Department of Mathematics & Computing, IIT (ISM) Dhanbad

Course Structure for 2 Year M.Tech. (Data Analytics) (W.E.F.: Academic Year 2019-2020)

Semester I

Course No.	**Course Name**	**L**	**T**	**P**	**C**
MCC531	Advanced Data Structures & Algorithms	3	0	0	9
MCC532	Fundamentals of Machine Learning	3	0	0	9
MCC533	Computer Oriented Numerical Methods	3	0	0	9
MCC534	Statistics in Decision Makings	3	0	0	9
MCC535	Soft Computing Techniques	3	0	0	9
MCC536	Advanced Data Structures & Algorithms-Practical	0	0	2	2
MCC537	Fundamentals of Machine Learning-Practical	0	0	2	2
MCC538	Computer Oriented Numerical Methods-Practical	0	0	2	2
	Total	**15**	**0**	**6**	**51**

Semester II

Course No.	**Course Name**	**L**	**T**	**P**	**C**
MCC539	Advanced DBMS	3	0	0	9
MCC540	Neural Networks and Deep Learning	3	0	0	9
MCD5##	D Elective 1	3	0	0	9
XXXXXXX	Open Elective 1	3	0	0	9
XXXXXXX	Open Elective 2	3	0	0	9
MCC541	Advanced DBMS-Practical	0	0	2	2
MCC542	Neural Networks and Deep Learning-Practical	0	0	2	2
	Total	**15**	**0**	**4**	**49**

Semester III

Course No.	Course Name	L	T	P	C
MCC597	Thesis Unit 1	0	0	0	36
	Total	**0**	**0**	**0**	**36**

Semester IV

Course No.	Course Name	L	T	P	C
XXXXXX	Open Elective 3	3	0	0	9
MCD5##	D Elective 2	3	0	0	9
MCC598	Thesis Unit -2	0	0	0	18
	Total	**6**	**0**	**0**	**36**

Course No.	D Elective 1	L	T	P	C
MCD531	Cryptography	3	0	0	9
MCD532	Data Mining	3	0	0	9
MCD533	Numerical Optimization	3	0	0	9
MCD534	Sports Analytics	3	0	0	9
MCD535	Bioinformatics	3	0	0	9
MCD536	Advanced Algorithms for Graphs	3	0	0	9
MCD537	Design of Experiments	3	0	0	9
	D Elective 2				
MCD538	Time Series Analysis	3	0	0	9
MCD539	Big data	3	0	0	9
MCD540	Biostatistics	3	0	0	9
MCD541	GPU Computing	3	0	0	9
MCD542	Video Analytics	3	0	0	9
MCD543	Missing Data Analysis in Survey Sampling	3	0	0	9
MCD544	Software Testing	3	0	0	9

Course No.	Open Elective **	L	T	P	C
MCO531	Stochastic Processes	3	0	0	9
MCO532	Advanced Multivariate Analysis	3	0	0	9
MCO533	Numerical Linear Algebra	3	0	0	9

****** Any one out of this list or from Departmental Electives or from other Departments may be opted subject to the offered by Departments.

MCC531 **Advanced Data Structures and Algorithms** **L-T-P: 3-0-0**

Prerequisite:
Objective: To provide the practical implementation of algorithms using efficient data structures
Outcome: To expose students the various methods of designing techniques.

Course Content	**Unit I** **4 Lectures** Review of Basic Concepts, Asymptotic Analysis of Recurrences. **Unit II** **6 Lectures** Randomized Algorithms. Randomized Quicksort, Analysis of Hashing algorithms. Algorithm Analysis Techniques - Amortized Analysis. Application to Splay Trees. **Unit III** **10 Lectures** External Memory ADT - B-Trees. Priority Queues and Their Extensions: Binomial heaps, Fibonacci heaps, applications to Shortest Path Algorithms. Partition ADT: Weighted union, path compression, Applications to MST **Unit IV** **10 Lectures** Algorithm Analysis and Design Techniques. Dynamic Programming-Bellman-Ford, Greedy Algorithms. Network Flows-Max flow, min-cut theorem, Ford-Fulkerson, Edmonds-Karp algorithm, Bipartite Matching **Unit V** **9 Lectures** NP-Completeness and Reductions, Cook's theorem, Satisfiability. Beyond NP-completeness, Introduction to different algorithms paradigms
Learning Outcome	**Unit I:** Understanding of asymptotic notations and basic mathematical preliminaries used in analyzing algorithms. **Unit II:** This unit will help the students in understanding the randomized algorithms and amortized analysis of running time. **Unit III:** This will help in understanding the concepts and use of B-trees and various heaps. This will help in understanding the concepts and use of B-trees and various heaps. Understanding of union and path data structures and their applications to MST. **Unit IV:** To understand different algorithm designing techniques. This will help students in knowing the concepts of network flows and their use in designing various graph algorithms. **Unit V:** To know the concept of NP-completeness. Help the students in understanding what can be done beyond NP-completeness.
Text Books	1. T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein: Introduction to algorithms, PHI, 3rd Edition, 2010
Reference Books	1. M. A. Weiss, Data Structures and Algorithm Analysis , Addison-Wesley, 2nd Edition, 2013 2. A. V. Aho, J. E. Hopcroft, J. D. Ullman, Data Structures and Algorithms, Addison-Wesley, 1st Edition, 1982 3. S. S. Skiena: The Algorithm Design Manual, Springer, 2nd Edition, 2008 4. J. Kleinberg, E. Tardos, Algorithm Design, Addison-Wesley, 1st Edition, 2005

MCC532 **Fundamentals of Machine Learning** **L-T-P: 3-0-0**

Prerequisite:
Objective: All Data Analytics applications are depended on machine learning techniques. This course will provide the students an exposure about how to use machine learning techniques in Data Analytics.
Outcome: Students will learn the use of Machine Learning in Data Analytics.

Course Content	**Unit I** **8 Lectures** Classification/Regression techniques such as Naive Bayes', decision trees, SVMs **Unit II** **9 Lectures** Boosting/Bagging and linear and non-linear regression, logistic regression, maximum likelihood estimates, regularization, basics of statistical learning theory **Unit III** **9 Lectures** Perceptron rule, multi-layer perceptron, backpropagation, brief introduction to deep learning models **Unit IV** **6 Lectures** Dimensionality reduction techniques like PCA, ICA and LDA **Unit V** **7 Lectures** Unsupervised learning: Clustering, Gaussian mixture models, Some case studies
Learning Outcome	**Unit I:** This unit will help students to understand different classification techniques like: Naive Bayes', decision trees, SVMs. **Unit II:** This unit will help students to understand the Application of different regression techniques in Machine learning. **Unit III:** This unit will help students to get the concept of supervised learning like: ANN and their different application in classification, prediction and other areas of Machine Learning. **Unit IV:** This unit will help students to get the concept of different dimensionality reduction techniques which will be useful for Big data analysis. **Unit V:** This unit will help students to get the concept of different unsupervised learning.
Text Books	1. Andreas C. Müller and Sarah Guido, Introduction to Machine Learning with Python: A Guide for Data Scientists.
Reference Books	1. Kevin P. Murphy and Francis Bach, Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) 2. Shai Shalev-Shwartz and Shai Ben-David, Understanding Machine Learning: From Theory to Algorithms

MCC533 **Computer Oriented Numerical Methods** **L-T-P: 3-0-0**

Prerequisite:
Objective: Student will get knowledge of advanced Numerical Methods and Programming Language Python.
Outcome: Student will able to apply Numerical Methods for solving Real World problems. Student will be able to analyze problems in terms of accuracy and precision

Course Content	**Unit I** **6 Lectures** Introduction to Numerical Python: Basics of Python, Array, Elementary programming concepts, Functions and Modules, libraries like Numpy, Scipy, Matplotlib, IEEE Floating Point Standard, Precision, Arithmetic Accuracy and Errors. **Unit II** **7 Lectures** Advanced Methods for Interpolation: Central Difference Interpolation Formulae, Cubic Splines. **Unit III** **6 Lectures** Numerical Integration: Adaptive Integration, Guassian Quadrature, Multiple Integrals. Solution of Integral Equations. **Unit IV** **6 Lectures** Numerical Solution of linear and Nonlinear Equations: Introduction to methods for solving sparse linear systems, tridiagonal systems, solution of system of nonlinear equations, iterative methods, complex roots of nonlinear equations. **Unit V** **14 Lectures** (a) Solution of Ordinary Differential Equations (ODEs): Explicit Runge-Kutta Methods, Adaptive Stepsize Control, Single Step Methods for Stiff Initial Value Problems, Multistep Methods, Extrapolation methods, Shooting Method, Boundary value problems. (b) Numerical Solutions of Partial Differential Equations (PDEs): Solutions of Elliptic, Parabolic and Hyperbolic PDEs.
Learning Outcome	**Unit I:** To learn the basics of Numerical Python **Unit II:** To understand the advanced methods for Interpolation **Unit III:** Be able to perform Numerical Integration **Unit IV:** Be able to solve linear and Nonlinear Equations using numerical methods **Unit V:** To understand methods for solving Ordinary Differential Equations (ODEs). To understand methods for solving Partial Differential Equations (PDEs).
Text Books	1. Programming Python by Mark Lutz, O'Reilly Media, 2016
Reference Books	1. Numerical Python, by Robert Johansson, Springer Nature (2015) 2. Numerical Methods in Engineering with Python by Jaan Kiusalaas, Cambridge University Press (2013) 3. Numerical Methods for Engineers and Scientists by Joe D. Hoffman, Steven Frankel, CRC Press, 2001 4. Applied Numerical Analysis by Curtis F. Gerald and Patrick O. Wheatley, Pearson Education India, 2004 5. Numerical Analysis by Richard L. Burden, J. Douglas Faires, Cengage Learning, 2015

MCC534 **Statistics in Decision Makings** **L-T-P: 3-0-0**

Prerequisite:
Objective: To equip the students with statistical tools applicable in decision making process.
Outcome: After completion of this course, students will be equipped with the knowledge of analyzing the data generated in real life problems with the help of important statistical tools and techniques.

Course Content	**Unit I** **11 Lectures** Introduction and philosophy of Uncertainty (Basic probability), Basic discrete and continuous distributions: Binomial, Poisson, Normal, Log-normal, Exponential and Bivariate Normal distribution. **Unit II** **10 Lectures** Theory of estimation: Criteria of a good estimator, Point and Interval Estimations, Bayesian Estimation. **Unit III** **8 Lectures** Theory of Hypothesis Testing: Test of mean, variance, proportions, correlation co-efficient, regression coefficient under normality assumption, test of independence of attribute. **Unit IV** **3 Lectures** Non-parametric tests for non-normal population: run test, sign test, Mann-Whitney Wilcoxon U-tests. **Unit V** **7 Lectures** Analysis of variance: One-way and Two-way; Design of Experiments: CRD, RBD and LSD with their applications.
Learning Outcome	**Unit I:** Gives the idea about fundamentals of probability theory with some important discrete and continues probability distributions. **Unit II:** Gives the idea about classical and Bayesian estimation procedures applicable in real life problems. **Unit III:** Gives the idea about testing of hypothesis and important tests applicable in data analytics. **Unit IV:** Gives the idea about non-parametric tests applicable for the situations where normality assumption is doubtful. **Unit V:** Gives the idea about basic analysis of variance and basic design of experiments applicable in most of the industry oriented problems.
Text Books	1. Fundamentals of Mathematical Statistics, Sultan Chand and Sons, by Gupta S.C. and Kapoor, V. K.
Reference Books	1. Introduction to the Theory of Statistics, Tata McGraw-Hill, New Delhi. by Mood M., Graybill F.A. and Boes D.C. 2. An Introduction to Probability Theory and Mathematical Statistics, John Wiley. by Rohatgi, V.K. and Saleh. 3. Linear Statistical Inference and its Applications, Wiley Eastern Ltd, New Delhi. by Rao C. R. 4. Mathematical Statistics, 3rd Ed, Books and Allied (P) Ltd, Kolkata, by Parimal Mukhopadhyay. 5. Theory of Point Estimation, 2nd Ed, Wiley Eastern Ltd. by Lehmann, E.L and Casella G. 6. Testing Statistical Hypotheses, 3rd Ed, Springer. by Lehmann, E.L and Joseph P. Romano. 7. Fundamentals of Applied Statistics, Sultan Chand and Sons, by Gupta S.C. and Kapoor,V. K.

MCC535 **Soft Computing Techniques** **L-T-P: 3-0-0**

Prerequisite:
Objective: To provide exposure to theory application of soft computing techniques in Data Analytics.
Outcome: This course will provide the students an exposure about how to use computing techniques in Data Analytics.

Course Content	**Unit I** **11 Lectures** Fuzzy sets: Membership functions, Basic operations, Fuzzy relations, Defuzzification, Fuzzy inference, Fuzzy rule based system. **Unit II** **11 Lectures** Genetic Algorithm: Working principle, Cross over mutation, roulette wheel selection, tournament selection, population, binary encoding and decoding for any optimization problem, Multi objective GAs, Concepts on non-domination, Tournament selection, Crowding distance operator, Ranking. **Unit III** **9 Lectures** Rough Sets: Lower and upper approximations, Discernibility matrix, Accuracy of Approximations. **Unit IV** **8 Lectures** Hybridization of soft computing tools like Neuro-fuzzy, Rough fuzzy, Rough-Fuzzy-GA, Ant Colony based optimization.
Learning Outcome	**Unit I:** This unit will help students to understand Fuzzy logics. **Unit II:** This unit will help students to understand the Application of Genetic algorithm. **Unit III:** This unit will help students to apply the concept of rough sets and will learn to deal with big data. **Unit IV:** Students will be able to apply different soft computing techniques in this unit.
Text Books	1. G.J. Klir and B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, 1995.
Reference Books	1. M. Mitchell, An Introduction to Genetic Algorithms, MIT Press, 2000. 2. R. L. Haupt and S.E. Haupt, Practical Genetic Algorithms, John Willey & Sons, 2002.

MCC536 **Advanced Data Structures and Algorithms-Practical** **L-T-P: 0-0-2**

Prerequisite:
Objective: To provide the practical implementation of algorithms using efficient data structures
Outcome: Students will learn how to implement different data structures and Algorithms using C or C++.

Course Content	1. Implementation of randomized quicksort algorithm 2. Implementation of hash functions and associated algorithms 3. Implementation of operations on splay trees 4. Implementation of operations on binary heaps 5. Implementation of operations on Fibonacci heaps 6. Implementation on operations on B-Trees 7. Implementation of operations on partition ADT and union-find data structures 8. Implementation of Bellman-Ford algorithm 9. Implementation of Ford-Fulkerson algorithm 10. Implementation of Edmonds-Karp algorithm
Text Books	1. T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein: Introduction to algorithms, PHI, 3rd Edition, 2010
Reference Books	1. M. A. Weiss, Data Structures and Algorithm Analysis , Addison-Wesley, 2nd Edition, 2013 2. A. V. Aho, J. E. Hopcroft, J. D. Ullman, Data Structures and Algorithms, Addison-Wesley, 1st Edition, 1982 3. S. S. Skiena: The Algorithm Design Manual, Springer, 2nd Edition, 2008 4. J. Kleinberg, E. Tardos, Algorithm Design, Addison-Wesley, 1st Edition, 2005

MCC537 **Fundamentals of Machine Learning-Practical (will be carried out Using R/Python)** **L-T-P: 0-0-2**

Prerequisite:
Objective: Data Structures is the basic course of Computer Science. It is required in every field of Computer Science. Objective of this course is to impart knowledge of Data Structures.
Outcome: Students will learn how to implement machine learning techniques using R/Python

Course Content	1. Application of Regression Analysis in Data Analytics. 2. Application of Decision Tree and SVM in Data Analytics 3. Application of Nonlinear Regression in Data Analytics 4. Problems related to Logistic Regression in Data Analytics 5. Application of Statistical Learning in Data Analytics 6. Application of applications of Neural Network and deep Learning in Data Analytics 7. Application of PCA in Data Analytics 8. Application of ICA and LDA in Data Analytics 9. Application of Supervised Learning in Data Analytics 10. Application of Clustering Techniques in Data Analytics
Text Books	1. B. Lantz, Machine Learning with R, Packt Publishing Limited, 2013
Reference Books	1. P. Mueller and L Massaron, Machine Learning (in Python and R), John Wiley & Sons, 2016

MCC538 **Computer Oriented Numerical Methods-Practical** **L-T-P: 0-0-2**

Prerequisite:
Objective: Student will get an exposer about the implementation of Numerical Methods using Python.
Outcome: Student can use Python to Solve Various Numeral Methods using its libraries like Nympy, Scpy, matplotlib

Course Content	1. Basics of Python programming. 2. Implementation of Python Functions and Modules,Use of libraries like Numpy, Scipy, Matplotlib. 3. Python programme for Hermite interpolation, Chebyshev Polynomials, Chebyshev Series,Cubic and B-Splines. 4. Python programme for Adaptive Integration, Guassian Quadrature, Multiple Integrals and Solution of Integral Equations. 5. Python programme for Numerical Solution of linear Equations. 6. Python programme for Solution of Nonlinear Equations. 7. Python programme Solution of Ordinary Differential Equations (ODEs): Explicit Runge-Kutta Methods. 8. Python programme for Multistep Methods, Extrapolation methods. 9. Python programme for Shooting Method, Boundary value problems. 10. Python programme for Numerical Solutions of Partial Differential Equations (PDEs).
Text Books	As provided with Code: MCC5303
Reference Books	As provided with Code: MCC5303

MCC539 **Advanced DBMS** **L-T-P: 3-0-0**

Prerequisite:
Objective: To provide knowledge of advanced DBMS.
Outcome: They will able design data bases used for Data Analytics.

Course Content	**Unit I** **9 Lectures** Object oriented model: Nested relations, modelling nested relations as object model, extension of SQL, object definition and query language (ODL, OQL), object relational database model, storage and access methods. Active databases, Advanced trigger structures, SQL extensions. **Unit II** **11 Lectures** Security and Integrity: Discretionary and mandatory access control; Facilities in SQL, access control models for RDBMS and OODBMS. Distributed Database: Basic Structure, fragmentation algorithms, trade-offs for replication, query processing, recovery and concurrency control; Multi-database systems; Design of Web Databases. **Unit III** **11 Lectures** Data Mining and Warehousing: Association Rule algorithms, algorithms for sequential patterns; Clustering and classification in data mining; Basic structure of a data warehouse; Extension of ER Model, materialistic view creation. **Unit IV** **4 Lectures** On line analytical processing and data cube. Deductive databases, recursive query construction, logical database design and data log. **Unit V** **4 Lectures** One or more of the following topics: (i) Temporal database, (ii) Multimedia database, (iii) Text retrieval and mining, (iv) Web mining, and (v) Any topic of current interest.
Learning Outcome	**Unit I:** At the end of this unit, students will be able to write complex data querries in SQL on relational database models, set triggers, and learn about data storage ance access methods. **Unit II:** At the end of this unit, students will be able to write their first web application based on a distributed database. **Unit III:** At the end of this unit, students will be able store and access data from various database storage systems and mine ana analyze databases for making business decisions. **Unit IV:** At the end of this unit, students will be able design deductive and logical databases and write recursive queries. **Unit V:** At the end of this unit, students will be well-versed with the topics that are selected for teaching.
Text Books	1. Database System Concepts, Korth, Silberschatz and Sudarshan, McGraw Hill.
Reference Books	1. Database: Principles, Programming, Performance, P. O'Neil, Morgan Koffman 2. Principles of Database and Knowledge-Base Systems, J.D. Ullman, Computer Science

MCC540 **Neural Networks & Deep Learning** **L-T-P: 3-0-0**

Prerequisite:
Objective: All Data Analytics applications are depended on Neural Network and Deep Learning.
Outcome: This course will provide the students an exposure about how to use Neural Network and Deep Learning in Data Analytics.

Course Content	**Unit I** **8 Lectures** Introduction: Various paradigms of earning problems, Perspectives and Issues in deep learning framework, review of fundamental learning techniques. **Unit II** **8 Lectures** Feedforward neural network: Artificial Neural Network, activation function, multi-layer neural network. **Unit III** **11 Lectures** Training Neural Network: Risk minimization, loss function, backpropagation, regularization, model selection, and optimization. Conditional Random Fields: Linear chain, partition function, Markov network, Belief propagation, Training CRFs, Hidden Markov Model, Entropy. **Unit IV** **5 Lectures** Deep Learning: Deep Feed Forward network, regularizations, training deep models, dropouts, Convolutional Neural Network, Recurrent Neural Network, Deep Belief Network. **Unit V** **7 Lectures** Probabilistic Neural Network: Hopfield Net, Boltzman machine, RBMs, Sigmoid net, Autoencoders. Deep Learning research: Object recognition, sparse coding, computer vision, natural language processing. Deep Learning Tools: Caffe, Theano, Torch.
Learning Outcome	**Unit I:** This unit will help students to understand deep learning. **Unit II:** This unit will help students to understand the concept of ANN. **Unit III:** This unit will help students to get the concept different training methods of ANN. **Unit IV:** This unit will help students to get the concept of different Deep Learning techniques. **Unit V:** This unit will help students to get the concept of different types of ANN.
Text Books	1. M. Nielsen, Neural Networks and Deep Learning, Determination Press, 2015
Reference Books	1. Charu C. Aggarwal, Neural Networks and Deep Learning: A Textbook, Springer, 2018.

MCO531 **Stochastic Processes** **L-T-P: 3-0-0**

Prerequisite:
Objective: The objective of the course will be to give idea to the students about various Stochastic Processes.
Outcome: This course will be useful for analysis of different financial market data, Business data.

Course Content	**Unit I** **9 Lectures** Definition and classification of general stochastic processes. Markov Chains: definition, transition probability matrices, classification of states, limiting properties. **Unit II** **11 Lectures** Chains with Discrete State Space: Poisson process, birth and death processes. Renewal Process: Renewal Process when Time is discrete, Renewal Process when Time is continues, Renewal Theory and Analysis. **Unit III** **9 Lectures** Markov Process with Continuous State Space: Introduction to Brownian motion, Wiener Process, Differential equation of Wiener Process, Kolmogorov Equations. **Unit IV** **6 Lectures** Markov Decision Process, Branching Process **Unit V** **4 Lectures** Congestion Process: Queuing Process, M/M/1 Queue
Learning Outcome	**Unit I:** This unit will help students to understand basics of Stochastics Processes and Markov Chain. **Unit II:** This unit will help students to understand different type of Point processes and their applications. **Unit III:** This unit will help students to get the concept different continuous State processes with application in finance. **Unit IV:** This unit will help students to get the concept of Markov decision process and its applications. **Unit V:** This unit will help students to get the concept of Queuing theory and its application.
Text Books	1. Stochastic Processes by J. Medhi, New Age International Publication.
Reference Books	1. Elements of Applied Stochastic Processes by U.N. Bhat, John Wiley and Sons. 2. Probability and Statistics with Reliability, Queuing, and Computer Science Applications by K.S. Trivedi, Prentice Hall of India.

MCO532 **Advanced Multivariate Analysis** **L-T-P: 3-0-0**

Prerequisite:
Objective: This course will demonstrate the properties of multivariate distributions and their applications.
Outcome: **S**tudents will learn about the application of Multivariate Analysis techniques in Data Analytics.

Course Content	**Unit I** **7 Lectures** Review of multivariate normal distribution and its properties, distributions of linear and quadratic forms, tests for partial and multiple correlation coefficients and regression coefficients and their associated confidence regions. Wishart distribution. **Unit II** **7 Lectures** Construction of tests, likelihood ratio principles, inference on mean vector, Hotelling's T^2. MANOVA. Inference on covariance matrices **Unit III** **12 Lectures** Discriminant analysis. Principal component analysis, factor analysis and clustering. Unified approach for constructions of probability distributions on $\boldsymbol{R^p}$ **Unit IV** **6 Lectures** Dimension reduction techniques: Principal component and generalized canonical variable analysis –constructions and related inference problems. **Unit V** **7 Lectures** Large *p*–small *n* problems in testing of multi parameter hypotheses: Tests for the mean vector in $\boldsymbol{N_p}(\mu,\Sigma)$, null and non-null asymptotic distributions of their test statistics.
Learning Outcome	**Unit I:** Demonstrate the properties of multivariate distributions such as multivariate normal, Wishart distribution etc. and their applications **Unit II:** Inference on parameters of the multivariate normal distributions and constructions of different tests, MANOVA etc. **Unit III:** Understanding discriminant analysis and principal component analysis for analyzing multivariate data. **Unit IV:** Use principal component analysis effectively for data exploration and data dimension reduction for high dimensional data **Unit V:** Students will learn the theory of statistical inference for high dimensional data analysis.
Text Books	1. T. W. Anderson (1984), An Introduction to Multivariate Statistical Analysis. 2nd Ed. John Wiley C. E. Leiserson, R. L. Rivest, C. Stein: Introduction to algorithms, PHI, 3rd Edition, 2010.
Reference Books	1. R. A. Johnson and D. W. Wichern (2013), Applied Multivariate Statistical Analysis. 6th Ed. Pearson 2. C. R. Rao (2002), Linear Statistical Inference and its Applications. 2nd Ed. Wiley 3. M. S. Srivastava and C. G. Khatri (1979), An Introduction to Multivariate Statistics, Elsevier North Holland, Inc., New York 4. R. J. Muirhead (2009). Aspects of Multivariate Statistical Theory. 2nd Ed. Wiley-Interscience.

MCC541 **Advanced DBMS-Practical** **L-T-P: 0-0-3**

Prerequisite:
Objective: Advanced DBMS is the important course Data Analytics.
Outcome: Students will learn how to use and design Data Base in Data Analytics.

Course Content	1. Accessing the database 2. Basic, intermediate, and advanced SQL 3. Introduction to Python database toolbox 4. Introduction to Git 5. Database access from a programming language 6. Database metadata access from a programming language 7. Create a webpage connected to a database server 8. Create functions to generate HTML 9. Term projects: self-conceived or assigned by the instructor 10. Presentation of the term projects
Text Books	1. Database System Concepts, Korth, Silberschatz and Sudarshan, McGraw Hill.
Reference Books	1. Database: Principles, Programming, Performance, P. O'Neil, Morgan Koffman 2. Principles of Database and Knowledge-Base Systems, J.D. Ullman, Computer Science.

MCC542 **Neural Networks & Deep Learning-Practical** **L-T-P: 0-0-2**

Prerequisite:
Objective: Neural Network & Deep Learning Practical is important course Data Analytics.
Outcome: Students will learn how to use Neural Network & Deep Learning in Data Analytics.

Course Content	1. Implementation of different activation functions to train Neural Network. 2. Implementation of different Learning Rules. 3. Implementation of Perceptron Networks 4. Implementation of Adeline network for system identification. 5. Implementation of Madeline network 6. Pattern matching using different rules. 7. Project related to application of machine learning in healthcare 8. Project related to application of machine learning in business analysis 9. Project related to application of machine learning in sports analytics 10. Project related to application of machine learning in Time Series Analysis & Forecasting.
Text Books	1. S. N. Sivanandam and S. N. Deepa, Introduction to neural networks using Matlab, 2016.
Reference Books	1. Simon Haykin, Neural Networks and Learning Machines, PHI, 2008.

MCO533 **Numerical Linear Algebra** **L-T-P: 3-0-0**

Prerequisite:
Objective: The objective of the course will be to give idea to the students about Numerical Linear Algebra.
Outcome: Students will get an exposure to Matrix Iterative Methods for Analysis Real Life System.

Course Content	**Unit I** **6 Lectures** LU factorization, Pivoting, Cholesky decomposition, Iterative refinement, QR factorization, Gram-Schmidt orthogonalization. **Unit II** **7 Lectures** Projections, Householder reflectors, Givens rotation, Singular Value Decomposition, Rank and matrix approximations, image compression using SVD, Least squares and least norm solution of linear systems **Unit III** **7 Lectures** Pseudoinverse, normal equations, Eigenvalue problems, Gershgorin theorem, Similarity transform, Eigenvalue & eigenvector computations and sensitivity, Power method, Schur decomposition, Jordan canonical form, QR iteration with & without shifts. **Unit IV** **7 Lectures** Hessenberg transformation, Rayleigh quotient, Symmetric eigenvalue problem, Jacobi method, Divide and Conquer, Computing the Singular Value Decomposition, Golub-Kahan-Reinsch algorithm, Chan SVD algorithm, Generalized SVD, Generalized and Quadratic eigenvalue problems. **Unit V** **12 Lectures** Generalized Schur decomposition (QZ decomposition), Iterative methods for large linear systems: Jacobi, Gauss-Seidel and SOR, convergence of iterative algorithms. Krylov subspace methods: Lanczos, Arnoldi, MINRES, GMRES, Conjugate Gradient and QMR, Pre-conditioners, Approximating eigenvalues and eigenvectors.
Learning Outcome	**Unit I:** Understanding the concept of factorization of matrix into two product of two matrices. **Unit II:** Student will able to understand the SVD and its application in real life problem. **Unit III:** Solving eigenvalue and eigenvector problem using numerical linear algebraic technique. **Unit IV:** Student will learn algorithm for the computation of singular value computation. **Unit V:** Students will able to solve the large linear system of equations using Krylov subspace methods.
Text Books	1. B.N. Dutta, Numerical Linear Algebra and Applications, SIAM, 2010.
Reference Books	1. R. Bellman, Introduction to Matrix Analysis, SIAM, 1997 2. R.S. Varga, Matrix Iterative Analysis, Springer, 2000.

MCD531 **Cryptography** **L-T-P: 3-0-0**

Prerequisite:
Objective: To understand the classical and modern cryptosystems for secure encryption and decryption.
Outcome: The students will be able to understand the basic idea of encryption and decryption schemes.

Course Content	**Unit I** **7 Lectures** Brief introduction to number theory, Euclidean algorithm, Euler's totient function, Fermat's theorem and Euler's generalization, Chinese Remainder Theorem, primitive roots and discrete logarithms, Quadratic residues, Legendre and Jacobi symbols. **Unit II** **8 Lectures** Cryptography and cryptanalysis, classical cryptosystems, concept of block and stream ciphers, private and public key cryptography. Encryption standard: DES and differential and linear cryptanalysis, Advanced encryption standards **Unit III** **6 Lectures** RSA public key cryptosystems: RSA system, primality testing. Diffe-Hellman key exchange system. Massey-Omura cryptosystem for message transmission. **Unit IV** **6 Lectures** Other public key cryptosystems: El Gamal public key cryptosystem, algorithms for discrete log problem, Knapsack public key cryptosystems. **Unit V** **12 Lectures** Digital signature and hash functions: El Gamal signature scheme, digital signature standard, onetime undeniable and fail-stop signatures, computationally collision-free hash functions, extending hash functions, examples of hash functions. Introduction to elliptic curves, basic facts, elliptic curve cryptosystems
Learning Outcome	**Unit I:** Students will learn the basics of number theory. **Unit II:** Students will be able to understand classical and public key encryption and decryption techniques. **Unit III:** Students will learn RSA cryptosystems and Massey-Omura cryptosystem for message transmission. **Unit IV:** Students will learn El Gamal and Knapsack cryptosystems and discrete log problem. **Unit V:** The students will be able to understand different signature schemes and properties of hash functions and their applications. Students will be able to understand basics of elliptic curves and their applications in designing cryptosystems
Text Books	1. N. Koblitz, A Course in Number Theory and Cryptography, 2nd Edition, 1994.
Reference Books	1 J. Hoffstein, J. Pipher, J. H. Silverman, An Introduction to Mathematical Cryptography, Springer, 2008. 2. J. Buchmann, Introduction to Cryptography, 2nd Edition, Springer, 2012. 3. D. Stinson Cryptography Theory and Practice, 3rd edition, Chapman & Hall, CRC, 2005.

MCD532 **Data Mining** **L-T-P: 3-0-0**

Prerequisite:
Objective: To provide exposure to theory of Data Mining.
Outcome: Students will get an idea about how to apply Data Mining techniques in Data Analytics.

Course Content	**Unit I** **8 Lectures** Introduction: Introduction to data mining and knowledge discovery from databases. Scalability issues of data mining algorithms**.** **Unit II** **11 Lectures** Introduction to Data warehousing: General principles, modelling, design, implementation, and optimization. **Unit III** **9 Lectures** Data preparation: Pre-processing, sub-sampling, feature selection. Associations, dependence analysis, correlation, rule generation- a priori algorithm, FP Trees etc. and evaluation. **Unit IV** **11 Lectures** Cluster analysis and Outlier Detection, Temporal and spatial data mining: Mining complex types of data. Advanced topics: High performance computing for data mining, distributed data mining.
Learning Outcome	**Unit I:** This unit will help students to understand basics of Data Mining and Scalability issues of data mining algorithms. **Unit II:** This unit will help students to understand the concept of Data warehousing. **Unit III:** This unit will help students to understand the concept of Data preparation techniques. **Unit IV:** This unit will help students to understand the concept of Cluster analysis and Outlier Detection, Temporal and spatial data mining: Mining complex types of data. Advanced topics: High performance computing for data mining, distributed data mining.
Text Books	1. J. Han, M.Kamber and J. Pie, Data Mining: Concepts and Techniques, Elsevier, 2012
Reference Books	1. M. J. Zaki and W. Meira, Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, 2014 2. E. Frank, I. H. Witten and M. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, 2017.

MCD533 **Numerical Optimization** **L-T-P: 3-0-0**

Prerequisite:
Objective: The course deals with the basic idea of mathematical programming (Linear and Nonlinear).
Outcome: The student will get an exposer on how to work out the computational implementation of a numerical algorithm for solving Linear and Nonlinear Programming Problem and do presentations.

Course Content	**Unit I** **8 Lectures** Linear Programming: Review of various techniques of linear programming problems. **Unit II** **12 Lectures** Unconstraint Optimization: One dimensional- Fibonacci methods, Golden Section, Bisection and Newton Method; Higher dimensional- Random search methods, Nelder and Meed method, Steepest descent method, Grid search method, Conjugate gradient methods **Unit III** **10 Lectures** Constraint Optimization: Kuhn Tucker conditions, Zoutendijk's method of feasible directions, Sequential programming approach, Cutting plane methods. Rosen's gradient projection method **Unit IV** **9 Lectures** Penalty Function Methods: Basic approach of the penalty function method, Exterior penalty function method, Interior penalty function method, Augmented Lagrange multiplier method.
Learning Outcome	**Unit I:** Students will get the basic idea of linear programming. **Unit II:** Students will learn new methods to find the solution of the constrained optimization problem. **Unit III:** In this unit, students will be able to obtain the optimality conditions for the constraint optimization problems. **Unit IV:** This unit will help the student to convert the given constrained optimization problems into an equivalent unconstrained optimization problem.
Text Books	1. Nocedaa, J. and Wright, S. J.: "Numerical Optimization", Springer series in Operations Research, Springer-Verlag, 2006.
Reference Books	1 Chandra, S., Jayadeva and Mehra, A.: "Numerical Optimization with Applications", Narosa, 2009. 2. Rao, S. S.: "Engineering Optimization: Theory and Practice", John Wiley, 2009. 3. Bazaraa, M. S., Sherali, H. D. and Shetty, C. M.: "Nonlinear Programming Theory and Algorithms", Second Edition, John Wiley and Sons, 1993.

MCD534 **Sports Analytics** **L-T-P: 3-0-0**

Prerequisite:

Objective: At present it is a very important area in Data Analytics. This course will provide exposure to theory as well as applications related to Sports Analytics

Outcome: The student will get an exposer on how to using Data Analytics Techniques in Sorts Analytics.

Course Content	**Unit I** **6 Lectures** An Introduction to Analytics and Data, The Data Ecosystem **Unit II** **10 Lectures** The Data Game, Strategic Talent Management Analytics, Analytics in Sport Marketing, Choosing performance metrics for players and teams. **Unit III** **9 Lectures** Big data in sports, Predicting outcomes of games, tournaments, and seasons, Machine learning in sports, Evaluating performance through profit and loss **Unit IV** **10 Lectures** Analytics in Digital Marketing, Sport Finance by the Numbers, Sport Law by the Numbers, Sports Telecom, Media and Technology. **Unit V** **4 Lectures** Manufacturing/Production Analytics, Event Management by the Numbers, Facility Management Analytics
Learning Outcome	**Unit I:** This unit will help students to understand basics of Sports Analytics. **Unit II:** This unit will help students to understand Talent Management Analytics, Analytics in Sport Marketing, Choosing performance metrics for players and teams. **Unit III:** This unit will help students to get the concept Big data in sports, Predicting outcomes of games, tournaments, and seasons, Machine learning in sports, Evaluating performance through profit and loss. **Unit IV:** This unit will help students to get the concept of Sport Finance by the Numbers, Sport Law by the Numbers, Sports Telecom, Media and Technology. **Unit V:** This unit will help students to get the concept of Production Analytics, Event Management by the Numbers, Facility Management Analytics.
Text Books	1. Thomas W. Miller, Sports Analytics and Data Science: Winning the Game with Methods and Models, O'Reilly Publication, 2015.
Reference Books	1 G. Fried and C. Mumcu, Sport Analytics, Routledge, Taylor & Francis Group, 2016. 2. F. Green, Winning With Data: CRM and Analytics for the Business of Sports, Routledge, Taylor & Francis Group, 2018. 3. Jim Albert, Mark E. Glickman, Tim B. Swartz and Ruud H. Koning, Handbook of Statistical Methods and Analyses in Sports, CRC Press, 2016.

MCD535 **Bioinformatics** **L-T-P: 3-0-0**

Prerequisite:
Objective: This course will provide knowledge on Bioinformatics.
Outcome: **S**tudents will learn about the application of Bioinformatics in Data Analytics.

Course Content	**Unit I** **9 Lectures** Biological Databases: Organisation, searching and retrieval of information, accessing global bioinformatics resources using internet links. **Unit II** **10 Lectures** Nucleic acids sequence assembly, restriction mapping, finding simple sites and transcriptional signals, coding region identification, RNA secondary structure prediction **Unit III** **10 Lectures** Similarity and Homology, dotmatrix methods, dynamic programming methods, scoring systems, multiple sequence alignments, evolutionary relationships, genome analysis **Unit IV** **10 Lectures** Protein physical properties, structural properties – secondary structure prediction, hydrophobicity patterns, detection of motifs, structural database (PDB). Genome databases, Cambridge structure database, data mining tools and techniques, Structural Bioinformatics.
Learning Outcome	**Unit I:** This unit will help students to understand how to Organize searching and retrieval of information, from Biological Database. **Unit II:** This unit will help students to understand about Nucleic acids sequence assembly, restriction mapping, finding simple sites and transcriptional signals, coding region identification, RNA secondary structure prediction. **Unit III:** This unit will help students to understand the concept of evolutionary relationships, genome analysis. **Unit IV:** This unit will help students to understand the concept of detection and prediction of different Protein Database and Genome Database.
Text Books	1. Gribkov, M., and Devereux, J. (Eds), Sequence Analysis Primer, Stockton Press, 1991.
Reference Books	1. Mount, D.W., Bioinformatics: Sequence and Genome Analysis, Cold. Spring Harbor Laboratory Press, 2001. 2. Baxevanis, A.D., and Ouellette, B.F.F. (Eds), Bioinformatics: A practical guide to the analysis of the genes and proteins, Wiley-Interscience, 1998.

MCD536 **Advanced Algorithms for Graphs** **L-T-P: 3-0-0**

Prerequisite:

Objective: This course will provide knowledge on Graph Theory.

Outcome: **S**tudents will learn about the application of Graph Algorithms in various areas of Data Analytics.

Course Content	**Unit I** **2 Lectures** Basic definitions, Graph representations: adjacency matrix, incidence matrix, adjacency list **Unit II** **8 Lectures** Shortest path problems: Single source SP problem, SP tree, Ford's labelling method, labelling and scanning method, efficient scanning orders— topological order for acyclic networks, shortest first search for non-negative networks (Dijkstra), BFS search for general networks, correctness and analysis of the algorithms; All pair SP problem— Edmond-Karp method, Floyd's algorithm and its analysis **Unit III** **9 Lectures** Flows in Networks: Basic concepts, maxfiow-mincut theorem, Ford and Fulkerson augmenting path method, integral flow theorem, maximum capacity augmentation, Edmond-Karp method, Dinic's method and its analysis, Malhotra-Kumar-Maheswari method and its analysis, Preflow- push method (Goldbergarj Ian) and its analysis; Better time bounds for simple networks. Minimum cost flow: Minimum cost augmentation and its analysis **Unit IV** **11 Lectures** Matching problems: Basic concepts, bipartite matching— Edmond's blossom shrinking algorithm and its analysis; Recent developments. Planarity*:* Basic fact about planarity, polynomial time algorithm. **Unit V** **9 Lectures** Graph isomorphism: Importance of the problem, backtrack algorithm and its complexity, isomorphism complete problems, polynomial time algorithm for planar graphs, group theoretic methods. Applications of graphs in various fields like telecommunications, networking, image processing, pattern recognition, graph cut algorithms
Learning Outcome	**Unit I:** Students will learn the basic definitions on graphs and its representation on a computer. **Unit II:** This unit will help the students in understanding the various shortest path problems and the efficient algorithms associated with those. **Unit III:** This will help in various advanced algorithms in network flows. **Unit IV:** Students will learn the concepts of matchings in graphs and efficient algorithms to compute the matching in a graph. **Unit V:** Students will learn the planarity and the algorithm on testing planarity of a given graph. Students will learn the one of the difficult problem the graph isomorphism problem. Students will learn the application of graphs in different areas.
Text Books	1. T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein: Introduction to algorithms, PHI, 3rd Edition, 2010.
Reference Books	1. D. Junknickel: Graphs, Networks, and Algorithms, Springer, 4th Edition, 2013 2. M. C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, Annals of Discrete Maths. 57, Elsevier, 2nd Edition, 2004 3. C. M. Hoffman: Group Theoretic Algorithms and Graph Isomorphisms, Springer-Verlag, Berlin, 1982 4. C. H. Papadimitriou, K. Stiglitz: Combinatorial Optimization: Algorithms and Complexity, PHI, 1997.

MCD537 **Design of Experiments** **L-T-P: 3-0-0**

Prerequisite:

Objective: In Statistics designing of an experiment means to decide how the measurements or observations should be taken during the experiment so that a valid statistical analysis in economical way may be provided.

Outcome: During the scientific experiments or industrial studies huge data are being generated and to draw meaningful conclusions for the data, design of experiments help in providing the relevant results and interpretations.

Course Content	**Unit I** **5 Lectures** Analysis of variance one way and two-way (with m observations per cell) classifications **Unit II** **11 Lectures** Basic principles of design of experiments, CRD, RBD and LSD and their analysis, estimation of missing observations. **Unit III** **11 Lectures** Factorial experiments: 2^2, 2^3, 3^2 and 3^3 experiments, confounding in 2^3 factorial experiment. **Unit** **7 Lectures** Balanced Incomplete Block Design (BIBD), relation between their parameters. Intra and Inter block analysis of BIBD. **Unit V** **5 Lectures** Split plot and simple lattice designs**.**
Learning Outcome	**Unit I:** Gives the idea about analyzing the variations creep in the data due to various factors for complicated situations. **Unit II:** Gives the idea about fundamentals of design of experiments and some basic designs. **Unit III:** Gives the idea about factorial experiments, where different factors are considered at different levels. **Unit IV:** Gives the idea about incomplete block designs arises in many practical situations. **Unit V:** Gives the idea about split plots and lattice designs when some factors are harder to vary than other factors.
Text Books	1. Design and Analysis of Experiments, 1986, 2nd Ed. by M. N. Das & N. C. Giri, Wiley Eastern Ltd. New Delhi.
Reference Books	1. Design and Analysis of Experiments by D. C. Montgomery, 1984, 2nd Ed., John Wiley & Sons, New York. 2. Fundamentals of Applied Statistics by S. C. Gupta and V. K. Kapoor, 1993, 3rd Ed., Sultan Chand, New Delhi. 3. Experimental Designs, (1992 2nd Ed.) , William G. Cochran, Gertrude M. Cox, Wiley

MCD538 **Time Series Analysis** **L-T-P: 3-0-0**

Prerequisite:

Objective: The objective of the course will be to give idea to the students about Time Series Analysis.

Outcome: This course will be useful for analysis of different time series data, modelling forecasting using the time series data.

Course Content	**Unit I** **12 Lectures** Discrete parameter stochastic processes, strong and weak stationary, autocovariance and autocorrelation. Periodogram and correlogram analysis. Linear Time Series Models: Stationary and Non Stationary Models, model identifications, parameter estimation and forecasting. **Unit II** **9 Lectures** Conditional Heteroscedastic models and their applications. **Unit III** **8 Lectures** Multivariate Time Series Analysis and their Applications **Unit IV** **6 Lectures** PCA and Factor Models and their applications **Unit V** **4 Lectures** Transfer Function models: identification, fitting and application.
Learning Outcome	**Unit I:** This unit will help students to understand basics of Time Series, techniques for identifying different Time Series models. **Unit II:** This unit will help students to understand Heteroscedastic models and their applications. **Unit III:** This unit will help students to get the concept Multivariate Time Series Analysis and their Applications. **Unit IV:** This unit will help students to get the concept of PCA and factor models and their use in Data Analytics. **Unit V:** This unit will help students to get the concept of Transfer Function models with its application in Data Analytics.
Text Books	1. Time Series Analysis, Forecasting and Control by Box and Jenkins, Pearson Education.
Reference Books	1. Introduction to Time Series and Forecasting by P.J. Brockwell and R.A. Davis, Springer. 2. Analysis of Financial Time Series by Ruey S. Tsay 3. Time Series Analysis and Its Applications with R Examples by R.H. Shumway & D.S. Stoffer.

MCD539 **Big data** **L-T-P: 3-0-0**

Prerequisite:
Objective: This course will provide exposure to theory of Big data.
Outcome: Students will able to use Big data analysis in Data Analytics.

Course Content	**Unit I** **8 Lectures** Basic Statistics and R, basic statistical concepts with a brief review of R. **Unit II** **9 Lectures** Relationships and Representations, Graph Databases. Introduction to Spark 2.0 and Hadoop. **Unit III** **9 Lectures** Language processing, Analysis of Streaming Data, Applications of ML Library, Basic Neural Network and Tensor Flow, Advance Tensor Flow **Unit IV** **13 Lectures** Assessing Quality of Big Data Analysis, Analysis of Images, OCR Applications, Analysis of Speech Signal, Page Rank like Search systems, Analysis of Streaming Data with Tensor Flow, VoltDB, Data Flow Engines and other memory databases.
Learning Outcome	**Unit I:** This unit will help students to understand of R and its use in Statistical Analysis. **Unit II:** This unit will help students to understand Relationships and Representations, Graph Databases. **Unit III:** This unit will help students to get the concept Language processing, Analysis of Streaming Data, **A**pplications of ML Library, Basic Neural Network and Tensor Flow, Advance Tensor Flow. **Unit IV:** This unit will help students to get the idea about the applications of Big data in various areas like: Image analysis, Speech and Signal analysis, analysis of Data flow engine, etc.
Text Books	1. H. Karau, Learning Spark: Lightning-Fast Big Data Analysis, O' Rielly Media Inc., 2015.
Reference Books	1. EMC Education Services, Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, John Willey & Sons, 2015. 2. T. Harkness, Big Data: Does Size Matter?, Bloomsbury Publication, 2015.

MCD540 **Biostatistics** **L-T-P: 3-0-0**

Prerequisite:

Objective: During the process of life sciences, medicinal and clinical experiments, precious data are being generated, which need careful and valid statistical analysis for drawing the meaningful conclusions.

Outcome: The course content of this paper has been finalized so that it may be helpful to the students who are intending to join the health sector or pharmaceutical industries.

Course Content	**Unit I** **7 Lectures** Statistical Methods in Clinical Trials: Introduction and its phases I, II, III and IV, statistical designs-fixed sample trials design and Sequential design, Randomization, Blinding. **Unit II** **6 Lectures** Biological Assays, Feller's theorem. Dose-response relationships-qualitative and quantitative response. **Unit III** **11 Lectures** Data editing and transformations, Transformation in general: logarithmic, square root and power transformations; transformations for proportions: angular, probit and logit transformations. Outlying observations: box plot, M-estimators. Test for normality - p-p plot and q-q plot and Kolmogorov-Smirnov test **Unit IV** **7 Lectures** Categorical Data Analysis: Categorical response data, logistic regression-odds ratio, Wald's statistic, logistic regression and its diagnostics memory databases. **Unit V** **8 Lectures** Repeated Measures ANOVA: One Way and Two Classified Data, Epidemiological study designs and its analysis
Learning Outcome	**Unit I:** Gives the idea about fundamentals and different approaches of survey sampling. **Unit II:** Gives the idea about advanced sampling schemes and important estimators. **Unit III:** Gives the elaborative methods for using the auxiliary information at estimation stage in survey sampling. **Unit IV:** Gives the idea about the methods of constructing unbiased estimators in survey sampling. **Unit V:** Provides the concept of successive sampling and introduces some natural estimators.
Text Books	1. Fundamentals of Biostatistics, ANE Books, India by Rastogi, V.B.
Reference Books	1. Biostatistical Analysis, Pearson by Jerrold H. Zar. 2. Biostatistics: A Foundation for Analysis in the Health Sciences, 10th Edition by Daniel W. W and CrossL. C. 3. Biostatistics and Epidemiology – A Primer for Health and Biomedical professionals, 3rd Ed, Springer by Sylvia Wasserthial and Smoller.

MCD541 **GPU Computing** **L-T-P: 3-0-0**

Prerequisite:

Objective: Understand GPU Architecture: CUDA Device, Grid, Blocks, threads, Steaming Multiprocessors (SMs), Warps and Memory organization.

Outcome: Develop Parallel programming skills on the GPU with CUDA: Data-parallel programming for High Performance Computing (HPC) and Optimize CUDA Application Programmes.

Course Content	**Unit I** **9 Lectures** GPU and CUDA programming Architectures : Understanding Parallelism with GPU, device memories and data transfer, kernel functions. Threads, Blocks, Grid, Warps: thread organization, synchronization and transparent scalability, thread assignment, thread scheduling and latency tolerance. **Unit II** **9 Lectures** Memories: Importance of Memory Access Efficiency, Shared Memory, Global Memory, Constant Memory and Texture Memory, a Strategy for Reducing Global Memory Traffic, Memory as a Limiting Factor to Parallelism, Concepts of tiling. Performance Considerations: More on Thread Execution, Global Memory Bandwidth, Dynamic Partitioning of SM Resources. **Unit III** **7 Lectures** Floating Point Considerations: Floating-Point Format, Representable Numbers, Special Bit Patterns and Precision, Arithmetic Accuracy and Rounding, Algorithm Considerations. Streams and Multi GPU Solutions: Atomic Operations, Single Stream, Multiple Streams, GPU Work Scheduling, Zero-Copy Host Memory, Portable Pinned Memory **Unit IV** **7 Lectures** General purpose GPU computing with PyCUDA and PyOpenCL: An overview of OpenCL, Important OpenCL concepts and Basic Program Structure, NumbaPro. **Unit V** **7 Lectures** GPU Computing Applications- A Case Study in Machine Learning (8 Lectures): Convolution Neural Networks, Convolution Layer, Reduction of Convolution Layer to Matrix Multiplication, cuDNN Library.
Learning Outcome	**Unit I:** To understand GPU and CUDA programming Architectures Parallelism. To understand thread organization, synchronization. **Unit II:** To know the concept of Memories and their Importance as a Limiting Factor to Parallelism To understand Performance concepts. **Unit III:** To learn the basics of Floating-Point Format, Representable Numbers. To understand the Arithmetic Accuracy and Rounding concepts. To know the Streams and Multi GPU Solutions concepts **Unit IV:** To learn General purpose GPU computing with PyCUDA and PyOpenCL. **Unit V:** To apply the GPU Computing for in Case Study in Machine Learning.
Text Books	1. CUDA by Example: An Introduction to General-Purpose GPU Programming, by Jason Sanders, Edward Kandrot, publisher Addison-Wesley Professional, 2010.
Reference Books	1. Programming Massively Parallel Processors: A Hands-on Approach by David B. Kirk, Wen-mei W. Hwu, Elsevier, 2010 2. Hands-On GPU Programming with Python and CUDA: Explore high-performance parallel computing with CUDA by Dr. Brian Tuomanen, Packt Publishing Ltd, 2018

MCD542 **Video Analytics** **L-T-P: 3-0-0**

Prerequisite:
Objective To provide exposure to theory as well as practical systems used in Video Analytics.
Outcome: The students will learn the how to use Video Analytics to analyse Image data.

Course Content	**Unit I** **8 Lectures** Introduction to Digital Image and Video Processing, Background Modeling. **Unit II** **9 Lectures** Object Detection and Recognition, Image and Motion Features, Multi Object Tracking. **Unit III** **11 Lectures** Trajectory Analysis, Activities and Events, Anomaly Detection, Compressed Domain Video Analytics, Multi Camera Surveillance, Camera Coordination, Video Indexing **Unit IV** **11 Lectures** Mining and Retrieval. Deep learning for Vision and Image Processing: CNN, RNN, Vision and Language: Image captioning, Visual Q & A.
Learning Outcome	**Unit I:** This unit will help students to understand basics Video processing. **Unit II:** This unit will help students to understand Object Detection and Recognition, Image and Motion Features, Multi Object Tracking. **Unit III:** This unit will help students to get the concept Trajectory Analysis, Activities and Events, Anomaly Detection, Compressed Domain Video Analytics, Multi Camera Surveillance, Camera Coordination, Video Indexing. **Unit IV:** This unit will help students to get the concept of Mining and Retrieval. Deep learning for Vision and Image Processing: CNN, RNN, Vision and Language: Image captioning, Visual Q & A.
Text Books	1. Richard Szeliski, Computer Vision: Algorithms and Applications, Springer 2010.
Reference Books	1. Forsyth, D.A., and Ponce, J., Computer Vision: A Modern Approach, Pearson Education, 2003.

MCD543 **Missing data analysis in Survey Sampling** **L-T-P: 3-0-0**

Prerequisite:

Objective: The aim of this paper is to teach the students about various techniques which deals with how to manage with the incomplete survey data for providing the desired estimates

Outcome: The students will learn how to using missing data analysis techniques in Data Analytics.

Course Content	**Unit I** **3 Lectures** Basics of survey sampling: sampling frame, sampling design, basic principles of survey sampling. **Unit II** **13 Lectures** Types of sampling: Probability sampling and non-probability sampling, Probability sampling: Simple random sampling, Stratified random sampling, Systematic sampling, Cluster and stage sampling. **Unit III** **8 Lectures** Use of auxiliary information at estimation stage, Ratio, product, difference and regression methods of estimation and their properties **Unit IV** **9 Lectures** Introduction of Missing data: Reasons and types of missing data; Techniques to handle the missing data: Imputation method; Mean, Ratio and Regression methods of imputations; Hansen and Hurwitz (1946) technique **Unit V** **6 Lectures** Sensitive variable; Randomized response techniques– Warner's, Simmon's and Two Stage response techniques**.**
Learning Outcome	**Unit I:** Gives the idea about fundamentals of survey sampling. **Unit II:** Gives the idea about important sampling schemes applicable in most of the practical situations. **Unit III:** Gives the methods of using the auxiliary information at estimation stage in survey sampling. **Unit IV:** Gives the idea about handling the missing data problems in survey sampling arise due to non-response. **Unit V:** Gives the idea about tackle the problems of non-response in survey sampling when it occurs due to sensitive nature of characteristics.
Text Books	1. Sampling Techniques, 3rd Ed, Wiley Eastern Ltd. by Cochran, W.G.
Reference Books	1. Sampling Theory of Surveys with Applications, IASRI New Delhi, 1984 Ed. by Sukhatme P V., Sukhatme B. V. and Sukhatme S., and Ashok C. 2. Sampling Theory and Methods, Statistical Publishing Society, Calcutta. by Murthy, M.N. 3. Advanced Sampling – Theory with Applications, Kluwer Publications by Sarjinder Singh. 4. Sampling Theory, Narosa Publications, New Delhi by Desraj and Chandhok P.

MCD544 **Software Testing** **L-T-P: 3-0-0**

Prerequisite:

Objective: The aim of this paper is to provide exposure to art of software testing

Outcome: Student will get an idea about how to generate test tools and how to automate the testing strategies.

Course Content	**Unit I** **11 Lectures** Introduction to Software Testing: Fundamentals of Verification and Testing, Review of software development models, Test Metrics, Software Testing Principles, Testing and Debugging, Software Quality, Requirement Behaviour and Correctness, Fundamentals of Test Process, The Tester's Role in a Software Development Organization **Unit II** **6 Lectures** Static Testing: Structured examination, Control flow & Data flow, Determining Metrics. **Unit III** **6 Lectures** Dynamic Testing: Black Box Testing, Black Box Testing, Gray Box Testing, Intuitive and Experience Based Testing. **Unit IV** **9 Lectures** Test Management: Test Organization, Test Planning, Test Strategies, Levels of Testing, Testing Tools Automation of Test Execution: Types of test Tools, Selection and Introduction of Test Tools. **Unit V** **7 Lectures** Testing Object Oriented Software: Introduction to Object Oriented testing concepts, Differences in Object Oriented testing, testing Object Oriented systems**.**
Learning Outcome	**Unit I:** This unit will help students to understand fundamentals of Software Testing, concept of Software Quality, different metrics related to software testing. **Unit II:** This unit will help students to understand about Static testing techniques of Software. **Unit III:** This unit will help students to understand about Dynamic testing techniques of Software. **Unit IV:** This unit will help students to get the concept of Test Management: Test Organization, Test Planning, Test Strategies, Levels of Testing, Testing Tools Automation. **Unit V:** This unit will help students will learn about different methods about testing of Object Oriented Software.
Text Books	1. G. J. Myers, C. Sandler and T. Badgett, The Art of Software Testing, John Willey & Sons, 2015.
Reference Books	1. E. Dustin, T. Garrett and B. Gauf, Implementing Automated Software Testing: How to Save Time and Lower Costs While Raising Quality, Pearson, 2009. 2. L. Tamres, Introducing Software Testing, Pearson, 2002.