SRI SATHYA SAI INSTITUTE OF HIGHER LEARNING

(Deemed to be University)

Syllabus for
M.Sc. (Data Science and Computing)

(Effective from the batch 2022-23 onwards)

Prasanthi Nilayam – 515 134

Sri Sathya Sai District, Andhra Pradesh, Ph.: 08555 287239; Fax: 286919

Website: www.sssihl.edu.in ; Email: registrar@sssihl.edu.in

M.Sc. (Data Science and Computing)

INTRODUCTION

Data Science has grown to be a domain of scientific study due to the deluge of data generated and acquired through various means. Data driven scientific discovery has contributed a lot to scientific investigation. Important contributions of data acquisition, visualization and analytics with tools from Machine Learning is seen in domains like Business Intelligence, Financial services, Climate Modeling, Weather forecasting, Medical, Chemo, Bio, Onco informatics etc., and the list goes on.

This programme is designed specifically for graduates in Computer Science and Computer Applications. Having Degrees like B.Sc.(Computer Science), BCA and B.Tech./B.E. in Computer Science.

In order to equip the students to continue with higher studies in academic disciplines for Ph.D., the candidates undergoing the course should also be comfortable to take the National Level qualifying examinations like UGC NET, GATE etc.

Programme Specific Objectives:

The proposed syllabus aims to achieve the following objectives:

1. To produce good human beings who are skilled in learning from data with fair and ethical means to produce meaningful applications for societal harmony and goodness.
    a. Through courses like Awareness and Moral Classes to impart righteous living
    b. Through Integral Education to live with everyone and care for society
2. To produce manpower trained to understand the science in learning from data and the needed computing skills to develop practical solutions.
    a. Through foundational courses in mathematics, statistics and computer science like Applied Linear Algebra, Optimization Techniques, Inferential Statistics, Computer Organization & Design and Design and Analysis of Algorithms.
    b. Through core courses in Data Science like Machine Learning, Deep Learning, Artificial Intelligence and Natural Language Processing.

    c. Through software labs like Data Engineering and lab components for various courses.
    d. Through implementation of a data science and computing software project.
    e. Through specialized elective courses like Machine Learning Operations, Reinforcement Learning etc.
3. To train young minds to be industry ready. This is achieved by offering courses like Software Lab in Data Engineering, Reinforcement Learning, Machine Learning Operations etc.
4. To develop programming and problem solving skills. This is achieved by offering Lab components for different subjects.

Programme Specific Outcomes:

Upon the completion of the programme, a student must be

- Grounded in the roots of morality and ethics, ready to serve society
- Balanced in theoretical knowledge and practical skill of data science and computing to draw insights from data
- Able to take up research or a higher academic degree in Data Science or Computer Science
- Able to do the industry role of Data Analyst or Data Engineer or Data Scientist
- Able to develop AI solutions for selected real world problems from data
- Able to appreciate the core values and philosophy of Sri Sathya Sai Institute of Higher Learning.
- Able to imbibe Core values in life and lead the life as propounded by Bhagawan Sri Sathya Sai Baba.

The course structure and syllabus provides foundational, core, advanced and working knowledge in Mathematics, Statistics and Computer Science.

All the subjects are to be awarded 4 credits except Software Lab for Data Visualization.

For some of the subjects the credits are split between Theory and Practical based on the necessity. For 1 credit of practical 2 periods are allocated.

A few subjects are purely practical as they are intended to improve programming skill of the students in a specific language or platform. Eight periods are allotted for a four credits practical course.

In order to facilitate development of skill in problem solving and to provide exposure to applications of the concepts learnt in a given Theory subject a facility for

Tutorial/Practical is also provided within the curriculum. One or two periods per week is provided for Tutorial/Practical for every subject based on the requirement.

In order to cater to specialization, elective courses are provided in the areas of Reinforcement Learning, Machine Learning Operations (MLOps), Information Retrieval, Combinatorial Graph Theory, Robotics, Topological Data Analysis (TDA) etc. All electives are of 4 credits. Based on necessity the credits may be split between Theory and Practical.

* * *

DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE
SCHEME OF INSTRUCTION AND EVALUATION

M.Sc. (Data Science and Computing)

(Effective from 2022-23 batch and onwards)

Paper Code	Title of the Paper	Credits	Hours	Modes of Evaluation	Types of Papers	Maximum Marks
Semester I						
MDSC-101	Applied Linear Algebra	3	3	IE2	T	100
MDSC-101(P)	Practicals: Applied Linear Algebra	1	2	I	P	50
MDSC -102	Inferential Statistics	3	3	IE2	T	100
MDSC-102(P)	Practicals: Inferential Statistics	1	2	I	P	50
MDSC -103	Optimization Techniques	3	3	IE2	T	100
MDSC-103(P)	Practicals: Optimization Techniques	1	2	I	P	50
MDSC -104	Computer Organization and Architecture	4	4	IE2	T	100
MDSC -105	Design and Analysis of Algorithms	4	4	IE2	T	100
MDSC-106	Software Lab for Data Visualization	2	4	I	P	50
PAWR-100	Awareness Course – I: Education for Life	1	2	I	T	50

		23 Credits	**29** Hours			**750** Marks

Semester II						
MDSC-201	Regression Methods	**3**	3	IE2	T	100
MDSC-201(P)	Practicals: Regression Methods	**1**	2	I	P	50
MDSC-202	Multivariate Statistical Analysis	**3**	3	IE2	T	100
MDSC-202(P)	Practicals: Multivariate Statistical Analysis	**1**	2	I	P	50
MDSC-203	Machine Learning	**3**	3	IE2	T	100
MDSC-203(P)	Practicals: Machine Learning	**1**	2	I	P	50
MDSC-204	Big Data Analytics	**4**	4	IE2	T	100
MDSC-205	Software Lab in Data Engineering	**4**	8	I	P	100
MDSC-206	Mini Project	**2**	4	I	PW	50***
PAWR-200	Awareness Course – II: God, Society and Man	**1**	2	I	T	50
		23 Credits	**33** Hours			**750** Marks

Semester III						
MDSC-301	Stochastic Processes	**3**	3	IE2	T	100
MDSC-301(P)	Practicals: Stochastic Processes	**1**	2	I	P	50
MDSC -302	Deep Learning	**3**	3	IE2	T	100
MDSC -302 (P)	Practicals: Deep Learning	**1**	2	I	P	50
MDSC-303	Natural Language Processing	**3**	3	IE2	T	100
MDSC-303(P)	Practicals: Natural Language Processing	**1**	2	I	P	50
MDSC-304	Cloud Computing	**3**	3	IE2	T	100
MDSC-304(P)	Practicals: Cloud Computing	**1**	2	I	P	50
MDSC-403	Project Interim Review*	–	10	I	PW	50*
PAWR-300	Awareness Course –III: Guidelines for Morality	**1**	2	I	T	50
		17 Credits	**32** Hours			**700*** Marks

Semester IV						
MDSC -401	Elective - I	4	4**	IE2	T	100**
MDSC-402	Elective - II	4	4**	IE2	T	100**
MDSC-403	Project*	10	22	E2	PW	150*
PAWR-400	Awareness Course –IV: Wisdom for Life	1	2	I	T	50
		19 Credits	**32**** Hours			**400**** Marks

	GRAND TOTAL	**82** Credits	**123**** Hours			**2600**** Marks

Notes:

1. (*) Project work MDSC-403 will commence in 3rd semester and continue to 4th semester with the allocation of 50 Marks in third semester and 150 marks in the fourth semester towards the project work.

2. (*) For students undertaking projects (MDSC-403), the evaluation will be based on three components, viz.
    a. A preliminary review of an interim report in respect of the project work at the end of 3rd semester will be conducted for 50 marks and the marks allocated will be carried forward to 4th semester MDSC-403 for overall grading.
    b. A project Viva voce by a committee constituted by the Head of the Department as per regulations will be conducted for 50 marks in the 4th semester.
    c. An E2 type evaluation of the project report at the end of 4th semester will be for 100 marks.

3. (*)Total marks for the project will be 200 marks against total credits of 10 accounted for in 4th semester.

4. A number of electives have been identified and listed. These courses are identified with a special code. All these subjects are also allocated 4 credits each.

5. (**) Elective courses may have the credits split between Theory and Practical based on the chosen treatment of the subject and its requirement. Accordingly, the number of periods allocated for the subject (Theory + Practical) will vary. That will influence the total number of hours allocated for the subject and the total marks for the semester too.

6. The choice of electives being offered in each semester is at the discretion of the Head of the Department.

7. (***) The Mini-Project (MDSC-206) will be undertaken during the second semester by the candidate. This could be based on an internship (taken online/on-campus) with an industry or a field work etc., with a mentoring faculty from the department. Students will be asked to make a presentation along with a submission of the report of the work done towards the end of the second semester i.e., within one week before the last working day of the semester. This will be evaluated internally by a panel of minimum two faculty of the department constituted by HoD/Associate HoD. Total marks for the mini-project would be for 50 marks which will be equally distributed between the presentation and the report submitted.

Indicator	Legend
IE1	CIE and ESE ; ESE single evaluation
IE2	CIE and ESE ; ESE double evaluation
I	Continuous Internal Evaluation **(CIE)** only Note: 'I' does not connote 'Internal Examiner'
E	End Semester Examination **(ESE)** only Note: 'E' does not connote 'External Examiner'
E1	ESE single evaluation
E2	ESE double evaluation

Continuous Internal Evaluation (CIE) & End Semester Examinatio
(ESE)

Indicator	Legend
T	Theory
P	Practical
V	Viva voce
PW	Project Work
D	Dissertation

PS: Please refer to guidelines for 'Modes of Evaluation for various types of papers', and 'Viva voce nomenclature & scope and constitution of the Viva voce Boards'.

List of Electives:

1. **MDSC-AI:** Artificial Intelligence [4 Credits]
2. **MDSC-CGT:** Combinatorics and Graph Theory [4 credits]
3. **MDSC-RL:** Reinforcement Learning [3 credits]
   MDSC-RL(P): Practicals: Reinforcement Learning [1 credit]
4. **MDSC-MLO**: Machine Learning Operations [3 Credits]
   MDSC-MLO(P): Practicals: Machine Learning Operations [1 credit]
5. **MDSC – ATS:** Applied Time Series Analysis [4 credits]
6. **MDSC – IR**: Information Retrieval [4 credits]
7. **MDSC – NS:** Network Security [4 credits]
8. **MDSC –IoT:** Internet of Things [3 credits]
   MDSC – IoT (P): Practicals: Internet of Things Lab [1credit]
9. **MDSC – TDA:** Topological Data Analysis [4 credits]
10. **MDSC – LSP:** Linux System Programming [4 credits]
11. **MDSC-DS**: Distributed Systems [4 Credits]

Semester I

[MDSC-101] - Applied Linear Algebra		3 Credits
	Course Objective: The course focuses on iterative techniques for solving large sparse linear systems of equations which typically stem from the discretization of partial differential equations. In addition, computation of eigenvalues, least square problems and error analysis will be discussed.	
	Course Outcome : Develop the skill set to 1. explain and fluently apply fundamental linear algebraic concepts such as matrix norms, eigen- and singular values and vectors; 2. estimate stability of the solutions to linear algebraic equations and eigenvalue problems; recognize matrices of important special classes, such as normal, unitary, Hermitian, positive definite and select efficient computational algorithms based on this classification;	
Unit	**Topic**	**No. of Periods**
1	Review of Vectors and Matrices : Vector Addition, Linear Combination, Inner Product, Orthogonality, Norm, Cauchy-Schwarz inequality, Matrix addition & multiplication, Column space, Linear Independence, Rank of a Matrix, Gaussian Elimination, Determinant, Inverse, Adjoint, Cofactor, Null space, Rank-Nullity theorem	5
2	Applications of Matrices : Electric Circuits, Traffic flow, Graph theory, Social Networks, Dominance directed graph, Influential node	5
3	Applications of Linear Transformations : Linear Transformations, Gaussian Random Variable, Linear Transformation on Gaussian Random Vectors, Gaussian classification in Machine Learning	5
4	Applications of Eigenvalues and Eigenvectors : Introduction to Eigenvalues & Eigenvectors, Eigenvalue decomposition, Positive semi-definite matrices, Principal Component Analysis, Eigenfaces	8
5	Applications of Least Squares Solution : Orthogonality, Gram-Schmidt orthogonalization, Least Squares solution, Projection matrix, Least Norm solution, Pseudo-inverse, Rank Deficient Matrices, Linear Regression, Polynomial Fitting	8
6	Applications of Linear Minimum Mean Square Error (LMMSE) : Introduction to LMMSE, LMMSE estimate & Covariance Matrix, LMMSE estimation in Linear Systems, Auto Regression, Recommender System	8

Total: 39 Periods
Key Text Gilbert Strang, Linear Algebra and its applications, 4th Edition, Thomson Brooks/Cole, 2005.
REFERENCES 1. Stephen Boyd & Lieven Vandenberghe, Introduction to Applied Linear Algebra - Vectors, Matrices and Least Squares, Cambridge University Press, 2018 2. Philip N. Klein, Coding the Matrix - Linear Algebra through Applications to Computer Science, Newtonian Press, 2013

[MDSC-101(P)]- Practicals: Applied Linear Algebra 1 Credits		
	Course Objective: To apply algorithms used in Applied Linear Algebra using a programming language like python	
	Course Outcome: 1. Ability to understand and develop linear algebra related functions 2. To analyze and discover characteristics of a dataset	
Unit	**Topic**	**No. of Periods**
1	The following programs are to be implemented in basic python - Dot Product of Vectors, Matrix Multiplication with basic python - Determinant of a Matrix, Matrix Inverse with basic python - Rank of a Matrix with basic python - Gaussian Elimination with basic python	10
2	The following are to be implemented using numpy - Eigenvalues, Eigenvectors, SVD - Eigenfaces - Linear Regression - Matrix Exponential	16
		Total: 26 Periods

Key Text: Leo Chin, Tanmay Dutta, NumPy Essentials, Packt Publishing, 2016
REFERENCES Umit Mert Cakmak, Mert Cuhadaroglu, Mastering numerical computing with Numpy, Packt Publishing, 2018

[MDSC-102] – Inferential Statistics			**3 Credits**
	Course Objectives: To enable students ● To understand and make inferences based on relations found in a sample of the given population ● To understand and appreciate individual statistical test and it's nuances of working ● Acquire an understanding of the concepts of sampling distribution, statistical reliability and hypothesis testing, as well as the principles and procedures of the various tests of significance		
	Course Outcome : Develop the skill set to 1. Write a computer program to carry out data analyses 2. Interpret the output of statistical analysis 3. Set up and perform hypothesis tests, interpret p-values, and report the results of the analysis in a way that is interpretable for the public		
Unit	**Topic**	**Details**	**No. of Periods**
1	Elements of Random Variables	Random Variables: Univariate, Bivariate random variables; Expectation, Variance of a random variable; Conditional Expectation, Covariance and Correlation; Moment Generating Functions, independence of random variables and the reproductive property of certain distributions; Special distributions: Binomial, exponential, Gamma and Normal distributions; Transformation of random variable: $aX + b$,X^2, e^{tX}, and $log\ X$; Convergence of Random variables: Convergence in distribution or in probability, Weak Law of Large Numbers and Central Limit Theorem	11
2	Estimation	An overview of statistical inference, Methods of point estimation: Maximum Likelihood Estimation, Method of Moments; Uniformly Minimum Variance Unbiased Estimators (UMVUE), Cramer-Rao Inequality and Decision-Theoretic Approach to Estimation; Confidence Intervals and Confidence Regions	14

3	Testing Hypotheses	Formulation of Testing Hypotheses, Neyman-Pearson Fundamental Lemma, Exponential Type Families, Uniformly Most Powerful Tests for Some Composite Hypotheses and applications; Likelihood Ratio Tests and its applications: Contingency Tables and Goodness-of-Fit Test; Decision-Theoretic Approach to Testing Hypotheses and relationship between Testing Hypotheses and Confidence Regions.	14
			Total: 39 Periods

Key Text(s):

George G. Roussas, An Introduction to Probability and Statistical Inference, Second Edition, Academic Press, 2015

Chapters: 3.1, 3.3, 4.1-4.3, 5, 6.1, 7-12.

References:

1. Paul G. Hoel, Sidney C. Port, Charles J. Stone, Introduction to Statistical Theory, Houghton Mifflin Company, BOSTON, 1971
2. G.Casella, R.L.Berger, Statistical Inference, Second Edition, Duxbury Advanced Series, 2001
3. Vijay K. Rohatgi, A. K. Md. Ehsanes Saleh, An Introduction to Probability and Statistics, Second Edition, A Wiley-Interscience Publication, John Wiley & Sons, Inc - 2001

[MDSC-102(P)] - Practicals: Inferential Statistics 1 Credit

Course Objective: To train the students on implementing different statistical tests in python.

Course Outcome: Develop the skill set to

1. Visualize the statistical techniques to explore data
2. suggest different statistical tests on a given real world data
3. write a computer program to carry out data analyses

Unit	Topic	Details	No. of Periods
1	Review of Python Libraries	Loading and reading of datasets; Fundamentals in NumPy, Pandas and Matplotlib	6

2	Descriptive Statistics	Mean, median, mode, variance and correlation matrix; Basic plots: Dot plot, Box-plot, Histograms, and Scatter plots; Distributions: Binomial, Gamma, Normal; Simulation of Central Limit Theorem	6
3.	Statistical Tests	Normality Tests: Shapiro-Wilk Test, D'Agostino's; $K^2 -$ Test and Anderson-Darling Test; Statistical Hypothesis Tests: Student's t-test, Paired Student's t-test, Chi-square test and F-test and Analysis of Variance Test (ANOVA)	6
4	Data Analysis	Different cases studies can be done using Datasets;	8
			Total: 26 Periods

Key Texts:

1. Phuong Vo.T.H , Martin Czygan, Ashish Kumar, and Kirthi Raman, Python: Data Analytics and Visualization, Published by Packt Publishing Ltd - 2017
2. NumPy User Guide, Retrieved July 7, 2022 from https://numpy.org/doc/stable/numpy-user.pd
3. Pandas User Guide, Retrieved July 7, 2022 from https://pandas.pydata.org/docs/pandas.pdf

	[MDSC-103] – Optimization Techniques 3 Credits
	Course Objective: ● To model and discuss the documented real-world applications. ● Study of mathematical programming algorithms. ● Apply the mathematical results and numerical techniques of optimization theory to concrete optimization problems
	Course Outcome: Develop the skill set to 1. Translate a real world problem statement to mathematical formulation of a specific type of optimization problem 2. Understand, identify and solve optimization problems using relevant technique

Unit	Title	Contents	No. of Periods
1	Linear Programming Modeling and their solutions	Introduction to Linear Programming Problem (LPP), Modeling of LPP, Graphical method, simplex method, Artificial Starting Solution Methods, Special cases: degeneracy, alternative optima, unbounded and infeasible solution, and Graphical sensitivity analysis	6
2	Duality and Post Optimal Analysis	LP-Duality, Primal-Dual Relationships, Economic Interpretation of duality, Dual Simplex and Generalized Simplex algorithms, and Post Optimal Analysis	6
3	Advanced Linear Programming	Simplex method fundamentals, Revised Simplex Method, Bounded-Variable Algorithm, Duality, Parametric programming	7
4	Integer Programming	Formulation and Applications, Cutting Plane Algorithm, and Branch and Bound Method.	6
5	Classical Optimization Techniques	Unconstrained problems: Necessary and sufficient conditions, and The Newton-Raphson Method; Constrained problems: equality constraints - Jacobi method and Lagrangean method; Inequality constraints - KKT conditions	7
6	Nonlinear Programming	Unconstrained algorithms: Direct search, Gradient methods; Constrained algorithms: separable, quadratic, chance constrained programs and Linear combination method	7
			Total: 39 Periods

Key Text(s):
Hamdy A.Taha, Operations Research - An Introduction, 10th Edition, Pearson Education, 2017

Chapters: 2.1, 2.2, 3.1-3.5 and 3.6.1, 4, 7, 9, 20, and 21

References:

1. L.R.Foulds, Optimization Techniques-An Introduction, 1st Edition, Springer-Verlag New York Inc., 1981
2. Edwin K.P.Chong, S.H.Zak, An Introduction to Optimization, 4th Edition, John Wiley & Sons, Inc., 2001
3. Boyd, Stephen, and Lieven Vanderberghe, Convex Optimization, Cambridge, UK: Cambridge University Press, 2004
4. Hillier, Lieberman, Introduction to Operations Research, Seventh Edition, The McGraw-Hill, 2001

	[MDSC-103 (P)] – Practicals: Optimization Techniques 1 Credit		
	Course Objective: To introduce the ability to program for different optimization techniques		
	Course Outcome: Develop the skill set to 1. solve various problems based optimization techniques using Python and Excel-solver 2. write algorithms for optimization methods such as Newton-Raphson, Gradient method and their variants 3. visualize the curves and surfaces in python environment 4. use python library *scipy.optimize*		
Unit	**Title**	**Contents**	**No. of Periods**
1	Review of Python Objects	Python list, dictionary and loops, functions; Basic plotting of curves and contour plots and 3-D plots	4
2	Solutions of LPP	Solving LP problems using Excel solver; Sensitivity Analysis using Excel solver; Solving LP problems using *scipy.optimize.linprog*	8
3	Optimization Algorithms	Implementation of Newton-Raphson method; Implementation of Gradient Ascent and Gradient Descent methods	4
4	Python Library *scipy.optimize*	Univariate function minimization; Constrained and Unconstrained minimization of multivariate scalar function; Global Optimization; Least-Square Optimization; Custom minimizers	10
			Total: 26 Periods

Key Text(s):

1. SciPy Reference Guide, Retrieved July 7, 2022 from https://docs.scipy.org/doc/scipy-1.7.1/scipy-ref-1.7.1.pdf

   Section 2.4 (Optimization - *scipy.optimize*)

2. Chistian Hill, Learning Scientific Programming with Python, Second Edition, Cambridge University Press, 2020
3. Hamdy A.Taha, Operations Research- An Introduction, 10^{th} Edition, Pearson Education - 2017

[MDSC-104] – Computer Organization and Architecture 4 Credits			
	Course Objective: To study and understand the basics of computer organization and architecture (CPU, memory, I/O).		
	Course Outcome : Develop the skill to 1. Evaluate the merits and pitfalls in computer performance measurements 2. Evaluate impact of ISA on cost/performance of computer design. 3. Suggest enhancement in the performance by exploiting Instruction Level Parallelism 4. Understand memory hierarchy and its impact on computers performance		
Unit	**Title**	**Contents**	**No. of Periods**
1	Introduction	Performance, the Power Wall, the Switch from Uniprocessors to Multiprocessors, Historical Perspective.	6
2	Instruction Set Design	Operations of the Computer Hardware, Operands of the Computer Hardware, Signed and Unsigned Numbers, Representing Instructions in the Computer, Logical Operations, Instructions for Making Decisions, Supporting Procedures in Computer Hardware, MIPS Addressing for 32-Bit Immediates and Addresses, Parallelism and Instructions.	12
3	Arithmetic for Computers	Addition and Subtraction, Multiplication, Division, Floating Point representation, Computer Arithmetic.	8
4	The Processor	Logic Design Conventions, Building a Datapath, Pipelining, Pipelined Datapath and Control, Data Hazards: Forwarding vs. Stalling, Control Hazards, Exceptions	12
5	Memory Hierarchy	The Basics of Cache, Measuring and Improving Cache Performance, Virtual Memory, A Common Framework for Memory Hierarchies, Parallelism and Memory Hierarchies: Cache Coherence	14
			Total: 52 Periods
Key Text(s): David A. Patterson, and John L. Hennessy, Computer Organization and Design: The Hardware/Software Interface, Fourth Edition, Elsewhere Publications, 2011			
Chapters: 1, 2, 3, 4, 5			
REFERENCE BOOKS: 1. Randal E. Bryant and David R. O'Hallaron, Computer Systems: A Programmer's Perspective, Second Edition, Prentice Hall, 2011 2. John P. Hayes, Computer Architecture And Organization, McGraw Hill, 1998			

[MDSC-105] – Design and Analysis of Algorithms 4 Credits

Course objectives: To train the student to be able to

- Develop problem solving skills by analyzing various problems and to learn the techniques for implementation.
- Analyze the asymptotic performance of algorithms
- Write rigorous correctness proofs for algorithms.

Course outcome: develop the skill to

1. Analyze and identify the algorithm of a specific type such as Greedy, Divide and Conquer etc
2. Implement computer program for an algorithm based on different problem solving methods
3. Discriminate between different problem solving approaches
4. Analyze worst-case running times of algorithms based on asymptotic analysis and justify the correctness of algorithms

Unit	Title	Unit Contents	No. of Periods
1	Introduction	Algorithm, Algorithm Specification, and Performance Analysis. Randomized Algorithms. Basic Data Structure: Stacks and Queues, Trees, Dictionaries, Priority Queues, Sets and disjoint Set Union, Graphs.	8
2	Divide and Conquer	Binary search, Finding MIN and MAX, Merge sort, Quick sort, Selection, Strassen's Matrix Multiplication, convex Hull.	6
3	The Greedy method	Knapsack problem, Tree vertex splitting, Job Sequencing with deadlines, minimum cost spanning Trees, optimal merge patterns, single source shortest path.	8
4	Dynamic Programming	General Method, Multistage Graph, All pairs shortest path, single source shortest path, Optimal Binary Search Trees, 0/1 Knapsack, reliability design, the traveling salesperson problem.	8
5	Basic traversal and Search Techniques	Techniques for Binary Trees, graphs, spanning trees, DFS	6
6	Backtracking	General Method, 8-queens problem, sum of subsets, Graph coloring, Hamiltonian cycles, Knapsack problems. Branch and Bound: the general method, 0/1 Knapsack problem, TSP	8
7	NP-Hard and NP-Complete Problems	Basic concepts, Cooks theorem, NP-Hard graph problems, NP-Hard Scheduling problem, NP-Hard code generation problems, some simplified NP-Hard problems	8
			Total: 52 Periods

Key Text:
E Horowitz, S Sahani S Rajasekaran, Fundamentals of Computer Algorithms, Second Edition, Universities Press, 2008
Chapters: 1,2,3,4,5,6,7 and 11.

Reference Texts:

1. Alfred V. Aho and John E. Hopcraft, and Jeffrey D. Ullman, The Design Analysis of Computer Algorithms, Pearson, 1974
2. Thomas H. Cormen, Charles E. Leiserson, R.L. Rivest, Introduction to Algorithms, Prentice Hall of India Publications, New-Delhi, 2008
3. Sara Baase and Allen Van Gelder, Computer Algorithms: Introduction to Design and Analysis, Third Edition, Pearson education (Singapore) Pvt. Ltd, New Delhi, 2000
4. Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, The Design and Analysis of Computer Algorithms, Pearson Education (Singapore) Pvt. Ltd New Delhi. 2012

[MDSC-106] – Software Lab for Data Visualization* 2 Credits			
Course Objective: ● To make the student learn different visualization techniques for projecting the data			
Course Outcome : Develop the skill set to 1. distinguish qualitative and quantitative data 2. make inferences out of the different plots 3. develop insight of the data			
Unit	**Title**	**Unit Contents**	**No. of Periods**
1	Numbers that Summarize the data	Measures of Average, Measures of Variance, Measures of correlation and Measures of Ratio.	14
2	Fundamental Variations of Graphs	A Brief History of Graphs; Graphical means for Qualitative data; Visual attributes: Lines, Histograms, Bar plot, Scatter plot and Box-Plots and inferences through examples	12
3	General Design Principles for Communication	Organizing, Highlighting, Integration, Table Design, General Graph Design, Multi-Variable display;	14
4	Case Studies	Given a case study, following aspects can covered: - make data storytelling with visualization - business aspects of the problem and inferences	16
	Total		56
	***For implementation: Instructor can use any visualization framework Matplotlib/Seaborn/Plotly, ggplot2, Tableau, PowerBI, or any other**		

Key Text:

1. Stephen Few, Show me the numbers: Designing tables and graphs to enlighten, Second Edition, Analytic Press, 2012
   Chapters - 1 to 11
2. Cole Nussbaumer Knaflic, Storytelling with Data, John Wiley & Sons, Inc. - 2015
3. Matplotlib Documentation, Retrieved June 30, 2022 from https://matplotlib.org/3.5.1/plot_types/index.html
4. Seaborn Documentation, Retrieved June 30, 2022 from https://seaborn.pydata.org/introduction.html

SEMESTER-II

	[MDSC-201] – Regression Methods 3 Credits		
	Course Objectives: ● To teach students the standard Regression methods in statistics for determining the relationships between variables ● To develop the skill to use statistical relationships to forecast future observations ● To teach Regression models that are used to predict and forecast future outcomes		
	Course Outcome : Develop the skill set to 1. develop a deeper understanding of the linear regression model and its applications 2. diagnose and apply corrections to problems with the generalized linear model found in real data		
Unit	**Title**	**Contents**	**No. of Periods**
1	Simple Linear Regression	Model, Least Squares Estimation, Hypothesis Testing, Interval Estimation, Prediction of new observations, Coefficient of Determination, Regression through Origin, Estimation by Maximum Likelihood, Application examples;	8
2	Multiple Linear Models	Models, Estimation of model parameters, Hypothesis Testing, Confidence Intervals, Prediction of new observations, Hidden Extrapolation, Standardized Regression Coefficients, MultiCollinearity, Application examples.	8
3	Model Adequacy Checking	Residual Analysis, Press Statistic, Detection and treatment of Outliers, Lack of fit	7
4	Model Adequacy Correction	Variance stabilizing transformations, Transformations to Linearize, Analytical methods for selecting a transformation, Generalized and Weighted Least squares	8
5	Generalized Linear Models	Logistic Regression, Poisson Regression, Generalized Linear Model	8
		Total : **39 Periods**	
Key Text(s):			

Douglas C. Montgomery, Elizabeth A. Peck and G. Geoffrey Vining, Introduction to Linear Regression Analysis, 5th Edition, Wiley, 2012 Chapters: 1 - 5, 13
References: 1. Norman Draper and Harry Smith, Applied Regression Analysis, Third Edition, John Wiley & Sons, Inc., 1998 2. STAT 501: Regression Methods, Retrieved Jul 5, 2022 from https://online.stat.psu.edu/stat501/

	[MDSC-201 (P)] – Practicals: Regression Methods 1 Credit		
	Course Objective: To introduce different statical techniques from Regression Methods in R programming		
	Course Outcome: Develop the skill set to 1. Perform basic data analysis using R 2. Develop linear models and evaluate for given real world datasets using R		
Unit	**Title**	**Contents**	**No. of Periods**
1	tydiverse	Review the data structure in R, Loading and indexing the Data, Data analysis in R using tydiverse package	10
2	Linear Models for Regression	Simple Linear Regression, Multiple Linear Regression, Check multicollinearity of the data, Residual Analysis	8
3	Linear Models for Classification	Logistic Regression and Poisson Regression; Evaluation of Models; Confusion matrix and ROC Curves	8
			Total: 26 Periods
Key Text(s): 1. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer Science, New York, 2013 https://www.statlearning.com/ Retrieved on August 08, 2022. 2. Hadley Wickham & Garrett Grolemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, O'Reilly Media, Inc., 2017			

[MDSC-202] – Multivariate Statistical Analysis 3 Credits			
	Course Objective: To learn multivariate statistical methods that uncover surprising but valid linkages between variables and explain and predict their measured values.		
	Course Outcome : Develop the skill set to 1. Select appropriate methods of multivariate data analysis 2. Detect outliers and cleaning multivariate data; 3. Perform several statistical tests on multivariate data;		
Unit	**Title**	**Contents**	**No. of Periods**
1	Matrix Algebra and Optimization	Matrix and vector algebra, positive-definite matrices, spectral decomposition theorem; random vectors and matrices, mean vectors and covariance matrices; matrix inequalities and maximization	5
2	Sample Geometry and Random Sampling	Random samples, Sample mean vector, Sample Covariance and correlation matrices; Generalized variance and total variance; Sample values of Linear combination of variables	4
3	The Multivariate Normal Distribution	Multivariate normal density and its properties; maximum likelihood estimators of the parameters and their sampling distributions, Wishart Distribution; Assessing the Assumption of Normality; Detecting Outliers and Cleaning Data;	8
4	Testing Hypothesis	Tests of hypothesis about the mean vector of normal distribution, Hotelling's T^2- statistics and Likelihood Ratio Tests; Confidence Regions and Simultaneous Confidence Statements; Large Sample Inferences about a Population Mean Vector, Comparing Mean Vectors from Two Populations, Simultaneous Confidence Intervals for Treatment Effects, Testing for Equality of Covariance Matrices;	8
5	Advanced Multivariate Statistical Techniques	Principal Component Analysis: Population and Sample Principal Components, Summarizing Sample Variance by Principal Components; Graphing the Principal Components; Canonical Correlation Analysis: Canonical Variates and Canonical Correlations, Interpreting the Population Canonical Variables; and	14

		the Sample Canonical Variates and Sample Canonical Correlations Discriminant Analysis: Bayes rule and Classification problem, Classification for Two Multivariate Normal Populations, Evaluating Classification Functions	
Total: 39 Periods			
Key Text(s): 1. Richard Johnson and Dean Wichern, Applied Multivariate Statistical Analysis, 6th Edition, Pearson Publications, 2007 Chapters: 2, 3, 4, 5.1-5.5, 6.1-6.3, 6.5-6.6, 8.1-8.4, 10.1-10.4, 10.6, 11.1-11.4			
References: 1. Anderson T. W., An Introduction to Multivariate Statistical Analysis, Wiley, 2003. 2. Kshirsagar, A. M., Multivariate Analysis, Marcel Dekker, 1972. 3. STAT-505: Applied Multivariate Statistical Analysis, Retrieved Jul 5, 2022 from https://online.stat.psu.edu/stat505/			

[MDSC-202(P)] – Practicals: Multivariate Statistical Analysis 1 Credit			
	Course Objective: To introduce different statical techniques from Multivariate Statistical Analysis in R programming		
	Course Outcome: Develop the skill set to 1. Write R program to carry out multivariate data analysis; 2. Develop different models like PCA, Discriminant Analysis, and Correlation Analysis using R		
Unit	**Title**	**Contents**	**No. of Periods**
1	Matrix Algebra through R	vectors, matrices and their operations, Computing eigenvalues and eigenvectors for given matrix, Positive definite matrices and computing square-root of a positive definite matrix	4
2	ggplot2	- Loading and reading multivariate datasets in R - Visualize plots: Dot plots, Box-plots, Histograms, and scatter plots - Perform Exploratory Data Analysis for multivariate Data - QQ and Chi-Square plots	10

		- Transforming the data	
3	Modeling	- Perform MANOVA - Perform Principal Component Analysis - Perform Canonical Correlation Analysis - Perform LDA and QDA for Classification of two populations	12
			Total: 26 periods

Key Text(s):

1. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer Science, New York, 2013 Retrieved on August 08, 2022 from https://www.statlearning.com/

2. Hadley Wickham & Garrett Grolemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, O'Reilly Media, Inc., 2017

[MDSC-203] – Machine Learning 3 Credits

Course Objective:

- To introduce students to the basic concepts and techniques of Machine Learning.
- To become familiar with regression methods, classification methods, clustering methods.
- To become familiar with Dimensionality reduction Techniques

Course Outcome : Develop the skill set to

1. Gain knowledge about basic concepts of Machine Learning
2. Identify machine learning techniques suitable for a given problem
3. Solve the problems using various machine learning techniques
4. Apply Dimensionality reduction techniques.

Unit	Title	Contents	No. of Periods
1	Introduction	Machine Learning: Introduction, Types of machine learning, supervised learning-Basics, Reinforcement Learning	2
2	Regression Models	Linear Regression, Multivariate Regression, Subset Selection, Shrinkage Methods, Principal Component Regression, Partial Least squares Linear Classification, Logistic Regression, Linear Discriminant Analysis	8

3	Support Vector Machine	Perceptron, Support Vector Machines, Neural Networks - Introduction, Early Models, Perceptron Learning, Backpropagation, Initialization, Training & Validation, Parameter Estimation - MLE, MAP, Bayesian Estimation	8
5	Decision Trees	Decision Trees, Regression Trees, Stopping Criterion & Pruning loss functions, Categorical Attributes, Multiway Splits, Missing Values, Decision Trees - Instability Evaluation Measures Bootstrapping & Cross Validation, Class Evaluation Measures, ROC curve, MDL, Ensemble Methods - Bagging, Committee Machines and Stacking, Boosting	8
6	Ensemble Techniques	Gradient Boosting, Random Forests, Multi-class Classification, Naive Bayes, Bayesian Networks Undirected Graphical Models, HMM, Variable Elimination, Belief Propagation	8
7	Clustering	Partitional Clustering, Hierarchical Clustering, Birch Algorithm, CURE Algorithm, Density-based Clustering, Gaussian Mixture Models, Expectation Maximization	5
			Total: 39 Periods

Key Text(s):

1. Trevor Hastie, Robert Tibshirani, and Jerome H. Friedman, The Elements of Statistical Learning, Second Edition, Springer, 2009 Retrieved on July 05, 2022 from https://hastie.su.domains/Papers/ESLII.pdf
2. Chrisopher M Bishop, Pattern Recognition and Machine Learning, Springer, 2006
3. Rogers and Girolami, A First Course in Machine Learning, Chapman and Hall/CRC, 2015

References:

1. Barber, Bayesian Reasoning and Machine Learning, Cambridge University Press, 2012
2. Hal Daumé III, A Course in Machine Learning, e-Edition, 2012
3. Mitchell, Machine Learning, McGraw Hill, 1997
4. CS229-Machine Learning, Stanford University, Retrieved June 30, 2022 from https://see.stanford.edu/Course/CS229

	[MDSC-203(P)] – Practicals: Machine Learning 1 Credit		
	Course Objective: To implement basic ML algorithms in python		
	Course Outcome : Develop the skill set to 1. Mastering in python library scikit-learn by implementing several ML algorithms 2. Design application using machine learning techniques		
Unit	**Title**	**Contents**	**No. of Periods**
1	Introduction	- Review numpy, pandas and matplotlib - Implementation of Linear algorithms, Non-Linear Algorithms and ensembling algorithms, bias-variance tradeoff, from scratch in python	10
2	Python Library scikit-learn	Develop different ML algorithms using scikit-learn	16
			Total: 26 Periods

Key Text(s):

1. Jason Brownlee, Master Machine Learning Algorithms, Discover How They Work and Implement them from Scratch, Machine Learning Mastery, eBook, 2017 Retrieved Jun 30, 2022 from GitHub - -- Jason Brownlee Master Machine Learning Algorithms
2. Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'reilly, 2019
3. Scikit-learn Documentation, retrieved June 30, 2022 from scikit-learn Tutorials — scikit-learn 1.0.2 documentation

[MDSC-204] – Big Data Analytics 4 Credits

Course Objective:

- To provide an overview of an exciting growing field of big data analytics.
- To teach the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.
- To enable students to have skills that will help them to solve complex real-world problems in decision support.

Course Outcome: Students will be able to

1. Understand the key issues in big data management and its associated applications in intelligent business and scientific computing.
2. Interpret business models and scientific computing paradigms, and apply software tools for big data analytics.
3. Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc.

S. No	UNIT	Topics	No. of Periods
1	Introduction to Big Data	Big Data - Why and Where?, Characteristics of Big Data and Dimensions of Scalability, Getting value out of Big Data, Foundations for Big Data Systems and Programming	4
2	Similarity Algorithms	Near-Neighbor search, Shingling, Similarity preserving summary, Locality sensitive functions, Distance measures, Locality sensitive hashing and its applications to different distance measures, Applications of Locality sensitive hashing	10
3	Streaming Data	Stream Data model, Sampling data in a stream, Filtering streams, Counting distinct elements in a stream, Application of stream algorithms in counting.	10
4	Link Analysis	Page Rank, Computation of Page Rank, Topic sensitive page rank, Link spam.	6
5	Frequent Item sets	Market-Basket model, A-priori algorithm, Larger datasets in main memory, Limited pass algorithms, Counting frequent sets in a stream	10
6	Social Network Graphs	Clustering, Discovering of communities, Partitioning, Finding overlapping communities, Simrank, Counting Triangles, Neighborhood properties	12
			Total: 52 Periods

Key Text(s):

Anand Rajaraman, Mining Massive Datasets, Stanford University Press, 2014

Chapters: 2.1 - 2.3, 3.1 - 3.7, 4.1 - 4.6, 5.1 - 5.4, 6.1 - 6.4, 10

	[MDSC-205] – Software Lab in Data Engineering 4 Credits		
	Course Objective: Build, monitor and manage real time data pipelines to create data engineering infrastructure using open source projects		
	Course Outcome : Develop the skill set to perform extract, transform and load data pipelines that forms the foundation of data engineering		
Unit	**Title**	**Contents**	**No. of Periods**
1	Python for Apache Spark	Overview of Variables & Data Types, Conditionals & Loops, Functions & Packages, Collections & Classes	16
2	Spark Architecture	Introduction, Databricks platform	8
3	Spark SQL	Introduction, Joins & Temporary views, Higher Order functions, Use cases	24
4	Spark ML	Introduction, Components, Basics of MLFlow, Basics of AutoML, Basics of Feature Store, Use case	32
5	Spark Streaming	Introduction, Batch & Streaming engines, Big Data Ecosystem, Use case	24
			Total: 104 Periods

Key Text(s):

1. Jules S. Damji, Brooke Wenig, Tathagata Das, Denny Lee, Learning Spark - Lightning Fast Data Analysis, Second Edition, O'Reilly, 2015
2. Nick Pentreath, Machine Learning with Spark, PACKT Publishing, 2015

SEMESTER-III

<table>
<tr><th colspan="4">[MDSC-301] - Stochastic Processes 3 Credits</th></tr>
<tr><td></td><td colspan="3">Course Objective: Stochastic models are among the most widely used tools in operations research and management science. Stochastic processes and applications can be used to analyze and solve a diverse range of problems arising in production and inventory control, resource planning, service systems, computer networks and many others.</td></tr>
<tr><td></td><td colspan="3">Course Outcomes: Develop the skill set to
1. elucidate the power of stochastic processes and their range of applications;
2. demonstrate essential stochastic modeling tools including Markov chains and Gaussian processes;
3. formulate and solve problems which involve setting up stochastic models
4. to solve different techniques from stochastic processes using R</td></tr>
<tr><th>Unit</th><th>Title</th><th>Contents</th><th>No. of Periods</th></tr>
<tr><td>1</td><td>Introduction to Markov Chains</td><td>Stochastic Processes, Markov Chains, Transition Probabilities, Limiting and Stationary distributions, Irreducible Markov chain, Periodicity, Ergodicity and Time Reversibility of Markov chains, Regeneration and strong Markov property, Probability Generating Functions, Extinction of Branching processes and Markov Chain and Monte Carlo: Hasting Algorithm, Gibbs Sampler</td><td>7</td></tr>
<tr><td>2</td><td>Poisson Process</td><td>Arrival, Interarrival times, Infinitesimal probabilities, Thinned, and Spatial Poisson processes.</td><td>8</td></tr>
<tr><td>3</td><td>Continuous Markov chains</td><td>Infinitesimal Generator, Long-Term Behavior, Time-Reversibility, Birth and Death Process, Queuing Theory, Subordinated Poisson process;</td><td>12</td></tr>
<tr><td>4</td><td>Brownian Motion</td><td>Brownian Motion, Random Walk, and Gaussian process;
Transformations and properties, Variations and applications; Martingales;</td><td>12</td></tr>
<tr><td colspan="4" align="right">Total: 39 Periods</td></tr>
<tr><td colspan="4">Key Text(s):
Robert P.Dobrow, Introduction to Stochastic Process with R, John Wiley & Sons, Inc., 2016
Chapters: 1, 2.1-2.5, 3.1-3.8, 4, 5.1-5.3, 6, 7, 8
Retrieved July 5, 2022 from https://people.carleton.edu/~rdobrow/stochbook/</td></tr>
<tr><td colspan="4">References:
1. Paul G. Hoel, Sidney C. Port and Charles J. Stone, Stochastic Process, Houghton Mifflin Company, BOSTON, 1972.
2. Olga Korosteleva, Stochastic Process with R, An Introduction, First Edition, CRC Press, 2022
3. J Medhi, Stochastic Process, Third Edition, New Academic Science Limited, 2012
4. Ross, S., Stochastic Processes, second edition, John Wiley, 1996.</td></tr>
</table>

5. Goswami, A. and Rao, B. V., A Course in Applied Stochastic Processes, Hindustan Book Agency, 2006.

	[MDSC-301(P)] - Practicals: Stochastic Processes 1 Credit		
	Course Objective: To implement different techniques from Stochastic Processes in R		
	Course Outcomes: Develop the skill set to 1. demonstrate essential stochastic modeling tools including Markov chains and Gaussian processes in R; 2. formulate and solve problems which involve setting up stochastic models with R		
Unit	**Title**	**Contents**	**Periods**
1	Implementation in R	● Simulation of a Markov chain ● Simulation of Gambler's Ruin ● Compute the Higher Transition Probabilities ● Computing Limiting and Stationary distributions ● Computing an Expected Return Time ● The following simulation can be done: ○ Eat, Play, Sleep ○ Brownian Motion ○ Random Walk ○ Gaussian process	26
			Total: 26 Periods

Key Text(s):

Robert P.Dobrow, Introduction to Stochastic Process with R, John Wiley & Sons, Inc., 2016
Chapters: 1, 2.1-2.5, 3.1-3.8, 4, 5.1-5.3, 6, 7, 8
Retrieved July 5, 2022 from https://people.carleton.edu/~rdobrow/stochbook/

References:

Olga Korosteleva, Stochastic Process with R, An Introduction, First Edition, CRC Press, 2022

[MDSC-302] – Deep Learning 3 Credits			
	Course Objective: To introduce the prominent deep learning architectures used in the industry. To impart a strong understanding of the training mechanisms. To expose students to Advanced topics like Attention based models, Deep Generative models		
	Course Outcome : Through this course, students will be able to: 1. Develop deep learning solutions to standard AI related problems 2. Have a strong understanding of the internal workings of deep learning architectures		
Unit	**Title**	**Contents**	**No. of Periods**
1	An Introduction to Neural Networks	McCulloch Pitts Neuron, Perceptrons, Perceptron Learning Algorithm and Convergence, Multilayer Perceptron (MLPs), Representation Power of MLPs, Sigmoid Neurons, Gradient Descent, FeedForward Neural Networks, Backpropagation.	8
2	Gradient Descent Strategies	Gradient Descent (GD), Momentum Based GD, Nesterov Accelerated GD, Stochastic GD, AdaGrad, RMSProp and Adams	5
3	Embedding and Representation Learning	Review of Principal Component Analysis, Autoencoders and relation to PCA, Regularization in autoencoders, Denoising autoencoders, Sparse autoencoders, Contractive autoencoders	5
4	Regularization for Deep Neural Networks	Bias variance Tradeoff, L^2- regularization, Early stopping, Dataset augmentation, Parameters sharing and tying, Injecting noise at input, Ensemble methods, Dropout, Greedy Layer Wise Pre-training, Better activation functions, Better weight initialization methods, Batch Normalization	5
5	Convolutional Neural Networks	The basic structure of convolutional Network: Padding, Strides, ReLU Layer, Pooling, and Fully Connected Layers; Case studies of Convolutional Architectures: AlexNet, ZFNet, VGGNet, GoogLeNet, ResNet; Visualizing Convolutional Neural Networks, Guided Backpropagation.	5
6	Recurrent Neural Networks	Recurrent Neural Networks, Backpropagation through time (BPTT), Vanishing and Exploding Gradients, Truncated BPTT, GRU, LSTMs	5
7	Advanced topics in Deep Learning	Encoder Decoder Models, Attention Mechanism, Attention over images, Variational autoencoders, Generative Adversarial Networks (GANs)	6
			Total: 39 Periods
Key Text(s):			

1. Ian Goodfellow, YoshuaBengio, Aaron Courville, Deep Learning, The MIT Press, 2016
2. Nikhil Buduma, Nicholas Locascio, Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms, O'ReillyMedia, 2017
3. Charu C. Agarwal, Neural Networks and Deep Learning: A Textbook, Springer, 2018

References:

1. Duda, R.O., Hart, P.E., and Stork, D.G, Pattern Classification, Wiley-Interscience, Second Edition, 2001
2. Theodoridis, S. and Koutroumbas, K., Pattern Recognition. Fourth Edition, Academic Press, 2008
3. Russell, S. and Norvig, N. Artificial Intelligence: A Modern Approach, Prentice Hall Series in Artificial Intelligence. 2003
4. Bishop, C. M. Neural Networks for Pattern Recognition. Oxford University Press. 1995.
5. Hastie, T., Tibshirani, R. and Friedman, J. The Elements of Statistical Learning. Springer. 2001.
6. Koller, D. and Friedman, N. Probabilistic Graphical Models. MIT Press. 2009.

	[MDSC-302(P)] – Practicals: Deep Learning 1 Credit		
	Course Objective: ● To build Deep Learning models using PyTorch		
	Course Outcome Develop the skill set to 1. implement several deep learning algorithm from scratch 2. Prototype solutions in Python		
Unit	**Title**	**Contents**	**Periods**
1	PyTorch Basics	Why PyTorch?, Basics of Tensors: Tensors, indexing, size, offset, slicing, dtypes, moving tensors to GPU Representation of real-world data using tensors: image data, tabular data, and text data Download datasets, Dataset class, Dataset Transformation, Normalizing the data, Create your own dataset using *dataloader*	12
2	Working with pretrained models	Pretrained models: AlexNet, ResNet, Artistic Style, CycleGAN etc	4

3	The mechanics of Learning	Gathering some data, Visualizing the data, Iterating to fit the model, Normalizing inputs, pytorch autograd, pytorch.nn module	6
4	Case studies	1. Classification of images in CIFAR10 dataset 2. Classification model to detect suspected tumors 3. Sentiment Analysis	4
			Total: 26 Periods

Key Text(s):

1. Eli Stevens, Luca Antiga, and Thomas Viehmann, Deep Learning with PyTorch, Manning Publications Co., 2020
2. PyTorch documentation, Retrieved July 05, 2022 from https://pytorch.org/docs/stable/index.html
3. PyTorch For Deep Learning, freecodecamp.org, retrieved July 05, 2022 from https://www.youtube.com/watch?v=GIsg-ZUy0MY

[MDSC-303] – Natural Language Processing 3 Credits

	Course Objective: To impart knowledge in classification of documents, retrieving and extracting information from documents, identifying important documents available as unstructured text.		
	Course Outcome Develop the skill set in NLP to use them in practical situations.		
Unit	**Title**	**Contents**	**No. of Periods**
1	Introduction to NLP	Regular Expressions, Words, Corpora, Text Normalization, Edit Distance	4
2	N-gram Language Models	N-grams, Evaluating Language Models, Sampling sentences, Generalization, Smoothing, Entropy & Perplexity	4
3	Text Classification	Naive Bayes classifier, Optimizing for sentiment analysis, Naive Bayes Language Model, Naive Bayes for other text classification tasks, Evaluation,	7

		Cross-validation, Logistic Regression, Multinomial Logistic Regression, Cross-entropy loss, Gradient Descent, Regularization	
4	Vector Semantics	Lexical Semantics, Vector Semantics, Words & Vectors, Cosine Distance, TF-IDF, Pointwise Mutual Information, Word2Vec, Visualizing embeddings, Semantic properties of embeddings, Evaluation vector models	6
5	Sequence Labeling	Part-of-Speech Tagging, Named Entity Tagging, HMM PoS tagging, Conditional Random Fields, Evaluation	4
6	Sequence Processing using Deep Learning	Feedforward Networks for NLP, Neural Language Models, RNN Language Models, Stacked and Bidirectional RNNs, LSTM, Transformers, Contextual generation and Summarization	7
7	Machine Translation and Encoder-Decoder Models	Language Divergence, Encoder-Decoder Model, Encoder-Decoder with RNN, Attention, Beam Search, Encode-Decoder with Transformers, Bidirectional Transformer Encoders, Transfer Learning, Evaluation	7
			Total: 39 Periods

Key Text(s):
Dan Jurafsky and James Martin, An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, Speech and Language Processing, Third Edition, Prentice Hall, 2022.

References:

1. Jacob Eisenstein, Introduction to Natural Language Processing, MIT Press, 2019

2. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python - Analyzing Text with the Natural Language Toolkit, O’reilly First edition, 2011

[MDSC-303(P)] –Practicals: Natural Language Processing* 1 Credit			
	Course Objective: Implement NLP techniques using a programming language like python		
	Course Outcome Develop the skill set to implement NLP solutions for real life problems.		
Unit	**Title**	**Contents**	**Periods**
1	Document similarity	Implement document similarity using different measures	3
2	Bigram and Trigram models	Implement vector models with Bigrams and Trigrams	4
3	Naive-Bayes & Logistic Regression	Implement Naive-Bayes and Logistic regression for document classification	3
4	PoS Tagging	Implement Parts-of-Speech tagging	4
5	Encoder-Decoder model	Implement Encoder-Decoder model with RNN	6
6	Basic Machine Translation	Implement machine translation of phrases in one language to other	6
			Total: 26 Periods
***All implementation using PyTorch/TensorFlow**			
Key Text(s): Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python - Analyzing Text with the Natural Language Toolkit, O'reilly, First edition, 2011			

[MDSC-304] – Cloud Computing 3 Credits			
Course Objective: To make students understand the core concepts of virtualization, cloud storage: key-value/NoSQL stores, cloud networking, fault-tolerance cloud using PAXOS, peer-to-peer systems, classical distributed algorithms such as leader election, time, ordering in distributed systems, distributed mutual exclusion, distributed algorithms for failures and recovery approaches, emerging areas of big data and many more.			
Course Outcome : Upon completing this course, 1. students will have intimate knowledge about the internals of cloud computing and how the distributed systems concepts work inside clouds. 2. working knowledge on the current industry systems such as Apache Spark, Google's Chubby, Apache Zookeeper, HBase, MapReduce, Apache Cassandra, Google's B4, Microsoft's Swan			
Unit No.	**Unit Title**	**Unit Contents**	**No. of Periods**
1	Introduction; Principles of Parallel and Distributed Computing	Cloud computing at a glance; Historical Developments; building Cloud computing environment; computing platforms and Technologies Principles of Parallel and Distributed Computing: Eras of Computing; parallel Vs. distributed computing; elements of distributed computing; technologies of Distributed computing	7
2	Virtualization and Cloud Computing Architecture	Characteristics of virtualized environments; virtualization techniques; virtualization and cloud computing; pros and cons of virtualization; examples. Cloud Reference model; Types of clouds; cloud economics; open challenges	5
3	Aneka: Cloud application Platform	Overview; anatomy of the Aneka container; building Aneka clouds; cloud programming and management	5
4	Concurrent Computing and High-Throughput Computing and Map Reduce Programming	Introducing parallelism; programming with threads; multithreading with Aneka; applications; Task Computing; task based Application Model; Task based Programming; Data Intensive Computing; Technologies; Aneka Map Reduce Programming	8

5	Cloud Platforms in Industry and Cloud Applications	Amazon Web services; Google App Engine; Microsoft Azure; Cloud scientific Applications; Business and Consumer Applications	6
6	Advanced Topics in Cloud Computing and Cloud Security	Energy Efficiency Clouds; Market based management clouds; Federated Clouds; Third Party Cloud Services; Infrastructure Security: Network level security, Host level security, and Application level security; Data security and Storage	8
			Total: 39 Periods

Key Text(s):
Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi, Mastering C*loud Computing, MGH-* 2013

Chapters: 1,2,3,4,5,6,7,8,9,10,11

REFERENCE BOOKS

1. Rajkumar Buyya, James Broberg, Andrzej M. Goscinski,*Cloud Computing:Principles and Paradigms, Wiley,* 2011
2. Barrie Sosinsky*, Cloud Computing Bible, Wiley-India,* 2010
3. Nikos Antonopoulos, Lee Gillam*, Cloud Computing: Principles, Systems and Applications, Springer,* 2012
4. Ronald L. Krutz, Russell Dean Vines*, Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Wiley-India,* 2010

	[MDSC-304(P)]–Practicals: Cloud Computing 1 Credit
	Course Objective: To make students understand the core concepts of virtualization, cloud storage: key-value/NoSQL stores, cloud networking, fault-tolerance cloud using PAXOS, peer-to-peer systems, classical distributed algorithms such as leader election, time, ordering in distributed systems, distributed mutual exclusion, distributed algorithms for failures and recovery approaches, emerging areas of big data and many more.

	Course Outcome Develop the skill set to 1. students will have intimate knowledge about the internals of cloud computing and how the distributed systems concepts work inside clouds. 2. working knowledge on the current industry systems such as Apache Spark, Google's Chubby, Apache Zookeeper, HBase, MapReduce, Apache Cassandra, Google's B4, Microsoft's Swan		
Unit	**Title**	**Contents**	**No. of Periods**
1	Implementations	Implementation of algorithms /exercises from different units in the syllabus in the Lab.	26
			Total: 26 Periods
Key Text(s): Rajiv Misra, Cloud Computing and Distributed Systems-IIT Patna Retrieved July 05, 2022 from https://nptel.ac.in/courses/106104182			

List of Electives:

1. **MDSC-AI:** Artificial Intelligence [4 Credits]
2. **MDSC-CGT:** Combinatorics and Graph Theory [4 credits]
3. **MDSC-RL:** Reinforcement Learning [3 Credits]
   MDSC-RL(P): Practicals-Reinforcement Learning [1 Credit]
4. **MDSC-MLO**: Machine Learning Operations [3 Credits]
   MDSC-MLO(P): Practicals: Machine Learning Operations [1 credit]
5. **MDSC – IR**: Information Retrieval [4 credits]
6. **MDSC – NS:** Network Security [4 credits]
7. **MDSC –IoT:** Internet of Things [3 credits] and
   MDSC – IoT (P): Practicals: Internet of Things Lab [1credit]
8. **MDSC – ATS:** Applied Time Series Analysis [4 credits]
9. **MDSC – TDA:** Topological Data Analysis [4 credits]
10. **MDSC – LSP:** Linux System Programming [4 credits]
11. **MDSC-DS:** Distributed Systems [4 credits]

[MDSC-AI] – Artificial Intelligence 4 Credits

Course Objective:

- To understand an autonomous agent
- To study a wide variety of search methods that agents can employ for problem solving.

Course Outcome : Students will be able to

1. Understand a problem from autonomous agent's point of view
2. Solve a real-world problem using various search methods to arrive at a human like decision
3. Write inference rules based on logic that can make decisions.

Unit	Title	Contents	No. of Periods
1	Introduction	Introduction: History, Can Machines think?, Turing Test, Winograd Schema Challenge, Language and Thought, Wheels & Gears, Philosophy, Mind, Reasoning, Computation, Dartmouth Conference, The Chess Saga, Epiphenomena.	10
2	Searching Concepts	State Space Search: Depth First Search, Breadth First Search, Depth First Iterative Deepening Heuristic Search: Best First Search, Hill Climbing, Solution Space, TSP, Escaping Local Optima, Stochastic Local Search	10
3	Search Algorithms	Population Based Methods: Genetic Algorithms, SAT, TSP, emergent Systems, Ant Colony Optimization, Finding Optimal Paths: Branch & Bound, A*, Admissibility of A*, Informed Heuristic Functions, Space Saving Versions of A*: Weighted A*, IDA*, RBFS, Monotone Condition, Sequence Alignment, DCFS, SMGS, Beam Stack Search	12
4	Game theory and advanced search Algorithms	Game Playing: Game Theory, Board Games and Game Trees, Algorithm Minimax, Alpha Beta and SSS*, Automated Planning: Domain Independent Planning, Blocks World, Forward & Backward Search, Goal Stack Planning, Plan Space Planning, Problem Decomposition: Means Ends Analysis, Algorithm Graph plan, Algorithm AO*.	10
5	Logical Inference	Rule Based Expert Systems: Production Systems, Inference Engine, Match-Resolve-Execute, Rete Net Deduction as Search: Logic, Soundness, Completeness, First Order Logic, Forward Chaining, Backward Chaining, Constraint Processing: CSPs, Consistency Based Diagnosis, Algorithm Backtracking, Arc Consistency, Algorithm Forward Checking	10
			Total: 52 Periods

Key Text(s): Deepak Khemani. A First Course in Artificial Intelligence, McGraw Hill Education (India), 2013
References: 1. Stefan Edelkamp and Stefan Schroedl, Heuristic Search: Theory and Applications, Academic Press, 2011 2. John Haugeland, Artificial Intelligence: The Very Idea, A Bradford Book, The MIT Press, 1985. 3. Pamela McCorduck, Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence, Second Edition, A K Peters/CRC Press; Second Edition, 2004. 4. Zbigniew Michalewicz and David B. Fogel. How to Solve It: Modern Heuristics, Second Edition, Springer; 2004 5. Judea Pearl, Heuristics: Intelligent Search Strategies for Computer Problem Solving, Addison-Wesley, 1984. 6. Elaine Rich and Kevin Knight, Artificial Intelligence, Tata McGraw Hill, 1991. 7. Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, 3rd Edition, Prentice Hall, 2009. 8. Eugene Charniak, Drew McDermott, Introduction to Artificial Intelligence, Addison-Wesley, 1985. 9. Patrick Henry Winston, Artificial Intelligence, Addison-Wesley, 1992.

[MDSC-CGT] - Combinatorics and Graph Theory 4 Credits			
	Course Objectives: ● To get introduced to the elementary principles of Combinatorics ● To understand the fundamental concepts of graph theory.		
	Course Outcome: Develop the skill set to evaluate some real time problems using concepts of graph theory and combinatorics.		
Unit	**Title**	**Contents**	**No. of Periods**
1	Elements of Graph Theory	Graphs, Special Types of Graphs, Graphs and Matrices, Graph Models and Distance, Coloring of Graphs: Chromatic Number and Chromatic Polynomial	6
2	Elements of Trees	Trees, Properties of Trees, Spanning and Counting of Trees, Trails, Circuits, Paths, Cycles and Planarity: Regular Polyhedra and Kuratowski's Theorem	12
3	Advanced Topics in Graph Theory	Matchings: Hall's Theorem and SDRs, The Konig-Egervary Theorem and Perfect Matching Ramsey Theory: Classical and Exact Ramsey Numbers	10
4	Basics in Combinatorics	Binomial, Multinomial Coefficients, The Pigeonhole Principle, The principle of Inclusion and Exclusion, Generating Functions, Partitions, Special Numbers: Stirling, Bell and Eulerian Numbers	10
5	Advanced Topics in Combinatorics	Polya's Theory of Counting: Permutation Groups, Burnside Lemma, The Cycle Index, Polya's Enumeration Formula; The Gale-Shapley Algorithm; Sylvester Problem and Convex Polygons	14
			Total: 52 Periods
Key Text(s): John M. Harris, Jeffry L.Hirst, and M.J. Mossinghoff, Combinatorics and Graph Theory, Second Edition, Springer 2008			
References: 1. Sebastian M.Cioaba, M. Ram Murthy, A First Course in Graph Theory and Combinatorics, Hindustan Book Agency-2009 2. Narsing Deo, Graph theory with Applications to Engineering and Computer Science, Prentice Hall-1974			

[MDSC-RL] - Reinforcement Learning 3 Credits			
	Course Objectives: ● To introduce basic mathematical functions in Reinforcement Learning ● To discuss various deep learning architectures to approximate Q-value functions		
	Course Outcomes: Develop skill set to 1. solve real-world problems using reinforcement learning techniques like MDP and Bellman Equation 2. Approximate Q-value functions using Deep Neural Networks		
Unit	**Title**	**Contents**	**No. of Periods**
1	Introduction to Reinforcement Learning	Reward functions and Determining a good reward function, State and Action; The Markov Decision Process (MDP), Bellman Equation: Estimating the value function and Q-function; Application of Dynamic Programming to solve Bellman Equation, Value iteration and Policy Iteration methods.	10
2	Temporal Difference Learning and Q-Learning	Challenges with Classical DP, Model-Based and Model-Free Approaches, Temporal Difference Learning, SARSA, Q-Learning, Explore vs Exploit.	6
3	Deep Q-Networks	Review of Deep Learning: Feed-Forward Neural Networks, Activation Functions, Loss Functions, and Optimizers in Deep Learning, Convolutional Neural Networks; The DQN Algorithm: Experience Replay, Target Q-Network, Clipping Rewards and Penalties; Double DQN, Dueling DQN;	8
4	Policy-Based Reinforcement Learning	Policy-Based approaches, Difference between value-based and policy-based approaches, The REINFORCE Algorithm, Methods to Reduce Variance in REINFORCE Algorithm.	8
5	Actor-Critic Models	Actor-Critic method and DQN, Advantage Actor-Critic architecture, Asynchronous Advantage Actor-critic (A3C) architecture and Synchronous Advantage Actor-critic (A2C) architecture	7
			Total: 39 Periods
Key Text(s): 1. Mohit Sewak, Deep Reinforcement Learning: Frontiers of Artificial Intelligence, Springer Nature Singapore Pvt Ltd. 2019 2. Richard S. Sutton and Andrew G. Barto, Reinforcement Learning, The MIT Press, 2018			
References:			

1. Phil Winder, Reinforcement Learning, O'Reilly Media, Inc., 2021

<table>
<tr><th colspan="4">[MDSC-RL(P)]- Practicals: Reinforcement Learning 1 Credit</th></tr>
<tr><td></td><td colspan="3">Course Objectives: To implement different Reinforcement Learning algorithms in python</td></tr>
<tr><td></td><td colspan="3">Course Outcomes: Develop skill set to
1. solve real-world problems using reinforcement learning techniques like MDP and Bellman Equation
2. Approximate Q-value functions using Deep Neural Networks</td></tr>
<tr><th>Unit</th><th>Title</th><th>Contents</th><th>No. of Periods</th></tr>
<tr><td>1</td><td>Python Implementations</td><td>- Grid-World Problem
- Value iteration to solve Grid-World Problem
- Policy Iterations to solve Grid-World Problem
- Define Q-Learning agents
- Testing the agent implementation
- Define DQN and Double DQN Agents
- A3C architecture
- Latest real-world applications from Reinforcement Learning</td><td>26</td></tr>
<tr><td colspan="4" align="right">Total: 26 Periods</td></tr>
<tr><td colspan="4">Key Text(s):
1. Mohit Sewak, Deep Reinforcement Learning: Frontiers of Artificial Intelligence, Springer Nature Singapore Pvt Ltd. 2019
2. Richard S. Sutton and Andrew G. Barto, Reinforcement Learning, The MIT Press, 2018</td></tr>
<tr><td colspan="4">References:
1. Phil Winder, Reinforcement Learning, O'Reilly Media, Inc., 2021</td></tr>
</table>

[MDSC-MLO] – Machine Learning Operations 3 Credits			
	Course Objective: Maintenance of the Machine learning models in the production environment has become one the challenges for modern organizations using technology as well as business. This course is designed to have a basic understanding of the ML model life cycle in the production environment.		
	Course Outcomes: Develop the skill set to 1. Understand the MLOps as a discipline in the industry 2. Learn the ML- life cycle 3. Example based learning on ML models		
Unit	**Title**	**Contents**	**No. of Periods**
1	MLOps: What and Why	Why Now and Challenges, People of MLOps, Key MLOps Features.	10
2	MLOps: How	Developing Models, Preparing for Production, Deploying to Production, Monitoring and Feedback Loop, Model Governance.	15
3	MLOps: Examples	Consumer Credit Risk Management, Marketing Recommendation Engines,Consumption Forecast	14
			Total : 39 Periods
Key Text(s): Mark Trveil and the Dataiku Team, Introducing MLOps: How to Scale Machine Learning in the Enterprise, O'Reilly, 2020 Retrieved Jun 30, 2022 from https://itlligenze.com/uploads/5/137039/files/oreilly-ml-ops.pdf			

<table>
<tr><th colspan="4">[MDSC-MLO(P)] – Practicals: Machine Learning Operations 1 Credits</th></tr>
<tr><td></td><td colspan="3">Course Objective: To earn the development of some containers</td></tr>
<tr><td></td><td colspan="3">Course Outcomes: Develop the skill set to implement
1. Web Applications for ML Models
2. Containerization</td></tr>
<tr><td>Unit</td><td>Title</td><td>Contents</td><td>No. of Periods</td></tr>
<tr><td>1</td><td>Review of ML Models</td><td>Basic Scripts-PYTHON-Model Zoo</td><td>6</td></tr>
<tr><td>2</td><td>Web Application</td><td>Converting the ML into a Web App using Flask</td><td>6</td></tr>
<tr><td>3</td><td>Containerization</td><td>Docker and Docker swarm</td><td>6</td></tr>
<tr><td>4</td><td>Scaling & Orchestration</td><td>Docker and Travis CI, AWS, and Google Kubernetes</td><td>8</td></tr>
<tr><td colspan="4" align="right">Total: 26 Periods</td></tr>
<tr><td colspan="4">Key Text:

Sandeep Giri, MLOps - Complete Hands-On Guide with Case Study, Article, August 23, 2021.

Retrieved July 5, 2022 from
https://cloudxlab.com/blog/mlops-machine-learning-operations-a-complete-hands-on-guide-with-case-study/</td></tr>
</table>

<table>
<tr><td colspan="4">[MDSC-IR] - INFORMATION RETRIEVAL 4 Credits</td></tr>
<tr><td></td><td colspan="3">Course Objective:
● The main objective of this course is to present the basic concepts in information retrieval and more advanced techniques of multimodal based information systems.
● Word statistics, Vector space model (relevance feedback, query expansion, document normalization, document re-ranking), evaluation of retrieval, generalized VSM, latent semantic indexing, Web retrieval, data fusion, meta search, multimodal retrieval, applications.</td></tr>
<tr><td></td><td colspan="3">Course Outcome: Student will be able to
1. Understand the underlined problems related to IR and
2. Acquired the necessary experience to design, and implement real applications using Information Retrieval systems.</td></tr>
<tr><td>Unit No.</td><td>Title</td><td>Contents</td><td>No. of Periods</td></tr>
<tr><td>1</td><td>INTRODUCTION</td><td>Boolean retrieval, The term vocabulary and postings lists, Dictionaries and tolerant retrieval</td><td>10</td></tr>
<tr><td>2</td><td>Indexing</td><td>Index construction, Index compression</td><td>12</td></tr>
<tr><td>3</td><td>Scoring</td><td>Scoring, term weighting & the vector space model, Computing scores in a complete search system</td><td>10</td></tr>
<tr><td>4</td><td>Evaluation and Query Expansion</td><td>Evaluation in information retrieval, Relevance feedback & query expansion</td><td>10</td></tr>
<tr><td>5</td><td>Classification</td><td>Text classification & Naive Bayes, Vector space classification</td><td>10</td></tr>
<tr><td colspan="4" align="right">Total: 52 Periods</td></tr>
<tr><td colspan="4">Key Text: Manning, Raghavan and Schutze, Introduction to Information Retrieval, 2009,
Freely Downloadable http://nlp.stanford.edu/IR-book/information-retrieval-book.html</td></tr>
<tr><td colspan="4">Chapters: 1 to 9, 13, 14</td></tr>
</table>

[MDSC-NS] - NETWORK SECURITY 4 Credits

Course Objective:

- To explore the design issues and working principles of various authentication protocols, PKI standards and various secure communication standards including Kerberos, IPsec, and SSL/TLS and email.
- To develop the ability to use existing cryptographic utilities to build programs for secure communication..

Course Outcome:

1. Apply the knowledge of cryptographic checksums and evaluate the performance of different message digest algorithms for verifying the integrity of varying message sizes.
2. Apply different digital signature algorithms to achieve authentication and design secure applications
3. Understand network security basics, analyze different attacks on networks and evaluate the performance of firewalls and security protocols like SSL, IPSec, and PGP.
4. Analyze and apply system security concepts to recognize malicious code.

Unit No.	Unit Title	Unit Contents	No. of Periods
1	Introduction	Computer Security Concepts, The OSI Security Architecture, Security Attacks, Security Services, Security Mechanisms, A Model for Network Security	4
2	Symmetric Encryption	Symmetric Encryption Principles, Symmetric Block Encryption Algorithms, Random and Pseudorandom Numbers, Stream Ciphers and RC4, Cipher Block Modes of Operation	10
3	Message Authentication and Hash Functions	Approaches to Message Authentication, Secure Hash Functions, Message Authentication Codes	4
4	Public Key Cryptography	Public-Key Cryptography Principles, Public-Key Cryptography Algorithms, Digital Signatures	6
5	Key Distribution and User Authentication	Kerberos, X.509 Certificates, Public-Key Infrastructure	6
6	Cloud Security	Cloud Security Risks and Countermeasures, Data Protection in the Cloud, Cloud Security as a Service	4
7	Transport-Level Security	Web Security Considerations, Secure Sockets Layer (SSL), Transport Layer Security (TLS), HTTPS, Secure Shell (SSH)	6
8	Electronic Mail Security	Pretty Good Privacy (PGP), S/MIME	6
9	IP Security	IP Security Overview, IP Security Policy, Encapsulating Security Payload, Combining Security Associations	6

	Total: 52 Periods
Key Text: William Stallings, Cryptography and Network Security : Principles and Practice, Fifth Edition, Pearson Education Inc, 2018	
Chapters: 1.1-1.6, 2.1-2.5, 3.1-3.6, 5.1-5.6, 6.1-6.5, 9.1-9.4, 10.1-10.4, 11.1-11.6, 12.1-12.6, 13.1-13.4, 14.1-14.5, 15.1-15.3, 16.1-16.5, 18.1-18.3, 19.1-19.5	
REFERENCE BOOKS: 1. Richard R. Brooks, Introduction to Computer and Network Security: Navigating Shades of Gray, 1st Edition, 2013. 2. Charlie Kaufman, Radia Perlman and Mike Speciner, Network Security: Private Communication in a public world, Second Edition, Prentice Hall PTR, 2002, ISBN 0-13-046019	

[MDSC- IOT] - Internet of Things 3 Credits

Course Objective:
Students will be explored to the interconnection and integration of the physical world and cyberspace. They are also able to design & develop IOT Devices.

Course Outcome: Students will be able to

1. Understand the application areas of IOT
2. Realize the revolution of Internet in Mobile Devices, Cloud & Sensor Networks
3. Understand building blocks of Internet of Things and characteristics.

Unit No.	Unit Title	Unit Contents	No. of Periods
1	Introduction	What is the Internet of Things? : History of IoT, About IoT, Overview and Motivations, Examples of Applications, Internet of Things Definitions and Frameworks : IoT Definitions, IoT Architecture, General Observations, ITU -T Views, Working Definition, IoT Frameworks, Basic Nodal Capabilities	4
2	FUNDAMENTAL IoT MECHANISMS AND KEY TECHNOLOGIES	Identification of IoT Objects and Services, Structural Aspects of the IoT, Environment Characteristics, Traffic Characteristics, Scalability, Interoperability, Security and Privacy, Open Architecture, Key IoT Technologies, Device Intelligence, Communication Capabilities, Mobility Support, Device Power, Sensor Technology, RFID Technology, Satellite Technology,	4
3	RADIO FREQUENCY IDENTIFICATION TECHNOLOGY	RFID: Introduction, Principle of RFID, Components of an RFID system, Issues EPCGlobal Architecture Framework: EPCIS & ONS, Design issues, Technological challenges, Security challenges, IP for IoT, Web of Things. Wireless Sensor Networks:	6

		History and context, WSN Architecture, the node, Connecting nodes, Networking Nodes, Securing Communication WSN specific IoT applications, challenges: Security, QoS, Configuration, Various integration approaches, Data link layer protocols, routing protocols and infrastructure establishment.	
4	RESOURCE MANAGEMENT IN THE INTERNET OF THINGS	Clustering, Software Agents, Clustering Principles in an Internet of Things Architecture, Design Guidelines, and Software Agents for Object Representation, Data Synchronization. Identity portrayal, Identity management, various identity management models: Local, Network, Federated and global web identity, user -centric identity management, device centric identity management and hybrid -identity management, Identity and trust.	10
5	INTERNET OF THINGS PRIVACY, SECURITY AND GOVERNANCE	Vulnerabilities of IoT, Security requirements, Threat analysis, Use cases and misuse cases, IoT security tomography and layered attacker model, Identity establishment, Access control, Message integrity, Non-repudiation and availability, Security model for IoT.	6
6	BUSINESS MODELS FOR THE INTERNET OF THINGS	Business Models and Business Model Innovation, Value Creation in the Internet of Things , Business Model Scenarios for the Internet of Things. Internet of Things Application : Smart Metering Advanced Metering Infrastructure, e-Health Body Area Networks, City Automation, Automotive Applications, Home Automation, Smart Cards,	9
			Total: 39 Periods

Key Text:

1. Daniel Minoli, Building the Internet of Things with IPv6 and MIPv6: The Evolving World of M2M Communications, First Edition, Willy Publications, 2013.
2. Dieter Uckelmann, Mark Harrison and Florian Michahelles, Architecting the Internet of Things, Springer-Verlag Berlin Heidelberg 2011
   E-copy available: Architecting the Internet of Things (archive.org)
3. Parikshit. Ahalle and Poonam N. Railkar, “Identity Management for Internet of Things”, River Publishers, 2015

Reference Books

1. Hakima Chaouchi, The Internet of Things Connecting Objects to the Web, Willy Publications, 2010
2. Olivier Hersent, David Boswarthick, Omar Elloumi, The Internet of Things: Key Applications and Protocols, Second Edition,WillyPublications, 2015
3. Daniel Kellmereit, Daniel Obodovski, The Silent Intelligence: The Internet of Things, Lightning Source Inc; First Edition, 2014
4. Fang Zhaho, Leonidas Guibas, Wireless Sensor Network: An information processing approach, Elsevier, 2005

[MDSC- IOT(P)] – Internet of Things 1 Credit			
	Course Objective: Students will be explored to the interconnection and integration of the physical world and cyberspace. They are also able to design & develop IOT Devices.		
	Course Outcome: Students will be able to implement the concepts learned		
Unit No.	**Unit Title**	**Unit Contents**	**No. of Periods**
1	Internet of things	Overview, technology of the internet of things, enchanted objects, Design principles for connected devices, Privacy, Web thinking for connected devices; Writing Code: building a program and deploying to a device, writing to Actuators, Blinking Led, Reading from Sensors, Light Switch, Voltage Reader, Device as HTTP Client,HTTP,PushVersusPull; Pachube, Netduino, Sending HTTP Requests—the Simple Way, Sending HTTP, Requests—the Efficient Way	13
2	HTTP	Device as HTTP Server, Relaying Messages to and from the Netduino, Request Handlers, WebHtml, HandlingSensorRequests, Handling Actuator Requests; Going Parallel: Multithreading, Parallel Blinker, prototyping online components, using an API, from prototypes to reality, business models, ethics, privacy, disrupting control, crowdsourcing	13

Total: 26 Periods

Key Text:

1. Daniel Minoli, Building the Internet of Things with IPv6 and MIPv6: The Evolving World of M2M Communications, First Edition, Willy Publications, 2013..
2. Dieter Uckelmann, Mark Harrison and Florian Michahelles, Architecting the Internet of Things, Springer-Verlag Berlin Heidelberg 2011
   E-copy available: [Architecting the Internet of Things (archive.org)]

Reference Books

1. Hakima Chaouchi, The Internet of Things Connecting Objects to the Web, Willy Publications, 2010
2. Olivier Hersent, David Boswarthick, Omar Elloumi, The Internet of Things: Key Applications and Protocols, Second Edition,WillyPublications, 2015

[MDSC-ATS]- Applied Time Series Analysis   4 Credits

Course Objective: To learn, understand patterns and apply statistical methods for the analysis of data that have been observed over time.

Course Outcome : Develop skill set to

1. understand how Time Series data differs from other data types and what components are likely in a given set of Time Series data.
2. communicate effectively on the results of Time Series models and forecasts in a concise manner.
3. make informed decisions on future prospects using Time Series models and forecasts.

Unit	Topic	Details	No. of Periods
1	Introduction to time Series Data	Characteristics of time series data: The nature of Time Series Data, Time Series Statistical Models, Measure of dependence: Autocorrelation and Cross-Correlation, stationary Time Series, Estimation of Correlation, Vector-valued and Multidimensional Series Data; Classical Regression, Exploratory Data Analysis, and Smoothing in the context of Time series Data	12
2	Seasonal Models	Autoregressive Moving Average Models, Difference Equations, Autocorrelation and Partial Autocorrelation, Forecasting, Estimation, Integrated Models for Nonstationary Data, Building ARIMA Models, and Multiplicative Seasonal ARIMA Models	12
3	Spectral Analysis	Cyclical Behavior and Periodicity, The Spectral Density, Periodogram and Discrete Fourier Transform, Nonparametric Spectral Estimation, Parametric Spectral Estimation, Multiple Series and Cross-Spectra, Linear Filters, Dynamic Fourier Analysis and Wavelets, Lagged Regression Models, Signal Extraction and Optimum Filtering, Spectral Analysis of Multidimensional Series	16
4	Advanced Topics	Long Memory ARMA, Fractional Differencing, Unit Root Testing, GARCH Models, Threshold Models, Regression with Autocorrelated Errors, Transfer Function Modeling, and Multivariate ARMAX Models	12
			Total: 52 periods

Key Text(s):

1. Robert H. Shumway, David S. Stoffer, Time Series and It's Applications: With R Examples, Third Edition, Springer, 2011

   Chapters: 1-5

References:

1. Jonathan D.Cryer, Kung-Sik Chan, Time Series Analysis: With Applications in R, Second Edition, Springer Texts in Statistics, Springer, 2008
2. Peter J. Brockwell, Richard A. Davis, Introduction to Time Series and Forecasting, Third Edition, Springer Texts in Statistics, Springer, 2016

[MDSC - TDA] – Topological Data Analysis 4 Credits

Course Objective: To understand complex datasets, where complexity arises from not only the massiveness of the data, but also from the richness of the features. The objective of this subject is to enable the students to become familiar with the new methods in Topological Data Analysis (TDA), from theory, algorithm and application perspectives.

Course Outcome: Student will be able to
1. infer high dimensional structure from low dimensional representations and convert data sets into topological objects.
2. pursue new research directions in the field of TDA and integrate advanced TDA techniques with other areas of data science such as data mining, machine learning, computer graphics, and data visualization.

Unit	Title	Contents	No. of Periods
1	Introduction	Graphs, connected components, topological space, manifold, point clouds.	12
2	Homology	Simplicial Complexes, Convex Set Systems, Delaunay Complexes and Alpha Complexes, Homology Groups, Relative Homology	12
3	Persistent homology	Persistent Homology, Efficient Implementations, Extended Persistence.	12
4	Persistence topology of data	Barcodes, Example of Natural image, Persistence Landscapes: Norms, Convergence, Confidence Intervals, and Stability of Persistence Landscapes, Statistical Inference using Landscapes	16
			Total: 52 Periods

Key Text(s):

1. Edelsbrunner, Herbert, Computational topology : An Introduction, AMS, 2010. Chapters: I, III, IV, VII.

2. Robert Ghrist. Barcodes: The persistent topology of data. Bulletin of the American Mathematical Society (AMS), 45(1): 61–75, 2008.
3. Peter Bubenik, Statistical Topological Data Analysis using Persistence Landscapes, J. of Machine Learning Research 16 (2015), 77-102.

Reference(s):

1. Frédéric Chazal and Bertrand Michel, An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists, 2017.
2. G. Carlsson, Topology and Data, Bulletin of the American Mathematical Society Volume 46(2), 2009.

[MDSC- LSP] – Linux System Programming 4 Credits

Course Objective:

- To learn the different set of system calls for the Linux Operating System
- To understand how the Linux OS manages files, processes and memory
- To implement inter-process communication using different mechanisms

Course Outcome: Student will be able to

1. Assimilate the internal abstractions of any Operating System
2. Utilize the insights gained from how these abstractions were implemented and apply them in other areas of work

Unit	Title	Topics	No. of Periods
1	Introduction	System Calls, Library Functions, Standard C Library, Error handling	8
2	File Management	Overview, File Operations (open, read, write, lseek, close), Atomicity, File Descriptors - relation to open files and duplication, File I/O variations (pread, pwrite, readv, writev), File truncation	8
3	Process Management	Process concept, Process Memory Layout, Virtual Memory Management, Stack Frames, Command line arguments, Environment Variables, Process - Creation, Termination, Execution and Monitoring	12
4	Memory Management	Heap and Stack Memory allocation, Memory Mapping - Creation, Unmapping, File mapping, Synchronization, Anonymous Mapping	12
5	Inter-Process Communication	Signals, Pipes, FIFO, POSIX Semaphores - Named Semaphore and Semaphore Operations, POSIX Shared Memory - Creation, Usage & Removal	12

Total: 52 Periods
Key Text: Michael Kerrisk, The Linux Programming Interface, First Edition, No Starch Press, 2010 Chapters: 3 (3.1 - 3.4), 4 (4.1 - 4.7), 5 (5.1 - 5.8), 6 (6.1 - 6.7), 7, 20, 24, 25, 26, 27, 44 (44.1 - 44.4, 44.6 - 44.8), 49 (49.1 - 49.5, 49.7), 53 (53.1 - 53.3), 54 (54.1 - 54.4)
Reference Texts: Robert Love, Linux System Programming, Second Edition, O'Reilly, 2014

[MDSC-DS] – Distributed Systems 4 Credits

Course Objectives:

- To provide students with contemporary knowledge in parallel and distributed systems
- To equip students with skills to analyze and design parallel and distributed applications
- To provide master skills to measure the performance of parallel and distributed algorithms

Course Outcomes: Students will be able to

1. Apply the principles and concepts in analyzing and designing the parallel and distributed system
2. Reason about ways to parallel problems
3. Gain an appreciation on the challenges and opportunities faced by parallel and distributed systems
4. Understand the middleware technologies that support distributed applications such as RPC, RMI and object based middleware
5. Improve the performance and reliability of distributed and parallel programs

Unit	Description	No. of Periods
1	Characterization Of Distributed Systems: Introduction, Examples of Distributed Systems, Trends In Distributed Systems, Focus On Resource Sharing, Challenges, Case Study: The World Wide Web. System Models: Physical Models, Architectural Models, Fundamental Models	10

2	Networking And Internetworking: Types Of Network, Network Principles, Internet Protocols, Case Studies: Ethernet, Wifi And Bluetooth. Interprocess Communication: The API For The Internet Protocols, External Data Representation And Marshaling, Multicast Communication, Network Virtualization: Overlay Networks, Case Study: MPI	10
3	Remote Invocation: Request-Reply Protocols, Remote Procedure Call, Remote Method Invocation, Case Study: Java RMI Indirect Communication: Group communication, Publish-subscribe systems, Message queues, Shared memory approaches Web Services: Web services, Service descriptions and IDL for web services, A directory service for use with web services, XML security, Coordination of web services, applications of web services.	10
4	Coordination And Agreement: Distributed mutual exclusion, Elections Coordination and agreement in group communication, Consensus and related problems Name Services: Name services and the Domain Name System, Directory services, Case study: The Global Name Service, Case study: The X.500 Directory Service. Time And Global States: Clocks, events and process states , Synchronizing physical clocks , Logical time and logical clocks, Global states, Distributed debugging	11
5	Distributed Transactions: Flat and nested distributed transactions, Atomic commit protocols, Concurrency control in distributed transactions, Distributed deadlocks. Replication: System model and the role of group communication, Fault-tolerant services, Case studies of highly available services: The gossip architecture, Bayou and Coda, Transactions with replicated data Mobile And Ubiquitous Computing: Association, Interoperation, Sensing and context awareness, Security and privacy, Adaptation, Case study: Cooltown	11
	Total: 52 Periods	
	Key Text(s): George Coulouris, Jean Dollimore, Tim Kindberg, Gordon Blair, Distributed Systems-Concepts and Design, Addison Wesley, 2012 Chapters: I - VI	
	References: 1. A. Taunenbaum, Distributed Systems: Principles and Paradigms, Pearson, 2006 2. G. Coulouris, J Dollimore, and T Kindberg, Distributed Systems: Concepts and Design, 5th Edition, Pearson Education, 2012	

***

APPENDIX

S. No.	Paper Code	Paper Title	Employability sector	Entrepreneurial and other skills imparted by the course	Relevance of the course to local, national, regional and global developmental needs
1	MDSC-101; MDSC-101(P)	Applied Linear Algebra; Practicals: Advance Linear Algebra	Academia, Industry	Problem solving, critical thinking	National programme on AI, Global AI needs etc. (scientists, software developers)
2	MDSC-102; MDSC-102(P)	Inferential Statistics; Practicals: Inferential Statistics	Academia, Industry	Problem solving, critical thinking	National programme on AI, Global AI needs etc. (scientists, software developers)
3	MDSC-103; MDSC-103(P)	Optimization Techniques; Practicals: Optimization Techniques	Academia, Industry	Problem solving, critical thinking	National programme on AI, Global AI needs etc. (scientists, software developers)
4	MDSC-104	Computer Organization and Architecture	Industry	Analysis	At all levels
5	MDSC-105	Design and Analysis of Algorithms	Academia, Industry	analysis, design, cognition	National / international (must need capability for computer scientists, software engineers)

6	MDSC-106	Software Lab for Data Visualization	Industry	Problem solving, Data Interpretation Skills	National programme on AI, Global AI needs etc. (scientists, software developers)
7	MDSC-201; MDSC-201(P)	Regression Methods; Practicals: Regression Methods	Academia, Industry	Analysis, Problem Solving	National programme on AI, Global AI needs etc. (scientists, software developers)
8	MDSC-202; MDSC-202(P)	Multivariate Statistical Analysis; Practicals: Multivariate Statistical Analysis	Academia, Industry	Analysis, Problem Solving	National programme on AI, Global AI needs etc. (scientists, software developers)
9	MDSC-203; MDSC-203(P)	Machine Learning; Practicals: Machine Learning	Industry	Modeling, Problem solving, Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
10	MDSC-204	Big Data Analytics	Industry	Analysis	National / international (scientists)
11	MDSC-205	Software Lab in Data Engineering	Industry	Problem Solving, Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
12	MDSC-301; MDSC-301(P)	Stochastic Processes; Practicals:Stochastic Processes	Academia, Industry	Analysis, Problem Solving	National/International (scientists)

13	MDSC-302; MDSC-302(P)	Deep Learning; Practicals: Deep Learning	Industry	Modeling, Problem Solving	National programme on AI, Global AI needs etc. (scientists, software developers)
14	MDSC-303; MDSC-303(P)	Natural Language Processing; Practicals: Natural Language Processing	Industry	Modeling, Problem Solving	National programme on AI, Global AI needs etc. (scientists, software developers)
15	MDSC-304; MDSC-304(P)	Cloud Computing; Practicals: Cloud Computing	Industry	Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
16	MDSC-305	Seminar	Academia, Industry	Problem Solving, Training towards research orientation, critical thinking, communication, presentation	At all levels
17	MDSC-403	Project/Dissertation	Academia, Industry	Modeling/design, problem solving, communication, critical thinking, analytical reasoning	At all levels
18	MDSC-404	Comprehensive Viva Voce	Academia, Industry	Critical Thinking, communication, memory	At all levels

19	MDSC-AI	Artificial Intelligence	Industry	Problem Solving, Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
20	MDSC-CGT	Combinatorics and Graph Theory	Academia, Industry	Analysis, Problem Solving	National/International (scientists)
21	MDSC-RL; MDSC-RL(P)	MDSC-RL: Reinforcement Learning; Practicals: Reinforcement Learning	Industry	Modeling, Problem Solving, Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
22	MDSC-MLO; MDSC-MLO(P)	MDSC-MLO: Machine Learning Operations; Practicals: Machine Learning Operations	Industry	Problem Solving, Analysis, Deploying models	National programme on AI, Global AI needs etc. (scientists, software developers)
23	MDSC-ATS	Applied Time Series Analysis	Academia, Industry	Modeling, Problem Solving, Forecasting	National programme on AI, Global AI needs etc. (scientists, software developers)
24	MDSC-IR	Information Retrieval	Academia, Industry	Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
25	MDSC-NS	Network Security	Academia, Industry	Analysis	National/ International (scientists)

26	MDSC-IoT; MDSC-IoT (P)	Internet of Things; Practicals: Internet of Things Lab	Industry	Problem Solving, Analysis	National programme on AI, Global AI needs etc. (scientists, software developers)
27	MDSC-TDA	Topological Data Analysis	Industry	Problem solving, Analysis	National/ International (scientists)
28	MDSC-LSP	Linux System Programming	Industry	Analysis	National/ International (scientists)
29	MDSC-DS	Distributed Systems	Academia, Industry	Analysis	National/ International (scientists)