	Seminar	C	0	0	4	2
2	22M19CI492	Project, Part I	C	0	0	24	12
		Total Credit					14

Summary

Semester	Cr
I	23
I	23
III	14
IV	16
Total	**76**

Departmental Electives

DE	Course Code	Course Name	C/E	L	T	P	Cr
DE-I	22M1WCI134	Data Storage Technologies - and Data	E	3	0	0	3
DE-I	22M1WCI131	Data Warehousing Mining -	E	3	0	0	3
DE-II	22M1WCI136	Data Visualization	E	3	0	0	3
DE-II	22M1WCI132	Artificial Intelligence Techniques	E	3	0	0	3
DE- III	22M1WCI133	Introduction to Statistical Learning -	E	3	0	0	3
DE- III	22M1WCI135	Cryptography and Information System security	E	3	0	0	3
DE-IV	22M1WCI231	Advanced Computational Techniques	E	3	0	0	3
DE-IV	22M1WCI232	Medical Image Analysis	E	3	0	0	3
DE-V	22M1WCI233	Knowledge-Based AI: Cognitive Systems	E	3	0	0	3
DE-V	22M1WCI234	Social and Information Network Analysis	E	3	0	0	3
DE-VI	22M1WCI235	Reinforcement Learning	E	3	0	0	3
DE-VI	22M1WCI236	Natural Language Processing	E	3	0	0	3

COURSE CONTENTS
MTech Computer Science & Engineering with specialization in Data Science

Approved in Academic Council held on 28 July 2021

Advanced Data Structure

COURSE CODE: 22M11CI111
COURSE CREDITS: 3
CORE/ELECTIVE: CORE
: 3-0-0
–

Prerequisites: Data Structures, Algorithms

Course Objective

The field of data structures is similar in spirit to algorithms: what's the fastest way to solve problem X. But data structures take a different tact, focusing on ways to organize a typically long-lived corpus of data so that queries can be answered really fast. A typical example are web search engines, which store sophisticated indexes that support a variety of queries (e.g., what documents contain word X) quickly. In contrast to most algorithmic problems, "linear time" is too slow--you can't afford to scan through the entire web to answer a single query--and the goal is typically logarithmic or even constant time per query. Also, space tends to become a more important issue, because the data corpus is usually big, and you don't really want a data structure that is larger than (or even as big as) the data itself.

Course Outcome (CO)

S No	**Course outcomes (Compiler Design) (**10M11CI213**)**	**Level of Attainment**
CO-1	Student will learn dynamic optimality:(Splay tree, Skip list, tango)	Familiarity
CO-2	Student will learn geometric data structures:(KD tree, Quad Tree, R tree, ray shooting)	Familiarity
CO-3	Student will learn string data structures:(Trie, Compact trie, patricia, suffix tree)	Assessment
CO-4	Student will learn integer data structures: (Binomial heap, fibonacci heap, soft heap)	Usage
CO-5	Student will learn dictionary data structures: (Cuckoo hashing, Bloom filter, Inverted index)	Assessment
CO-6	Student will learn external memory model:(I/O complexity, Lazy update, B tree, Buffer tree)	Usage
CO-7	Student will learn cache-oblivious:(Large matrix multiplication, block tree)	Assessment
CO-8	Student will learn distributed data structures:(DHT, DBST, trees)	Assessment
CO-9	Student will learn stream data structures:(Sampling, sketching, fingerprint, wavelet, histogram)	Usage

Learning Outcome

The students shall deepen the generic skills to design and implement ADTs, non linear data structures and algorithms for a broad-based set of computing problems in various domains.

Course Contents/Lecture Plan

Sr. No	Theory Component	Lecture
1.	**Dynamic optimality**: binary search trees, analytic bounds, splay trees, geometric view, greedy algorithm, independent rectangle, Wilber, and Signed Greedy lower bounds; key-independent optimality; O(lglg*n*)-competitive Tango trees	**6**
2.	**Dictionary data structures**: Cuckoo hashing, Bloom filter, Inverted index	**6**
3.	**String data structures**: suffix tree, suffix array, Inverted index	**3**
	T1	
4.	**Integer data structures**: Binomial heap, fibonacci heap, soft heap and their applications to MST and shortest path problems	4
5.	**Round-robin MST:** Tarzan's implementation using fibonacci heap and Amortized complexity	**2**
6.	**Succinct data structures:** k-ary tree, multisets, rank/select	**2**
7.	**Geometric data structures**: orthogonal range queries, range trees, interval trees, KD tree, Quad Tree, R tree, ray shooting	**6**
	T2	
8.	**Nearest neighbour search:** approximation and locality sensitive hashing	**2**
9.	**External memory / cache-oblivious**: I/O complexity, Lazy update, B tree, Buffer tree and cache oblivious implementation of Large matrix multiplication and block tree	**4**
10.	**Distributed data structures:** hash table, trees, stack and lists	**5**
11.	**Stream data structures:** synopsis, sketches, histogram, fingerprint, wavelets, sliding windows etc	**4**
	T3	
	Total Lectures	**42**

Schedule : Lectures: 3 hours/week.

Evaluation Components

S.No	Exam	Marks	Duration	Coverage/Scope of Examination
1	Test -1	15	1 hr.	Syllabus covered upto T - 1
2	Test -2	25	1 hr 30 min.	Syllabus covered upto T - 2
3	Test - 3	35	2 hr.	Entire Syllabus
4	Assignments, Quizzes, Attendance	25	Entire Semester	Tutorials / Assignment - 10 Quizzes -10 Attendance - 5.

Teaching Methodology

The course will use the mixed technique of interactive lectures, guided case studies, literature survey, regular assignments and project work. In addition to the material covered in the class, students will be required to explore study, evaluate, present and implement several data structures and their applications. Teaching in this course is designed to engage the students in active and experiential learning by taking a problem solving and design oriented approach with special emphasis on real world applications. Lectures will be highly interactive and work oriented. Students will have to work individually as well as in groups inside as well as outside the class. Students are expected to carry out lot of design and programming oriented project work. Each student is also expected to choose a topic from the list provided. There should be a report component, on the order of 10 pages, let's say in the range 5-20 pages, a code component regarding some dedicated application and they will give a presentation about their project.

Text Book(s) / Reference Book(s)

1) Peter Brass : Advanced Data Structure
2) Jeff Vitter: External Memory Data structures
3) Muthukrishnan: Synopsis data structures

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO10	PO11	PO12	Avg
CO1	3	3	2	2	2	2	1	1	2	2	2	2	2
CO2	3	3	2	2	2	1	1	2	2	2	1	2	1.91
CO3	3	3	2	2	2	2	2	1	2	2	1	2	2
CO4	3	3	2	2	3	1	2	2	2	2	1	2	2.08
CO5	3	3	2	2	2	2	3	2	2	2	2	2	2.25
CO6	3	3	2	2	2	3	2	1	2	2	2	2	2.16
CO7	3	3	3	3	3	1	2	2	2	2	1	2	2.08
CO8	3	3	2	2	3	2	2	1	2	2	1	2	2.08
CO9	3	3	3	3	2	1	2	2	2	3	2	2	2.33
Avg	3	3	2.22	2.22	2.33	1.67	1.88	1.55	2	2.11	1.44	2	

Mathematical Foundations for Data Science

COURSE CODE: 22M11MA111

COURSE CREDITS: 3

CORE/ELECTIVE:

L-T-P: 3-0-0

Pre-requisite: Knowledge of mathematical concepts in undergraduate level

Course Objectives:

1. To introduce students to the fundamental mathematical concepts of multivariate calculus and linear algebra required for data analysis & machine learning.
2. To provide an understanding for the Data science students on inferential statistical concepts to include sampling, estimation, parametric & non-parametric hypothesis testing, correlation & multiple regression.
3. To extend and familiarize the students with the concepts of Stochastic models with Bayesian inference.

Course Outcomes:

S. No.	urse Outcomes	Level of Attainment
CO-1	Demonstrate calculus understanding of concepts of multivariable	Familiarity
CO-2	Understand the use of eigenvalues, eigenvectors in Matrix factorizations and vector space concepts.	Assessment
CO-3	Demonstrate proficiency on computing probabilities of discrete and continuous probability distributions and their modeling applications.	Assessment
CO-4	Develop appropriate optimization models for some applications in Data science	Usage

Course Contents:

Unit	Contents	Lectures required
1	Functions of several variables - Directional Derivative and Gradient vectors - Lagrange Multipliers. Matrix factorizations: Singular value decomposition (SVD); Basics of *vector space;* Projections: hyperplanes; Least-Squares problem.	**10**
2	Probability distributions of joint random variables; conditional expectation, covariance and correlation; standard univariate and multivariate distributions. Correlation & linear regression, classification by logistic regression; sampling distributions; Point & interval estimation, maximum likelihood estimation; concept of hypothesis testing: confidence Intervals; tests of significance for comparing two population means and variances. Non-parametric tests: Mann-Whitney test, Kolmogorov-Smirnov test (KS test).	**18**
3	Classification of random processes; Markov chains (discrete & continuous);Bayesian inference: Gibbs Sampling and data augmentation, Metropolis-Hastings independence sampling algorithm; Introduction to Optimization; Unconstrained optimization - Gradient Descent, Conjugate Gradient Descent; Constrained optimization - KKT conditions.	**14**
Total Lectures		**42**

Suggested Text Book(s):

1. G. Strang: ``*Introduction to Linear Algebra*", Wellesley-Cambridge Press, Fifth edition, USA, 2016.
2. Oliver C. Ibe: ``Fundamentals of applied probability and random processes", Academic press, 2005.
3. Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright: `` *Optimization for machine learning*",Mit Press, 2012.
4. Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining: ``Introduction to Linear Regression Analysis", John Wiley and Sons, Inc. NY, 2003.
5. Sheldon M. Ross, Stochastic Processes, Wiley, 1995.
6. W. R. Gilks, S. Richarson and D. Spiegelhalter, ``Markov chain Monte Carlo methods in Practice", Chapman and Hall, 1996.

Suggested Reference Book(s):

1. S. M. Ross: ``Introduction to Probability Models", Academic Press.
2. David G. Luenberger: ``Optimization by Vector Space Methods", John Wiley & Sons (NY), 1969.
3. Cathy O'Neil and Rachel Schutt: ``Doing Data Science", O'Reilly Media, 2013.
4. Larry A. Wasserman, ``All of Statistics: A Concise Course in Statistical Inference", Springer Texts in Statistics Springer, New York, 2004.
5. Samprit Chatterjee and A. S. Hadi, ``Regression Analysis by Example", 4th Ed., John Wiley and Sons, Inc, 2006.

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3.	T-3	35	2 Hours	Entire Syllabus
4.	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance - 5

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

Course outcomes (Parallel and Distributed Algorithms)	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Average
CO-1	2	2	2	2	2	1	1	1	2	2	2	2	
CO-2	2	3	3	3	3	1	1	1	2	2	1	2	
CO-3	2	2	2	2	3	1	1	1	2	2	1	2	
CO-4	2	3	3	3	2	1	1	1	2	3	2	2	
Average													

Approved in Academic Council held on 28 July 2021

Introduction to Data Science

COURSE CODE: 22M11CI112
COURSE CREDITS: 3
CORE/ELECTIVE: CORE
L-T-P: 3-0-0

Pre-requisite: Basic knowledge of high school mathematics and reasonable programming experience.

Course Objectives:

- Asking the correct questions and analyzing the raw data.
- Modelling the data using various complex and efficient algorithms.
- Visualizing the data to get a better perspective.
- Understanding the data to make better decisions and find the final result.
- Basic understanding of python and machine learning algorithms

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understanding the basics of data science	Familiarity
CO-2	Using versatile and flexible languages (Python and R programming) for supporting data science	Usage
CO-3	Using data processing for collecting and manipulating the data into the usable and desired form	Usage
CO-4	Using data visualization to easily access the huge amount of data in visuals	Usage
CO-5	Using statistics to collect and analyze the numerical data in a large amount and finding meaningful insights from i	Usage
CO-6	Understanding linear algebra to represent, model, synthesize and summarize the complex data	Usage

Course Contents:

Unit	Contents	Lectures required
1	**Data Science an Introduction:** Computer Science, Data Science, and Real Science, What is Data Science? Need for Data Science, Data Science Components, Tools for Data Science, Data Science Lifecycle, Applications of Data Science.	**2**
2	**Python and R Programming for Data Science for Data Science:** Introduction to Python Programming (Python Basics,Python Data Structures**,** Python Programming Fundamentals**,** Working with Data in Python**,** Working with NumPy, Pandas, SciPy, and Matplotlib). Introduction to R Programming (R basics, Data structures in R, R Programming fundamentals, Working with Data in R, Stings, and Dates in R, Discover R's packages to do graphics).	**5**

3	**Data Processing:** Data Operations, Data cleansing, Processing CSV Data, Processing JSON Data, Processing XLS Data, Relational databases, NoSQL Databases, Date and Time, Data Wrangling, Data Aggregation, Reading HTML Pages, Processing Unstructured Data, Word tokenization, Stemming and Lemmatization.	**5**
4	**Data Visualization:** Chart Properties, Chart Styling, Box Plots, Heat Maps, Scatter Plots, Bubble Charts, 3D Charts, Time Series, Geographical Data, Graph Data	**6**
5	**Statistical Data Analysis:** Measuring Central Tendency, Measuring Variance, Normal Distribution, Binomial Distribution, Poisson Distribution, Bernoulli Distribution, P- Value, Correlation, Chi-square Test, Linear Regression	**8**
6	**Supervised and Unsupervised Learning**: Classification and Regression, Generalization, over fitting and under fitting, Supervised Machine learning algorithms, Types of unsupervised learning, preprocessing and scaling, clustering	**6**
7	**Working with text data**: types of data represented as strings, representing text data as a bag of words, stop words, rescaling, advanced tokenization, stemming and lemmatization	**5**
Total lectures		**42**

Suggested Text Book(s):

- Data Science from Scratchby Joel Grus
- Data Science for Dummiesby Lillian Pierson and Jake Porway
- An Introduction to Statistical Learningby Gareth James, Daniela Witten, et al.
- Introduction to machine learning with python: Andreas C. Muller and Sarah Guido

Suggested Reference Book(s):

- An Introduction to Probability and Statistics by V.K. Rohatgi & A.K. Md. E. Saleh, Wiley, (2008), 3rd ed.
- Introduction to Probability Theory and Statistical Inference by H.J. Larson, John Wiley & Sons, (2005) 3rd ed.

Other useful resource(s):

- https://nptel.ac.in/courses/110/106/11010604/
- https://onlinecourses.nptel.ac.in/noc18_cs28/ preview

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance - 5

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

Course outcomes (Foundation for Data Science and Visualization)	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Average
CO-1	2	2	2	3	2	1	1	1	3	2	2	1	1.8
CO-2	3	3	3	3	3	1	1	2	2	2	1	1	2.1
CO-3	2	2	2	2	1	1	2	2	2	2	1	2	1.8
CO-4	3	3	3	3	2	1	1	3	2	3	3	2	2.4
CO-5	3	2	3	3	1	1	2	1	1	3	2	1	1.9
CO-6	2	2	1	2	2	2	3	2	1	2	3	1	1.9
Average	2.5	2.3	2.3	2.7	1.8	1.2	1.7	1.8	1.8	2.3	2	1.3	2.0

Advanced Data Structure Lab

COURSE CODE: 22M17CI171
COURSE CREDITS: 1
CORE/ELECTIVE: CORE
L-T-P: 0-0-1

Pre-requisites: None

Course Objectives:

1. Develop problem solving ability using Programming
2. Develop ability to design and analyze algorithms
3. Introduce students to data abstraction and advanced data structures
4. Develop ability to design and evaluate advanced data structures
5. Apply advanced data structure concepts to various examples and real life applications

Course outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understanding Dictionary data structures	Familiarity
CO-2	Student will learn string data structures:(Trie, Compact trie, patricia, suffix tree)	Usage
CO-3	Student will learn Integer data structures	Assessment
CO-4	Student will learnDistributed data structures: hash table, trees, stack and lists.	Assessment
CO-5	Student will learn Stream data structures:synopsis, sketches, histogram, fingerprint, wavelets, sliding windows etc	Usage
CO-6	Student will learn Geometric data structures: orthogonal range queries, range trees, interval trees, KD tree, Quad Tree, R tree, ray shooting	Usage

List of Experiments:

S No	Description	Hours
1	Implement the Dictionary data structures	2
2	Implement thestring data structures: Trie, Compact trie, patricia, suffix tree	4
3	Implement Binomial heap	2
4	Implement insertion, deletion and display operations in Binary Search Trees	2
5	Implement insertion, deletion and display operations in Red Black Trees	2
6	Implement insertion, deletion and display operations in B- Trees	2
7	Implement Huffman Coding Algorithm.	4
8	Implement Distributed data structures: hash table, trees, stack and lists.	4
9	Implement string matching algorithm.	2
10	Implement Quad trees	4
		28

Approved in Academic Council held on 28 July 2021

Evaluation Scheme:

1	Mid Sem. Evaluation	20 Marks
2	End Sem. Evaluation	20 Marks
3	Attendance	15 Marks
4	Lab Assessment	45 Marks
	Total	100 marks

Suggested Books/Resources:

1. Peter Brass : Advanced Data Structure
2. Jeff Vitter: External Memory Data structures
3. Muthukrishnan: Synopsis data structures
4. Corman et al : Introduction to Algorithms, 3rd Edn., 2009
5. http://www.nptel.iitm.ac.in/video.php?subjectId=106102064, last accessed Mar 13, 2014.
6. http://www.cs.auckland.ac.nz/~jmor159/PLDS210/ds_ToC.html, last accessed Mar 13, 2014.
7. http://courses.cs.vt.edu/csonline/DataStructures/Lessons/index.html, last accessed Mar 13, 2014.
8. Link to topics related to course:
   a. http://cse.iitkgp.ac.in/~pallab/pds16/pds16.htm
   b. https://onlinecourses.nptel.ac.in/programming101/preview

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

CO/PO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**	**Average**
CO-1	3	2	3	2	2	3	3	3	2	3	2	2	2.5
CO-2	3	3	3	2	3	3	3	3	2	3	2	3	2.8
CO-3	3	3	3	2	2	3	3	3	3	3	2	2	2.7
CO-4	3	3	3	3	3	3	3	2	2	3	3	3	2.8
CO-5	3	3	3	2	2	2	3	3	3	3	2	2	2.6
CO-6	3	3	3	3	3	3	3	2	2	3	3	3	2.8
Average	3	2.83	3	2.33	2.5	2.83	3	2.67	2.6	3	2.33	3	

Introduction to Data Science Lab

COURSE CODE: 22M17CI172
COURSE CREDITS: 02
CORE/ELECTIVE: CORE
L-T-P: 0-0-2

Pre-requisite: Python/ R Programming and Machine Learning

Course Objectives:

- Asking the correct questions and analyzing the raw data.
- Modelling the data using various complex and efficient algorithms.
- Visualizing the data to get a better perspective.
- Understanding the data to make better decisions and find the final result.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understanding the basics of data science	Familiarity
CO-2	Using versatile and flexible languages (Python and R programming) for supporting data science	Usage
CO-3	Using data processing for collecting and manipulating the data into a usable and desired form	Usage
CO-4	Using statistics to collect and analyze the numerical data in a large amount and finding meaningful insights from i	Usage
CO-5	Understanding linear algebra to represent, model, synthesize and summarize the complex data	Usage
CO-6	Using data visualization to easily access the huge amount of data in visuals	Usage

List of Experiments

S No	Description	Hours
1	Write a Python/R program to create a vector of a specified type and length. Create a vector of numeric, complex, logical, and character types of length 6. Write a Python/R program to add two vectors of integer type and length 3. Write a Python/R program to create a list containing a vector, a matrix, and a list and remove the second element Write a Python/R program to create a list containing a vector, a matrix, and a list and update the last element.	**03**
2	Write Python/R programs to solve the following tasks in both of them. Read numbers from a file, and print them out in sorted order. Read a text file, and count the total number of words. Read a text file, and count the total number of distinct words. Read a file of numbers, and plot a frequency histogram of them	**03**

3	**Statistical Data Analysis:** Measuring Central Tendency, Measuring Variance, Normal Distribution, Binomial Distribution, Poisson Distribution, Bernoulli Distribution, P-Value, Correlation, Chi-square Test, Linear Regression	**03**
4	**Data Visualization** Construct a revealing visualization of some aspect of your favorite data set, using: A well-designed table. A dot and/or line plot. A scatter plot. A heatmap. A bar plot or pie chart. A histogram. A data map.	**03**
5	**Data Visualization** Create ten different versions of line charts for a particular set of (x, y) points. Which ones are best and which ones worst? Explain why	**02**
6	**Data Visualization** Construct scatter plots for sets of 10, 100, 1000, and 10,000 points. Experiment with the point size to find the most revealing value for each data set	**02**
7	**Supervised Learning** K-Nearest Neighbors, Linear models, Naïve Bayes classifiers,Decision trees, Neural Networks	**03**
8	**Unsupervised Learning** Principal component analysis, Non-negative matrix factorization, K-means clustering, Agglomerative clustering, Comparison and evaluating clustering algorithms	**03**
9	**Working with text data** Rescaling data with tf-idf,clustering topic modeling and document	**02**
10	**Wrapping up:** Approaching a machine learning problem, Testing production systems, machine learning frameworks and packages, probabilistic programming, neural networks, scaling to larger datasets	**04**
Total Lab hours		**28**

Suggested Text Books:

- Python Data Science Hand Book JakeVanderPlas
- Introduction to machine learning with python: Andreas C. Miller and Sarah Guido
- Mastering Python for Data Science - by Samir Madhavan.
- R Programming for Data Science by Roger D. Peng

Reference Books:

- Python Data Science: Hands-on Learning for Beginners Kindle Editionby Travis Booth (Author)
- Data Science with R: A Step By Step Guide with Visual Illustrations & Examples by Andrew Oleksy
- R for Data Science: Import, Tidy, Transform, Visualize, and Model Data by Hadley Wickham (Author), Garrett Grolemund

EvaluationScheme:

1	Mid Sem. Evaluation	20 Marks
2	End Sem. Evaluation	20 Marks
3	Project-1	10 Marks
4	Lab Assessment	45 Marks
5	Lab Attendance	5 Marks
	Total	100 marks

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

Course outcomes (Data Science and Visualization Lab)	**PO-1**	**PO-2**	**PO-3**	**PO-4**	**PO-5**	**PO-6**	**PO-7**	**PO-8**	**PO-9**	**PO-10**	**PO-11**	**PO-12**	**Average**
CO-1	2	2	2	3	2	1	1	1	3	2	2	1	1.8
CO-2	3	3	3	3	3	1	1	2	2	2	1	1	2.1
CO-3	2	2	2	2	1	1	2	2	2	2	1	2	1.8
CO-4	3	3	3	3	2	1	1	3	2	3	3	2	2.4
CO-5	3	2	3	3	1	1	2	1	1	3	2	1	1.9
CO-6	2	2	1	2	2	2	3	2	1	2	3	1	1.9
Average	2.5	2.3	2.3	2.7	1.8	1.2	1.7	1.8	1.8	2.3	2	1.3	2.0

Mathematical Foundations for Data Science Lab

COURSE CODE: 22M17MA171
COURSE CREDITS: 2
CORE/ ELECTIVE: Elective
L-T-P: 0-0-2

Pre-requisite: Mathematical Foundations for Data Science (Theory)

Course Objectives:

1. To introduce students to the fundamental mathematical concepts of multivariate calculus and linear algebra required for data analysis & machine learning.
2. To provide an understanding for the Data science students on inferential statistical concepts to include sampling, estimation, parametric & non-parametric hypothesis testing, correlation & multiple regression.
3. To extend and familiarize the students with the concepts of Stochastic models with Bayesian inference.

Course Outcomes:

S. No.	Course Outcomes	Level of Attainment
CO-1	Understand the theoretical concepts of calculus of several variables.	Familiarity
CO-2	Understand the concept of matrix factorization and notion of vector space and their applications.	Assessment & Usage
CO-3	Understand random variables and their distributions in modeling random phenomenon	Assessment & Usage
CO-4	Understand optimization. And solve min/max problems of constrained	Assessment & Usage

List of Experiments

S. No.	Description	Hours
1	Find the gradient and Hessian for the given function of several variables. Determine the extrema of an unbounded f(x, y). Are the point(s) you have found maxima, minima, or saddle points?	2
2	Compute calculations of singular-value decomposition.	2
3	Generate two vectors x and y of n independent N(0,1) random variables. Illustrate their joint probability distribution through graph: y versus x	2
4	Generate a NxK matrix X, with K = 256 realizations of the stochastic process x[n], n = 1, 2, . . . , N. For a given value of N, and let each x[n] be N(0,1). Plot the ensemble average.	2
5	Write a program to demonstrate *Binomial* for the Given parameters.	2
6	Generate n random variables from N(0,1) and U(0,1) distributions respectively. Plot the histograms.	2
7	Calculate regression coefficients and the equation of regression line for the given data.	
8	Simulate the *distribution of the mean* for specified distribution – normal; obtain the parameters; Display histogram.	2

Approved in Academic Council held on 28 July 2021

9	From the frequency distribution table for a given date, write a program to calculate mean and standard deviation.	2
10	Hypothesis testing for two population means for the given data.	2
11	Non-parametric test: Kolmogorov-Smirnov normality test.	2
12	Implement the conjugate gradient descent algorithm for the given problem.	2
13	Obtain a Bayesian estimate of the mean and variance of the given data and compare it to the MLE.	
14	Write a code to solve the given constrained optimization problem with inequality Constraints.	2
	Total Lab hours	**28**

Suggested Resources:

a. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, ``Mathematics for Machine Learning,” Cambridge University Press, 2019.
b. Steven J. Leon and Lisette de Pillis, ``Linear Algebra with Applications,” Pearson, 10th ed, 2019.
c. Gene H. Golub and Charles F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, 4th Ed., 2013.
d. Andreas Antoniou, Wu-Sheng Lu, ``Practical Optimization: Algorithms and Engineering Applications,” Springer US, 2021.
e. S. L. Miller and D. G. Childers, Probability and Random Processes With Applications to Signal Processing and Communications. Elsevier Academic Press, Burlington, MA, 2004.
f. Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining: ``Introduction to Linear Regression Analysis”, John Wiley and Sons, Inc. NY, 2003.
g. W. R. Gilks, S. Richarson and D. Spiegelhalter, ``Markov chain Monte Carlo methods in Practice”, Chapman and Hall, 1996.

Evaluation Scheme:

1	Mid Sem. Evaluation	20 Marks
2	End Sem. Evaluation	20 Marks
3	Attendance	15 Marks
4	Lab Assessment	45 Marks
	Total	100 marks

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

CO/PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO10	PO11	PO12	Average
CO1	3	3	3	3	2	2	1	1	1	1	1	1	**1.83**
CO2	3	3	3	3	3	1	1	1	1	1	1	3	**2.00**
CO3	3	3	2	3	2	3	2	1	1	1	2	1	**2.00**
CO4	3	3	3	2	3	2	1	1	1	1	1	1	**1.83**
CO5	2	2	3	3	3	3	1	1	1	1	1	1	**1.83**
C06	2	3	3	3	2	2	2	2	2	2	2	2	**2.25**
Average	**2.67**	**2.83**	**2.80**	**2.80**	**2.60**	**2.20**	**1.20**	**1.00**	**1.00**	**1.00**	**1.20**	**1.40**	

Soft Computing

COURSE CODE: **22M11CI211**
COURSE CREDITS: 3
CORE/ELECTIVE: CORE
L-T-P: 3-0-0

Pre-requisite: None

Course Objectives:

1. Understanding basics of soft computing.
2. Learning fuzzy logic.
3. Learning neural networks.
4. Learning genetic algorithms.
5. Applying fuzzy logic, neural networks and genetic algorithm.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Basics of soft computing and understanding fuzzy logic.	Familiarity
CO-2	Learn and analyze neural networks.	Assessment
CO-3	Learn and analyze genetic algorithms.	Assessment
CO-4	Apply important algorithmic design paradigms and method soft computing.	Usage

Course Contents:

Unit	Contents	Lectures required
1	Introduction to Soft Computing and Neural Networks: Evolution of Computing: Soft Computing Constituents, From Conventional AI to Computational Intelligence: Machine Learning Basics	6
2	Fuzzy Logic: Fuzzy Sets, Operations on Fuzzy Sets, Fuzzy Relations, Membership Functions: Fuzzy Rules and Fuzzy Reasoning, Fuzzy Inference Systems, Fuzzy Expert Systems, Fuzzy Decision Making	8
3	Neural Networks: Machine Learning Using Neural Network, Adaptive Networks, Feed forward Networks, Supervised Learning Neural Networks, Radial Basis Function Networks: Reinforcement Learning, Unsupervised Learning Neural Networks, Adaptive Resonance architectures, Advances in Neural networks	10
4	Genetic Algorithms: Introduction to Genetic Algorithms (GA), Applications of GA in Machine Learning: Machine Learning Approach to Knowledge Acquisition.	5
5	Matlab/Python Lib: Introduction to Matlab/Python, Arrays and array operations, Functions and Files, Study of neural network toolbox and fuzzy logic toolbox, Simple implementation of Artificial Neural Network and Fuzzy Logic Recent Trands in deep learning, various classifiers, neural networks and genetic algorithm. Implementation of recently proposed soft computing techniques.	13
Total lectures		**42**

Approved in Academic Council held on 28 July 2021

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance – 5

1. Textbook(s): 1. Jyh:Shing Roger Jang, Chuen, Tsai Sun, EijiMizutani, Neuro: Fuzzy and Soft Computing, Prentice Hall of India, 2003.
2. References 1. George J. Klir, Bo Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall, 1995.

Course Outcomes (COs) contribution to the ProgrammeOutcomes(POs)

Course outcomes	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO- 10	PO- 11	PO- 12	Average
CO-1	2	2	2	2	2	1	1	1	2	2	2	2	**1.9**
CO-2	2	3	3	3	3	1	1	1	2	2	1	2	**2**
CO-3	2	2	2	2	3	1	1	1	2	2	1	2	**1.9**
CO-4	2	3	3	3	2	1	1	1	2	3	2	2	**2.1**
Average	**2**	**2.5**	**2.5**	**2.5**	**2.5**	**1**	**1**	**1**	**2**	**2.2**	**1.5**	**2**	

Deep Learning Techniques

COURSE CODE: **22M11CI212**
COURSE CREDITS: 3
CORE/ELECTIVE: CORE
L-T-P: 0-0-3

Pre-requisite: None

Course Objectives:

1. understand complexity of Deep Learning algorithms and their limitations
2. understand modern notions in data analysis oriented computing;
3. be capable of confidently applying common Deep Learning algorithms in practice
4. and implementing their own;
5. be capable of performing distributed computations;
6. be capable of performing experiments in Deep Learning using real-world data.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understand the concepts of Tensor Flow, its main functions, operations and the execution pipeline	Familiarity
CO-2	Implement deep learning algorithms, understand neural networks and traverse the layers of data abstraction which will empower the student to understand data more precisely	Assessment
CO-3	Learn topics such as convolutional neural networks, recurrent neural networks, training deep networks and high-level interfaces	Assessment
CO-4	Understand the language and fundamental concepts of artificial neural networks	Usage

Course Contents:

Unit	Contents	Lectures required
1	**Introduction to TensorFlow:** Computational Graph, Key highlights, Creating a Graph, Regression example, Gradient Descent, TensorBoard, Modularity, Sharing Variables, Keras**Perceptrons**: What is a Perceptron, XOR Gate	**8**
2	**Activation Functions :**Sigmoid,ReLU, Hyperbolic Fns,Softmax **Artificial Neural Networks** : Introduction, Perceptron Training Rule, Gradient Descent Rule	**8**
3	**Gradient Descent and Backpropagation**: Gradient Descent, Stochastic Gradient Descent, Backpropagation, Some problems in	**8**
	ANN **Optimization and Regularization** :Overfitting and Capacity, Cross Validation, Feature Selection, Regularization, Hyperparameters	
4	**Introduction to Convolutional Neural Networks**: Introduction to CNNs, Kernel filter, Principles behind CNNs, Multiple Filters, CNN applications **Introduction to Recurrent Neural Networks**: Introduction to RNNs, Unfolded RNNs, Seq2Seq RNNs, LSTM, RNN applications	**10**
5	**Deep Learning applications**: Image Processing, Natural Language Processing, Speech Recognition, Video Analytics	**8**
Total lectures		**42**

Approved in Academic Council held on 28 July 2021

Suggested Text Book(s):

1. Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016.

Suggested Reference Book(s):

2. 1. Bishop, C. ,M., Pattern Recognition and Machine Learning, Springer, 2006.
3. 2. Yegnanarayana, B., Artificial Neural Networks PHI Learning Pvt. Ltd, 2009.
4. 3. Golub, G.,H., and Van Loan,C.,F., Matrix Computations, JHU Press,2013.
5. 4. Satish Kumar, Neural Networks: A Classroom Approach, Tata McGraw-Hill Education,
6. 2004

Other useful resource(s):

1. https://nptel.ac.in/courses/106/106/106106184/
2. https://cse.iitkgp.ac.in/~adas/courses/dl_spr2020/syllabus.html
3. https://www.edx.org/learn/deep-learning
4. http://introtodeeplearning.com/

Evaluation Scheme:

S. No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance - 5

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

Course outcomes	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Average
CO-1	2	2	2	2	2	1	1	1	2	2	2	2	**1.8**
CO-2	2	3	3	3	3	1	1	1	2	2	1	2	**2**
CO-3	2	2	2	2	3	1	1	1	2	2	1	2	**1.8**
CO-4	2	3	3	3	2	1	1	1	2	3	2	2	**2.1**
Average	**2**	**2.5**	**2.5**	**2.5**	**2.5**	**1**	**1**	**1**	**2**	**2.3**	**1.5**	**2**	**1.9**

Big Data Analytics

COURSE CODE: **22M11CI213**

COURSE CREDITS: 3

CORE/ELECTIVE: ELECTIVE

L-T-P: 0-0-3

Introduction

This course introduces basic technology (algorithms, architectures, systems) and advanced research topics in connection with large-scale data management and information extraction techniques for big data. The course will start by introducing the fundamentals of Big data and cover modern distributed database systems and algorithms and Big data systems adopted in industry and science applications. Distributed storage and parallel processing and architectures that support data analytics will be examined, and students will learn how to implement a distributed data processing system. The course will also cover critical topics in mining and knowledge discovery of big data, with applications in social analytics, cyber security, and information networks, among others that are already in public eye.

Course Objectives (Post-conditions) Knowledge objectives:

1. Understand the need for Bid Data Analytics
2. Master the concepts of large scale file systems and map reduce framework
3. Master the concepts of mining data streams
4. Master the concepts of Link analysis and frequent item sets discovery from Big data
5. Master the concepts of clustering for streams and parallelism.

Application objectives:

1. Understand and apply the Big Data Analysis techniques to real world problems including social analytics, cyber security, and information networks, among others that are already in public eye.
2. Being able to describe and apply the Data Analytics lifecycle to Big Data projects including set up of smart cities.

Expected Student Background (Preconditions)

1. Good knowledge of Statistics
2. Working knowledge of databases
3. Good knowledge of data structures and algorithms

Topics Outline:

Serial Number	Topics	Hours
1	Introduction to Big Data: Big data time line, Why this topic is relevant now? Is big data fad? Where using big data makes a difference? Introduction to statistical modeling and machine learning, Ordinary data processing versus big data processing: Challenges and opportunities	3
2	Map Reduce and the New Software Stack: Distributed File Systems, Map Reduce, Algorithms Using Map Reduce, Complexity Theory for Map Reduce	3
3	Mining Data Streams: The Stream Data Model, Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream, Estimating Moments and Windowing, Decaying Windows	5

4	Link Analysis: Page Rank and Efficient Computation of Page Rank, Topic-Sensitive Page Rank, Link Spam, Hubs and Authorities	5
5	Frequent Item sets from Big Data: The Market-Basket Model, Market Baskets and the A-Priori Algorithm, Handling Larger Datasets in Main Memory, Limited-Pass Algorithms, Counting Frequent Items in a Stream	7
6	Clustering for Big Data: Introduction to Clustering Techniques, Hierarchical Clustering, Clustering in Non-Euclidean Spaces, Clustering for Streams and Parallelism	8
7	Mining Social Network Graphs: Social Networks as Graphs, Clustering of Social-Network Graphs, Direct Discovery of Communities, Partitioning of Graphs, Finding Overlapping Communities, Neighborhood Properties of Graphs	6
8	Recommendation Systems: A Model for Recommendation Systems, Content-Based Recommendations, Collaborative Filtering and Dimensionality Reduction	5
	Total Lectures	42

References

1. Anand Rajaraman and Jeffery David Ullman, Mining of Massive Datasets, Cambridge University Press, 2012
2. Jared Dean, Big Data, Data Mining and Machine Learning, Wiley Big data Series, 2014
3. Judith Hurwitz, Alan Nugent, Fern Halper and Marica Kaufman, Big Data for Dummies, Wiley Press, 2013

Evaluation Scheme:

S No	Examination	Marks
1	T1	15
2	T2	25
3	T3	35
4	* Internal Marks	25

*Internal Marks Breakdown:
Assignments 9 marks (3x3) Quizzes 12 marks (3x4) Regularity 4 Marks

Soft Computing Lab

COURSE CODE: **22M17CI271**
COURSE CREDITS: 1
CORE/ELECTIVE: CORE
L-T-P: 0-0-1

Pre-requisite: Basic concepts of Programming

Course Objectives:

1. Understanding basics of soft computing.
2. Learning fuzzy logic.
3. Learning neural networks.
4. Learning genetic algorithms.
5. Applying fuzzy logic, neural networks and genetic algorithm.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Basics of soft computing and understanding fuzzy logic.	Familiarity
CO-2	Learn and analyze neural networks.	Assessment
CO-3	Learn and analyze genetic algorithms.	Assessment
CO-4	Apply important algorithmic design paradigms and method soft computing.	Usage

List of Experiments

S No	Description	Hours
1	Understanding Tensor flow	**2**
2	Understanding keras	**2**
3	Implement Union, Intersection, complement and difference operations on Fuzzy sets.	**2**
4	Create Fuzzy relation by Cartesian product of any two Fuzzy sets and perform Max- Min composition of any two Fuzzy re	**2**
5	Build Logistic Regression Classifier using Neural Networks	**2**
6	Build Deep neural network for classification	**2**
7	Build neural network for Regression	**2**
8	Build a classification model using different parameter initialization techniques.	**2**
9	Build classification model using Mini Batch gradient and Stochastic Gradient techniques.	**4**
10	Implement Genetic algorithm.	**4**
Total Lab hours		**24**

Approved in Academic Council held on 28 July 2021

Suggested/Resources:

1. Textbook(s): 1. Jyh:Shing Roger Jang, Chuen, Tsai Sun, EijiMizutani, Neuro: Fuzzy and Soft Computing, Prentice Hall of India, 2003.
2. References 1. George J. Klir, Bo Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall, 1995.

Evaluation Scheme:

1	Mid Sem. Evaluation	20 Marks
2	End Sem. Evaluation	20 Marks
3	Attendance	15 Marks
4	Lab Assessment	45 Marks
	Total	100 marks

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

CO/PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO1 0	PO1 1	PO1 2	Average
CO1	3	3	3	3	2	2	1	1	1	1	1	1	**1.83**
CO2	3	3	3	3	3	1	1	1	1	1	1	3	**2.00**
CO3	3	3	2	3	2	3	2	1	1	1	2	1	**2.00**
CO4	3	3	3	2	3	2	1	1	1	1	1	1	**1.83**
Average	**2.67**	**2.83**	**2.80**	**2.80**	**2.60**	**2.20**	**1.20**	**1.00**	**1.00**	**1.00**	**1.20**	**1.40**	

Deep Learning Techniques Lab

COURSE CODE: **22M17CI272**
COURSE CREDITS: 2
CORE/ELECTIVE: CORE
L-T-P: 0-0-2

Pre-requisite: None

Course Objectives:

1. Apply knowledge of mathematics, science, and engineering fundamentals to solve problems in Computer science and Engineering.
2. Identify, formulate and analyze complex problems.
3. Design system components or processes to meet the desired needs within realistic constraints for the public health and safety, cultural, societal and environmental considerations.
4. Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data for valid conclusions.
5. Select and apply modern engineering tools to solve the complex engineering problem.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO1	Understand the characteristics and types of artificial neural network and remember working of biological Neuron and Artificial Neural Network.	Familiarity
CO2	Apply learning algorithms on perceptron and apply back propagation learning on Neural Network.	Assessment
CO3	Apply Feedback NN and plot a Boltzmann machine and associative memory on various application	Assessment
CO4	Apply different types of auto encoders with dimensionality reduction and regularization.	Assessment
CO5	Design Convolutional Neural Network and classification using Convolutional Neural Network.	Usage
CO6	Solve sequence learning problem and implement long short term memory and gated recurrent units	Usage

List of Experiments

S No	Description	Hours
1	To Write a program to implement Perceptron.	**2**
2	To write a program to implement AND OR gates using Perceptron.	**2**
3	To implement Crab Classification using pattern net	**2**
4	Write a program to implement classification of linearly separable Data with a perceptron	**2**

5	To study Long Short Term Memory for Time Series Prediction	**2**
6	To study Convolutional Neural Network and Recurrent Neural Network	**2**
7	To study ImageNet, GoogleNet, ResNet convolutional Neural Networks	**2**
8	To study the use of Long Short Term Memory / Gated Recurrent Units to predict the stock prices based on historic data	**2**
9	Recurrent Neural Networks	**2**
10	Deep Generative Models	**2**
11	Linear Regression and Logistic Regression	**2**
12	Decision Tree	**2**
13	SVM (Support Vector Machine)	**2**
14	Naive Bayes	**2**
15	kNN (k- Nearest Neighbors)	**2**
16	K-Means	**2**
17	Random Forest	**2**
18	Gradient Boosting Algorithms	**2**
Total Lab hours		**36**

Suggested Text Book(s):

1. Goodfellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016.

Suggested Reference Book(s):

1. Bishop, C. ,M., Pattern Recognition and Machine Learning, Springer, 2006.
2. Yegnanarayana, B., Artificial Neural Networks PHI Learning Pvt. Ltd, 2009.
3. Golub, G.,H., and Van Loan,C.,F., Matrix Computations, JHU Press,2013.
4. Satish Kumar, Neural Networks: A Classroom Approach, Tata McGraw-Hill Education,
5. 2004

Other useful resource(s):

1. https://nptel.ac.in/courses/106/106/106106184/
2. https://cse.iitkgp.ac.in/~adas/courses/dl_spr2020/syllabus.html
3. https://www.edx.org/learn/deep-learning
4. http://introtodeeplearning.com/

Evaluation Scheme:

1	Mid Sem. Evaluation	20 Marks
2	End Sem. Evaluation	20 Marks
3	Attendance	15 Marks
4	Lab Assessment	45 Marks
	Total	100 marks

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

CO/PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO1 0	PO1 1	PO1 2	Average
CO1	3	3	3	3	2	2	1	1	1	1	1	1	**1.83**
CO2	3	3	3	3	3	1	1	1	1	1	1	3	**2.00**
CO3	3	3	2	3	2	3	2	1	1	1	2	1	**2.00**
CO4	3	3	3	2	3	2	1	1	1	1	1	1	**1.83**
CO5	2	2	3	3	3	3	1	1	1	1	1	1	**1.83**
CO6	2	3	3	3	2	2	2	2	2	2	2	2	**2.25**
Average	**2.67**	**2.83**	**2.80**	**2.80**	**2.60**	**2.20**	**1.20**	**1.00**	**1.00**	**1.00**	**1.20**	**1.40**	

Big Data using Hadoop Lab

COURSE CODE: **22M17CI273**

COURSE CREDITS: 1

CORE/ELECTIVE: ELECTIVE

L-T-P: 0-0-2

Pre-requisites: An understanding in practicalities of big data and Hadoop to any programming language (Preferably,C)

Course Objective

The lab course provides a complete description of the inner working of big data and Hadoop. The main focus is on the design of big data techniques. The course also aims to convey the language specifications, use of file management tasks, and map-reduce behind the design of algorithms. It builds an understanding of various techniques like map-reduce, pig Latin scripts, etc.

Course Outcomes (COs):

S No	Course outcomes	Level of Attainment
CO-1	Perform setting up and Installing Hadoop in various operating modes	Familiarity
CO-2	Implement file management tasks in Hadoop	Familiarity
CO-3	Running a Map-Reduce Paradigm	Computational skills
CO-4	Run Pig then write Pig Latin scripts	Technical skills

List of Experiments

S No	Topic	Hours
1	Implement the following data structures in Java: Linked list, stacks, queues, Set and Map	**2**
2	Perform setting up and Installing Hadoop in its three operating modes: Standalone, Pseudo distributed, Fully distributed Use web-based tools to monitor your Hadoop setup.	**4**
3	Implement the following file management tasks in Hadoop: • Adding files and directories Retrieving files Deleting files Hint: A typical Hadoop workflow creates data files (such as log files) elsewhere and copies them into HDFS using one of the above command-line utilities.	**4**
4	Run a basic Word Count Map Reduce program to understand Map Reduce Paradigm	**4**
5	Write a Map Reduce program that mines weather data. Weather sensors collecting data every hour at many locations across the globe gather a large volume of log data, which is a good candidate for analysis with Map Reduce since it is semi-structured and record-oriented	**4**
6	Implement Matrix Multiplication with Hadoop Map Reduce	**2**
7	Install and Run Pig then write Pig Latin scripts to sort, group, join, project, and filter your data	**4**
8	Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, functions, and indexes	**4**
	Total Lab Hours	**28**

Approved in Academic Council held on 28 July 2021

Evaluation Scheme

S No	Exam	Coverage/Scope of Examination	Marks
1	Mid Term Test	Viva and Written Exam	20
2	End Term Test	Viva and Written Exam	20
3	Lab Records		15
4	Teacher Assessment	(Quality and quantity of experiment performed, learning laboratory skills)	30
5	Attendance and Participation in lab		9+6=15
	Total		**100**

Reference Books:

- Jain, Vinay Kumar. Big Data and Hadoop. Khanna Publishing, 2017.

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

Course outcomes	PO - 1	PO - 2	PO - 3	PO - 4	PO - 5	PO - 6	PO - 7	PO - 8	PO - 9	PO - 10	PO - 11	PO - 12	Average
CO-1	3	2	2	2	1	3	3	2	1	3	2	1	2.1
CO-2	3	3	2	2	1	3	3	1	3	3	3	1	2.3
CO-3	3	3	2	2	2	2	2	1	2	3	3	2	2.3
CO-4	3	3	3	2	2	2	3	2	2	1	2	2	2.3
CO-5	3	3	2	1	1	3	3	1	3	3	3	1	2.3
Average	3	2.8	2.2	1.8	1.4	2.6	2.8	1.4	2.2	2.6	2.6	1.4	2.2

Elective Courses

Approved in Academic Council held on 28 July 2021

Advanced Mobile and Distributed Computing

COURSE CODE: 18M1WCI115
COURSE CREDITS: 03
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

Pre-requisite Course: Students are expected to have a solid grasp of the fundamentals of computer system including a basic understanding of the operation of the computer, especially CPU.

Course Objectives (Post-conditions) Knowledge objectives:

1. It focuses on leading system architecture, high speed interconnects, and programming models that have been used for parallel and distributed computing environments.
2. This course will cover fundamental principles, advanced algorithms, and engineering tradeoffs in building large-scale parallel and distributed computing systems, as well as the high speed interconnects that bring them all together.
3. Students will gain an in-depth understanding of research and development in HPC, and their impact on computational sciences.

Application objectives:

1. Grasp a thorough understanding on the advances of technologies, system architecture and communication architecture that propelled the growth of parallel and distributed computing systems
2. 2 Accomplish a good understanding of principles and practices in high speed interconnects
3. Gain an appreciation on the challenges and opportunities faced by parallel systems, and cloud computing environments

Course Outcomes:

S No	Course outcomes	Level of Attainment
CO-1	You will Understand Hardware Architecture - Symmetric Multiprocessing , Distributed Shared Memory, Multicomputers	Familiarity
CO-2	You will Understand Software Architecture - Client server architecture , 3-tier architecture , N-tier architecture, Peer-to-peer	Familiarity
CO-3	Cluster computing	Familiarity
CO-4	Grid computing	Familiarity
CO-5	Virtualisation and Cloud Computing	Technical skills
CO-6	Recent trends in processor technologies - Superscalar processors, Multi-core processors, Embedded processors	Familiarity

Course Contents / Lecture Plan:

S No	Topic	Hrs
1	Hardware Architecture - Symmetric Multiprocessing ,Distributed Shared Memory, Multicomputers	7
2	Software Architecture - Client server architecture , 3-tier architecture , N-tier architecture, Peer-to-peer	7

3	Cluster computing	7
4	Grid computing	7
5	Virtualisation and Cloud Computing	7
6	Recent trends in processor technologies – Superscalar processors, Multi-core processors, Embedded processors	7
	Total	42

Evaluation Scheme:

S.No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Assignments, Quizzes, and Attendance	25	Entire Semester	Assignment (3) - 10 Quizzes (2) - 10 Attendance - 5

Text Book(s):

1. Distributed Systems : Concepts and Design - George Coulouris, Jean Dollimore, and Tim Kindberg. 2. Introduction to Reliable Distributed Programming - Rachid Guerraoui and Louis Rodrigues, Springer-Verlag, Berlin, Germany, 2006. 3. Distributed Systems: Principles and Paradigms - Andrew Tanenbaum and Maarten

Reference Book(s)

A. Research Papers

Course Material:

Lecture presentations and assignments will be posted on the student resource from time to time.

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

S No	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Weightage
CO-1	H	M	M	H	M	H	H	H	L	H	M	M	81%
CO-2	H	H	H	M	L	H	H	H	M	H	M	M	83%
CO-3	H	H	H	L	M	H	M	H	M	H	M	H	83%
CO-4	H	H	H	M	M	M	H	M	M	M	H	H	83%
CO-5	H	H	H	M	M	M	M	M	L	M	M	H	75%
CO-6	H	H	H	H	M	H	H	M	M	H	H	M	89%
	100%	96%	96%	78%	3%	89%	93%	85%	59%	81%	85%	85%	

Artificial Intelligence Techniques

COURSE CODE: 22M1WCI132
COURSE CREDITS: 03
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

Introduction

The course introduces the students with a number of modern meta-heuristic techniques taken from the area of natural computation for solving hard optimization problems. The student will be encouraged to solve real world optimization problems using the learnt techniques and to do a guided research on methodological issues of the techniques. Topics are likely to be drawn from the following list: planning, scheduling, games, search, reasoning (constraint- based, model-based, spatial, temporal), knowledge representation, decision-making under uncertainty, reinforcement learning, agents, and foundations.

Knowledgeobjectives:

1. Gain both a wide and a deep knowledge of the Topics(s) taught in the current instance of the course.
2. Improve their skills at navigating through, and critically examining, the scientific literature on the taught Topics(s).

Application objectives:

The course will teach you to apply the basic principles, models, and algorithms of AI to recognize, model, and solve problems in the analysis and design of information systems. Through home assignments you will be able to analyze the structures and algorithms of a selection of techniques related to searching, reasoning, machine learning, and language processing.

Expected Student Back ground (Preconditions)

Students are expected to have a sound understanding of artificial intelligence techniques and the fundamentals of computing before commencing this course. These include planning, knowledge representation, advanced search techniques, and heuristic approaches to problem solving.

Topics Outline:

S No	Topics	Hrs
1	Logic and Artificial Intelligence	7
2	Automated planning and scheduling	6
3	Optimization Techniques in Artificial Intelligence	10
4	Uncertain Knowledge and Reasoning	11
5	Expert Systems and Artificial Intelligence	8
6	Swarm Intelligence and Natural level processing	6
		48

References

1. Z. Michalewicz and D.B. Fogel, *How to Solve It: Modern Heuristics* ,Springer-Verlag, 2000.
2. E. Aartsand J.K. Lenstra, *Local Search in Combinatorial Optimization* ,Priceton University Press, 2003.
3. D. Corne, M. Dorigo,F. Glover, *New Ideas in Optimisation* , McGraw-Hill, 1999.
4. J.J. Schneider and S. Kirpatric, *Stochastic Optimization*, Springer-Verlag, 2006.

Approved in Academic Council held on 28 July 2021

Evaluation Scheme:

S No	Examination	Marks
1	T-1	15
2	T-2	25
3	T-3	35
4	*Internal Marks	25

CRYPTOGRAPHY AND INFORMATION SECURITY

COURSE CODE: 22M1WCI135

COURSE CREDITS: 3

CORE/ELECTIVE: ELECTIVE

L-T-P: 3-0-0

Pre-requisite Course: Computer Networks, Discrete Mathematics

Description & Rationale:

In this module, students will be provided with the contemporary knowledge of Cryptography techniques along with the security of information systems. It focuses on introducing the concepts, principles, techniques and methodologies required to design and assess the security of information exchange over complex networks, information systems and applications.

Objectives:

- To understand how cryptography can be used as an effective tool in providing assurance concerning privacy and integrity of information.
- To recognize the concept of encryption/decryption.
- To describe the different types of ciphers along with the identification of the characteristics of a good cipher.
- To provide skills to design security protocols for solving security problems.
- To develop skills necessary to help organizations in designing, testing and implementing well-planned information security measures for for information systems.

Course Outcomes:

S No	Course outcomes (Cryptography and Information System Security)	Level of Attainment
CO-1	To understand a variety of generic security threats and vulnerabilities, and identify &analyze particular security problems for a given application.	Familiarity
CO-2	To gain knowledge on the notions of various types cryptography techniques.	Familiarity
CO-3	To understand and analyze the various cryptography techniques, their principles and applications.	Familiarity and Assessment
CO-4	To identify and analyze the applications of security techniques and technologies in solving real life security problems.	Usage
CO-5	To understand, analyze and evaluate the security of information systems.	Familiarity, Usage and Assessment
CO-6	To analyze and interpret the mechanisms to provide security for communicating information.	Assessment

Course Contents / Lecture Plan:

S No	Topics	Hrs
1	**Introduction:** Security Attacks: Motives, vulnerabilities, Defense strategies and techniques, Various Attacks- DoS, DDoS, Session Hijacking and Spoofing, Phishing, Buffer Overflow, Format String Attacks, SQL Injection, Malicious Software, Prevention and Detection, Data Protection, Response, Recovery and Forensics.	4
2	**Basics of Cryptography:** Symmetric Cipher Model, Substitution Techniques, Transportation Techniques, Other Cipher Properties- Confusion, Diffusion, Block andStream Ciphers.	4
3	**Secret Key Cryptography:** Data Encryption Standard (DES), Strength of DES, Block Cipher, Design Principles and Modes of Operations, Triple DES.	6
4	**Public Key Cryptography:** Principles of Public Key Cryptosystems, RSA Algorithm, Diffie- Hellman Key Exchange algorithm	5
5	**Cryptographic Hash Functions:** Applications of Cryptographic Hash Functions, Secure Hash Algorithm, Message Authentication Codes – Message Authentication Requirements and Functions, HMAC, Digital signatures, Digital Signature Schemes, Authentication Protocols, Digital Signature Standards. Authentication Applications-Kerberos, Key Management and Distribution, X.509 Directory Authentication service, Public Key Infrastructure.	7
6	**Electronic Mail Security**-Pretty Good Privacy, S/MIME**, Operating System Protection**-Memory and Address protection, File Protection Mechanism, User Authentication. **Database Security-**Security Requirement, Reliability and Integrity, Sensitive data, Multilevel Databases.	6
7	**IP Security:** Overview, Architecture, Authentication Header, Encapsulating SecurityPayload, Combining security Associations, Internet Key Exchange, WebSecurity: Web Security Considerations, Secure Sockets Layer andTransport Layer Security, Electronic Payment.	5
8	**Intrusion Detection Systems and Firewalls** Intruders, Intrusion Detection, Password Management, Firewalls- Need, Characteristics, Types of Firewalls, Placement of Firewalls, Firewall Configuration, Trusted systems.	5
	Total	**42**

Tools and Technologies: None

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Tutorials / Assignments, Quizzes, Attendance	25	Entire Semester	Assignment (5) - 9 Quizzes (5) - 12 Attendance - 4

Text Book(s):

T1: 'Cryptography and Network Security- Principles and Practices' by William Stallings, 8th Edition, Prentice Hall Publication

T2: 'Network security and Cryptography' by Bernard Menezes, Ist Edition, Cengage Learning Publication

T3: 'Computer Security- Principles and Practice' by William Stallings, Ist Edition, Pearson Education

Reference Book(s):

R1: 'Network Security Essentials' by William Stallings, 4th Edition, Pearson Publication R2: 'Applied Cryptography' by Bruce Schneier, Edition 2001, Wiley & Sons Inc

R3: 'Cryptography and Network', 2nd edition, by Behrouz A Fourouzan, Debdeep Mukhopadhyay, TMH.

Web Resources: https://onlinecourses.nptel.ac.in/noc18_cs07/preview
https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=8013 https://link.springer.com/journal/145

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

S No	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Weightage
CO-1	H	M	M	H	M	H	H	H	L	H	M	M	81%
CO-2	H	H	H	M	L	H	H	H	M	H	M	M	83%
CO-3	H	H	H	L	M	H	M	H	M	H	M	H	83%
CO-4	H	H	H	M	M	M	H	M	M	M	H	H	83%
CO-5	H	H	H	M	M	M	M	M	L	M	M	H	75%
CO-6	H	H	H	H	M	H	H	M	M	H	H	M	89%
	100%	96%	96%	78%	3%	89%	93%	5%	59%	81%	85%	85%	

Data Warehousing and Data Mining

COURSE CODE: 22M1WCI131

COURSE CREDITS: 03

CORE/ELECTIVE: ELECTIVE

L-T-P: 3-0-0

Pre-requisites: Data Structures, Compilers, Operating Systems, Computer Networks, Machine Learning and Genetic Algorithms.

Course Objective:

1. To describe the concept of Data warehouse & its attributes
2. To study different data warehouse models, architectures and implementation
3. To understand the basic concept of data mining and its functionality
4. To understand the concept of classification techniques and its implementation
5. To understand the concept of association rules, different techniques and implementation details
6. To understand the concept of cluster analysis, anomaly detection and its usage and implementation details

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	To describe the concept of Data Warehouse & its attributes	Assessment
CO-2	To study different data warehouse models, architectures and implementation	Assessment
CO-3	To understand the basic concept of data mining and its functionality	Assessment
CO-4	To understand the concept of classification techniques and its implementation	Assessment
CO-5	To understand the concept of association rules, different techniques and implementation details	Assessment
CO-6	To understand the concept of cluster analysis, anomaly detection and its usage and implementation details	Usage

Course Contents:

Unit	Contents	Lectures required
1	Introduction: Concepts of Data Warehouse and Data Mining including its functionalities, stages of Knowledge discovery in database(KDD) , Setting up a KDD environment, Issues in Data Warehouse and Data Mining, Application of Data Warehouse and Data Mining	5
2	Architecture: DBMS vs. Data Warehouse, Data marts, Metadata, Multidimensional data model, Data Cubes, Schemas for Multidimensional Database: Stars, Snowflakes and Fact Constellations, Data Warehouse Architecture, Distributed and Virtual Data Warehouse, Data Warehouse Manager, OLTP, OLAP, MOLAP, HOLAP, types of OLAP, servers	8

3	Introduction: Data Mining, Motivation, Challenges, Origins of Data Mining, Data Mining Tasks, Data: Types of Data, Data Quality, Data Pre-processing, Measures of Similarity and Dissimilarity, Exploring Data: Iris Data Set, Summary Statistics, Visualization, OLAP and Multi dimensional Data Analysis	**5**
4	Classification: Basic Concepts and Preliminaries, Approach to Solving a Classification Problem, Decision Tree Induction, Model Over fitting, Evaluating Performance of Classifier Alternative Techniques: Rule-Based Classifier, Nearest Neighbour Classifiers, Artificial Neural Network (ANN), Support Vector Machine (SVM), Ensemble Methods, Class Imbalance Problem, Multiclass Problem	**8**
5	Association Analysis: Basic Concepts and Problem Definition, Frequent Item set Generation, Rule Generation, Representation of Frequent Item sets, FP-Growth Algorithm, Evaluation of Association Patterns, Handling Categorical Attributes , Handling Continuous Attributes, Handling a Concept Hierarchy, Sequential Patterns, Subgraph Patterns	**8**
6	Cluster Analysis: Basic Concepts, Characteristics of Data, Clusters, Partitional Clustering, Agglomerative Hierarchical Clustering, Prototype-Based Clustering, Density-Based Clustering, Graph-Based Clustering, Cluster Evaluation Anomaly Detection: Preliminaries, Statistical Approaches, Proximity-Based Outlier Detection, Density-Based Outlier Detection, Clustering-Based Techniques	**8**
Total lectures		**42**

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Tutorials / Assignments, Quizzes, Attendance	25	Entire Semester	Assignment (5) - 9 Quizzes (5) - 12 Attendance - 4

Suggested Text Book(s):

1. Data Mining Concepts and Techniques J. Han and M. Kamber Morgan Kaufmann, 2006, ISBN 1-55860-901-6
2. Introduction to Data Mining Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson Education (Addison Wesley), 0-321-32136-7

Suggested Reference Book(s):

1. Mining Massive data sets Anand Rajaram, Jure Leskovec and Jeff Ullman Cambridge University Press
2. https://onlinecourses.nptel.ac.in/noc18_cs14/preview
3. https://www.coursera.org/specializations/data-mining.
4. https://www.futurelearn.com/courses/data-mining-with-weka

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

Course outcomes	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	eightage
CO-1	3	2	1	2	2	2	2	2	2	1	1	2	1.8
CO-2	3	3	3	3	3	3	2	3	3	1	1	3	2.6
CO-3	3	3	3	3	2	2	1	2	3	2	3	3	2.5
CO-4	3	1	2	1	2	1	1	2	1	2	2	2	1.7
CO-5	3	3	3	3	2	3	1	2	3	2	3	3	2.6
CO-6	3	2	1	2	2	1	2	2	2	1	1	2	1.8
Weightage	3	2.3	2.2	2.3	2.2	2	1.5	2.3	2.2	1.5	1.8	2.5	

Data Visualization

COURSE CODE: 22M1WCI136
COURSE CREDITS: 03
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

Pre-requisite: Basic knowledge of high school mathematics and reasonable programming experience

Course Objectives:

- Asking the correct questions and analyzing the raw data.
- Modeling the data using various complex and efficient algorithms.
- Visualizing the data to get a better perspective.
- Understanding the data to make better decisions and finding the final result.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understanding the basics of data science	Familiarity
CO-2	Using versatile and flexible languages (Python and R programming) for supporting data science	Usage
CO-3	Using data processing for collecting and manipulating the data into usable and desired form	Usage
CO-4	Using data visualization to easily access the huge amount of data in visuals	Usage
CO-5	Using statistics to collect and analyze the numerical data in a large amount and finding meaningful insights from i	Usage
CO-6	Understanding linear algebra to represent, model, synthesize, and summarize the complex data	Usage

Course Contents:

Unit	Contents	Lectures required
1	Data Science an Introduction: Computer Science, Data Science, and Real Science, What is Data Science? Need for Data Science, Data Science Components, Tools for Data Science, Data Science Lifecycle, Applications of Data Science	**6**
2	Python and R Programming for Data Science for Data Science: Introduction to Python Programming (Python Basics, Python Data Structures, Python Programming Fundamentals, Working with Data in Python, Working with NumPy, Pandas, SciPy, and Matplotlib). Introduction to R Programming (R basics, Data structures in R, R Programming fundamentals, Working with Data in R, Stings and Dates in R,)	**6**

3	Data Processing: Data Operations, Data cleansing, Processing CSV Data, Processing JSON Data, Processing XLS Data, Relational databases, NoSQL Databases, Date and Time, Data Wrangling, Data Aggregation, Reading HTML Pages, Processing Unstructured Data, Word tokenization, Stemming and Lemmatization	**8**
4	Statistical Data Analysis: Measuring Central Tendency, Measuring Variance, Normal Distribution, Binomial Distribution, Poisson Distribution, Bernoulli Distribution, P- Value, Correlation, Chi-square Test, Linear Regression	**8**
5	Linear Algebra: Visualizing Matrix Operations (Matrix Addition, Matrix Multiplication, Applications of Matrix Multiplication, Identity Matrices and Inversion, Matrix Inversion and Linear Systems, Matrix Rank) Factoring Matrices (Why Factor Feature Matrices, LU Decomposition and Determinants), Eigen values and Eigenvectors (Properties of Eigen values, Computing Eigen values), Eigen value Decomposition (Singular Value Decomposition, Principal Components Analysis)	**8**
6	Data Visualization: Chart Properties, Chart Styling, Box Plots, Heat Maps, Scatter Plots, Bubble Charts, 3D Charts, Time Series, Geographical Data, Graph Data	**6**
Total lectures		**42**

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Tutorials / Assignments, Quizzes, Attendance	25	Entire Semester	Assignment (5) - 9 Quizzes (5) - 12 Attendance - 4

Suggested Text Book(s):
- Data Science from Scratch by Joel Grus
- Data Science for Dummies by Lillian Pierson and Jake Porway
- An Introduction to Statistical Learning by Gareth James , Daniela Witten , et al.

Reference Book(s):
- An Introduction to Probability and Statistics by V.K. Rohatgi & A.K. Md. E. Saleh, Wiley, (2008), 3rd ed.
- Introduction to Probability Theory and Statistical Inference by H.J. Larson, John Wiley & Sons, (2005) 3rd ed.

Other useful resource(s):
- https://nptel.ac.in/courses/110/106/110106064/
- https://onlinecourses.nptel.ac.in/noc18_cs28/preview

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

Course outcomes	PO- 1	PO- 2	PO- 3	PO- 4	PO- 5	PO- 6	PO- 7	PO- 8	PO- 9	PO- 10	PO- 11	PO- 12	Weightage
CO-1	2	2	2	3	2	1	1	1	3	2	2	1	1.8
CO-2	3	3	3	3	3	1	1	2	2	2	1	1	2.1
CO-3	2	2	2	2	1	1	2	2	2	2	1	2	1.8
CO-4	3	3	3	3	2	1	1	3	2	3	3	2	2.4
CO-5	3	2	3	3	1	1	2	1	1	3	2	1	1.9
CO-6	2	2	1	2	2	2	3	2	1	2	3	1	1.9
Weightage	2.5	2.3	2.3	2.7	1.8	1.2	1.7	1.8	1.8	2.3	2	1.3	

Data Storage Technologies

COURSE CODE: 22M1WCI134
COURSE CREDITS: 03
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

The course focuses on designing, implementing, and administering storage for today's evolving organizations, getting under the hood of the technologies that enable performance, resiliency, availability, recoverability, and simplicity.

Course Outcomes:

S No	Course outcomes	Level of Attainment
CO-1	Students will learn about the Storage Media and Technologies	Familiarity
CO-2	To introduce students to the Memory Hierarchy and design the Hardware and Software for access	Assessment
CO-3	To study the Large Storages and Scalability issues	Assessment
CO-4	To analysis and discuss the Storage Partitioning, Storage System Design, Legacy Systems	Usage
CO-5	Students will analysis the Performance, Reliability, and Security issues	Usage
CO-6	Design and develop the model for any real world application using Data Storage Technologies	Usage

Course Contents / Lecture Plan:

S No	Topics	Hrs
1	Storage Media and Technologies: Magnetic, Optical and Semiconductor Media, Techniques for read/write Operations, Issues and Limitations	8
2	Usage and access Positioning in the Memory Hierarchy, Hardware and Software Design for Access, Performance issues	8
3	Large storages Hard Disks, Networked Attached Storage, Scalability issues	8
4	Storage architecture Storage Partitioning, Storage System Design, Caching, Legacy Systems	8
5	Storage area networks Hardware and Software Components, Storage Clusters/Grids. Storage QoS– Performance, Reliability, and Security issues	10
	Total	42

Approved in Academic Council held on 28 July 2021

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered uptoT-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Tutorials / Assignments, Quizzes, Attendance	25	Entire Semester	Assignment (2) -10 Quizzes (2) - 10 Attendance -5

TEXT BOOKS

The Complete Guide to Data Storage Technologies for Network-centric Computing Paper back– Import, Mar 1998 by Computer Technology Research Corporation

REFERENCE BOOKS

Data Storage Networking: Real World Skills for the Comp TIA Storage by Nigel Poulton,2014

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

S No	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Total
CO-1	2	2	2	3	2	1	1	1	3	2	2	1	1.8
CO-2	3	3	3	3	3	1	1	2	2	2	1	1	2.1
CO-3	2	2	2	2	1	1	2	2	2	2	1	2	1.8
CO-4	3	3	3	3	2	1	1	3	2	3	3	2	2.4
CO-5	3	2	3	3	1	1	2	1	1	3	2	1	1.9
CO-6	2	2	1	2	2	2	3	2	1	2	3	1	1.9
Average Score	2.5	2.3	2.3	2.7	1.8	1.2	1.7	1.8	1.8	2.3	2	1.3	2.0

Introduction to Statistical Learning

COURSE CODE: 22M1WCI133
COURSE CREDITS: 3
CORE/ELECTIVE: OPEN ELECTIVE
L-T-P: 3-0-0

Pre-requisite: Basics of programming and mathematics.

Course Objectives:

- To understand the fundamentals of statistical learning and, analyzing regression and classification.
- To understand and analyze pre-processing data and model selection.
- To understand and analyze non-linear and tree based regression.
- To understand and analyze Support Vector Classifier and Machines, and unsupervised learning.

Course Outcomes:

S. No.	Course Outcomes	Level of Attainment
CO-1	Understanding statistical learning and, analyzing regression and classification.	Familiarity and usage
CO-2	Analyzing resampling, model selection and regularization.	Usage
CO-3	Analyzing non-linear and tree based regression.	Usage
CO-4	Analyzing Support Vector Machines and unsupervised learning	Usage

Course Contents:

Unit	Contents	Lectures required
1	**Introduction to statistical learning:** Prediction accuracy, Model interpretability, Supervised and Unsupervised learning, Assessing model accuracy	**3**
2	**Regression and Classification:** Regression- Linear, MultiLinear and kNN, Classification-Logistic regression, Linear and quadratic discriminant analysis, and kNN	**6**
3	**Resampling, Model Selection and Regularization:** Cross- validation, the bootstrap, Subset selection, Shrinkage methods, Dimension reduction methods, considerations in high dimension	**7**
4	**Non-linear Regression:** Polynomial regression, Step functions, Regression splines, Smoothing splines, Local regression, Generalized additive methods, **Tree Based Regression:** the basics of decision trees, Bagging, Random forest and Boosting,	**7**
5	**Support Vector Machines:** Maximal margin classifier, Support Vector Classifier and Support Vector Machines, **Unsupervised Learning:** The challenges, Principal Component Analysis, Clustering methods	**6**
Total lectures		**42**

Approved in Academic Council held on 28 July 2021

Suggested Text Book(s):

- Gareth, James, Witten Daniela, Hastie Trevor, and Tibshirani Robert. *An introduction to statistical learning: with applications in R*. Spinger, 2013.

Suggested Reference Book(s):

1. Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*. Vol. 1, no. 10. New York: Springer series in statistics, 2001.
2. Kulkarni, Sanjeev, and Gilbert Harman. *An elementary introduction to statistical learning theory*. Vol. 853. John Wiley & Sons, 2011.
3. MELLO, RODRIGO F., and Moacir Antonelli Ponti. *Machine Learning: A Practical Approach on the Statistical Learning Theory*. Springer, 2018.

Evaluation Scheme:

S No	Ex am	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour	Syllabus covered upto T-1
2	T-2	25	1.5 Hours	Syllabus covered upto T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance – 3 Participation in class-2

Course Outcomes (COs) contribution to the Programmed Outcomes (POs)

Course Outcomes	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Average
CO-1	3	3	3	3	3	3	1	1	2	1	3	2	**2.33**
CO-2	3	3	3	3	3	3	2	1	2	1	3	2	**2.42**
CO-3	3	3	3	3	3	3	2	1	2	1	3	2	**1.42**
CO-4	3	3	3	3	3	3	1	1	2	1	3	2	**2.33**
Average	**3**	**3**	**3**	**3**	**3**	**3**	**1.5**	**1**	**3**	**1**	**3**	**3**	

Knowledge-Based AI: Cognitive Systems

COURSE CODE: 22M1WCI233
COURSE CREDITS: 3
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

Pre-requisite: Discrete Mathematics, Programming Languages

Course Objectives:

The course aims to familiarize the students with the basic concepts as well as with the state- of-the-art research literature in knowledge-based artificial intelligence. After successful completion of this course, students will be able to:

1. Understand concepts, methods, and prominent issues in knowledge-based artificial intelligence
2. Learn specific skills and abilities needed to apply those concepts to the design of knowledge-based AI agents
3. Understand the relationship between knowledge-based artificial intelligence and the study of human cognition.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understanding the basic concepts of knowledge-based AI & Cognitive Systems	Familiarity
CO-2	Understanding various fundamental concepts & techniques related to AI as well as cognitive systems.	Familiarity
CO-3	Learning & implementing various reasoning methods such as common sense reasoning, analogical reasoning,etc	Usage
CO-4	Designing & creating as well as metacognition knowledge	Assessment

Course Contents:

Unit	Contents	Lectures Required
1	**Introduction:** Introduction to KBAI & Cognitive Systems, Where Knowledge-Based AI fits into AI as a whole - Cognitive systems: what are they? - AI and cognition: how are they connected?	**5**
2	**Fundamentals:** Semantic Networks - Generate & Test - Means-Ends Analysis - Problem Reduction - Production Systems	**5**
3	**Common Sense Reasoning**: Frames - Understanding - Common Sense Reasoning - Scripts	**4**
4	**Planning & Learning:** Logic, Planning, Learning by Recording Cases –Incremental Concept Learning -Classification - Version Spaces & Discrimination Trees	**6**
5	**Analogical Reasoning:** Case-Based Reasoning - Explanation-Based Learning - Analogical Reasoning	**5**
6	**Visuospatial Reasoning:** Constraint Propagation - Visuospatial Reasoning	**5**
7	**Design & Creativity:** Configuration - Diagnosis - Design - Creativity	**6**
8	**Metacognition:** Learning by Correcting Mistakes - Meta- Reasoning - AI Ethics	**6**
	Total Lecture Hours	**42**

Approved in Academic Council held on 28 July 2021

Suggested Text Book(s):

- Russell, Stuart, and Peter Norvig. "Artificial intelligence: a modern approach." (2002).
- Vernon, David. Artificial cognitive systems: A primer. MIT Press, 2014.

Suggested Reference Book(s):

- Rothman, Denis. *Artificial Intelligence by Example: Develop machine intelligence from scratch using real artificial intelligence use cases*. Packt Publishing Ltd, 2018.

Other useful resources (s):

Link to NPTEL course contents:
- https://nptel.ac.in/courses/106/105/106105077/ Link to topics related tocourse:
- https://www.youtube.com/watch?v=0Alhb1HDsNw
- https://www.youtube.com/watch?v=JMUxmLyrhSk
- https://www.udacity.com/course/knowledge-based-ai-cognitive-systems--ud409

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 Hour.	Syllabus covered up to T-1
2	T-2	25	1.5 Hours	Syllabus covered up to T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance – 3 Participation in class-2

Course Outcomes (COs) contribution to the Program Outcomes (POs)

Course outcomes (Knowledge-Based AI: Cognitive Systems)	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Average
CO-1	2	2	2	2	2	1	1	1	2	2	2	2	1.8
CO-2	2	3	3	3	3	1	1	1	2	2	1	2	2
CO-3	2	2	3	3	3	1	1	1	2	2	1	2	1.9
CO-4	2	3	3	3	2	1	1	1	2	3	1	2	2
Average	2	2.5	2.8	2.8	2.5	1	1	1	2	2.3	1.3	2	1.9

Medical Image Analysis

COURSE CODE: 22M1WCI232
COURSE CREDITS: 03
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

This course focuses on classical and machine learning based techniques for medical image analysis.

Course Outcomes:

S No	Course outcomes (Medical Image Analysis)	Level of Attainment
CO-1	To introduce students to the outline the Medical Image Analysis.	Familiarity
CO-2	To interpret and analyse the images in a quantitative way.	Assessment
CO-3	To apply the machine Learning and Deep Learning Approaches for medical image analysis,	Assessment
CO-4	To apply the learned techniques for novel disease diagnosis and prognosis.	Usage

Course Contents / Lecture Plan:

S No	Topics	Hrs.
1	Introduction to image processing and medical imaging modalities, denoising and enhancement	5
2	Tissue and Cell Segmentation: clustering, active contours and level sets based approaches	5
3	Medical Image alignment: rigid and deformable registration	5
4	Fundus Image analysis, Retinal Vessel Segmentation	5
5	MRI image analysis and segmentation, 3D brain reconstruction from and MRI slices and analysis	5
6	Microscopic image analysis and interpretation	5
7	X-Ray and CT image segmentation, diagnosis and prognosis of various diseases	6
8	Correlation between different medical imaging modalities and conversions, augmenting clinical measurements with medical imaging modalities for diseases diagnosis and prognosis	6
	Total	42

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage/Scope of Examination
1	T-1	15	1 Hr.	Syllabus covered upto Mid Semester
2	T-2	25	1.5 Hrs	Full Syllabus
2	T-3	35	2 Hrs	Full Syllabus
3	Regularity, Assignment and Quiz	25	Entire Semester	Regularity-4, Assignments- 9, Quiz – 12

Text Books

1. Prince, J. L., & Links, J. M. (2006). Medical imaging signals and systems. Upper Saddle River, NJ: Pearson Prentice Hall.
2. Suetens, P. (2017). Fundamentals of medical imaging. Cambridge university press.

Self-Learning Material:

Medical Image Processing by Prof. Jeff Orchard at University of Waterloo: https://cs.uwaterloo.ca/~jorchard/cs473/CS473/Welcome.htm

Course Outcomes (COs) contribution to the Programme Outcomes(POs)

S No	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Total
CO-1	3	1	1	1	2	2	1	2	2	1	2	3	**1.8**
CO-2	3	2	1	2	2	2	1	2	2	1	3	2	**1.9**
CO-3	2	2	3	2	2	1	3	2	1	1	1	1	**1.8**
CO-4	2	2	3	2	2	1	3	2	1	1	1	1	**1.8**
Average Score	**2.5**	**1.8**	**2**	**1.8**	**2**	**1.5**	**2**	**2**	**1.5**	**1**	**1.8**	**1.8**	**1.8**

Natural Language Processing

COURSE CODE: 22M1WCI236
COURSE CREDITS: 03
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

Pre-requisite:

Knowledge of CS's Core and System Courses on Programming, Data Structure, Algorithms, Artificial Intelligence.

Course Objectives (Post-conditions) Knowledge objectives:

On completion of the course, you are expected to develop

a) Ability to analyze natural language phenomenon and create models mimicking human behavior.
b) Ability to design and build NLP application systems in real life based on societal needs.
c) Ability to evaluate NLP software systems.
d) Ability to develop software systems that learn from real life data
e) Ability to re-engineer components of a NLP software in a multi-lingual environment and evolving language scenario.
f) Ability to translate a specification into a design, and then realize that design practically, using an appropriate processing tools and methods.

Topics Outline:

S No	Topic	Hrs
1	Introduction, NLP applications, Ambiguities, Role of syntax, semantics, context, world-knowledge & pragmatics, Rule-based approach, vs. Statistical approach to analysis	3
2	Fundamental aspects of English grammar; Contrastive examples from Hindi, Linguistic resources, Morphological Analysis and synthesis,	4
3	Syntactic Analysis, TN, RTN, ATN, CFG, Probabilistic CFG,	7
4	Collocations, Multi-word expressions, Named entities,	4
5	Language modeling, Markov models, Hidden Markov Models,Parts of speech tagging,	8
6	Word Sense Disambiguation, Pronoun reference disambiguation,	5
7	Text similarity, Text categorization, Text Summarization,	3
8	Sentiment Analysis,	2
9	Information Retrieval, Question Answering,	3
10	Natural Language Generation, Machine Translation.	3
	Total	**42**

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage/Scope of Examination
1	T-1	15	1 Hr.	Syllabus covered upto Mid Semester
2	T-2	25	1.5 Hrs	Full Syllabus
2	T-3	35	2 Hrs	Full Syllabus
3	Regularity, Assignment and Quiz	25	Entire Semester	Regularity-4, Assignments- 9, Quiz – 12

References

1. Daniel Jurafsky& James H.Martin, " Speech and Language Processing", Pearson Education (Singapore) Pte. Ltd., 2002.
2. James Allen, "Natural Language Understanding", Pearson Education, 2003
3. C.D. Manning and H. Schutze, "Foundations of Statistical Natural Language Processing", MIT Press, 1999.
4. Daniel M. Bikel and ImedZitouni (ed.), "Multilingual Natural Language Processing Applications: From Theory to Practice", Pearson Education, 2013

Reinforcement Learning

COURSE CODE: 22M1WCI235
COURSE CREDITS: 3
CORE/ELECTIVE: ELECTIVE
L-T-P: 3-0-0

Pre-requisite: Probability & Statistics, Machine Learning

Course Objectives:

The course aims to familiarize the students with the basic concepts as well as with the state- of-the-art research literature in deep reinforcement learning. After successful completion of this course, students will be able to:

1. Structure of a reinforcement learning problem
2. Understand and apply basic RL algorithms for simple sequential decision-making problems in uncertain conditions.
3. Evaluate the performance of the solution
4. Interpret state-of-the-art RL research and communicate their results.

Course Outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understanding the basic concepts of Reinforcement Learning	Familiarity
CO-2	Understanding Probability Primer Techniques in detail.	Familiarity
CO-3	Learning & implementing various methods & techniques such as TD methods, Monte Carlo, Markov Decision Process, etc.	Usage
CO-4	Learning & understanding policy gradients along with its advantages and disadvantages.	Assessment

Course Contents:

Unit	Contents	Lectures Required
1	**Introduction:** Origin and history of Reinforcement Learning research. Its connections with other related fields and with different branches of machine learning.	**2**
2	**Probability Primer:** Axioms of probability, concepts of random variables, PMF, PDFs, CDFs, Expectation. Concepts of joint and multiple random variables, joint, conditional and marginal distributions. Correlation and independence.	**4**
3	**Markov Decision Process:** Introduction to RL terminology, Markov property, Markov chains, Markov reward process (MRP). Introduction to and proof of Bellman equations for MRPs along with proof of existence of solution to Bellman equations in MRP. Introduction to Markov decision process (MDP), state and action value functions, Bellman expectation equations, optimality of value functions and policies, Bellman optimality equations.	**6**

4	**Prediction & Control by Dynamic Programming:** Overview of dynamic programming for MDP, definition and formulation of planning in MDPs, principle of optimality, iterative policy evaluation, policy iteration, value iteration, Banach fixed point theorem, proof of contraction mapping property of Bellman expectation and optimality operators, proof of convergence of policy evaluation and value iteration algorithms, DP extensions.	**6**
5	**Monte Carlo methods for Model Free Prediction &Control :**Overview of Monte Carlo methods for model free RL, First visit and every visit Monte Carlo, Monte Carlo control, On policy and off policy learning, Importance sampling.	**6**
6	**TD Methods:** Incremental Monte Carlo Methods for Model Free Prediction, Overview TD(0),TD(1) and TD(λ), k-stepestimators, unified view of DP, MC and TD evaluation methods, TD Control methods - SARSA, Q-earning and their variants.	**6**
7	**Function Approximation Methods:** Getting started with the function approximation methods, Revisiting risk minimization, gradient descent from Machine Learning, Gradient MC and Semi-gradient TD(0) algorithms, Eligibility trace for function approximation, After states, Control with function approximation, Least squares, Experience replay in deep Q- Networks.	**6**
8	**Policy Gradients:** Getting started with policy gradient methods, Log-derivative trick, Naive REINFORCE algorithm, bias and variance in Reinforcement Learning, Reducing variance in policy gradient estimates, baselines, advantage function, actor-critic methods.	**6**
Total Lecture Hours		**42**

Suggested Text Book(s):

- Sutton, Richard S., and Andrew G. Barto. "Reinforcement learning: An introduction." (2011).
- Leon-Garcia, Alberto. "Probability, statistics, and random processes for electrical engineering." (2017).

Suggested Reference Book(s):

- Murphy, Kevin P. *Machine learning: a probabilistic perspective*. MIT press, 2012.

Other useful resources (s):

1. Link to NPTEL course contents:
- https://nptel.ac.in/courses/106/106/106106143/

2. Link to topics related tocourse:
- https://www.youtube.com/watch?v=JgvyzIkgxF0
- https://www.youtube.com/watch?v=LzaWrmKL1Z4

Evaluation Scheme:

S. No	**Exam**	**Marks**	**Duration**	**Coverage / Scope of Examination**
1	T-1	15	1 Hour.	Syllabus covered up to T-1
2	T-2	25	1.5 Hours	Syllabus covered up to T-2
3	T-3	35	2 Hours	Entire Syllabus
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes(2) -10 Attendance – 3 Participation in class-2

Course Outcomes (COs) contribution to the Program Outcomes (POs)

Course outcomes (Reinforcement Learning)	PO-1	PO-2	PO-3	PO-4	PO-5	PO-6	PO-7	PO-8	PO-9	PO-10	PO-11	PO-12	Average
CO-1	2	2	2	2	2	1	2	2	2	2	1	2	1.8
CO-2	2	3	3	3	3	1	3	2	2	2	1	2	2.3
CO-3	2	3	3	3	3	2	3	2	2	2	1	2	2.3
CO-4	2	3	3	3	2	1	2	3	2	3	2	2	2.3
Average	2	2.8	2.8	2.8	2.5	1.3	2.5	2.3	2	2.3	1.3	2	2.2

Social and Information Network Analysis

COURSE CODE: 22M1WCI234
COURSE CREDIT: 2
CORE/ELECTIVE: ELECTIVE
L-T-P: 2-0-0

Pre-requisites: Graph Theory, Probability

Course Objectives:

1. To analyze the large networks on models and algorithms that abstracts their basic properties.
2. To explore how to practically analyze large-scale network data and how to reason it through models for network structure and evolution
3. To do the friend prediction in online social networks
4. To identify the functional modules in biological networks and disease outbreak detection.

Course outcomes:

S No	Course Outcomes	Level of Attainment
CO-1	Understand what constitutes a social network	Familiarity
CO-2	Represent networks in graph-theoretic language	Assessment
CO-3	Design effective and reliable network research projects	Assessment
CO-4	Configure network and attribute data for standard software packages	Assessment
CO-5	Comprehend the reasons for using, and principles of, permutation tests	Assessment
CO-6	Apply centrality measures appropriately	Assessment

Course Contents:

	Topics	Topics Outcomes	No of Lectures
1	Introduction	1. Understand what constitutes a social network 2. Identify and describe different levels of analysis 3. Formulate problems in terms of network variables	3
2	Mathematic al Foundations	1. Represent networks in graph-theoretic language 2. Identify paths, walks, trails and components 3. Formulate networks in matrix terms 4. Compute and interpret multiplication of adjacency matrices	4
3	Research Design	1. Design effective and reliable network research projects 2. Identify sources and boundaries of network data 3. Understand and minimize the effects of data error	4
4	Data Collection	1. Identify sources of network data 2. Design effective network questionnaires 3. Mine archival and electronic sources for network data	3
5	Data Management	1. Configure network and attribute data for standard software packages 2. Apply elementary transformations to matrix data Extract and reconfigure network and attribute data	4

6	Multivariate Techniques Used in Network Analysis	1. Represent one- and two-mode data in a two- dimensional map 2. Cluster data into groups using hierarchical clustering 3. Correctly interpret the information contained in the clusters and maps	4
7	Visualization	1. Visualize networks with or without node attributes in a meaningful way 2. Embed edge characteristics in network diagrams 3. Represent network change over time graphically	4
8	Testing Hypotheses	1. Comprehend the reasons for using, and principles of, permutation tests 2. Formulate testable hypotheses at the dyadic, monadic and whole-network level 3. Understand when SIENA and exponential random graph models may be appropriate	4
9	Characterizing Whole Networks	1. Calculate and interpret cohesion measures in a whole network 2. Undertake a triad census 3. Compute and evaluate measures of transitivity, reciprocity and clustering	4
10	Centrality	1. Apply centrality measures appropriately 2. Interpret the results of a centrality analysis on undirected, directed and valued data 3. Understand the limitations and constraints of the standard centrality measures	4
11	Subgroups	1. Understand the similarities and differences of the main approaches in detecting cohesive subgroups 2. Select appropriate methods given the size and nature of the network 3. Perform a cohesive subgroup analysis on a variety of types of network	4
			Total=42

Evaluation Scheme:

S No	Exam	Marks	Duration	Coverage / Scope of Examination
1	T-1	15	1 hour.	Syllabus covered upto Test-1
2	T-2	25	1.5 hours	Syllabus covered upto Test-2
3	T-3	35	2 hours	Syllabus covered upto Test-3
4	Teaching Assessment	25	Entire Semester	Assignment (2) - 10 Quizzes (2) -10 Attendance - 5

Approved in Academic Council held on 28 July 2021

Suggested Text Book(s):

1. Analyzing Social Networks 1st Edition by Stephen P Borgatti, Martin G. Everett, Jeffrey C. Johnson, SAGE Publications Ltd; 1 edition

Suggested Reference Book(s):

1. Influence and Behavior Analysis in Social Networks and Social Media by Mehmet Kaya, Reda Alhajj, Lecture Notes in Social Networks, Springer International Publishing
2. Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining by Nitin Agarwal, Nima Dokoohaki, Serpil Tokdemir, Lecture Notes in Social Networks, Springer International Publishing

Course Outcomes (COs) contribution to the Programme Outcomes (POs)

Course outcomes	PO- 1	PO- 2	PO- 3	PO- 4	PO- 5	PO- 6	PO- 7	PO- 8	PO- 9	PO-10	PO- 11	PO- 12	Average
CO-1	3	3	3	2	2	3	2	2	2	3	1	3	2.4
CO-2	3	3	3	2	3	2	3	2	2	3	1	3	2.5
CO-3	3	3	3	2	2	3	1	2	3	3	1	3	2.4
CO-4	3	3	3	2	3	3	2	2	3	3	1	3	2.6
CO-5	3	3	3	2	3	3	2	2	3	3	1	3	2.6
CO-6	3	3	3	2	3	3	2	2	1	3	1	3	2.4
Average	3	3	3	2	2.7	2.8	2	2	2.3	3	1	3	