

M.Sc APPLIED STATISTICS AND DATA ANALYTICS
(Academic Year 2022-23 onwards)

Program Outcomes

- PO1 Knowledge in Statistics and Data Analytics:** Understand the basic concepts, fundamental principles and the scientific theories related to Statistics and Data Analytics.
- PO2 Abstract thinking:** Ability to absorb and understand the abstract concepts that lead to various advanced theories in mathematics and Statistics.
- PO3 Modelling and solving:** Ability in modelling and solving problems by identifying and employing the appropriate existing theories and methods.
- PO4 Advanced theories and methods:** Understand advanced theories and methods to design solutions for complex statistical problems in Data Science.
- PO5 Applications in Engineering and Sciences:** Understand the role of statistics and apply the same to solve the real life problems in various fields of study.
- PO6 Modern software tool usage:** Acquire the skills in handling scientific tools towards solving problems and solution analysis in Data Science.
- PO7 Environment and sustainability:** Understand the significance of preserving the environment towards sustainable development.
- PO8 Ethics:** Imbibe ethical, moral and social values in personal and social life leading to highly cultured and civilized personality. Continue to enhance the knowledge and skills in applied statistics and data analytics for constructive activities and demonstrate highest standards of professional ethics.
- PO9 Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
- PO10 Communication:** Develop various communication skills such as reading, listening, speaking and discussing which will help in expressing ideas and views clearly and effectively.
- PO11 Project management and Research:** Demonstrate knowledge, understand the scientific and management principles and apply these to one's own work, as a member/ leader in a team to manage projects and multidisciplinary research environments. Also use the research-based knowledge to analyse and solve advanced problems in data sciences.
- PO12 Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Curriculum
(effective from the academic year 2022-23)

Semester I

Course code	Course	L T P	Credit	ES
22MAT521	Data Structures and Algorithms	3 0 2	4	C
22MAT522	Introduction to Data Analytics with R Programming	3 0 2	4	E
22MAT523	Linear Algebra	3 0 2	4	A
22MAT524	Optimization Techniques	3 1 0	4	D
22MAT525	Probability Theory and Estimation	3 0 2	4	B
22MAT526	Python Programming	3 0 2	4	F
21CUL501	Cultural Education	2 0 0	P/F	G
	Total		24	

Semester II

Course code	Course	L T P	Credit	ES
22MAT527	Big Data Analytics and Hadoop	3 0 2	4	D
22MAT528	Database Management	3 0 2	4	F
22MAT529	Data Mining	3 1 0	4	E
22MAT530	Machine Learning	3 0 2	4	C
22MAT531	Multivariate Statistics and Regression Analysis	3 0 2	4	B
22MAT532	Statistical Inference and Design of Experiments	3 0 2	4	A
21AVP501	Amrita Value Programme	1 0 0	1	G
22AVP103	Mastery Over Mind	1 0 2	2	
	Total		27	

Semester III

Course code	Course	L T P	Credit	ES
22MAT621	SQC and Reliability Theory	3 0 2	4	A
22MAT620	Deep Learning	3 0 2	4	B
	Elective I	3 0 0	3	D
	Elective II	3 0 0	3	E
	Elective II	3 0 0	3	F
22MAT693 [@]	Live-in-Lab. [@] / Open Elective *	2 0 0	2	J
	Total		19	

Semester IV

Course code	Course	L T P	Credit	ES
22MAT695	Dissertation		10	P
	Total		10	

Total credits for the programme: 80

[@] Course code for Live in Lab

ELECTIVES (any three)

Course code	Course	L T P	Credit	ES
22MAT731	Business Analytics	3 0 0	3	D/E
22MAT732	Categorical Data Analysis	3 0 0	3	D/E
22MAT733	Computational Biology	3 0 0	3	D/E
22MAT734	Computer aided drug designing	3 0 0	3	D/E
22MAT735	Demography and Actuarial Statistics	3 0 0	3	D/E
22MAT736	Healthcare Analytics	3 0 0	3	D/E
22MAT737	Market Analytics	3 0 0	3	D/E
22MAT738	Mining of Massive Datasets	3 0 0	3	D/E
22MAT739	Official Statistics	3 0 0	3	D/E
22MAT740	Parallel and Distributed Systems	3 0 0	3	D/E
22MAT741	Pattern Recognition	3 0 0	3	D/E
22MAT742	Queuing Theory	3 0 0	3	D/E
22MAT743	Reinforcement Learning	3 0 0	3	D/E
22MAT744	Sampling Techniques	3 0 0	3	D/E
22MAT745	Social Network Analytics	3 0 0	3	D/E
22MAT746	Special Distribution Functions	3 0 0	3	D/E
22MAT747	Stochastic Process	3 0 0	3	D/E
22MAT748	Survival Analysis	3 0 0	3	D/E
22MAT749	Taugchi Techniques	3 0 0	3	D/E
22MAT750	Thanking with Data	3 0 0	3	D/E

OPEN ELECTIVES PG

Course Code	Course Title	L – T – P	Cr.	ES
21OEL631	Advanced Statistical Analysis for Research	2 0 0	2	D/E
21OEL632	Basics of PC Software	2 0 0	2	D/E
21OEL633	Computer Hardware and Networking	1 0 1	2	D/E
21OEL634	Consumer Protection Act	2 0 0	2	D/E
21OEL635	Corporate Communication	2 0 0	2	D/E
21OEL636	Design Studies	2 0 0	2	D/E
21OEL637	Disaster Management	2 0 0	2	D/E
21OEL638	Essentials of Cultural Studies	2 0 0	2	D/E
21OEL639	Foundations of Mathematics	2 0 0	2	D/E
21OEL640	Foundations of Quantum Mechanics	2 0 0	2	D/E
21OEL641	Glimpses of Life through Literature	2 0 0	2	D/E
21OEL642	Information Technology in Banking	2 0 0	2	D/E
21OEL643	Knowledge Management	2 0 0	2	D/E
21OEL644	Marketing Research	2 0 0	2	D/E
21OEL645	Media for Social Change	2 0 0	2	D/E
21OEL646	Media Management	2 0 0	2	D/E

21OEL647	Object-Oriented Programming	2 0 0	2	D/E
21OEL648	Painting and Sculpture	1 0 1	2	D/E
21OEL649	Personal Finance	2 0 0	2	D/E
21OEL650	Principles of Advertising	2 0 0	2	D/E
21OEL651	Principles of Packaging	2 0 0	2	D/E
21OEL652	Scripting for Rural Broadcasting	1 0 1	2	D/E
21OEL653	Social Media Website Awareness	1 0 1	2	D/E
21OEL654	Theatre Studies	1 0 1	2	D/E
21OEL655	Writing for Technical Purposes	2 0 0	2	D/E
21OEL656	Yoga and Personal Development	1 0 1	2	D/E
21OEL657	Fundamentals of Legal Awareness	2 0 0	2	D/E

***One Open Elective** course is to be taken by each student, in the third semester, from the list of Open electives offered by the School.

@Students undertaking and registering for a Live-in-Lab project, can be exempted from registering for the Open Elective course in the third semester.

M.Sc Applied Statistics and Data Analytics
Syllabus
2022 Admissions onwards

SEMESTER-I

22MAT521

Data Structures and Algorithms

3 0 2 4

Course outcomes

CO1: Understand the various basic data types and trees.

CO2: Gain knowledge about standard data structures like Stack, Queue, list - array, linked list.

CO3: Know the importance of priority queue - Heaps; heap-based implementations; applications of heaps - sorting; Search Tree - Binary search tree.

CO4: To understand the basic concepts of time complexity and classes of time complexity.

CO5: To gain knowledge about some graph algorithms like shortest path algorithm and minimal spanning tree algorithms.

Unit I

Abstraction - Abstract data types; Data Representation; Elementary data types; Basic concepts of data Structures; Mathematical preliminaries - big-Oh notation; efficiency of algorithms; notion of time and space complexity; performance measures for data structures. ADT array - Computations on arrays - sorting and searching algorithms.

Unit-II

ADT Stack, Queue, list - array, linked list, cursor based implementations of linear structures. ADT Tree - tree representation, properties traversal of trees; ADT- Binary Trees – properties and algorithms .

Unit-III

ADT Priority Queue - Heaps; heap-based implementations; applications of heaps - sorting; Search Tree - Binary search tree; balanced binary search trees - AVL tree; Applications of Search Trees - TRIE; 2-3-4 tree; concept of B-Tree. ADT Dictionary - array based and tree based implementations; hashing - definition and application.

Unit-IV

Introduction to time complexity. Big-O, worst case complexity, polynomial classifications. Satisfiability, NP Complete and NP Hard (Definitions only).

Unit-V

Graphs algorithms: ADT- Data structure for graphs - Graph traversal- Transitive Closure- Directed Acyclic graphs - Weighted graphs – Shortest Paths - Minimum spanning tree – Greedy Methods for MST. Travelling salesman problem.

Text Books:

1. Goodrich M T, Tamassia R and Michael H. Goldwasser, “Data Structures and Algorithms in Python++”, Wiley publication, 2013.

Reference Books:

1. Goodrich M T and Tamassia R, “Data Structures and Algorithms in Java”, Fifth edition, Wiley publication, 2010.
2. Tremblay J P and Sorenson P G, “An Introduction to Data Structures with Applications”, Second Edition, Tata McGraw-Hill, 2002.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	3	2	2	2	2					1	1
CO2	3	3	3	2	2	2	2					1	1
CO3	3	3	3	2	2	3	2					1	1
CO4	3	2	3	2	2	2	2					1	1
CO5	2	2	2	1	1	2	1					1	1

22MAT522 Introduction to Data Analysis with R Programming

3 0 2 4

Course Outcomes

- CO1: Exploring and implementing exploratory data analytics
- CO2: Understanding correlation and regression and visualising them using R
- CO3: Understanding supervised learning through linear and logistic regressions
- CO4: Understanding and implementing classifiers for unsupervised data
- CO5: Exploring Massive data sets and implementing classification algorithms

Overview of R software, Data Frames, R Scripts, creating, importing/exporting and merging of datasets, creating matrices and basic matrix operations in R, 2d/3d plotting, programming in R (for, if else, do and while loops),.functions, creating report using R markdown.

Exploring data using R, Scatter plot, histogram, bar chart, pie chart, box plot, basic statistics computation (mean, median, variance etc.)

Generating random samples from standard distributions (such as Bernoulli, Poisson, Normal, Exponential etc.) and comparing theoretical pdfs/pmfs using histograms/frequency distributions, quantiles of sampling distributions (t, chi and F distribution).

Maximization/minimization of functions in R (some algorithm), MLE estimation. Polynomial fitting of scatter plot, introducing regression line, least squares estimates,

residual plots, testing normality of residuals (qqplot), goodness of fit measures and tests, testing of regression parameters, simulation of regression model, empirical distribution of least square estimator and its comparison with theoretical distribution.

Simulation of multivariate normal random vectors, estimation of mean and covariance matrix, eigen values and eigen vector of variance covariance matrix, spectral decomposition covariance matrix.

Generating dependent random variables with so

Text Books /Reference books:

1. James, G., Witten, D., Hastie, T., and Tibshirani, R., An introduction to statistical learning with applications in R, Springer, New York, 2013.
2. Wickham, H., Advanced R, CRC press, New York, 2015.
3. Wickham, H., and Grolemund, G., R for Data Science, O’Reilly Media Inc, Canada, 2017.
4. Data Science and big data analytics: Discovering, analyzing, visualizing and presentating data ,EMC Education Services,John Wiley 2015
5. John Hopcroft and Ravi Kannan, “Foundations of Data Science”, ebook, Publisher, 2013.
6. Amir D Aczel, Jayavel Soundarapandian, Complete Business Statistics, Seventh edition, McGraw Hill,NewDelhi

CO-PO Mapping:

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 5	PO 6	PO 7	PO 8	PO 9	PO1 0	PO1 1	PO1 2
CO 1	3	3	3	2	2	2	2					3	2
CO 2	3	3	3	2	2	2	2					3	2
CO 3	3	3	3	2	2	3	3					3	2
CO 4	3	2	3	2	2	2	3					3	2
CO 5	2	2	2	2	3	2	3					3	2

22MAT523

Linear Algebra

3 0 2 4

Course outcomes

CO-1: To understand the axioms in the definition of a vector space through examples; to understand Subspaces / Quotient Space / mappings and identify them; To familiarize the concept of basis and its relevance.

CO-2: To understand inner products and compute the angle/length of a vector. To construct the orthonormal basis.

CO-3: To familiarize the types of matrices, understand their properties and apply them in the real quadratic forms

CO-4: To understand the construction of matrices for a linear transformation in the triangular

CO 5 To understand the concepts of Eigen Values, Eigen Vectors & Diagonalization

Unit-I

Vector Spaces: Vector spaces - Sub spaces - Linear independence - Basis – Dimension.

Unit-II

Inner Product Spaces: Inner products - Orthogonality - Orthogonal basis - Gram Schmidt Process - Change of basis - Orthogonal complements - Projection on subspace - Least Square Principle.

Unit-III

Linear Transformations: Positive definite matrices - Matrix norm and condition number - QR-Decomposition - Linear transformation - Relation between matrices and linear transformations - Kernel and range of a linear transformation - Change of basis - Nilpotent transformations - Trace and Transpose, Determinants, Symmetric and Skew Symmetric Matrices, Adjoint and Hermitian Adjoint of a Matrix, Hermitian, Unitary and Normal Transformations, Self Adjoint and Normal Transformations, Real Quadratic Forms.

Unit-IV

Eigen values and Eigen vectors: Problems in Eigen Values and Eigen Vectors, Diagonalization, Orthogonal Diagonalization, Quadratic Forms, Diagonalizing Quadratic Forms, Conic Sections. Similarity of linear transformations - Diagonalisation and its applications - Jordan form and rational canonical form.

Unit-V

Quotient Space. Symmetric and Non-degenerate bilinear forms. Decompositions : LU,QR and SVD

Text Books

Howard Anton and Chris Rorres, “Elementary Linear Algebra”, Tenth Edition, John Wiley & Sons, 2010.

Reference Books:

1. Nabil Nassif, Jocelyne Erhel, Bernard Philippe, Introduction to Computational Linear Algebra, CRC press, 2015.
2. Gilbert Strang, “Linear Algebra and Its Applications”, Fourth Edition, Cengage, 2006.
3. Kenneth Hoffmann and Ray Kunze, Linear Algebra, Second Edition, Prentice Hall, 1971.
4. I. N. Herstein, ‘Topics in Algebra’, Second Edition, John Wiley and Sons, 2000.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	2	2	2	2	1					2	2

CO2	3	3	2	2	2	2	1					2	1
CO3	3	3	2	2	2	3	1					2	1
CO4	3	2	2	2	2	2	1					2	1
CO5	3	2	2	1	1	2	1					2	1

22MAT524

Optimization Techniques

3 1 0 4

CO1. Understand different types of Optimization Techniques in engineering problems.
CO2. Learn gradient based Optimizations Techniques in single variables as well as multi-variables (non-linear).
CO3. Understand the Optimality criteria for functions in several variables and learn to apply OT methods like Unidirectional search and Direct search methods.
CO4. Learn constrained optimization techniques. Learn to verify Kuhn-Tucker conditions and Lagrangian Method.
CO5. Understand and solve the integer linear programming and dynamic programming.

Course outcomes

Unconstrained optimization using calculus (Taylor's theorem, convex functions, coercive functions). Unconstrained optimization via iterative methods (Newton's method, Gradient/conjugate gradient based methods, Quasi-Newton methods). Constrained optimization (Penalty methods, Lagrange multipliers, Karush-Kuhn-Tucker conditions).

Introduction to Linear Programming: Lines and hyperplanes, Convex sets, Convex hull, Formulation of a Linear Programming Problem, Theorems dealing with vertices of feasible regions and optimality, Graphical solution. Simplex method (including Big M method and two-phase method), Dual problem, Duality theory, Dual simplex method, Revised simplex method.

Texts /References:

1. Beale, E.M.L., and Mackley, L. , Introduction to Optimization, John Wiley & Sons, Hoboken, 1988.
2. Chavatal, V., Linear Programming, W.H. Reeman and Company, New York, 1983.
3. Chong, E.P.K. and Zak, S.H., An Introduction to Optimization, 4th Edition, John Wiley & Sons, Hoboken, 2013.
4. Joshi, M.C., and Moudgalya, K., Optimization: Theory and Practice, Narosa, New Delhi, 2004.
 1. HamdyA.Taha(1987):OperationsResearch– An Introduction, 4/e, Prentice Hall of India ,Private Ltd, NewDelhi.
 2. S.S. Rao, “Optimization Theory and Applications”, Second Edition, New Age International (P) Limited Publishers, 1995.
 3. Kalyanmoy Deb, “Optimization for Engineering Design Algorithms and Examples”, Prentice Hall of India, New Delhi, 2004.

CO-PO Mapping:

	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO1	PO1	PO1
--	----	----	----	----	----	----	----	----	----	----	-----	-----	-----

	1	2	3	4	5	5	6	7	8	9	0	1	2
CO 1	3	3	3	2	2	2	2					1	1
CO 2	3	3	3	2	2	2	2					1	1
CO 3	3	3	3	2	2	3	2					1	1
CO 4	3	2	3	2	2	2	2					1	1
CO 5	2	2	2	1	1	2	1					1	1

22MAT525

Probability Theory and Estimation

3 0 2 4

Course outcomes

CO1: Understand the basics of probability, random variables and distribution functions.

CO2: Gain knowledge about standard statistical distributions and their properties

CO3: Know the importance of two dimensional random variables and correlation studies

CO4: To gain knowledge point estimation and properties

CO5: To gain knowledge about sampling distributions interval estimations.

Unit-I

Review of probability concepts - conditional probability- Bayes theorem.

Random Variable and Distributions: Introduction to random variable – discrete and continuous random variables and its distribution functions- mathematical expectations – moment generating function and characteristic function -

Unit-II

Standard distributions - Binomial, Multinomial, Poisson, Uniform, exponential, Weibull, Gamma, Beta, Normal. Mean, variance and applications of these distributions- Chebyshev's theorem and central limit theorem.

Unit-III

Joint, marginal and conditional probability distributions for discrete and continuous cases, stochastic independence, expectation of two dimensional random variables, conditional mean and variance, correlation and introduction to regression.

Unit-IV

Point estimation, properties, methods of estimating a point estimator, minimum risk estimators Sampling distributions of mean and variance, Central and Non-central distributions of t, F and Chi-Square distribution. Central limit theorem.

Unit-V

Interval estimation- Confidence interval for one mean, difference of two means, single proportion, difference of two proportions, single variance, ratio of two variances.

TEXT BOOKS /REFERENCE BOOKS:

4. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers and Keying Ye, Probability and Statistics for Engineers and Scientists, 8th Edition, Pearson Education Asia, 2007.
5. Douglas C. Montgomery and George C. Runger, Applied Statistics and Probability for Engineers , John Wiley and Sons Inc., 2005.
6. Ravichandran, J: Probability and Statistics for engineers, First Reprint Edition, Wiley India, 2012.
7. Hoel, P.G., Port, S.C., and Stone, C.J., Introduction to Probability Theory, Universal Book Stall, New Delhi, 1998.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	3	2	2	2	1					3	2
CO2	3	3	3	2	2	2	1					3	2
CO3	3	3	3	2	2	3	1					3	2
CO4	3	2	3	2	2	2	1					3	2
CO5	3	2	3	1	1	2	1					3	2

22MAT526

Python Programming

3 0 2 4

Course outcomes

CO-1: Understand the basic data types and string operations.
CO-2: Understand and apply various function calls in Python.
CO-3: Familiarise and implement boolean expressions, logical operators and executive statements.
CO-4: Execute the Python programme for tree traversals and search problems.
CO-5: Understand and apply the concepts of dictionaries and lists in Python programme.

Unit I

Introduction: History of Python, Need of Python Programming, Applications Basics of Python Programming, Running Python Scripts, Installing Python on Your Computer, Using the Terminal Command Prompt, IDLE, and Other IDEs, Variables, Assignment, Keywords, Input-Output, Indentation.Types, Operators and Expressions: Types - Integers, Strings, Booleans; Operators- Arithmetic Operators, Comparison (Relational) Operators, Assignment Operators, Logical Operators, Bitwise Operators, Membership Operators, Identity Operators, Expressions and order of evaluations Control Flow- if, if-elif-else, for, while, break, continue, pass .Case Study: An Investment Report and Approximating Square Roots.

Unit II

Data Structures: Lists - Operations, Slicing, Methods; Tuples, Sets, Dictionaries, Sequences Comprehensions. Case Study: Nondirective Psychotherapy

Unit III

Functions: Defining Functions, Calling Functions, Passing Arguments, Keyword Arguments, Default Arguments, Variable-length arguments, Anonymous Functions, Fruitful Functions (Function Returning Values), Scope of the Variables in a Function - Global and Local Variables.

Modules: Creating modules, import statement, from. Import statement, name spacing. Python packages: Introduction to PIP, Installing Packages via PIP, Using Python Packages. Text Files: Text Files and Their Format, Writing Text to a File , Writing Numbers to a File , Reading Text from a File , Reading Numbers from a File, Accessing and Manipulating Files and Directories on Disk. Case Study: Gathering Information from a File System

Unit IV

Data Gathering and Cleaning: Cleaning Data, Checking for Missing Values, Handling the Missing Values, Reading and Cleaning CSV Data, Merging and Integrating Data, Reading Data from the JSON Format, Reading Data from the HTML Format, and Reading Data from the XML Format.

Regular expressions: Character matching in regular expressions, Extracting data using regular expressions, Combining searching and extracting and Escape character. Case Study: Detecting the e-mail addresses in a text file.

Unit V

Popular Libraries for Data Visualization in Python: Matplotlib, Seaborn, Plotly, Geoplotlib, and Pandas. Data Visualization: Direct Plotting, Line Plot, Bar Plot, Pie Chart, Box Plot, Histogram Plot, Scatter Plot, Seaborn Plotting System , Strip Plot , Box Plot, Swarm Plot, Joint Plot , Matplotlib Plot , Line Plot Bar Chart ,Histogram Plot ,Scatter Plot , Stack Plot and Pie Chart.

Coding Simple GUI-Based Programs: Windows and Labels, Displaying Images, Command Buttons and Responding to Events, Viewing the Images of Playing Cards, Entry Fields for the Input and Output of Text, and Using Pop-up Dialog Boxes. Case Study: A GUI-Based ATM

Text Books:

1. Chun, W. (2006) Core python programming. Prentice Hall Professional.
2. Embarak, O. (2018). Data Analysis and Visualization Using Python: Analyze Data to Create Visualizations for BI Systems. Apress.
3. Lambert, K. A. (2011). Fundamentals of Python: First Programs. Cengage Learning.
4. Severance, C. (2013). Python for informatics: Exploring information. CreateSpace.

Reference Books

1. <https://www.w3schools.com/python>
2. Learning Python, Mark Lutz, Orielly
3. Python Programming: A Modern Approach, Vamsi Kurama, Pearson

4. VanderPlas, J. (2016). Python data science handbook: Essential tools for working with data. " O'Reilly Media, Inc.".

CO-PO Mapping:

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 5	PO 6	PO 7	PO 8	PO 9	PO1 0	PO1 1	PO1 2
CO 1	2	3	2	2	2	2	2					3	2
CO 2	2	3	2	2	2	2	2					3	2
CO 3	2	3	1	2	2	3	3					3	2
CO 4	2	2	1	2	2	2	3					3	2
CO 5	2	2	1	2	2	2	3					3	2

SEMESTER-II

22MAT527

Big Data Analytics and Hadoop

3 0 2 4

Course Outcomes

CO1: Understanding the concepts of Big Data

CO2: Understanding the aspects of managing, cleaning and sampling of Data

CO3: Understanding Hadoop architecture and implement Map Reduce concept

CO4: To understand the aspects of Data base management and Querying system

CO5: Understanding and executing HDFS using PIG and HIVE

Unit I

Introduction to Big Data: Types of Digital Data-Characteristics of Data – Evolution of Big Data - Definition of Big Data - Challenges with Big Data - 3Vs of Big Data - Non Definitional traits of Big Data - Business Intelligence vs. Big Data - Data warehouse and Hadoop environment.

Unit II

Big Data Analytics: Classification of analytics - Data Science - Terminologies in Big Data Data science process – roles, stages in data science project – working with data from files — exploring data – managing data – cleaning and sampling for modeling and validation. working with relational databases - NoSQL: Types of Databases – Advantages – NewSQL - SQL vs. NOSQL vs NewSQL.

Unit III

Introduction – distributed file system – Hadoop Components – Architecture – HDFS - algorithms using map reduce, Matrix-Vector Multiplication by Map Reduce – Hadoop - Understanding the Map Reduce architecture - Writing Hadoop Map Reduce Programs - Loading data into HDFS - Executing the Map phase - Shuffling and sorting - Reducing phase execution. Hadoop 2 (YARN): Architecture - Interacting with Hadoop Eco systems.

Unit IV

No SQL databases: Mongo DB: Introduction – Features - Data types - Mongo DB Query language - CRUD operations – Arrays - Functions: Count – Sort – Limit – Skip – Aggregate - Map Reduce. Cursors – Indexes - Mongo Import – Mongo Export. Cassandra: Introduction –

Features - Data types – CQLSH - Key spaces - CRUD operations – Collections – Counter – TTL
 - Alter commands - Import and Export - Querying System tables.

Unit V

Hadoop Eco systems: Hive – Architecture - data type - File format – HQL – SerDe - User defined functions - Pig: Features – Anatomy - Pig on Hadoop - Pig Philosophy - Pig Latin overview - Data types - Running pig - Execution modes of Pig - HDFS commands - Relational operators - Eval Functions - Complex data type - Piggy Bank - User defined Functions - Parameter substitution - Diagnostic operator.

Text Books / Reference Books:

1. Seema Acharya, Subhashini Chellappan, “Big Data and Analytics”, Wiley Publication, 2015.
2. Judith Hurwitz, Alan Nugent, Dr. Fern Halper, Marcia Kaufman, “Big Data for Dummies”, John Wiley & Sons, Inc., 2013.
3. Data Science and big data analytics : Discovering, analyzing , visualizing and presentating data ,EMC Education Services,John Wiley 2015
4. Tom White, “Hadoop: The Definitive Guide”, O’Reilly Publications, 2011.
5. Kyle Banker, “Mongo DB in Action”, Manning Publications Company, 2012.
6. Russell Bradberry, Eric Blow, “Practical Cassandra A developers Approach“, Pearson Education, 2014.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	3	2	2	2	2	2					2	2
CO2	2	2	2	2	2	2	2					2	2
CO3	2	2	2	2	2	3	2					2	2
CO4	2	2	1	2	2	2	2					1	2
CO5	1	2	1	1	1	2	2					1	2

22MAT528

Database Management

3 0 2 4

Course outcomes

CO1: Understand the basics database systems and rational models.

CO2: Gain knowledge about various design processes

CO3: To understand the intermediate SQL and advanced SQL.

CO4: To understand the normal forms.

CO5: To understand and apply the algorithms for decomposition using multi-values dependencies – PJNF and DKNF.

Unit I

Introduction to DBMS:Database System Vs File system, Database systems applications, Purpose of database systems - Data models. Relational models: Structure of relational databases – database schema keys – schema diagrams. Relational Query Languages – fundamental relational algebra operations – additional relational algebra operations. Introduction to SQL – Background – SQL data definition –structure of SQL queries – set operations – null values - aggregate functions – modifications to the database.

Unit II

Database design - overview of the design process – the entity-relationship model – constraints – entity-relationship diagrams – reduction to relation schemas - Entity-relationship design issues – weak entity sets – extended E-R features. Intermediate SQL: Nested subqueries - Join expression – Views – Transactions – integrity constraints – authorization. Advanced SQL – Accessing SQL from a program – functions and procedures – triggers.

Unit III

Relational database design – features of good relational designs – atomic domains and normal forms - 1NF, 2NF, 3NF, 4NF and BCNF – decomposition using functional dependencies - functional dependency theory – algorithm for decomposition -decomposition using multi-values dependencies – PJNF and DKNF. Over view of Transaction Management and Concurrency control

Text Book:

- 1) Silberschatz. A., Korth, H. F. and Sudharshan, S. “Database System Concepts”, 6th Edition, TMH, 2010

Reference Books

- 1) Elmasri, R. and Navathe, S. B. “Fundamentals of Database Systems”, 5th Edition, Addison Wesley, 2006
- 2) Date, C. J. “An Introduction to Database Systems”, 8th Edition, Addison Wesley, 2003.
- 3) Ramakrishnan, R. and Gehrke, J. “Database Management Systems”, 3rd Edition, McGrawHill, 2003.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	2	1					2	1
CO2	2	2	2	2	2	2	1					2	1
CO3	2	2	2	2	2	2	1					2	1
CO4	2	2	1	2	2	1	1					1	1
CO5	1	2	1	1	1	1	1					1	1

22MAT529

Data Mining

3 1 0 4

Course Objectives:

CO1: Learn data mining basic concepts and understand association rules mining.

CO2: Capable of grouping data using clustering techniques.

CO3: Able to identify the outliers of the given dataset.

CO4: Capable of minimizing dimensionality of the data with minimum loss of information.

CO5: Able to prioritize the web links and advertisements

Unit - I

Introduction to Data Mining: Introduction, What is Data Mining, Definition, KDD, Challenges, Data Mining Tasks, Data Preprocessing, Data Cleaning, Missing data, Dimensionality Reduction, Feature Subset Selection, Discretization and Binaryzation, Data Transformation; Measures of Similarity and Dissimilarity- Basics.

Unit - II

Association Rules: Problem Definition, Frequent Item Set Generation, The APRIORI Principle, Support and Confidence Measures, Association Rule Generation; APRIORI Algorithm. Bayesian Belief Networks and Additional Topics Regarding Classification.

Unit-III

Clustering: Problem Definition, Clustering Overview, Evaluation of Clustering Algorithms, Partitioning Clustering-K-Means Algorithm, K-Means Additional issues, PAM Algorithm; Hierarchical Clustering-Agglomerative Methods and divisive methods, Key Issues in Hierarchical Clustering, Strengths and Weakness;

Unit-IV

Outlier Detection: Outliers and Outlier Analysis -What Are Outliers?, Types of Outliers ,Challenges of Outlier Detection, Outlier Detection Methods, Statistical Approaches, Parametric Methods, Nonparametric Methods, Proximity-Based Approaches, Clustering-Based Approaches, Classification-Based Approaches, Mining Contextual and Collective Outliers.

Unit-V

Dimensionality Reduction: Principal-Component Analysis, Singular-Value Decomposition, and CUR Decomposition. Link Analysis: Page Rank, Efficient Computation of Page Rank, Topic-Sensitive Page Rank, Link Spam, Hubs and Authorities. Recommendation Systems: A Model for Recommendation Systems, Content-Based Recommendations, and the Netflix Challenge.

Text Books:

1. Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.
2. Rajaraman, A., & Ullman, J. D. (2011). Mining of massive datasets. Cambridge University Press.

Reference Books and websites:

1. <https://nptel.ac.in/courses/106/105/106105174/>
2. https://nptel.ac.in/content/storage2/nptel_data3/html/mhrd/ict/text/110105083/lec52.pdf
3. Ngo, T. (2011). Data mining: practical machine learning tools and technique, by ian h. witten, eibe frank, mark a. hell. ACM SIGSOFT Software Engineering Notes, 36(5), 51-52.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	3	2	2	2	2	2					2	2
CO2	2	2	2	2	2	2	2					2	2
CO3	2	2	2	2	2	3	2					2	2
CO4	2	2	1	2	2	2	2					1	2
CO5	1	2	1	1	1	2	2					1	2

Course Objectives:

CO1: To be able to formulate machine learning problems corresponding to different applications.

CO2: To understand a range of machine learning algorithms along with their strengths and weaknesses.

CO3 : To understand the basic theory underlying machine learning.

CO4: To be able to apply machine learning algorithms to solve problems of moderate complexity.

Unit-I

Introduction: Well-Posed Learning Problems, Designing a Learning System. The Inductive Learning Hypothesis and Concept Learning as Search (definitions). Machine Learning Basics: Learning Algorithms, Capacity, Over fitting and Under fitting , Hyper parameters and Validation Sets , No Free Lunch theorem, Estimators, Bias and Variance , Bayesian statistics, Supervised Learning Algorithms, Unsupervised Learning Algorithms, Stochastic Gradient Descent , Building a Machine Learning Algorithm and Issues in Machine Learning.

Unit-II

Decision Tree learning: Introduction, Decision tree representation, Appropriate problems for decision tree learning, The basic decision tree learning algorithm, Hypothesis space search in decision tree learning, Inductive bias in decision tree learning, Issues in decision tree learning. Implementation aspects of the Decision Tree and Classification Example.

Unit-III

Instance-Based Learning: Introduction, k-Nearest Neighbour Learning, Locally Weighted Regression, Radial Basis Functions, Case-Based Reasoning, Remarks on Lazy and Eager Learning.

Support Vector Machines: Optimal Separation: The Margin and Support Vectors, a Constrained Optimization Problem, Slack Variables for Non-Linearly Separable Problems. KERNELS: Choosing Kernels. The Support Vector Machine Algorithm and Multi-Class Classification. Case Study.

Unit-IV

Genetic Algorithms: Motivation, Genetic Algorithms, Elitism, Tournaments, and Niching, Using Genetic Algorithms and An illustrative Example. Hypothesis Space Search, Genetic Programming, Models of Evolution and Learning, Parallelizing Genetic Algorithms.

Unit-V

Reinforcement Learning: Introduction, the Learning Task, Q Learning, Non-Deterministic, Rewards and Actions, Temporal Difference Learning, Generalizing from Examples, Relationship to Dynamic Programming. Case Study.

Text Books:

1. Machine Learning – Tom M. Mitchell, - MGH
2. Machine Learning: An Algorithmic Perspective, Stephen Marsland, Taylor & Francis (CRC)
3. Haroon, D. (2017). Python Machine Learning Case Studies: Five Case Studies for the Data Scientist Apress.

Reference Books:

1. Harrington, P. (2012). Machine learning in action. Manning Publications Co..
2. Richard o. Duda, Peter E. Hart and David G. Stork, pattern classification, John Wiley & Sons
3. Machine Learning by Peter Flach , Cambridge.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	3	2	2	2	3					2	2
CO2	3	3	3	2	2	2	3					2	2
CO3	2	3	3	2	2	3	3					2	2
CO4	2	2	3	2	2	2	3					2	2

22MAT531

Multivariate Statistics and Regression Analysis

3 0 2 4

Course Outcomes

- CO1: To understand the basics of multivariate random variables and sampling distributions.
- CO2: To apply multivariate techniques for classification of distributions
- CO3: To understand the concept of PCA and its application in clustering analysis
- CO4: To gain knowledge on simple linear regression, estimation and testing of model parameters
- CO5: To gain knowledge on multiple linear and nonlinear regression and estimation of model parameters

Unit-I:

Multivariate Random variables and Distribution functions – Variance - covariance matrix – correlation - Bivariate normal distribution, Multivariate normal density and its properties - Definition of Wishart matrix and its properties, Mahalanobis Distance.

Unit-II:

Classification for two populations, classification with two multivariate normal populations, Fisher’s discriminant functions for discriminating several population.

Unit-III:

Principal components analysis, Dimensionality reduction, Factor Analysis- factor loadings using principal component analysis.

Unit-IV:

Simple Linear Regression- Properties, Least Squares Estimation of parameters, Hypothesis Tests in Simple Linear Regression, Interval estimation in simple linear regression, Coefficient of determination.

Unit-V:

Multiple Linear Regression: Estimation of model parameters. Nonlinear Regression models.

Text books/ Reference books:

1. Anderson, T. W. (1983): An Introduction to Multivariate Statistical Analysis. 3rdEd. Wiley.
2. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers and Keying Ye. Probability and Statistics for Engineers and Scientists, Eighth Edition, Pearson Education Asia, 2007.
3. Douglas C. Montgomery and Elizabeth A. Peck and G. Geoffrey Vining. Introduction to Linear Regression Analysis”, Third Edition, John Wiley & Sons, Inc
4. Casella, G. and Berger, R. Statistical Inference, 1st Edition, Duxbury Press, Pacific Grove, 2002.
5. Draper, N. and Smith, H. Applied Regression Analysis, 3rd Edition, John Wiley and Sons Series in Probability and Statistics, New York, 1998.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	3	2	2	3	1					3	2
CO2	3	3	3	2	2	3	1					3	1
CO3	3	3	3	2	2	3	1					2	1
CO4	3	2	3	2	2	2	1					2	1
CO5	3	2	3	2	2	2	1					2	1

22MAT532

Statistical Inference and Design of Experiments

3 0 2 4

CO1: To understand the concept of testing of hypothesis of various parameters using single sample and apply to engineering, science and business problems.

CO2: To know to apply goodness of fit tests and nonparametric tests

CO3: To understand statistical inference for two samples and apply to engineering, science and business problems.

CO4: To develop experiments and analyse the variance to conclude on the parameters of the population involved

CO5: To construct factorial experiments and to use for various real time problems

Unit I

Tests of Hypotheses for a Single Sample- Tests of Statistical Hypotheses, One-Sided and Two-Sided Hypotheses, *P*-Values in Hypothesis Tests, General Procedure for Hypothesis Tests, Tests on the Mean of a Normal Distribution, Variance Known- Tests on the Mean of a Normal

Distribution, Variance Unknown- Tests on the Variance and Standard Deviation of a Normal Distribution- Tests on a Population Proportion.

Unit II

Statistical Inference for Two Samples -Inference on the Difference in Means of Two Normal Distributions, Variances Known- Inference on the Difference in Means of two Normal Distributions, Variances Unknown- Paired *t*-Test- Inference on the Variances of Two Normal Distributions- Inference on Two Population Proportions.

Unit III

Goodness of Fit Tests –Distribution based Fitting- Categorical Data Analysis - Tests of Independence in Contingency Tables - The Kolmogorov–Smirnov Goodness of Fit Test for Continuous Data. Nonparametric Procedures -The Sign Test, The Wilcoxon Signed-Rank Test- The Runs Test for Randomness.

Unit IV

Design and Analysis of Single-Factor Experiments: The Analysis of Variance-Designing Engineering Experiments, Completely Randomized Single-Factor Experiment, The Random-Effects Model, Randomized Complete Block Design. Latin Square Design.

Unit V

Design of Experiments with Several Factors-Introduction, Factorial Experiments, Two-Factor Factorial Experiments, General Factorial Experiments, 2^k Factorial Designs.

Text Books /Reference Books

1. Douglas C. Montgomery and George C. Runger, Applied Statistics and Probability for Engineers, (2005) John Wiley and Sons Inc.
2. Sheldon. M. Ross : Probability and Statistics for Engineers and Scientists , McGraw-Hill , 2004.
3. Amir D Aczel, Soundarapandian Jayavel, Complete Business statistics, Boston : McGraw-Hill/Irwin, 2009
4. Ravichandran, J. Probability and Statistics for engineers, First Reprint Edition, Wiley India, 2012.
5. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers and Keying Ye, Probability and Statistics for Engineers and Scientists, 8th Edition, Pearson Education Asia, 2007.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	3	2	2	2	1					3	2
CO2	3	3	3	2	2	2	1					3	2
CO3	3	3	3	2	2	3	1					3	2
CO4	3	2	3	2	2	2	1					3	2
CO5	3	2	3	1	1	2	1					3	2

SEMESTER-III

22MAT621

SQC AND RELIABILITY THEORY

3 0 2 4

CO1 To understand the basic concepts of quality control and to construct variable and attribute control charts.

CO2 To understand and construct EWMA and CUSUM charts, analyse the process capability and Six Sigma quality metrics.

CO3 To gain knowledge about acceptance sampling methods and their properties

CO4 To gain knowledge about reliability and properties

CO5 To study reliability distributions and analyse reliability of systems and maintenance

Unit- I

Basic concept of quality control, process control and product control, Statistical process control, theory of control charts, Shewhart control charts for variables- R, s charts, attribute control charts - p, np, c, u charts, modified control charts.

Unit- II

OC and ARL curves of control charts, moving average control charts, EWMA charts, CUSUM charts, – two sided and one sided procedures – V – mask technique, process capability analysis, process capability indices – C_p and C_{pk} , Six Sigma quality metrics

Unit- III

Acceptance sampling for attributes, single sampling, double sampling, measuring performance of the sampling plans- OC, AQL, LTPD, AOQ, ATI curves.

Unit -IV

Introduction to Reliability and its needs; Different Approaches to Reliability Analysis, Application Areas, State Variable, Time to Failure, Failure Rate Function, Mean Time to Failure, Relationship between the Functions $F(t)$, $f(t)$, $R(t)$, and $z(t)$, Bath tub curve, Mean time to failure, Residual time

Unit - V

Parametric families of some common life distributions –Exponential, Weibull and Gamma and its characterization-Reliability estimation of parameters in these models. Fault Tree Analysis, Reliability Block Diagrams, Systems of Independent Components -System Reliability, Nonrepairable Systems, Quantitative Fault Tree Analysis, Reliability of Maintained Systems - Types of Maintenance, Downtime and Downtime Distributions, System Availability Assessment

Text Books/References:

1. Montgomery D. C. (2005) Introduction to Statistical Quality control, 5th edition, Wiley.
2. Ravichandran, J. Probability and Statistics for engineers, First Reprint Edition, Wiley India, 2012.
3. Schilling E. G. (1982) Acceptance Sampling in Quality Control, Marcel Decker.
4. Marvin Rausand and Arnljot Hoyland ,(2003): System Reliability Theory : Models, Statistical methods and applications, 2nd edition, John Wiley and Sons Inc., publications.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	2	2	2	2						3	1
CO2	3	2	2	2	2	2						3	1
CO3	3	2	2	2	2	3						3	1
CO4	3	2	1	2	2	2						3	1

CO5	2	2	2	2	2	2					2	1
-----	---	---	---	---	---	---	--	--	--	--	---	---

22MAT620

Deep Learning

3 0 2 4

Course outcomes

CO1: Understand the basics concepts of artificial neural networks.

CO2: Gain knowledge about activation functions and understand the multi layer neural network

CO3: Know the importance of regularization, bagging and ensemble methods

CO4: To gain knowledge convolution neural network and case studies

CO5: To gain knowledge about recurrent neural networks, adversarial neural networks, Spectral CNN and deep reinforcement learning

Unit-I

Biological neuron, idea of computational units, McCulloch – pitts unit and thresholding logic, linear perceptron, perceptron learning algorithm, convergence theorem for Perceptron learning algorithm, logistic regression, gradient descent.

Unit-II

Feed forward neural network, activation functions, non-linear activation functions. multi-layer neural network.

Unit-III

Practical aspects of deep Learning: training, testing, regularization –dataset augmentation, Noise robustness, multitask learning, bagging and other ensemble methods, dropout- generalization.

Unit-IV

Convolution neural networks, backpropagation convolutions and pooling – optimization algorithms: mini-batch gradient descent, - convolutional nets case studies using Keras/Tensorflow.

Unit-V

Neural network architectures – recurrent neural networks, adversarial neural networks Spectral CNN, self-organizing maps, restricted boltzmann machines, long short-term memory networks, deep meta learning - deep reinforcement learning.

Text Books / Reference Books

1. Ian Goodfellow, YoshuaBengio and Aeron Courville, Deep Learning, MIT Press,First Edition, 2016.
2. Gibson and Josh Patterson, Deep Learning A practitioner’s approach, Adam O’Reilly, First Edition, 2017.
3. Francois Chollet, Deep Learning with Python, Manning Publications Co, First Edition, 2018.
4. Bishop C.M.Neural Networks for Pattern Recognition, Oxford University Press,1995.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	3	2	2	2	2	2					2	2
CO2	2	2	2	2	2	2	2					2	2
CO3	2	2	2	2	2	3	2					2	2
CO4	2	2	1	2	2	2	2					1	2
CO5	1	2	1	1	1	2	2					1	2

21AVP501

Amrita Values Program

1001

Amrita University's Amrita Values Program (AVP) is a new initiative to give exposure to students to the richness and beauty of the Indian way of life. India is a country where history, culture, art, aesthetics, cuisine, and nature exhibit more diversity than anywhere else in the world. Amrita Values Programs emphasize making students familiar with the rich tapestry of Indian life, culture, arts, science, and heritage which has historically drawn people from all over the world. Post-graduate students shall have to register for any one of the following courses, in the second semester, which may be offered by the respective school.

Courses offered under the framework of the Amrita Values Program:

Art of Living through Amma

Amma's messages can be put into action in our life through pragmatism and attuning of our thought process in a positive and creative manner. Every single word Amma speaks, and the guidance received in matters which we consider trivial are rich in content and touches the very inner being of our personality. Life gets enriched by Amma's guidance, and She teaches us the art of exemplary life skills where we become witness to all the happenings around us keeping the balance of the mind.

Insights from the Ramayana

The historical significance of Ramayana, the first Epic in the world, influence of Ramayana on Indian values and culture, storyline of Ramayana, study of leading characters in Ramayana, influence of Ramayana outside India, misinterpretation of Ramayana by colonial powers and its impact on Indian life, relevance of Ramayana for modern times.

Insights from the Mahabharata

The historical significance of Mahabharata, the largest Epic in the world, influence of Mahabharata on Indian values and culture, storyline of Mahabharata, study of leading characters in Mahabharata, Kurukshetra War and its significance, importance of Dharma in society, message of the Bhagavad Gita, relevance of Mahabharata for modern times.

Insights from the Upanishads

Introduction: Sruti versus Smrti, overview of the four Vedas and the ten Principal Upanishads, the central problems of the Upanishads, ultimate reality, the nature of Atman, the different modes of consciousness, Sanatana Dharma and its uniqueness, The Upanishads and Indian Culture,

relevance of Upanishads for modern times, a few Upanishad Personalities: Nachiketas, Satyakama Jabala, Aruni, Shvetaketu.

Insights from Bhagavad Gita

Introduction to Bhagavad Gita, brief storyline of Mahabharata, context of Kurukshetra War, the anguish of Arjuna, counsel by Sri. Krishna, key teachings of the Bhagavad Gita, Karma Yoga, Jnana Yoga, and Bhakti Yoga, theory of Karma and Reincarnation, concept of Dharma, idea of the self and realization of the self, qualities of a realized person, concept of Avatar, relevance of Mahabharata for modern times.

Swami Vivekananda and his Message

Brief sketch of Swami Vivekananda's life, meeting with Guru, disciplining of Narendra, travel across India, inspiring life incidents, address at the parliament of religions, travel in the United States and Europe, return and reception in India, message to Indians about our duties to the nation.

Great Spiritual Teachers of India

Sri Rama, Sri Krishna, Sri Buddha, Adi Shankaracharya, Sri Ramanujacharya, Sri Madhvacharya, Sri Ramakrishna Paramahansa, Swami Vivekananda, Sri Ramana Maharshi, Mata Amritanandamayi Devi

Indian Arts and Literature:

The aim of this course is to present the rich literature, culture of ancient India, and help students appreciate their deep influence on Indian life, Vedic culture, the primary source of Indian culture, brief introduction, and appreciation of a few of the art forms of India, arts, music, dance, theatre, paintings, sculpture and architecture, the wonder language, Sanskrit, and ancient Indian Literature.

Importance of Yoga and Meditation in Life:

The objective of the course is to provide practical training in YOGA ASANAS with a sound theoretical base and theory classes on selected verses of Patanjali's Yoga Sutra and Ashtanga Yoga. The coverage also includes the effect of yoga on integrated personality development.

Appreciation of Kerala's Mural Art Forms:

A mural is any piece of artwork painted or applied directly on a wall, ceiling, or another large permanent surface. In the contemporary scenario, Mural painting is not restricted to permanent structures and is being done even on canvas. A distinguishing characteristic of mural painting is that the architectural elements of the given space are harmoniously incorporated into the picture. Kerala mural paintings are frescos depicting mythology and legends, which are drawn on the walls of temples and churches in South India, principally in Kerala. Ancient temples, churches, and places in Kerala, South India, display an abounding tradition of mural paintings mostly dating back to the 9th to 12th centuries CE when this form of art enjoyed Royal patronage. Learning Mural painting through the theory and practice workshop is the objective of this course.

Practicing Organic Farming

Life and nature are linked through the healthy practices of society for maintaining sustainability. When modern technological knowledge microorganisms are applied in farming using the traditional practice, we can avoid damage to the environment. The course will train the youth on modern practices of organic farming. Amma says "we have to return this land to the coming

generations without allowing even the slightest damage to happen to it.” Putting this philosophy to practice will bring about an awakening and enthusiasm in all to strive for good health and to restore the harmony in nature”

Ancient Indian Science and Technology

Science and technology in ancient and medieval India covered all the major branches of human knowledge and activities, including mathematics, astronomy, physics, chemistry, medical science and surgery, fine arts, mechanical, civil engineering, architecture, shipbuilding, and navigation. Ancient India was a land of sages, saints, and seers as well as a land of scholars and scientists. The course gives awareness of India's contribution to science and technology.

21CUL501

Cultural Education

2 0 0 P/F

Objective: Love is the substratum of life and spirituality. If love is absent life becomes meaningless. In the present world, if love is used as the string to connect the beads of values, life becomes precious, rare, and beautiful like a fragrant blossom. Values are not to be learned alone. They must be imbibed into the inner spirit and put into practice. This should happen at the right time when you have vitality and strength when your hearts are open.

The present course in value education is a humble experience-based effort to lead and metamorphosis the students through the process of transformation of their inner self towards achieving the best. Amma's nectarous words of wisdom and acts of love are our guiding principles. Amma's philosophy provides an insight into the vision of our optimistic future.

1. Invocation, Satsang, and Question-Answers
2. Values - What are they? Definition, Guiding Principles with examples Sharing own experiences
3. Values - Key to a meaningful life. Values in different contexts
4. Personality - Mind, Soul, and Consciousness - Q and A. Body-Mind-Intellect and the Innerpsyche Experience sharing
5. Psychological Significance of samskara (with e.g. From Epics)
6. Indian Heritage and Contribution and Q and A; Indian Ethos and Culture
7. Self-Discipline (Evolution and Practice) – Q and A
8. Human Development and Spiritual Growth - Q and A
9. Purpose of Life plus Q and A
10. Cultivating self-Development, Self-effort, and Divine Grace - their roles – Q and A.
11. Vedanta and Creation - Understanding a spiritual Master
12. Dimensions of Spiritual Education; Need for change Lecture – 1; Need for Perfection Lecture – 2
13. How to help others who have achieved less - Man and Nature Q and A, Sharing of experiences

REFERENCES

1. *Swami Amritaswaroopananda Puri - Awaken Children (Volume VII and VIII)*
2. *Swami Amritaswaroopananda Puri - Amma's Heart*
3. *Swami Ramakrishnanda Puri - Rising Along the Razor's Edge*

4. *Deepak Chopra - Book 1: Quantum Healing; Book 2: Alpha and Omega of God; Book 3: Seven Spiritual Rules for Success*
5. *Dr. A. P. J. Abdul Kalam - 1. Ignited Minds 2. Talks (CD)*
6. *Swami Ramakrishnananda Puri - Ultimate Success*
7. *Swami Jnanamritananda Puri - Upadesamritham (Trans: Malayalam)*
8. *Vedanta Kesari Publication - Values - Key to a meaningful life*
9. *Swami Ranganathananda - Eternal values for a changing society*
10. *David Megginson and Vivien Whitaker - Cultivating Self Development*
11. *Elizabeth B. Hurlock - Personality Development, Tata McGraw Hill*
12. *Swami Jagatmananda - Learn to Live (Vol.1 and 2), RK Ashram, Myslapore*

22AVP103

Mastery Over Mind (MAOM)

1-0-2 2

1. Course Overview

Master Over the Mind (MAOM) is an Amrita initiative to implement schemes and organise university-wide programs to enhance health and wellbeing of all faculty, staff, and students (UN SDG -3). This program as part of our efforts for sustainable stress reduction gives an introduction to immediate and long-term benefits and equips every attendee to manage stressful emotions and anxiety facilitating inner peace and harmony.

With a meditation technique offered by Amrita Chancellor and world-renowned humanitarian and spiritual leader, Sri Mata Amritanandamayi Devi (Amma), this course has been planned to be offered to all students of all campuses of AMRITA, starting off with all first years, wherein one hour per week is completely dedicated for guided practical meditation session and one hour on the theory aspects of MAOM. The theory section comprises lecture hours within a structured syllabus and will include invited guest lecture series from eminent personalities from diverse fields of excellence. This course will enhance the understanding of experiential learning based on university's mission: "Education for Life along with Education for Living", and is aimed to allow learners to realize and rediscover the infinite potential of one's true Being and the fulfilment of life's goals.

2. Course Syllabus

Unit 1 (4 hours)

Causes of Stress: The problem of not being relaxed. Need for meditation -basics of stress management at home and workplace. Traditions and Culture. Principles of meditation– promote a sense of control and autonomy in the Universal Human Value System. Different stages of Meditation. Various Meditation Models. Various practices of Meditation techniques in different schools of philosophy and Indian Knowledge System.

Unit 2 (4 hours)

Improving work and study performance. Meditation in daily life. Cultivating compassion and good mental health with an attitude of openness and acceptance. Research and Science of Meditation: Significance of practising meditation and perspectives from diverse fields like science, medicine, technology. philosophy, culture, arts, management, sports, economics,

healthcare, environment etc. The role of meditation for stress and anxiety reduction in one's life with insights based on recent cutting-edge technology. The effect of practicing meditation for the wholesome wellbeing of an individual.

Unit 3 (4 hours)

Communications: principles of conscious communication. Relationships and empathy: meditative approach in managing and maintaining better relationships in life during the interactions in the world, role of MAOM in developing compassion, empathy and responsibility, instilling interest, and orientation to humanitarian projects as a key to harness intelligence and compassion in youth. Methodologies to evaluate effective awareness and relaxation gained from meditation. Evaluating the global transformation through meditation by instilling human values which leads to service learning and compassion driven research.

TEXT BOOKS:

1. Mata Amritanandamayi Devi, "Cultivating Strength and vitality," published by Mata Amritanandamayi Math, Dec 2019
2. Swami Amritaswarupananda Puri, "The Color of Rainbow " published by MAM, Amritapuri.

REFERENCES:

1. Craig Groeschel, "Winning the War in Your Mind: Change Your Thinking, Change Your Life" Zondervan Publishers, February 2019
2. R Nagarathna et al, "New Perspectives in Stress Management "Swami Vivekananda Yoga Prakashana publications, Jan 1986
3. Swami Amritaswarupananda Puri "Awaken Children Vol 1, 5 and 7 - Dialogues with Amma on Meditation", August 2019
4. Swami Amritaswarupananda Puri "From Amma's Heart - Amma's answer to questions raised during world tours" March 2018
5. Secret of Inner Peace- Swami Ramakrishnananda Puri, Amrita Books, Jan 2018.
6. Mata Amritanandamayi Devi "Compassion :The only way to Peace:Paris Speech", MA Center, April 2016.
7. Mata Amritanandamayi Devi "Understanding and collaboration between Religions", MA Center, April 2016.
8. Mata Amritanandamayi Devi "Awakening of Universal Motherhood: Geneva Speech" M A center, April 2016.

3. Evaluation and Grading

Internal		External		Total
<i>Components</i>	<i>Weightage</i>		Practical (attendance and class participation) 60%	100%
Quizzes(based on the reading material)	20%	40%		
Assignments (Based on webinars and lecture series)	20%			

4. Course Outcomes (CO)

CO1: Relate to the causes of stress in one's life.

CO2: Experiment with a range of relaxation techniques

CO3: Model a meditative approach to work, study, and life.

CO4: Develop appropriate practice of MA-OM technique that is effective in one's life

CO5: Inculcate a higher level of awareness and focus.

CO6: Evaluate the impact of a meditation technique

***Programme Outcomes(PO)**(As given by NBA and ABET)

PO1: Engineering Knowledge

PO2: Problem Analysis

PO3: Design/Development of Solutions

PO4: Conduct Investigations of complex problems

PO5: Modern tools usage

PO6: Engineer and Society

PO7: Environment and Sustainability

PO8: Ethics

PO9: Individual & Team work

PO10: Communication

PO11: Project management & Finance

PO12: Lifelong learning

CO – PO Affinity Map

PO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PO 13	PO 14	PO 15
CO															
CO 1	3	3	3	2		-	2	3	-	3	-	3	-	-	-
CO 2	3	3	3	2	2	-	2	3	3	3	-	3	-	-	-

CO 3	3	3	2	2	2	2	2	3	3	3	-	3	-	-	-
CO 4	3	3	3	2	-	2	3	3	3	3	-	3	-	-	-
CO 5	3	2	2	2	-	2	-	3	2	2	-	2	-	-	-
CO 6	3	2	2	2	3	2	-	3	2	2	-	2	-	-	-

ELECTIVES

22MAT731

Business Analytics

3 0 0 3

Course Objectives:

1. Understanding the Role of Business Analyst and Data Science in business.
2. Understanding the basic concept of data management and data mining techniques
3. To understand the basic concept of machine learning
4. To understand the application of business analysis.

Introduction: What is business analytics? Historical Overview of data analysis, Data Scientist vs. Data Engineer vs. Business Analyst, Career in Business Analytics, What is data science, Why Data Science, Applications for data science, Data Scientists Roles and Responsibility

Data: Data Collection, Data Management, Big Data Management, Organization/sources of data, Importance of data quality, Dealing with missing or incomplete data, Data Visualization, Data Classification Data Science Project Life Cycle: Business Requirement, Data Acquisition, Data Preparation.

Introduction to Data Mining, The origins of Data Mining, Data Mining Tasks, OLAP and Multidimensional data analysis, Basic concept of Association Analysis and Cluster Analysis.

Application of Business Analysis: Retail Analytics, Marketing Analytics, Financial Analytics, Healthcare Analytics, Supply Chain Analytics.

Text Books:

1. Essentials of Business Analytics: An Introduction to the methodology and its application, Bhimasankaram Pochiraju, SridharSeshadri, Springer
2. Introduction to Machine Learning with Python: A Guide for Data Scientists 1st Edition, by Andreas C. Müller, Sarah Guido, O'Reilly
3. Introduction to Data Science, Laura Igual Santi Seguí, Springer

Reference Book:

1. Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson Education India
2. An Introduction to Business Analytics, Ger Koole, Lulu.com, 2019.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	2	-	-	-	-	1	1	
CO2	3	3	2	2	2	2	-	-	-	-	1	1	
CO3	2	2	3	2	2	2	-	-	-	-	1	1	

Course outcomes

CO1: Understand the Categorical Data-nominal and ordinal random variables.

CO2: Understand the two-way contingency tables.

CO3: Understand the three-way contingency tables.

CO4: Understand the basic of Generalized Linear Models

Categorical Data-nominal and ordinal random variables.

Two-way contingency tables: Table structure for two dimensions. Ways of comparing proportions. Measures of associations-odds ratio. Sampling distributions. Goodness-of-fit tests, testing of independence. Exact and large sample inference.

Three-way contingency tables, Partial associations, Cochran-Mantel-Haenszel methods. Conditional association and related inference.

Generalized Linear Models (GLMs): components of a GLM.

Logistic regression models for binary data, inference for logistic regression models, multiple logistic regression with qualitative predictors, exact inference for logistic regression, sample size and power of test.

Loglinear models for two-way and three-way contingency tables, inference for loglinear models, the connection between loglinear-logit regression models.

Multicategory logit models for nominal responses, cumulative logit models for ordinal responses.

Texts / References

1. Agresti, A., Categorical Data Analysis, 3rd Edition, Wiley, New York, 2013.
 2. Agresti, A., An Introduction to Categorical Data Analysis, 3rd Edition, Wiley, New York, 2019.
 3. Andersen, E.B., The Statistical Analysis of Categorical Data, Springer-Verlag, Berlin, 1994.
- Santner, T.J. and Duffy, D., The Statistical Analysis of Discrete Data, Springer-Verlag, New York, 1989.

CO-PO Mapping:

Course Outcomes:

CO1: To understand the basics of molecular modelling.

CO2: To understand the quantitative structure and activity relationship.

CO3: Understand and apply PCA in molecular design.

CO4: To understand important drug databases, designing Lipinski's rule of five.

Introduction to Molecular Modeling: Molecular Modeling and Pharmacoinformatics in Drug Design, Phases of Drug Discovery, Target identification and validation

Protein Structure Prediction and Analysis: Protein Structure prediction methods: Secondary Structure Prediction, Tools for Structure prediction; Protein structural visualization; Structure validation tools; Ramachandran Plot.

QSAR : Quantitative Structure and Activity Relationship - Historical Development of QSAR, Tools and Techniques of QSAR, Molecular Structure Descriptors.

Multivariate Statistical methods in QSAR -Principal Component Analysis (PCA) and Hierarchical Cluster Analysis(HCR). Regression analysis tools - Pincipal Component Regression (PCR), Partial Least Squares (PLS) - Case studies.

High Throughput / Virtual screening- Introduction, Basic Steps, Important Drug Databases, Designing Lipinski's Rule of Five, ADMET screening

Docking Studies- Target Selection, Active site analysis, Ligand preparation and conformational analysis, Rigid and flexible docking .

Molecular visualization tools: RasMol and Swiss-Pdb Viewer

Molecular docking tools: AutoDock and ArgusLab.

References/ Textbooks

1. Leach Andrew R., Valerie J. Gillet, An introduction to Chemoinformatics. Publisher: Kluwer academic , 2003. ISBN: 1402013477.
2. Gasteiger Johann, Handbook of Chemoinformatics: From Data to Knowledge (4 Volumes), 2003. Publisher: Wiley-VCH. ISBN:3527306803.
3. Opera Tudor I,Ed. , Chemoinformatics in drug discovery, Wiley-VCH Verlag,2005.
4. Bunin Barry A. Siesel Brian,Morales Guillermo,Bajorath Jürgen. Chemoinformatics: Theory, Practice, & Products Publisher:New York, Springer. 2006. ISBN: 1402050003.
5. Gasteiger Johann, Engel Thomas. Chemoinformatics: A Textbook. Publisher: WileyVCH; 1st edition. 2003. ISBN: 3527306811.

Kenneth M Merz, Jr, Dagmar Ringe, Charles H. Reynolds , Drug design: Structure and ligand based approaches (2010) publisher : Cmabridge University press

- CO1: Understand the basics of healthcare data analytics.
- CO2: Gain knowledge about phenotyping algorithms.
- CO3: Know the importance of clinical trials and prediction models.
- CO4: To gain knowledge pervasive health analysis.

Introduction to Healthcare Data Analytics- Electronic Health Records–Components of EHR- Coding Systems- Benefits of EHR- Barrier to Adopting HER Challenges- Phenotyping Algorithms. Challenges in Healthcare Data Analysis, Acquisition Challenges, Pre-processing, Transformation , Social Media Analytics for Healthcare.

Advanced Data Analytics for Healthcare : Review of clinical trials , Prediction Models. Statistical Prediction Models, Alternative Clinical Prediction Models, Survival Models, Predictive Models for Integrating Clinical and Genomic Data, Data Analytics for Pervasive Health, Fraud Detection in Healthcare, Pharmaceutical Discoveries and Clinical Decision Support Systems.

Text / References books :

1. Chandan K. Reddy and Charu C Aggarwal, “*Healthcare data analytics*”, Taylor & Francis, 2015
2. Hui Yang and Eva K. Lee, “*Healthcare Analytics: From Data to Knowledge to Healthcare Improvement*”, Wiley, 2016.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	1						2	1
CO2	2	2	2	2	2	1						2	1
CO3	2	2	2	2	2	1						2	1
CO4	2	2	1	2	2	1						2	1

- CO1: Understand the basics of business analytics.
- CO2: Gain knowledge about auto-correlations and time series analysis.
- CO3: To understand the linear time series models.
- CO4: To gain knowledge about sYule Walker estimation for AR processes.

Business Analytics Basics: Definition of analytics, Evolution of analytics, Need of Analytics, Business analytics vs business analysis, Business intelligence vs Data Science, Data Analyst Vs

Business Analyst, Business Analytics at the Strategic Level, Functional Level, Analytical Level, Data Warehouse Level. Market Segmentation Variables, Market Segmentation Types, Marketing Data Landscape, Analyzing the trend of data in Marketing– case studies.

Time series as a discrete parameter stochastic process, Auto - covariance, Auto-correlation functions and their properties. Exploratory time series analysis, Test for trend and seasonality, Exponential and moving average smoothing, forecasting based on smoothing.

Linear time series models: Autoregressive, Moving Average, autoregressive Moving Average models, Autoregressive Integrated Moving Average models. Estimation of ARMA models: Yule-Walker estimation for AR Processes, Maximum likelihood and least squares estimation for ARMA Processes.

Text / References Books:

1. Grigsby Gert H.N Laursen and Jesper Thorlund : *Business analytics for managers taking business intelligence beyond reporting*, second edition 2016.
2. Wayne L. Winston: *Marketing Analytics: Data-Driven Techniques with Microsoft Excel*, Wiley, 2014.
3. Mike Grigsby : *Marketing Analytics: A Practical Guide to Improving Consumer Insights Using Data Techniques*, Kogan Page; 2 edition ,2018
4. Mike Anderson, T.W : *The Statistical Analysis of Time Series*, John Wiley, New York, 1971.
5. Kendall, Sir Maurice and Ord, J.K. : *Time Series*, Edward Arnold, London, 1990.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	2	2	2	1						2	1
CO2	3	2	2	2	2	1						2	1
CO3	3	2	2	2	2	1						2	1
CO4	3	2	1	2	2	1						2	1

22MAT738 Mining of Massive Datasets

3 0 0 3

Course outcomes

CO1: Understand the basics of data mining and its limitations.

CO2: Gain knowledge about data mining streams.

CO3: Understand the clustering techniques for data mining.

CO4: Apply the dimensionality reduction algorithm for social network analysis.

Basics of Data Mining - computational approaches - statistical limits on data mining - MapReduce - Distributed File Systems . MapReduce . Algorithms using MapReduce . Extensions to MapReduce. Mining Data Streams: The Stream Data Model - Sampling Data in a Stream - Filtering Streams. Link analysis, Frequent itemsets, Clustering, Advertising on web, Recommendation system, Mining Social-Network Graphs, Dimensionality Reduction, Large-Scale Machine Learning.

Text / References Book

1. Jure Leskovec , Anand Rajaraman, Jeffrey David Ullman, Mining of Massive Datasets, Cambridge University Press, 2014.
2. Tom White, Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale O'Reilly Media; 4 edition , 2015.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	2	-	-	-	-	1	1	
CO2	3	3	2	2	2	2	-	-	-	-	1	1	
CO3	2	2	3	2	2	2	-	-	-	-	1	1	
CO4	3	3	3	2	2	2	-	-	-	-	1	1	

22MAT739

Official Statistics

3 0 0 3

Course outcomes

CO1: Understand the Indian statistical system.

CO2: Understand the CSO, NSSO and RGO.

CO3: Understand evaluation of performance of family welfare programmes and Statistics related to Industries.

CO4: To gain knowledge economic development.

Introduction to Indian Statistical systems- Role, function and activities of Central Statistical organization and State Agencies. Role of National Sample Survey Organization. General and special data dissemination systems. Scope and Contents of population census of India. statistics, their reliability and limitations. Role of Ministry of Statistics & Program Implementation (MoSPI), Central Statistical Office (CSO), National Sample Survey Office (NSSO), Registered General Office and National Statistical Commission.

Population growth in developed and developing countries, Evaluation of performance of family welfare programmes. Statistics related to Industries, foreign trade, balance of payment, cost of living, inflation, educational and other social statistics.

Economic development: Growth in per capita income and distributive justice indices of development, human development index. National income estimation- Product approach, income approach and expenditure approach. Measuring inequality in income: Gini Coefficient, Theil's measure; Poverty measurements: Different issues, measures of incidence and intensity; Combined Measures: Indices due to Kakwani, Sen etc.

Text Books:

1. Guide to Official Statistics (CSO) 1999.
2. Principles and Accommodation of National Population Census, UNEDCO
3. CSO (1989)a: National Accounts Statistics- Sources and Methods.
4. Guide to current Indian Official Statistics, Central Statistical Office, GOI, and New Delhi.<http://mospi.nic.in/>

CO 1 To understand Bayesian decision theory and its use
CO 2 To gain knowledge about Bayesian estimation methods
CO 3 To apply nonparametric techniques and linear discriminant functions
CO 4 To gain knowledge about nonmetric methods and algorithm independent machine learning
CO 5 To apply unsupervised learning and clustering

Course outcomes

Pattern recognition systems – the design cycle – learning and adaptation – Bayesian decision theory – continuous features – Minimum error rate classification – discriminant functions and decision surfaces – the normal density based discriminant functions. Bayesian parameter estimation – Gaussian case and general theory – problems of dimensionality – components analysis and discriminants- Nonparametric techniques – density estimation – Parzen windows – nearest neighborhood estimation – rules and metrics - decision trees – CART methods – algorithm-independent machine learning – bias and variance for regression and classification – resampling or estimating statistics- Unsupervised learning and clustering – mixture densities and identifiability – maximum likelihood estimates – application to normal mixtures – unsupervised Bayesian learning – data description and clustering – criterion functions for clustering – hierarchical clustering – k-means clustering.

Text Reference Book:

1. Richard O. Duda, Peter E. Hart and David G. Stork, “*Pattern Classification*”, Second Edition, 2003, John wily & sons.
2. Earl Gose, Richard Johnsonbaugh and Steve Jost, “*Pattern Recognition and Image Analysis*”, 2002, Prentice Hall of India.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	2	2	2	1						2	1
CO2	3	2	2	2	2	1						2	1
CO3	3	2	2	2	2	1						2	1
CO4	3	2	1	2	2	1						2	1
CO5	2	2	2	2	2	1						1	1

Course outcomes

CO1. Understand basic characteristics of Queueing models.
CO2. Learn basic concepts of Poisson models with single server.
CO3. Understand and apply the queueing modals for various problems..

Queuing Models: Basic characteristics of a Queueing Model – Role of Poisson and Exponential distributions, Stochastic Processes, Markov chains, Poisson Processes, Poisson Queuing Models with single server: Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, simple numerical problems on (M/M/1): (/FIFO) and (M/M/1): (N/FIFO) Models.

Poisson Queuing Models with multiple server: Descriptions of the model, Assumptions, Probability distributions for number of Units (steady state), waiting time distribution, simple numerical problems on (M/M/C): (/FIFO), (M/M/C): (N/FIFO) and (M/M/C): (C/FIFO) Models, M/M/G Models.

Text Books

1. Donald Gross&Carl M Harris(1998):Fundamentals of Queuing theory,John Wiley & Sons, Inc
2. Hamdy A.Taha(2006): Operations Research – An Introduction, 8/e , Prentice Hall of India Private Ltd., New Delhi

Reference Books

1. S.D.Sharma (2003)Operations Research , Kedar Nath Ram Nath & Co, Meerut, India
2. Kanthi Swarup, P.K.Gupta and Man Mohan (2004), Operations Research, Sultan Chand & Sons, New Delhi

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	1							1	
CO2	2	2	2	2	1							1	
CO3	1	2	1	2	1							1	

22MAT743 Reinforcement Learning

3 0 0 3

Course outcomes

CO1: Understand the basics of reinforcement learning. Its elements and limitations.

CO2: Understand the finite Markov decision process.

CO3: Understand the temporal difference learning and its advantages.

CO4: Understand the Sarsa maximization bias and double learning.

Introduction: Reinforcement Learning, Elements of Reinforcement Learning, Limitations and Scope, An Extended Example- Tic-Tac-Toe. Multi-armed Bandits: A k-armed Bandit Problem, Action-value Methods, The 10-armed Testbed, Incremental Implementation, Tracking a Nonstationary Problem, Optimistic Initial Values, Upper-Confidence-Bound Action Selection, Gradient Bandit Algorithms.

Finite Markov Decision Processes: The Agent–Environment Interface, Goals and Rewards, Returns and Episodes , Unified Notation for Episodic and Continuing Tasks, Policies and Value Functions, Optimal Policies and Optimal Value Functions, Optimality and Approximation.

Review of Markov process and Dynamic Programming.

Temporal-Difference Learning: TD Prediction, Advantages of TD Prediction Methods, Optimality of TD, Sarsa: On-policy TD Control, Q-learning: Policy TD Control. Expected Sarsa. Maximization Bias and Double Learning.

Text/ References Book:

1. Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 2018.
2. Sudharsan Ravichandiran, *Hand-on Reinforcement Learning with Python*, Packt Publications, 2018.
3. Sayon Dutta, *Reinforcement Learning with Tensor Flow: A beginner's guide*, Packt Publications, 2018.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	2	-	-	-	-	1	1	
CO2	3	3	2	2	2	2	-	-	-	-	1	1	
CO3	2	2	3	2	2	2	-	-	-	-	1	1	
CO4	3	3	3	2	2	2	-	-	-	-	1	1	

22MAT744**Sampling Techniques****3 0 0 3****Course outcomes**

CO1: Understand the preliminary concepts like schedules and pilot survey.

CO2: Gain knowledge about Simple random sampling with and without replacements, random number generation.

CO3: To understand Probability proportional to size sampling, estimates of these standard errors.

CO4: To gain knowledge on cluster sampling.

Preliminary concepts – schedules and questionnaires, pilot survey, non-sampling errors, use of random numbers. Simple random sampling with and without replacements, random number generation– estimates of population mean and population proportion and their standard errors, Probability proportional to size sampling, estimates of these standard errors. Stratified random sampling – estimates of sample statistic and estimates of their standard errors. Allocation of sample size in stratified random sampling. Linear and circular systematic sampling. Cluster sampling : Two stage sampling (equal first stage units). Ideas of ratio and regression estimators – only estimates of sample mean..

References

1. Cochran, W.G. : *Sampling Techniques*, 3rd Ed., Wiley Eastern. 1984
2. Murthy, M.N. : *Sampling Theory & Statistical Methods*, Statistical Pub. Society, Calcutta, 1977
3. Des Raj and Chandhok P. : *Sample Survey Theory*, Narosa Publishing House, 1988.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	1						2	1

CO2	2	2	2	2	2	1						2	1
CO3	2	2	2	2	2	1						2	1
CO4	2	2	1	2	2	1						2	1

22MAT745

Social Network Analytics

3 0 0 3

Course outcomes

Course Outcomes:

CO1: To understand the basics of social networks and its modelling.

CO2: To understand the fundamental of social data analytics.

CO3: Understand and apply the data mining concepts in social networks.

CO4: Carry out some case studies in social network analysis.

Unit 1 : Online Social Networks (OSNs)

Introduction - Types of social networks (e.g., Twitter, Facebook), Measurement and Collection of Social Network Data. Techniques to study different aspects of OSNs -- Follower-follower dynamics, link farming, spam detection, hashtag popularity and prediction, linguistic styles of tweets. Case Study: An Analysis of Demographic and Behaviour Trends using Social Media: Facebook, Twitter and Instagram

Unit 2: Fundamentals of Social Data Analytics

Introduction - Working with Social Media Data, Topic Models, Modelling social interactions on the Web – Agent Based Simulations, Random Walks and variants, Case Study: Social Network Influence on Mode Choice and Carpooling during Special Events: The Case of Purdue Game Day

Unit 3 : Applied Social Data Analytics

Application of Topic models, Information Diffusion, Opinions and Sentiments - Mining, Analysis and Summarization, Case Study: Sentiment Analysis on a set of Movie Reviews using Deep Learning techniques, Recommendation Systems, Language dynamics and influence in online communities, Community identification, link prediction and topical search in social networks, Case Study: The Interplay of Identity and Social Network: A Methodological and Empirical Study

Text and Reference Literature

1. Cioffi-Revilla, Claudio. *Introduction to Computational Social Science*, Springer, 2014.
2. Matthew A. Russell. *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, Github, and More*, 2nd Edition, O'Reilly Media, 2013.

3. Robert Hanneman and Mark Riddle. *Introduction to social network methods*. Online Text Book, 2005.
4. Jennifer Golbeck, *Analyzing the social web*, Morgan Kaufmann, 2013.
5. Claudio Castellano, Santo Fortunato, and Vittorio Loreto, *Statistical physics of social dynamics*, Rev. Mod. Phys. 81, 591, 11 May 2009.
6. S. Fortunato and C. Castellano, *Word of mouth and universal voting behaviour in proportional elections*, Phys. Rev. Lett. 99, (2007).
7. Douglas D. Heckathorn, *The Dynamics and Dilemmas of Collective Action*, American Sociological Review (1996).
8. Michael W. Macy and Robert Willer, *From factors to actors: Computational Sociology and Agent-Based Modeling*, Annual Review of Sociology Vol. 28: 143-166 (2002).
9. Nilanjan Dey Samarjeet Borah Rosalina Babo Amira Ashour, *Social Network Analytics - Computational Research Methods and Techniques, First Edition*, eBook ISBN: 9780128156414, Paperback ISBN: 9780128154588, Imprint: Academic Press, Published Date: 23rd November 2018

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	2	-	-	-	-	1	1
CO2	3	3	2	2	2	2	-	-	-	-	1	1
CO3	2	2	3	2	2	2	-	-	-	-	1	1
CO4	3	3	3	2	2	2	-	-	-	-	1	1

22MAT746

Special Distribution Functions

3 0 0 3

Course outcomes

CO1: To understand the basics of different beta distributions.

CO2: To understand the distribution like Empirical and Erlang.

CO3: Know the importance of inverse Gama distribution and generalized exponential distribution.

CO4: To understand the inverse Gaussian distribution and Lognormal distribution.

Inverted Beta Distribution, Noncentral Beta Distribution, Beta Binomial Distribution, Cauchy Distribution, Noncentral Chi-Squared Distribution, Dirichlet Distribution, Empirical Distribution Function, Erlang Distribution, Error Distribution, Generalized Exponential Distributions, Noncentral F-distribution, Inverted Gamma Distribution, Normal Gamma Distribution, Generalized Gamma Distribution, Inverse Gaussian (Wald) Distribution, Lognormal Distribution, Pareto Distribution, Power Function Distribution, Power Series (Discrete) Distribution, Wishart (Central) Distribution.

Text/Reference books

1. Catherine Forbes, Merran Evens, Nicholas Hastings and Brian Peacock. (2010). *Statistical Distributions*, Fourth Edition, Wiley & Sons Publication, USA.
2. Karl Bury (1999) : *Statistical distributions in Engineering* , Cambridge University Press.
3. Thomopoulos, Nick T (2017): *Statistical Distributions: Applications and Parameter Estimates*, Springer.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	3	2	2	2	2						2	1
CO2	2	2	2	2	2	2						2	1
CO3	2	2	2	2	2	3						2	1
CO4	2	2	1	2	2	2						2	1

22MAT747

Stochastic Process

3 0 0 3

Course outcomes

CO1 Understand the concepts of stochastic process, markov chains and classifical of states and chains.
CO2. Understand the markov process with discrete state space as poisson process and its properties with related theorems.
CO3. Understand the markov process with continuous state space as wiener process and its properties.
CO4. Understand the renewal process and related theorems.
CO5. Understand the concepts of branching process and Bellman-Harris process.

Random processes: General concepts and definitions - stationarity in random processes - strict sense and wide sense stationary processes - autocorrelation and properties- special processes – Poisson points, Poisson and Gaussian processes and properties , spectrum estimation , ergodicity, mean ergodicity, correlation ergodicity, Power spectrum density functions – properties, Markov process and Markov chain, transition probabilities, Chapman Kolmogrov theorem, limiting distributions classification of states.

Text Books:

1. J. Ravichandran, “*Probability and Random Processes for Engineers*”, First Edition, IK International, 2015
2. Douglas C. Montgomery and George C. Runger, *Applied Statistics and Probability for Engineers*, (2005) John Wiley and Sons Inc.

Reference Books:

1. A. Papoulis, and Unnikrishna Pillai, “*Probability, Random Variables and Stochastic Processes*”, Fourth Edition, McGraw Hill, 2002.
2. Scott L. Miller, Donald G. Childers, “*Probability and Random Processes*”, Academic press, 2012.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	2	2	2	1						2	1
CO2	3	2	2	2	2	1						2	1
CO3	3	2	2	2	2	1						2	1
CO4	3	2	1	2	2	1						2	1
CO5	2	2	2	2	2	1						1	1

22MAT748

Survival Analysis

3 0 0 3

Course outcomes

CO1: Understand the basics of survival distributions and its applications.

CO2: Understand the censoring schemes and its applications.

CO3: Know the importance of non-parametric methods.

CO4: To gain knowledge competing risk theory.

Survival Analysis: Functions of survival times, survival distributions and their applications
Censoring Schemes: Type I, Type II and progressive or random censoring with biological examples. Estimation of mean survival time and variance of the estimator for Type I and Type II censored data with numerical examples.

Non-parametric methods: Actuarial and Kaplan-Meier methods for estimating survival function and variance of the Estimator.

Competing Risk Theory: Indices for measurement of probability of death under competing risks and their inter-relations. Estimation of probabilities of death using maximum likelihood principle and modified minimum Chi-square methods.

References

1. Miller,R.G. *Survival analysis*, John Wiley ,1981
2. Collet, D. *Statistical analysis of life time data*,1984
3. Cox, D.R. and Oakes, D.: *Analysis of survival data*, Chapman & Hall, New York,1984
4. Gross, A.J. and Clark, V.A.: *Survival distribution: Reliability applications in the Biomedical sciences*, John Wiley and Sons,1975
5. Elandt-Johnson,R.E. Johnson, N.L. : *Survival models and data analysis*, John Wiley & sons.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
--	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------

Course outcomes

CO1: Understand the Data Classifications and qualifications. Scientific thinking.

CO2: Understand Data quality issues and data quality metrics.

CO3: Understand and apply the Ethics of data science

Data Classifications and qualifications. Scientific thinking. Creative and Logical thinking.

Complexities in data. Data quality issues and data quality metrics.

Ethics in data science.

References:

1. Longbing Cao, Data Science Thinking, The Next Scientific, Technological and Economic Revolution, Springer, 2018.
2. Max Shron, Thinking with Data, How to Turn Information Into Insights, O'Reilly Media, 2014.
3. Robert Stackowiak, Tracey Kelly, Design Thinking in Software and AI Projects, Apress, 2020.

CO-PO Mapping:

	PO1	PO2	PO3	PO4	PO5	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	2	2	2	2	2	-	-	-	-	1	1	
CO2	3	3	2	2	2	2	-	-	-	-	1	1	
CO3	2	2	3	2	2	2	-	-	-	-	1	1	