KALASALINGAM
ACADEMY OF RESEARCH & EDUCATION
(DEEMED TO BE UNIVERSITY)

Under sec. 3 of UGC Act 1956. Accredited by NAAC with "A" Grade

SCHOOL OF COMPUTING

Department of Computer Science and Information Technology

M.Sc., (Data Science)

(CBCS)

CURRICULUM AND SYLLABUS

REGULATIONS 2018

KALASALINGAM
ACADEMY OF RESEARCH & EDUCATION
(DEEMED TO BE UNIVERSITY)

Under sec. 3 of UGC Act 1956. Accredited by NAAC with "A" Grade

SCHOOL OF COMPUTING

Department of Computer Science and Information Technology

UNIVERSITY VISION

To be a Centre of Excellence of International Repute in Education and Research.

UNIVERSITY MISSION

To Produce Technically Competent, Socially Committed Technocrats and Administrators through Quality Education and Research

DEPARTMENT VISION

To become a Center of excellence offering quality education and innovation in Computer Science and Information Technology.

DEPARTMENT MISSION

1. To prepare the students to excel in the field of Computer Science and IT industry
2. To prepare the students to pursue higher studies and develop sustainable innovative solutions for the society.

KALASALINGAM
ACADEMY OF RESEARCH & EDUCATION
(DEEMED TO BE UNIVERSITY)
Under sec. 3 of UGC Act 1956. Accredited by NAAC with "A" Grade

SCHOOL OF COMPUTING

Department of Computer Science and Information Technology

M.Sc (Data Science)

Programme Educational Objectives (PEOs)

PEO1: Graduates will be capable to become leaders, equipped with managerial and analytical skills needed for data driven decision making

PEO2: Graduates are prepared to meet industry demand in the field of Data Science with proficiency in statistical methods and data analytics tools.

PEO3: Graduates will be engaged in lifelong learning and progress into research and development in Data Science and Big Data

Programme Outcomes (POs)

PO1: Understand the concepts that lead to various advanced theories in Mathematics, Statistics and Computer science.

PO2: Demonstrate methodologies to design solutions for complex data science problems.

PO3: Apply Mathematical sciences and recent technologies in Computer Science to solve real life problems

PO4: Acquire the skills in handling data analytics tools towards problem solving and solution analysis.

PO5: Enhance and spread knowledge in data science to society with ethical, moral and social values

PO6: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO7: Interpret data findings effectively to any audience, orally, visually and in written formats.

PO8: Apply research- based knowledge to analyse and manage projects in data science.

KALASALINGAM
ACADEMY OF RESEARCH & EDUCATION
(DEEMED TO BE UNIVERSITY)
Under sec. 3 of UGC Act 1956. Accredited by NAAC with "A" Grade

SCHOOL OF COMPUTING

Department of Computer Science and Information Technology

M.Sc., (Data Science)

(Those who joined 2019 onwards)

SCHEME OF INSTRUCTION

Course Code	Course Name	Course Type	L	T	P	C
MDS18R4001	Statistical Methods	T	3	1	0	4
MDS18R4002	Linear Algebra and its Applications	T	3	1	0	4
MDS18R4003	Foundations of Data Science	T	3	1	0	4
MDS18R4004	Data Structures and Algorithms	T	3	1	0	4
MDS18R4005	Advanced Database Management Systems	T	3	1	0	4
MDS18R4081	R Programming Lab	P	0	0	4	2
MDS18R4082	Data structures Lab	P	0	0	4	2
MDS18R4006	Statistical Modelling	T	3	1	0	4
MDS18R4007	Optimization Techniques	T	3	1	0	4
MDS18R4008	Machine Learning	T	3	1	0	4
MDS18R4009	Linear Predictive Analytics	T	3	1	0	4
MDS18R****	Elective-I	TP	3	1	0	4
MDS18R4083	Machine Learning Lab	P	0	0	4	2
MDS18R4084	Linear Predictive Analytics Lab	P	0	0	4	2
MDS18R4061	Python Programming	P	1	0	2	2
MDS18R5001	Big Data Analytics	T	3	1	0	4
MDS18R5002	Deep Learning	T	3	1	0	4
MDS18R5003	Cloud Computing	T	3	1	0	4
MDS18R5041	Research Methodology	T	3	1	0	2
MDS18R50**	Elective-II	TP	3	1	0	4
MDS18R5085	Deep learning Lab	P	0	0	4	2
MDS18R5062	Data Visualization Techniques	P	0	0	4	2
MDS18R5086	Seminar and Comprehensive Viva	P	0	0	4	2

MDS18R50**	Elective Paper-III (Online Courses)	T	4	0	0	4
MDS18R5098	Project and Viva Voce	Project	0	0	16	12

ELECTIVES

Course. Code	Course Name	Course Type	L	T	P	C
MDS18R4031	Information Security	TP	3	0	2	4
MDS18R4032	GPU Computing	TP	3	0	2	4
MDS18R4033	Natural Language Processing	TP	3	0	2	4
MDS18R4044	Internet of Things	TP	3	0	2	4
MDS18R4045	Mobile Application Development	TP	3	0	2	4
MDS18R4046	Recommendation System	TP	3	0	2	4
MDS18R4047	Data Analytics Tools	TP	3	0	2	4
MDS18R5031	Image and Video Analytics	TP	3	0	2	4
MDS18R5032	Web Data Analytics	TP	3	0	2	4
MDS18R5033	Business Analytics	TP	3	0	2	4
MDS18R5034	Text Analytics	TP	3	0	2	4
MDS18R5035	Social Media Mining	TP	3	0	2	4
MDS18R5036	Sentiment Analysis	TP	3	0	2	4
MDS18R5037	Health Care Data Analytics	TP	3	0	2	4

T-Theory TP-Theory with Practical L-Laboratory

Non-CGPA Courses

Sl. No.	Courses	Credit
1	NET/SET/JEST/GATE coaching classes*	1
	a) Pass in examination based on the coaching classes. Exam will be conducted by the department at the end of the third for coaching classes.	1
	b) Paper presentation in National/International Conferences/Seminars	1 1
	c) Participation in workshops (3 days)	1
	d) Participation in Guest Lecture (5 Nos.)	1
	e) Internship	1
	f) Foreign Language/National Language	

*80% attendance is compulsory in this category even if the student earns Non-CGPA credit as mentioned in the table.

The students should score minimum 2 credits for completing the

MDS18R4001	STATISTICAL METHODS	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES

To enable the students to understand about collection, presentation and analysis of data.

COURSE OUTCOMES

Upon successful completion for the course, students will be able to

CO1: understand Quantitative and Qualitative data

CO2: present the data using curves/graphics

CO3: understand different measures of dispersion

CO4: understand correlation with diagrammatic representation

CO5: Acquire the knowledge to use the regression analysis

UNIT- I

Introduction: Nature and scope of Statistics, limitations of statistics - Types of data: Concept of population and sample, primary and secondary data, quantitative and qualitative data, discrete and continuous data, cross-sectional and time series data. Diagrammatic and Graphical representation of data - Line diagram, bar diagram, pie diagram and sub-divided bar diagram, Frequency distribution and cumulative frequency distribution and their graphical representations, Frequency polygon, histogram, ogive, frequency curves.

UNIT-II

Measures of Location : Definition of location-Arithmetic mean, Median, Mode, Geometric mean and Harmonic mean for individual observations, discrete series and continuous series data and their properties.

UNIT-III

Measures of Dispersion : Definition of dispersion-Range, Quartile deviation, Mean deviation, Standard deviation, combined standard deviation, co-efficient of variation for individual observations, discrete series and continuous series data and their properties .

UNIT-IV

Measures of Skewness- Definition of skewness-Karl Pearson's coefficient of skewness, Bowley's coefficient of skewness, Kelly's coefficient of skewness.

Measure of Kurtosis-Definition of kurtosis- Co-efficient of skewness and Kurtosis based on moments.

UNIT -V

Bivariate data – scatter diagram, Correlation – Karl Pearson's coefficient of correlation- Rank correlation-Spearman's rank correlation coefficient.
Regression Analysis: Introduction - Uses of regression analysis - regression lines - regression equations of X on Y and Y on X - regression equation in terms of correlation table-Applications.

TEXT BOOKS:

1. Statistical Methods, S.P.Gupta, Sultan Chand and sons
2. Basic Statistics, R.Wilcox, Oxford University Press, 2009
3. Fundamentals of Statistics: Volume I, Goon A.M, Gupta M.K., Dasgupta B, World press, 1998

REFERENCES:

1. Murray R Spiegel and Larry J Stephens: Statistics, Schaum's Outline, Fourth edition, 2008
2. R.S.N. Pillai, Statistics, S. Chand Publishing Company Pvt Ltd , 1992

MDS18R4002	LINEAR ALGEBRA AND ITS APPLICATIONS	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES

To enable the students to understand the concept of linear algebra and its applications.

COURSE OUTCOMES

Upon successful completion for the course, students will be able to
CO1: understand the matrix operations
CO2: understand the vector spaces and subspaces, graphs and Networks
CO3: acquire the knowledge of determinant and applications
CO4: find the Eigen values and Eigenvectors of a matrix
CO5: enable to find maximum & minimum values of a function

UNIT-I

The Geometry of Linear Equations- An Example of Gaussian Elimination- Matrix Notation and Matrix Multiplication – Triangular Factors and Row Exchanges- Inverses and Transposes

UNIT-II

Vector Spaces: Vector Spaces and Subspaces - Solving Ax=0 and Ax=b – Linear Independence, Basis, and Dimension- The Four Fundamental Subspaces- Graphs and Networks- Linear Transformations

UNIT-III

Determinants: Introduction- Properties of the Determinant- Formulas for the Determinant- Applications of Determinants

UNIT-IV

Eigenvalues and Eigenvectors: Introduction- Diagonalization of a Matrix .- Difference Equations and Powers A^k - Differential Equations and e^{At} - Complex Matrices- Similarity Transformations

UNIT- V

Positive Definite Matrices: Minima, Maxima, and Saddle Points - Tests for Positive Definiteness- Singular Value Decomposition

TEXT BOOK:

1. Gilbert Strang(2006). Linear Algebra and Its Application, Fourth Edition, Academic Press.

REFERENCES:

1. David C. Lay, Steven R. Lay, Judi J. McDonald (2014). Linear Algebra and Its Applications, Pearson Education.
2. Peter D. Lax(2007). Linear Algebra and Its Applications, Second Edition, Wiley Publication

MDS18R4003	FOUNDATIONS OF DATA SCIENCE -	L	T	P	C
		3	1	0	4

COURSE OBJECTIVE

The main objective of the course is to give the students understanding various activities of data scientists and to train them in the usage of R software towards visualization and statistical analysis. At the end of the course students are expected to possess necessary skills in developing R packages.

COURSE OUTCOMES

Upon the completion of this course, the students will be able to

- **CO1:** Learn about the Data Evolution and understanding the data
- **CO2**: Understand the usage of basic tools in R for learning data science
- **CO3**: Managing Data Frames in R
- **CO4:** Learning the utilities of dlpyr package
- **CO5:** Managing graphic tools available in R

UNIT -I **12 Hours**

Introduction to Data Science - sources of data - different types of data and variables – Preprocessing of data – Handling Missing Values – Scaling Procedures – Dimension reduction

UNIT – II **12 Hours**

Introduction to R and R packages-understanding the functioning of R console – Data preparation using R data editor and other sources – saving, editing and reading data files and scripts – R operators, arithmetic, relational and logical operators – Structures in R.

UNIT- III **12 Hours**

Vectors and vector arithmetic Matrices. Matrix creations - manipulations, slicing and matrix operations including decompositions- Lists and their manipulations with exposure to apply functions of various types

UNIT- IV **12 Hours**

Data Frames: Creating Data frames using scripts – importing data from external sources like Excel spread sheets, text files and other sources - Exploring data frames-Accessing columns in a Data frame- slicing of data frame, operation on data frames using ***dplyr*** package.

UNIT V **12 Hours**

Data Visualization using base package and ggplot2 – Charts of various types – pie charts, histograms, bee sworm chart, frequency polygon, density curves – simple and multiple bar charts – box plot – scatter plot - creating subplots - Drawing inferences from plots and charts

Books for study and Reference:

1. Garrett Grolemund and Hadley Wickham (2016): R for Data Science, O'Reilly Media
2. Jeffrey S.Saltz,Jeffre M. Stanton(2018): An Introduction to Data Science, Sage Publications.
3. Nina Zumal, John Mount (2014): Practical Data science in R, Managing Publication Company
4. Peng,R.D.(2016) : R for Data Science, lulu.com
5. Purohit,S.G., Gore,S.D. and Dehmuk,S.R.(2019): Statistical Analysis using R,Narosha Publishing
6. Seema Acharya (2018) : Data Analytics using R, McGrawhill

R online help manual available in standard installation of R

MDS18R4004	DATA STRUCTURES ALGORITHMS	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES

This course aims to introduce a number of popular data structures and algorithms, along with the basic techniques in algorithm analysis. Students can understand common data structures and algorithms,, analyze the complexities of data structures and algorithms, choose appropriate data structures and algorithms for problem solving.

COURSE OUTCOMES:

CO1 : Perceive the role of algorithms and programming constructs as a systematic and efficient way of solving problems.

CO2 : Understand the fundamental concepts of Stack and Queue

CO3: Demonstrate various operations of heap, shortest path algorithms and single source shortest path algorithms

CO4: Understand and implement sorting and searching algorithms

CO5: Analyze the various problems and its efficiency

UNIT- I

Basics: Algorithm Specification - Data Abstraction - Performance Analysis - Arrays - Dynamic Allocated Arrays - Structures and Unions - Polynomials – Sparse Matrices- Representation of Multidimensional Arrays - Strings.

UNIT- II

Stacks and Queues: Stacks - Stacks Using Dynamic Arrays – Queues – Circular Queues Using Dynamic Arrays – Evaluation of Expressions – Multiple Stacks And QueuesLinked Lists: Singly Linked List And Chains - Representing Chains in C - Linked Stack AndQueues - Polynomials – Additional List Operations – Sparse Matrices - Doubly Linked List.

UNIT -III

Notion of an Algorithm - Fundamentals of Algorithmic Problem Solving - Important Problem Types - Fundamentals of the Analysis of Algorithm Efficiency - Analysis Framework - Asymptotic Notations and its properties - Mathematical analysis for Recursive and Non-recursive algorithms.

UNIT- IV

Computing a Binomial Coefficient - Warshall‟s and Floyd‟ algorithm - Optimal Binary Search Trees - Knapsack Problem and Memory functions. Greedy Technique- Prim‟s algorithm- Kruskal's Algorithm-Dijkstra's Algorithm-Huffman Trees.

UNIT -V

Limitations of Algorithm Power–Lower–Bound Arguments–Decision Trees–P, NP and NP–Complete Problems––Coping with the Limitations – Backtracking – n–Queens problem – Hamiltonian Circuit Problem – Subset Sum Problem–Branch and Bound – Assignment problem – Knapsack Problem – Traveling Salesman Problem– Approximation Algorithms for NP – Hard Problems – Traveling Salesman problem – Knapsack problem.

TEXT BOOKS:

1.. Ellis Horowitz, Sartaj Sahni and Anderson Freed (2009), Fundamentals of data structures in C,University Press
2. Ellis Horowitz, Sartaj Shani, Sanguthevar Rajasekaran (2013), Fundamentals of computer algorithms, Galgotias Publications private limited
3. Robert L Kruse (2008). Data Structures & Program Design, New Delhi: Prentice Hall
4. Anany Levitin, "Introduction to the Design and Analysis of Algorithms", Third Edition, Pearson Education, 2012.

REFERENCES:

1. Mark Allen Weiss(2012). Data Structures and Algorithm Analysis in C++; Pearson Education
2. Sartaj Sahni (2010). Data Structures, Algorithms, and Applications in C++; McGraw–Hill International Edition
3. Thomas H.Cormen, Charles E.Leiserson, Ronald L. Rivest and Clifford Stein, "Introduction to Algorithms", Third Edition, PHI Learning Private Limited, 2012.
4. Alfred V. Aho, John E. Hopcroft and Jeffrey D. Ullman, "Data Structures and Algorithms", Pearson Education, Reprint 2006.
5. Donald E. Knuth, "The Art of Computer Programming", Volumes 1& 3 Pearson Education, 2009.
6. Steven S. Skiena, "The Algorithm Design Manual", Second Edition, Springer, 2008

MDS18R4005	ADVANCED DATABASE MANAGEMENT SYSTEMS	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES:

The objective of the course is to present an introduction to database management systems, with an emphasis on how to organize, maintain and retrieve effective information from a DBMS.

COURSE OUTCOMES:

CO1 : Understand the fundamental elements of relational database management systems.
CO2 : Analyze the Indexing and Hashing techniques.
CO3: Analyze the concepts of Big data bases
CO4 : Understand the concepts of Distributed databases patterns and Consistency models
CO5 : Design data models to represent simple database application scenarios

UNIT–I

Introduction – Database concepts, Basic components of DBMS, sources of data – data models - hierarchical - network - XML and Stores – Relational Database Design: Anomalies ina Database-Functional Dependency - Lossless Join and Dependency - Preserving Decomposition - Third Normal Form- BoyceCodd Normal Form -Multivalued Dependency - Fourth Normal Form - Join Dependency - Project Join Normal Form -Domain Key Normal Form – SQL: Data Definition - Data Manipulation - Integrity Constraints-Views-PL/SQL.

UNIT-II

Indexing and Hashing - Query Processing - Transaction Processing - Concurrency Control and Recovery – Advanced Database Concepts and Emerging Applications: Distributed Databases - Object Oriented Databases – Object Relational Databases- Data mining and Data Warehousing -

UNIT-III

Big Data – Big Databases- SQL-NoSQL Tradeoffs-CAP Theorem-Eventual Consistency -NoSQL-database types - Document Oriented - Columnar - Graph - KeyValue Pair – NoSQLdatabase, design for performance / quality parameters, documents and information retrieval

UNIT-IV

Distributed Database Patterns— Distributed Relational Databases- Non-relational Distributed Databases- MongoDB – Sharing and Replication- HBase- Cassandra-Consistency Models— Types of Consistency- Consistency MongoDB- Hbase Consistency- Cassandra Consistency.

UNIT-V

Data Models and Storage- SQL- NoSQL APIs- Return SQL- Advance Databases—PostgreSQL-Riak- CouchDB- NEO4J- Redis- Future Databases— Revolution Revisited-Counter revolutionaries- Oracle HQ- Other Convergent Databases- Disruptive Database Technologies.

TEXT BOOKS

1.Abraham Silberchatz, Henry K.Forth, Sudharshan, “Database system Concepts” - (6th edition), McGraw Hill, 2010.
2.Guy Harrison, “Next Generation Databases”, Apress, 2015.
3.Eric Redmond, Jim R Wilson, “Seven Databases in Seven Weeks”, LLC. 2012.

REFERENCES

1.K. Pakhira, “Database Management System”, Phi Learning Pvt. Ltd., 2012

2. MongoDB: The Definitive Guide, 2nd Edition , Powerful and Scalable Data Storage, By Kristina Chodorow, Publisher: O'Reilly Media
3.MongoDB Basics - EelDavid Hows,Peter Membrey,coPlugge, Publisher Apress - Ebook(free) https://it-ebooks.info/book/4527/

MDS18R4081	R PROGRAMMING LAB	L	T	P	C
		0	0	4	2

Course Objectives:

The main objective of the course is to explore the power of R, statistical computing software for performing various tasks related to Data Science. At the end of the course, students will be able to perform several activities like exploratory data analysis, data visualization and building of their own packages

Course Outcomes:

Upon successful completion of this course, the student will be able to:

CO1: Perform calculations related to linear algebra.
CO2: Performing statistical calculations.
CO3: Creating charts/diagrams to visualize data.
CO4: Analyzing data frames using dplyr
CO5: effectively use ggplot package

List of Exercises.

1. Matrix construction and slicing
2. Solving a system of equations
3. Relationship between eigen values ,
   Trace and determinant of a matrix
4. QR decomposition
5. Pie chart with legends
6. Histogram with legends
7. Bar charts with legend
8. Multiple bar charts
9. Drawing Boxplot and identifying outliers, skewness and kurtic properties
10. Computing Basic Statistical Measures
11. Analysis of a data frame using dplyr package
13. Importing a csv file and performing statistical analysis
14. List creation and usage of lapply and sapply functions
15. User defined functions
16. ggplot usage

MDS18R4082	DATA STRUCTURES LAB	L	T	P	C
		0	0	4	2

Course Objectives

To develop skills to design and analyze simple linear and non linear data structures,students can identify and apply the suitable data structure for the given real world problem

Course Outcomes:

CO1: Identity the appropriate data structure for given problem
CO2: Have practical knowledge on the application of data structures
CO3: Implement linked list data structure to solve various problems.

CO4: Apply graph and tree traverse technique to various applications.
CO5: Implement Dijkstra's algorithm, Btrees and hash tables.

PROGRAMS

1. Factorial and Fibonacci series using recursion
2. Strassen's Matrix Multiplication
3. Array implementation of Linear and Binary search
4. Infix to postfix conversion
5. Singly Linked List operations
6. Binary tree traversals
7. Quick sort
8. Find the Shortest Path using Dijkstra's Algorithm – Greedy method
9. Knap sack problem using Dynamic programming
10. Travelling sales man problem and 8 queen's problem using backtracking

MDS18R4006	STATISTICAL MODELLING	L	T	P	C
		3	1	0	4

COURSE OBJECTIVE:

To enable the students to understand the concepts of theory of probability and s tatistics

COURSEOUTCOMES

Upon successful completion of this course, students will be able to
CO1: Understand the Axioms of probability
CO2: Understand the notion of probability distributions
CO3: Understanding the theory of sampling distribution and their usage in statistical testing
CO4: Learning the significance of nonparametric methods
CO5: Learning the use of analysis of variance.

Unit 1

Combinatorial Probability including Bayes theorem-Introduction to random variables

Unit 2

Probability distributions - Univariate Discrete and Continuous distributions : Binomial, Poisson, Uniform, Normal and Cauchy distributions (understanding the shapes and properties of distributions must be emphasized rather than focusing on mathematical aspects)- , Simulation from Univariate distributions

Unit 3

Notion of sampling from infinite populations, Meaning of Sampling Distributions, Elementary testing of hypothesis about mean- one sample and two sample problems, variance and equality of variances

Unit 4

Elementary Non Parametric Tests Covering Man Whitney U Test, Wilcoxson Signed Rank Test, Shapiro Wilk Test, Kolmogrov – Smirnov tests

Unit 5

One way and Two way Anova, Tukey HSD test, Kruskal Wallis Test

Book for Study and Reference:

V.K.Rohatgi,A.K. : An Introduction to Probability Theory and Mathematical Statistics, Wiley

REFERENCES:

1.GoonA.M.,GuptaM.K.&Dasgupta B,AnOutlineofStatisticalTheory(Vol-1),WorldPress ,1994.
2. ScymourLipschutz, Probability, Schaum's Outline, 1996.

MDS18R4007	OPTIMIZATION TECHNIQUES	L	T	P	C
		3	1	0	4

COURSE OBJECTIVE:

To enable the students to understand the optimization techniques and its importance in resource management.

COURSE OUTCOMES

Upon successful completion of this course, students will be able to

CO1: Formalize the LPP Model
CO2: Apply the transportation and assignment algorithm
CO3: Apply the shortest route problem
CO4: Understanding and applying dynamic programming models
CO5: Acquiring the knowledge of markovian and advanced queuing models

UNIT- I

Graphical solution of linear programming problems - LP solution space in equation form - Simplex method - Artificial starting solution.

UNIT -II

Introduction to transportation and assignment problems - The transportation algorithm - The Assignment model.

UNIT- III

Minimal spanning tree algorithm - Shortest route problem - Maximal flow model.

UNIT- IV

Elements of the dynamic programming model - Optimality principle - Examples of dynamic programming models and their solutions

UNIT -V

Single and multiple server Markovian queuing models – Steady state system size probabilities – Little's formula – Customer impatience – Priority queues – M/G/1 queuing system – P - K formula.

TEXT BOOK

1. Taha, H.A., Operations Research: An Introduction, 8th Edition, Pearson Education, 2011.

REFERENCES

1. Natarajan, A.M., Balasubramani, P., Tamilarasi, A., Operations Research, Pearson Education, Asia, 4th edition,2009.
2. Prem Kumar Gupta, Hira, D.S., Operations Research, 3rd Edition, Sultan Chand and Company Ltd, New Delhi, 11th edition,2007.

MDS18R4008	MACHINE LEARNING	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES:

The main objective of this course is to expose the students to some popular machine learning methods used in Industry. This course will pave foundation for introducing more modern methods being used Data Analytics industry

COURSE OUTCOMES

CO1: Understand the nature of data types and metrics
CO2: Understand the need for hierarchical clustering methods and various approaches of clustering.
CO3: Understand partitioning clustering and its applications in real life situations
CO4: Learning various types of classifiers and their application in different domains
CO4: Understand the usage of support vector machines.
CO5: Apply association analysis in real life situations.

Unit I: Introduction to Data Mining – Data Types – Distance Metrics- Various approaches of data mining.

Unit II: Hierarchical Clustering – Agglomerative Nesting – Divisive Analysis – Single Linkage – Complete Linkage – Average Linkage- Detailed study on Wards Clustering – Dendrograms

Unit III: Partitioning algorithms – k means clustering - implementation and issues – k medoids clustering - validity measures like entropy, purity, precision, recall, F measure and Silhouette coefficient and their interpretations.

Unit IV: Introduction to classification – Applications – Neighborhood based classification- k-nn classifier - Naïve Bayesian classification - Decision trees for categorical and interval valued features

Unit V: Support Vector Machines – need for kernels – multiclass svm – svm for classification, regression and novelty detection- tuning of parameters - Association Analysis – details of a priori algorithm – applications

Books for Study:

Kaufman.M. and Rousseeuw,P.J. (2005) : Finding Groups in Data: An Introduction to Cluster Analysis , John Wiley and Sons
Tan, P.N., Steinbach, M and Kumar, V. (2006) : Introduction to Data Mining, Pearson-Addison Wesly,

Books for References

Agarwal,C.C. (2015) : Data Mining: The Text Book, Springer
Berry, M.W. (2004) : Survey of Text Mining,Springer
Duda,R, Hart,P and Stork,D.(2001): Pattern Classication, John Wiley and Sons.
Han,J and Kamber,M and Pei,J. (2012): Data Mining : Concepts and Techniques, Margan Kaufmann Publishers
Pujari,A.K.(2010) : Data Mining Techniques, Universities Press (Hyderabad)
RajanChattamvelli (2016) : Data Mining Methods, Narosha Publishers
Zaki,M.J.andMeira,W. (2014) : Data Mining and Analysis: Fundamental Cocepts and Algorithms, Cambridge University Press.

MAT18R4009	LINEAR PREDICTIVE ANALYTICS	L	T	P	C
		3	1	0	4

COURSE OBJECTIVE:

To enable the students to learn various tools available in linear predictive analytics.

COURSE OUTCOMES

Upon successful completion of this course, students will be able to

CO1: Understand the role of linear models in prediction
CO2: Decide on Variable Selection and to know the application of modern regression tools
CO3: Apply various validation methods and detect abnormalities in the data set
CO4: Understand the application different discriminant procedures.
CO5: Use logistic regression and other statistical classifiers

Unit 1: Introduction to Predictive Analytics - Least-Square Method – Estimation and testing of intercept and regression coefficient – Confidence Intervals for Slope and Intercept Regression – Building confidence and prediction bands

Unit 2 : Multiple Linear Regression – Testing for regression coefficients and confidence intervals for regression coefficients in Multiple Linear Regression – R-square and adjusted R-square values- Variable Selection-Stepwise and Ridge Regression

Unit 3 : Outlier detection- Influential point, Cooks distance-Handling the problem of multicollinearity

Unit 4 : Classification, Linear Classifiers, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis

Unit 5 : Logistic Regression- CART- CHAID-their applications

Books for Study and Reference

1. Hastie, Tibshirani, Friedman, The Elements of Statistical Learning, Data Mining, Inference and Prediction, Second Edition, Springer Series in Statistics.
2. JohnsonR.A.andWichern,W(2001):AppliedMultivariateStatisticalAnalysis,Fifthedition, PrenticeHall.
3. Montgomery,D.G.,Peck,E.A. and Vining,G.G.(201) : Linear Regression Analysis,Wiley

MDS18R4083	MACHINE LEARNING LAB	L	T	P	C
		0	0	3	2

Course Objectives

The main objective of this course is to understand various types of data types one encounters in machine learning. In addition the students will be introduced the mathematical background as well as the applications of various machine learning algorithms.

Upon successful completion of this course, students will be able to

CO1: Understand the basics of machine learning concepts
CO2: Apply necessary preprocessing techniques and data transformation techniques. **CO3:** Apply various kind of data analytics functions to build models
CO4: Implement the code for regression/classification/clustering
CO5: Find the accuracy of the model build and explain the evaluation

List of Exercises

1. knn classification
2. Concept of neighbourhood for binary data
3. Tuning knn using e1071
4. knn for mixed data type
5. Naïve Bayesian classification
6. Laplacian smoothing
7. Decision tree for categorical data
8. SVM using different types of kernals
9. SVM for regression
10. Rule mining
11. Hierarchical clustering
12. Partitioning clustering
13. Image segmentation using k-means

MDS18R4084	LINEAR PREDICTIVE ANALYTICS LAB	L	T	P	C
		0	0	3	2

Course Objectives

The main objective of this course is to make students understand a complex situation involving several variables and come out with a model that can be used to explain the situation and perform prediction in a suitable manner.

Course Outcomes

Upon successful completion of this course, the student will be able to:

CO1: Perform simple regression analysis including prediction; test the significance of regression coefficients

CO2: Handle data sets with several regressors.

CO3: Use different evaluation criteria and model validation

CO4: Use predictive models based on statistical theory

CO5: Understand the prediction models related to different types of dependent variables

List of Exercises

1. Regression model - testing significance
2. Confidence and prediction bands
3. Multiple Linear Regression model
4. Ridge regression
5. Stepwise regression
6. Logistic regression
7. Variables selection
8. Linear discriminant analysis
9. Quadratic discriminant analysis

MDS18R4061	PYTHON PROGRAMMING	L	T	P	C
		1	0	2	2

Course Objectives

Upon successful completion of this course, students will be able to

CO1: Understand the basic features of python programming
CO2: write algorithm to solve the given problem.
CO3: Apply appropriate functions to read and preprocess the given dataset.
CO4: Implement the python code to analyze different kind of datasets
CO5: Show the analyzing results in the form of chart, histogram, plots, graph and explain them.

List of Programs

1. Implement a sequential search
2. Create a calculator program
3. Explore string functions
4. Implement Selection Sort
5. Implement Stack
6. Read and write into a file
7. Demonstrate usage of basic regular expression
8. Demonstrate use of advanced regular expressions for data validation.
9. Demonstrate use of List
10. Demonstrate use of Dictionaries
11. Create Comma Separate Files (CSV), Load CSV files into internal Data Structure
12. Write script to work like a SQL SELECT statement for internal Data Structure made in earlier exercise
13. Write script to work like a SQL Inner Join for an internal Data Structure made in earlier exercise

MDS18R5001	BIG DATA ANALYTICS	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES

- To understand the Big Data Platform and its Use cases
- To provide HDFS Concepts and Interfacing with HDFS
- To understand Map Reduce Jobs and apply analytics on Structured, Unstructured Data.

At the end of the course student will be able to

CO1 : Explain the big data management architecture and big data technology components
CO2: Understand the fundamentals of hadoop system and analyze data with hadoop
CO3: Understand MapReduce function and its types and apply MapReduce function to various datasets
CO4: Install spark and SCALA, Explain spark applications and basic features of SCALA
CO5: Write queries to access/update data from NoSQL, MongoDB, Hbase and RDBMS databases

UNIT –I

Introduction to big data: Introduction - Big Data– Characteristics of Big Data - Big data management architecture– Examining Big Data Types - Big Data Technology Components –- Big data analytics -Big data analytics examples – Web Data Overview - Web Data in Action

UNIT– II

Hadoop : Introduction - History of Hadoop – Hadoop Ecosystem– Analyzing data with Hadoop – Hadoop Distributed File System– Design – HDFS concepts – Hadoop filesystem -Data flow - Hadoop I / O – Data integrity - Serialization – Setting up a Hadoop cluster – Cluster specification –cluster setup and installation - YARN

UNIT– III

MapReduce: Introduction - Understanding Map, Reduce functions – Scaling out – Anatomy of a MapReduce Job Run – Failures - Shuffle and sort – Mapreduce types and formats – features - counters – sorting – Mapreduce Applications - Configuring and setting the environment – Unit test with MR unit– local test

UNIT–IV

Spark: - Installing spark - Spark applications, Jobs, Stages and Tasks -Resilient Distributed databases– Anatomy of a Spark Job Run - Spark on YARN– SCALA: Introduction– Classes and objects– Basic types and operators– built–in control structures– functions and closures– inheritance

UNIT–V

NoSQL Databases: Introduction to NoSQL– MongoDB: Introduction - Data types -- Creating,Updating and deleing documents –Querying - Introduction to indexing - Capped collections. Hbase:Concepts – Hbase Vs RDBMS – Creating records– Accessing data - Updating and deleting data -Modifying data– exporting and importing data.

TEXT BOOK:

1. Bill Franks (2012). Taming the Big Data Tidal wave, John Wiley & Sons
2. Tom White (2012). Hadoop : The Definitive Guide, Third Edition, O'Reilly Media
3. Martin Odersky, Lex Spoon,, Bill Venners(2010),Programming in Scala, Second Edition, Artima Press, California.

REFERENCES:

1. Boris lublinsky, Kevin t. Smith, Alexey , Yakubovich(2015). Professional Hadoop Solutions,Wiley
2. Chris Eaton, Dirk deroos et al.(2012). Understanding Big data, McGraw Hill
3. Min Chen (2014). Big Data : Related Technologies, Challenges and Future Prospects, Springer
4. Judith Hurwitz (2013). Big Data for Dummies, John Wiley & Sons

MDS185002	DEEP LEARNING	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES

- To acquire knowledge on the basics of neural networks.
- To implement neural networks using computational tools for variety of problems.
- To explore various deep learning algorithms.

COURSE OUTCOMES

Students will be able to:

CO1: Develop algorithms simulating human brain.

CO2: Implement Neural Networks in Tensor Flow for solving problems.

CO3: Explore the essentials of Deep Learning and Deep Network architectures.

CO4: Define, train and use a Deep Neural Network for solving real world problems that require artificial Intelligence based solutions.

CO5: Implement deep learning in various applications

UNIT-I

Basics of Deep leaning- Deep learning architectures: Convolutional Neural Networks :Neurons in Human Vision-The Shortcomings of Feature Selection-Vanilla Deep Neural Networks Don't Scale-Filters and Feature Maps-Full Description of the Convolutional Layer-Max Pooling-Full Architectural Description of Convolution Networks-Closing the Loop on MNIST with Convolutional Networks-Image Preprocessing Pipelines Enable More Robust Models-Accelerating Training with Batch Normalization-Building a Convolutional Network for CIFAR-10-Visualizing Learning in Convolutional Networks- Leveraging Convolutional Filters to Replicate Artistic Styles-Learning ConvolutionalFilters for Other Problem Domains-Training algorithms.

UNIT-II

Memory Augmented Neural Networks : Neural Turing Machines-Attention-Based Memory Access-NTM Memory Addressing Mechanisms-Differentiable Neural Computers-Interference-Free Writing in DNCs-DNC Memory Reuse-Temporal Linking of DNC Writes-Understanding the DNC Read Head-The DNC Controller Network- Visualizing the DNC in Action-Implementing the DNC in TensorFlow-Teaching a DNC to Read and Comprehend.

UNIT-III

Deep Reinforcement Learning: Deep Reinforcement Learning Masters Atari Games-What Is Reinforcement Learning?-Markov Decision Processes (MDP)-Explore Versus Exploit-Policy versus Value Learning-Pole-Cart with Policy Gradients-Q-Learning and Deep Q-Networks-Improving and Moving Beyond DQN.

UNIT-IV

Implementing Neural Networks in TensorFlow : What Is TensorFlow?-How Does TensorFlow Compare to Alternatives?-Installing TensorFlow-Creating and Manipulating TensorFlow Variables-TensorFlow Operations-Placeholder Tensors-Sessions in TensorFlow-Navigating Variable Scopes and Sharing Variables-Managing Models over the CPU and GPU-Specifying the Logistic Regression Model in TensorFlow-Logging and Training the Logistic Regression Model

UNIT-V

Applications: Deep learning for computer vision, Deep Learning Applications at the Enterprise Scale, Deep Learning Models for Healthcare Applications.

TEXT BOOK:

1. Nikhil Buduma, Nicholas Locascio, "Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms", O'Reilly Media, 2017.

REFERENCE:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, "Deep Learning (Adaptive Computation and Machine Learning series", MIT Press, 2017.

MDS18R5003	CLOUD COMPUTING	L	T	P	C
		3	1	0	4

OBJECTIVES:

This course gives the idea of cloud environment, building software systems and components that scale to millions of users in modern internet, Virtualization, Security related challenges and cloud concepts capabilities across the various cloud service models including Iaas,Paas,Saas, and developing cloud based software applications on top of cloud platforms.

COURSE OUTCOMES:

CO1: Understanding the key dimensions of the challenge of Cloud Computing and Services.
CO2: Articulate the main concepts, key technologies, strengths and limitations of cloud computing.
CO3: to Identify the architecture, infrastructure and delivery models of cloud computing.
CO4: Understand various performance criteria to evaluate the quality of the cloud architecture and advanced Technologies.
CO5: Explain the core issues of cloud computing such as security, privacy and interoperability.

UNIT- I

Evolution of Cloud Computing -System Models for Distributed and Cloud Computing - NIST Cloud Computing Reference Architecture –IaaS - On-demand provisioning - Elasticity in cloud -Egs of IaaS providers – PaaS - Egs. Of PaaS providers – SaaS - Egs. Of SaaS providers - Public , Private and Hybrid clouds.

UNIT- II

Basics of virtualization – Types of Virtualization –Implementation Levels of Virtualization – Virtualization Structures - Tools and Mechanisms – Virtualization of CPU, Memory, I/O Devices – Desktop virtualization - Server Virtualization.

UNIT- III

Architectural Design of Compute andStorage Clouds - Layered Cloud Architecture Development DesignChallenges – Inter Cloud Resource Management - Resource Provisioning and Platform Deployment - Global Exchange of Cloud Resources.

UNIT- IV

Parallel and Distributed programming-Paradigms - MapReduce, Twister and Iterative MapReduce - Hadoop, Library from Apache - Mapping Applications – Programming Support –
Google App Engine, Amazon AWS – Cloud Software Environments –Eucalyptus, Open nebula, OpenStack.

UNIT- V

Security Overview - Cloud Security-Challenges - Software-as-a-Service Security - Security Governance -Risk Management - Security Monitoring - Security Architecture Design -Data Security - Application Security - Virtual Machine Security.

TEXT BOOKS:

1. Distributed and Cloud Computing, From Parallel Processing to the Internet of Things by Kai Hwang, Geoffrey C Fox, Jack G Dongarra, Morgan Kaufmann Publishers, 2012.
2. Cloud Computing: Implementation, Management, and Security by John W.Rittinghouse and James F.Ransome : CRC Press 2010

REFERENCE BOOKS:

1. Cloud Computing, A Practical Approach by Toby Velte, Anthony Velte, Robert Elsenpeter: TMH, 2013
2. Cloud Application Architectures: Building Applications and Infrastructure in the Cloud: Transactional Systems for EC2 and Beyond (Theory in Practice
(O'Reilly)) by George Reese: O'Reilly
3. James E. Smith, Ravi Nair, Virtual Machines: Versatile Platforms for Systems and Processes, Elsevier/Morgan Kaufmann, 2005.
4. Katarina Stanoevska-Slabeva, Thomas Wozniak, Santi Ristol, "Grid and Cloud Computing - A Business Perspective on Technology and Applications",

MDS18R5041	RESEARCH METHODOLGY	L	T	P	C
		3	1	0	4

COURSE OBJECTIVES

- Explain what research is and what it is not, and the different definitions of research.
- Present some aspects of the debate about the nature of knowledge and the value of scientific method.
- Understand some basic concepts of research and its methodologies
- Identify appropriate research topics
- Select and define appropriate research problem and parameters
- Prepare a project proposal (to undertake a project)
- Organize and conduct research (advanced project) in a more appropriate manner
- Write a research report and thesis

COURSE OUTCOMES

Students will be able to:

CO1: To develop understanding of the basic framework of research process.

CO2: To develop an understanding of various research designs and techniques.

CO3: To identify various sources of information for data collection.

CO4: Gain a practical understanding of the various methodological tools used for social scientific research.

CO5: To prepare good research report.

UNIT- I

The features of scientific research -Building blocks of science in research -Concept of Applied and Basic research –Quantitative and Qualitative Research Techniques -Need for theoretical frame work -Hypothesis development -Hypothesis testing with quantitative data. Research design -Purpose of the study: Exploratory, Descriptive, Hypothesis Testing.

UNIT- II

Laboratory and the Field Experiment -Internal and External Validity -Factors affecting Internal validity. Measurement of variables -Scales and measurements of variables. Developing scales - Rating scale and attitudinal scales -Validity testing of scales -Reliability concept in scales being developed -Stability Measures.

UNIT- III

Interviewing, Questionnaires, etc. Secondary sources of data collection. Guidelines for Questionnaire Design -Electronic Questionnaire Design and Surveys. Special Data Sources: Focus Groups, Static and Dynamic panels. Review of Advantages and Disadvantages of various Data – Collection Methods and their utility. Sampling Techniques - Probabilistic and non–probabilistic samples.

UNIT- IV

Data Analysis-Factor Analysis -Cluster Analysis -Discriminant Analysis -Multiple Regression and Correlation -Canonical Correlation -Application of Statistical (SPSS) Software Package in Research.

UNIT –V

Purpose of the written report -Concept of audience -Basics of written reports. Integral parts of a report -Title of a report, Table of contents, Abstract, Synopsis, Introduction, Body of a report - Experimental, Results and Discussion -Recommendations and Implementation section -Conclusions and Scope for future work.

TEXT BOOKS

1. Donald R. Cooper and Ramela S. Schindler, Business Research Methods, Tata McGraw–Hill Publishing Company Limited, New Delhi, 2000
2.Uma Sekaran, Research Methods for Business, John Wiley and ons Inc., New York, 2000.

REFERENCES

1.C.R.Kothari, Research Methodology, Wishva Prakashan, New Delhi, 2001.
2. Donald H.McBurney, Research Methods, Thomson Asia Pvt. Ltd. Singapore, 2002.

3.G.W.Ticehurst and A.J.Veal, Business Research Methods, Longman, 1999.
4. Ranjit Kumar, Research Methodology, Sage Publications, London, New Delhi, 1999.

MDS18R5085	DEEP LEARNING LAB	L	T	P	C
		0	0	4	2

COURSE OBJECTIVES

- To acquire knowledge on the basics of neural networks.
- To implement neural networks using computational tools for variety of problems.
- To explore various deep learning algorithms.

COURSE OUTCOMES

Students will be able to:

CO1: Develop algorithms simulating human brain.
CO2: Implement Neural Networks in Tensor Flow for solving problems.
CO3: Explore the essentials of Deep Learning and Deep Network architectures.
CO4: Define, train and use a Deep Neural Network for solving real world problems that require artificial Intelligence based solutions.
CO5: Implement deep learning in various applications

MDS18R5062	DATA VISUALIZATION TECHNIQUES	L	T	P	C
		0	0	4	2

COURSE OBJECTIVES

- To explore the fundamental concepts of data pre-processing, extraction, cleaning, annotation, integration.
- To understand the various information visualization techniques.
- An understanding of the key techniques and theory used in visualization, including data models, graphical perception and techniques for visual encoding and interaction.
- Understand why visualization is an important part of data analysis
- Understand the components involved in visualization design.

COURSE OUTCOMES

Students will be able to:

CO1: Design and create data visualizations.
CO2: Gain the knowledge about D3.
CO3: Use knowledge of chart library, generate chart and animate the graphs.
CO4: Apply the various operations on data source.
CO5: Understand and apply principles of data visualization.

UNIT -I

Context of data visualization - Definition, Methodology, Visualization design objectives. Key Factors - Purpose, visualization function and tone, visualization design options - Data representation, Data Presentation, Seven stages of data visualization, widgets,data visualization tools. Mapping – Time series –Connections and correlations – Scatter plot maps – Trees, Hierarchies and Recursion - Networks and Graphs, Info graphics.

UNIT- II

Introduction to D3 – Fundamental Technology –Drawing with data - Scales - Axes - Updates, Transition and Motion - Interactivity – Layouts -Geomapping - Exporting- Data to create Visualizationwith SVG – SVG - Styling CSS -Shapes - SVG Properties - SVG Text – Drawing - Transformations - Building Chart with SVG (Scalable Vector Graphics) – Shaping Web Pages - Selections - Attributes - Chaining Methods-Data Joins – Sizing - scales - axes - Loading - Filtering - Interactive Charts - Buttons using Data Join - Transition using Key

UNIT- III

D3-BASED REUSABLE CHART LIBRARY: Introduction to D3 - Setup and Deployment - Generate Chart - Customize Chart - How to Use APIs - Customize Style - Building Real time and Live Updating animated graphs with C3.

UNIT- IV

TABLEAUE INTRODUCTION: Environment Setup - Navigation - File & Data Types. DATA SOURCE: Custom Data View - Extracting Data - Fields Operations - Editing Meta Data - Data Joining - Data Blending. Worksheets

UNIT -V

TABLEAUE CHARTS: Bar Chart - Line Chart - Pie Chart - Scatter Plot - Bubble Chart -Gantt Chart - Histograms – Waterfall Charts. ADVANCED: Dashboard - Formatting -Forecasting - Trend Lines

TEXT BOOK:

1. Ben Fry, “Visualizing Data”, O‟Reilly Media, Inc., 2007.

REFERENCES

1. Scott Murray, “Interactive data visualization for the web”, O‟Reilly Media, Inc., 2013.
2. Ritchie S. King – Visual Storytelling with D3 - An Introduction to Data Visualization with D3,Addison-Wesley-Data Analytic Series,ISBN 10: 0321933176
4. Elijah Meeks , D3.js in Action, Second Edition: Data visualization with JavaScript, Publisher: Manning Publications, 2017 , ISBN: 9781617294488

ELECTIVES

MDS18R4031	INFORMATION SECURITY	L	T	P	C

		3	0	2	4

COURSE OBJECTIVES

- To select appropriate techniques to tackle and solve problems in the discipline of information security management;
- To know why security and its management are important for any modern organization;

COURSE OUTCOMES

Upon successful completion of this course, students will be able to

1. Develop an understanding of security policies (such as authentication, integrity and confidentiality), as well as protocols to implement such policies in the form of message exchanges.
2. To know the aspects of risk management
3. To study the critical need for ensuring Information Security in Organizations
4. To know the legal, ethical and professional issues in Information Security
5. To know the technological aspects of Information Security

UNIT- I

Introduction to Information Security: Introduction- - Security - Critical Characteristics of Information -
NSTISSC Security Model - Components of an Information System - Security Components Approaches to Information Security Implementation - The Systems Development Life Cycle - The Security Systems Development Life Cycle

UNIT- II

The Need for Security: Business Needs First - Threats - Attacks - Secure Software Development.Risk management: Overview - Risk Identification - Risk Assessment- Risk Control Strategies -Selecting a Risk Control Strategy - Quantitative Versus Qualitative Risk Control Practices Planning for Security: Information Security Policy, Standards and Practices - The Information Security Blue print

UNIT- III

Security Technology: Firewalls and VPNS - Intrusion Detection, Access Control, and other Security Tools - Intrusion Detection and Prevention Systems - Honey Pots - Honey Nets, and Padded Cell Systems - Scanning and Analysis Tools - Access Control Devices

UNIT- IV

Cryptography: Foundations of Cryptology– cipher methods– cryptographic algorithms - cryptographic tools. Physical Security: Physical access controls - fire security and safety - failure of supporting utilities and structural collapse - Interception of data - mobile and portable systems - special considerations for physical security threats

UNIT- V

Implementation of Information Security: Information Security Project Management –Technical Topics of Implementation – Non Technical aspects of Implementation Information Security Maintenance: Security Management Models – The Maintenance Model – Digital forensics

TEXT BOOK:

1. Michael E. Whitman and Herbert J. Ma ttord (2014). Principles of Information Security, 5/e, Cengage Learning,Indian edition

REFERENCES:

1. Charles A.Sennewald(2011). Effective Security Management, 5/e,Elsevier
2. Dhiren R. Patel(2008). Information Security: Theory and Practice, Prentice Hall of India PvtLtd
3. S.M. Bhaskar, S.I. Ahson(2008). Information Security: A Practical Approach, Alpha Science
4. Gerald L.Kovacich(2003). Information System Security Officer's Guide, Butterworth Hinemann

MDS18R4032	GPU COMPUTING	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES

The goal is to convey a deep understanding of GPU architecture and APIs (OpenGL, GLSL, CUDA) with important practical applications. The goal is an understanding of both the traditional use of GPUs for rendering graphics, as well as the use of GPUs for general purpose computations (GPGPU), or GPU Computing.

COURSE OUTCOMES

Upon successful completion of this course, students will be able to

1 To understand the architecture and programming of GPUs (Graphics Processing Units).

2 To understand the traditional use for rendering graphics, as well as the use of GPUs for general purpose computations (GPGPU), or GPU Computing.

3 To study the traditional use of GPUs for graphics and visualization, as well as their use for general purpose computations (GPGPU).

4 To know the GPU many-core hardware architectures, shading and computer programming languages and APIs.

5 To study the linear algebra computations, alternative and future architectures.

UNIT I

Introduction: History, Graphics Processors, Graphics Processing Units,GPGPUs. Clock speeds, CPU / GPU comparisons, Heterogeneity, Accelerators, Parallel programming, CUDA OpenCL /OpenACC,Hello World Computation Kernels, Launch parameters, Thread hierarchy, Warps / Wavefronts, Thread blocks / Workgroups, Streaming multiprocessors, 1D / 2D/ 3D thread mapping, Device properties, Simple Programs

UNIT II

Memory: Memory hierarchy, DRAM / global, local / shared, private / local,textures, Constant Memory, Pointers, Parameter Passing, Arrays and dynamic Memory, Multi-dimensional Arrays, Memory Allocation, Memory copying across devices, Programs with matrices, Performance evaluation with different memories

UNIT III

Synchronization: Memory Consistency, Barriers (local versus global), Atomics,Memory fence. Prefix sum, Reduction. Programs for concurrent Data Structures such as Worklists, Linked-lists. Synchronization across CPU and GPU **Functions**: Device functions, Host functions, Kernels functions, Using libraries (such as Thrust), and developing libraries.

UNIT IV

Support: Debugging GPU Programs. Profiling, Profile tools, Performance aspects **Streams**: Asynchronous processing, tasks, Task-dependence, Overlapped data transfers, Default Stream, Synchronization with streams. Events, Event-based-Synchronization - Overlapping data transfer and kernel execution, pitfalls.

UNIT V

Case Studies: Image Processing, Graph algorithms, Simulations, Deep Learning **Advanced topics**: Dynamic parallelism, Unified Virtual Memory, Multi-GPU processing, Peer access, Heterogeneous processing

REFERENCES

1. Programming Massively Parallel Processors: A Hands-on Approach; David Kirk, Wen-mei Hwu;Morgan Kaufman; 2010 (ISBN: 978-0123814722)

CUDA Programming: A Developer's Guide to Parallel Computing with GPUs; Shane Cook :Morgan Kaufman; 2012 (ISBN: 978-0124159334)

MDS18R4033	NATURAL LANGUAGE PROCESSING	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES

This course introduces the fundamental concepts and techniques of natural language processing (NLP). Students will gain an in-depth understanding of the computational properties of natural languages and the commonly used algorithms for processing linguistic information. The course examines NLP models and algorithms using both the traditional symbolic and the more recent statistical approaches.

COURSE OUTCOMES

Upon successful completion of this course, students will be able to

1 To introduce students the challenges of empirical methods for natural language processing (NLP) applications.

2 To introduce basic mathematical models and methods used in NLP applications to formulate computational solutions.

3 To provide students with the knowledge on designing procedures for natural language resource annotation and the use of related tools for text analysis and hands-on experience of using such tools.

4 To introduce students work in information retrieval, information extraction, and knowledge discovery using different natural language resources.

5 To give an overview of the major technologies in speech recognition and synthesis including tools for acoustic analysis and hands-on experience of using such tools

UNIT I

OVERVIEW AND LANGUAGE MODELING

Overview: Origins and challenges of NLP-Language and Grammar-Processing Indian Languages-NLP Applications-Information Retrieval. Language Modeling: Various Grammar- based Language Models-Statistical Language Model.

UNIT II

WORD LEVEL AND SYNTACTIC ANALYSIS

Word Level Analysis: Regular Expressions-Finite-State Automata-Morphological Parsing-Spelling Error Detection and correction-Words and Word classes-Part-of Speech Tagging. Syntactic Analysis: Context-free Grammar-Constituency- Parsing-Probabilistic Parsing.

UNIT III

SEMANTIC ANALYSIS AND DISCOURSE PROCESSING

Semantic Analysis: Meaning Representation–Lexical Semantics– Ambiguity–Word Sense Disambiguation. Discourse Processing: cohesion–Reference Resolution– Discourse Coherence and Structure.

UNIT IV

NATURAL LANGUAGE GENERATION AND MACHINE TRANSLATION

Natural Language Generation: Architecture of NLG Systems– Generation Tasks and Representations– Application of NLG. Machine Translation: Problems in Machine Translation–Characteristics of Indian Languages– Machine Translation Approaches–Translation involving Indian Languages.

UNIT V

INFORMATION RETRIEVAL AND LEXICAL RESOURCES

Information Retrieval: Design features of Information Retrieval Systems–Classical, Non–classical, Alternative Models of Information Retrieval – valuation Lexical Resources: World Net–Frame Net–Stemmers–POS Tagger– Research Corpora.

TEXT BOOKS

1. Tanveer Siddiqui, U.S. Tiwary, "Natural Language Processing and Information Retrieval", Oxford University Press, 2008.

REFERENCES

1 Daniel Jurafsky and James H Martin, "Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition", 2 nd Edition, Prentice Hall, 2008.
2. James Allen, "Natural Language Understanding", 2nd edition, Benjamin /Cummings publishing company, 1995.

MDS18R4044	INTERNET OF THINGS	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES

Explore the interconnection and integration of the physical world and the cyber space to Design & Develop IOT Devices.

COURSE OUTCOMES

Upon successful completion of this course, students will be able to

1 Able to realize the revolution of Internet in Mobile Devices, Cloud & Sensor Networks.

2 Able to understand building blocks of Internet of Things and characteristics.

3 To know the aspects of sensors and activators.

4 To study the Arduino and Raspberry packages.

5 To know the various Real time applications of IoT

UNIT- I

Introduction-Characteristics - Physical design - Protocols-Logical design - Enabling technologies - IoT levels-Domain specific IoTs - IoT vs M2M

UNIT- II

IoT systems management - IoT design methodology-Specifications - Integration and Application Development

UNIT- III

Sensors and activators - Communication modules - Zigbee-RFID-Wi-Fi-Power sources.

UNIT -IV

Platform - Arduino/Intel Galileo/Raspberry Pi- Physical device - Interfaces - Programming - APIs/Packages - Web services.

UNIT -V

Various Real time applications of IoT-Connecting IoT to cloud-Cloud storage for IoT-Data Analytics for IoT- Software & Management Tools forIoT.

TEXT BOOKS

1. Arshdeep Bahga, Vijay Madisetti, "Internet of Things-A hands-on approach", Universities Press,2015.

REFERENCES

1. Manoel Carlos Ramon, —Intel® Galileo and Intel® Galileo Gen 2: API Features and Arduino Projects for Linux Programmers• , Apress,2014.
2. Marco Schwartz, —Internet of Things with the Arduino Yun• , Packt Publishing,2014

MDS18R4045	MOBILE APPLICATION DEVELOPMENT	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES

Mobile application development frameworks; Architecture, design and engineering issues, techniques, methodologies for mobile application development.

COURSE OUTCOMES

1 To study an Application models of mobile application frameworks.

2 To know the User-interface design for mobile applications.

3 To understand how to managing application data and how to Integrating with cloud services in Mobile applications

4 To understand Integrating networking, the OS and hardware into mobile-applications.

5 To understand Publishing, deployment, maintenance and management.

UNIT I

Getting Started With Android Programming - What is Android ?, Android SDK installation and configuration, Anatomy of an Android application, Activities, Fragments and Intents-Understanding Activities, Linking Activities using intents, Fragments, Calling Built in applications using intents, Displaying Notifications.

UNIT II

The Android User Interface- Understanding the components of a screen, Adapting to display orientation, Managing changes to screen orientation, Creating the user interface programmatically, Listening for UI notifications, Designing User Interface with Views- Using basic views, Using Picker Views, Understanding Specialized fragments,

UNIT III

Data Persistence - Saving and Loading User Preferences, Persisting Data to Files, Creating and sing Databases. Content Providers - Sharing Data in Android, Using a Content Provider, Creating Your Own Content Providers.

UNIT IV

Messaging - SMS Messaging, Sending Email. Location-Based Services - Displaying Maps, Getting Location Data, Monitoring a Location.

UNIT V

Networking - Consuming Webservices using HTTP, Consuming JSON Services, Sockets Programming, Developing Android Services - Creating Your Own Services, Establishing Communication between a service and an activity, Binding Activities to Services, Understanding Threading, Publishing Android Applications.

TEXT BOOKS

1. Wei-Meng Lee,"Beginning Android 4 Application Development", Wrox publications, 2012

REFERENCES

1. The Android Developer's Cookbook: Building Applications with the Android SDK James Steele, Nelson to Addison Wesley Publications 2010 First Edition.
2. Professional Android Application Development. Reto Meier, Wrox publications, 2009, Second Edition

MDS18R5031	IMAGE AND VIDEO ANALYTICS	L	T	P	Credit
		3	0	2	4

COURSE OBJECTIVE

The goal of machine vision is to develop methods that enable a machine to "understand" or analyze images and videos. This course will address the research issues towards developing algorithms that can perform high-level visual recognition tasks on real-world images and videos. This course will review and discuss current approaches to high-level visual recognition problems, such as background modeling, object recognition and categorization, tracking, scene understanding, human motion understanding, etc.

COURSE OUTCOMES

CO1: To have a better knowledge about videos
CO2: To enrich students with data analytics
CO3:To understand the video content analysis
CO4: To expose the student to various applications and case studies of Video analytics.

UNIT I INTRODUCTION TO BIG DATA AND DATA ANALYSIS

Introduction to Big Data Platform – Challenges of Conventional Systems – Web Data – Evolution of Analytic Scalability – Analytic Processes and Tools – Analysis Vs Reporting – Modern Data Analytic Tools – Data Analysis: Regression Modeling – Bayesian Modeling – Rule Induction.

UNIT II MINING DATA STREAMS

Introduction to Stream Concepts – Stream Data Model And Architecture – Stream Computing – Sampling Data in a Stream – Filtering Streams – Counting Distinct Elements in a Stream–Estimating Moments – Counting Oneness in a Window – Decaying Window – Real Time Analytics Platform(RTAP) Applications – Case Studies.

UNIT III VIDEO ANALYTICS

Introduction – Video Basics – Fundamentals for Video Surveillance – Scene Artifacts – Object Detection and Tracking: Adaptive Background Modelling and Subtraction – Pedestrian Detection and Tracking – Vehicle Detection and Tracking – Articulated Human Motion Tracking in Low Dimensional Latent Spaces.

UNIT IV BEHAVIOURAL ANALYSIS AND ACTIVITY RECOGNITION

Event Modelling – Behavioural Analysis – Human Activity Recognition – Complex Activity Recognition – Activity modeling using 3D shape - Video summarization – shape based activity models – Suspicious Activity Detection.

UNIT V HUMAN FACE RECOGNITION AND GAIT ANALYSIS

Introduction: Overview of Recognition algorithms – Human Recognition using Face: - Face Recognition from still images – Face Recognition from video – Evaluation of Face Recognition Technologies – Human Recognition using gait: HMM Framework for Gait Recognition – View Invariant Gait Recognition – Role of Shape and Dynamics in Gait Recognition.

REFERENCES

1. Michael Berthold, David J.Hand, "Intelligent Data Analysis", Springer, 2007.
2. Anand Rajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", Cambridge University Press, 2012.
3. Yunqian Ma, Gang Qian, "Intelligent Video Surveillance: Systems and Technology", CRC Press (Taylor and Francis Group), 2009.
4. Rama Chellappa, Amit K.Roy– Chowdhury, Kevin Zhou.S, "Recognition of Humans and their Activities using Video", Morgan & Claypool Publishers, 2005.

MDS18R	WEB DATA ANALYTICS	L	T	P	C
		4	0	0	4

UNIT- I

Introduction : Web Analytics 2.0 – Clickstream– multiple outcome analysis–experimentation and testing– voice of customer – competitive intelligence– the tactical shift –Optimal strategy for choosing web analytics

UNIT –II

Clickstream analysis: Metrics–Eight critical web metrics–web metrics demystified –strategically aligned tactics for impactful web –Web analytics report–Foundational analytical strategies–clickstream analysis made actionable–challenges

UNIT- III

Measuring Success–Actionable Outcome KPIs– Moving beyond conversion rates– Micro and macro conversion–Measuring success for a non –ecommerce website– Leveraging qualitative data: Surveys–Web enabled emerging user research options

UNIT- IV

A/B Testing – Multivariate testing–Actionable testing ideas–Controlled experiments–Competitive intelligence analysis–CI data source, types, secrets– website traffic analysis–Search and keywordanalysis– audience identification and segmentation analysis

UNIT –V

Emerging analytics: Social. mobile, video: Measuring social web – the data challenge– analyzing mobile customer experiences–measuring the success of blogs– quantifying the impact of Twitter - Analyzing the performance of videos

TEXT BOOKS

1. Avinash Kaushik (2010) , Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity, 1 st Edition, Wiley Publishing.
2. Dietmar Jannach, Markus Zanker(2011), Recommender system–An introduction, Cambridge University Press
3. Bing Liu(2012), Sentiment Analysis and opinion mining, Morgan and claypool Publishing

REFERENCES

1.Eric Enge, Stephan Spencer, Jessie Stricchiola, The Art of SEO: Mastering Search Engine Optimization, 3 rd Edition.
2. Kristina Halvors, Content Strategy for the Web, 1 st Edition.

MDS18R4033	BUSINESS ANALYTICS	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES:

The main objective of this course is making the students to know about various analytics done in business related to consumer behaviour and different sectors of business like banking, finance, health care etc.

COURSE OUTCOMES:

CO1: Able to understand the challenges in health care industry and top health care adaptations
CO2: Gain knowledge in Banking, Finance and Insurance
CO3: Understand analytics tools in telecommunications
CO4: Analyze the consume behavior in retail industry
CO5: Perform case study on leading e–commerce industries

UNIT– I

Healthcare analytics - Introduction – Potential contributions – Challenges of healthcare industry – current and future state of healthcare analytics - top healthcare analytics adaptations

UNIT– II

Banking and Finance: Systems of Banking - Commercial Banking - New Financial Services:Insurance Services - Types of Insurance - Housing Finance.

UNIT– III

Telecommunication: Introduction – End–User Needs and Demands– Telecom Business

UNIT– IV

Retail analytics - Understanding the new consumer - Marketing in a consumer – driven era – Managing the brand to drive loyalty

UNIT– V

Case studies: Walmart, Netflix, Facebook, Uber, Amazon, Kaggle

TEXT BOOK:

1. Dwight McNeill(2013). A Framework for Applying Analytics in Healthcare: What Can Be Learned from Best Practices in Banking , Retail, Politics and Sports, Pearson Education

REFERENCES:

1.Gomez Clifford(2011). Banking and Finance Theory Law and practice, PHI Learning
2. Anders Olsson(2004). Understanding Changing Telecommunications, Wiley Publications
3. Jennifer LeClaire, Danielle Dahlstrom, Vivian Braun. Business analytics in Retail for dummies, 2 nd IBM Limited edition
4.. Alistair Croll (2013) Lean analytics: Use Data to Build a Better Startup faster, O Reilly Publishers
5. Bernard Marr (2016). Big Data in Practice - How 45 successful companies used big data

MDS18R5034	TEXT ANALYTICS	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES:

The main objective of this course is making the students to know about various analytics done in texts available in Internet and social media with knowing prior knowlwdgw about Text Mining

COURSE OUTCOMES:

CO1: Understand Text Mining algorithms ,pre–processing in textual information
CO2: Demonstrate about Text categorization using classification and clustering
CO3: Understand about information retrieval usin key word search through unsupervised algorithms
CO4: Demonstrate about Probabilristic models and mixture models in text mining
CO5: Understand Visualization approaches for Text Analytics

UNIT– I

Text Mining – Definition–General Architecture–Algorithms– Core Operations : Distributions– Frequentand near Frequent Sets– Associations - Isolation of Interesting Patterns - Analyzing Document Collection over Time - Using Background Knowledge for text mining–Pre–processing–

Textual information to numerical vectors –Collecting documents–document standardization–tokenization–lemmatization–vector generation for prediction–sentence boundary determination – evaluation performance

UNIT –II

Text Categorization –Definition –Document Representation – Feature Selection –Decision TreeClassifiers –Rule– based Classifiers – Probabilistic and Naive Bayes Classifiers – Linear Classifiers–Classification of Linked and Web Data – Meta–Algorithms–Clustering –Definition–Vector Space Models – Distance–based Algorithms–Word and Phrase–based Clustering –Semi–Supervised Clustering –Transfer Learning

UNIT –III

Information retrieval and text mining–keyword search–nearest –neighbor methods–similarity–web–based document search–matching–inverted lists–evaluation. Information extraction–Architecture – Co–reference – Named Entity and Relation Extraction–Template filling and database construction–Applications.Inductive –Unsupervised Algorithms for Information Extraction. Text Summarization Techniques –Topic Representation –Influence of Context –Indicator representations –Pattern Extraction –Apriori Algorithm – FP Tree algorithm

UNIT– IV

Probabilistic Models for Text Mining –Mixture Models –Stochastic Processes in Bayesian Nonparametric Models –Graphical Models –Relationship Between Clustering, Dimension Reduction and Topic Modeling –Latent Semantic Indexing –Probabilistic Latent Semantic Indexing –Latent Dirichlet Allocation–Interpretation and Evaluation –Probabilistic DocumentM.Sc Data Clustering and Topic Models –Probabilistic Models for Information Extraction –Hidden Markov Models – Stochastic Context–Free Grammars – Maximal Entropy Modeling –Maximal Entropy Markov Models –Conditional Random Fields

UNIT– V

Visualization Approaches –Architectural Considerations –Common Visualization Approaches for text mining –Example–Mining Text Streams –Text Analytics in Social Media –Opinion Mining and Sentiment Analysis –Document Sentiment Classification – Opinion Spam Detection –Text Mining Applications and Case studies

TEXT BOOK:

1.Sholom Weiss, Nitin Indurkhya, Tong Zhang, Fred Damerau “The Text MiningHandbook:Advanced Approaches in Analyzing Unstructured Data,”,Springer, paperback 2010

REFERENCES:

1.Ronen Feldman, James Sanger–“The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data”–Cambridge University press, 2006.
2.Charu C. Aggarwal ,ChengXiangZhai,Mining Text Data, Springer; 2012 .

MDS18R4035	SOCIAL MEDIA MINING	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES:

The main objective of this course is making the students to gain knowledge in Social media mining and know about social media communities.

COURSE OUTCOMES:

CO1: Understand the basic concepts of social media and challenges in social media mining
CO2: Demonstrate about network models in social media
CO3: Undestand about social media communities and community algorithms
CO4: Understand about diffusion models in social network
CO5: Understand about recommendation algorithms and community behavior

UNIT- I

Social Media Mining – Introduction - Atoms - Molecules - Interactions - Social Media mining Challenges – Graphs – Basics - Nodes - Edges - Degree of Distribution– Types -Directed - Undirected - Weighted – Graph Connectivity – Tress and Forests - Bipartite graphs - Complete Graphs - Sub graphs - Planar Graphs – Graph Representation – Graph Traversal Algorithms - Shortest path algorithms Dijkstra‟s – Spanning tree algorithms - Prims – Bipartite matching –Ford–Fulkerson algorithm

UNIT- II

Network Models - Measures - Node : Eigen Centrality - Page Rank - Group Measures - Betweenness centrality – group degree centrality, centrality, and group – Closeness centrality – Node Linking Behavior – Transitivity and reciprocity – Linking Analysis – Cluster coefficient - Jaccard – Case Study : –Modelling small networks with real world model

UNIT–III

Social media Communities - Social Communities - Member based Detection - Node degree -Node Similarity - Node reachability – Group Based detection methods – balanced - robust –modular - dense – hierarchical – Spectral Clustering : Balanced Community algorithm Community Evolution – Evaluation.

UNIT–IV

Social Network - Information Diffusion - Types – herd behavior – information cascades diffusion of innovation - epidemics - Diffusion Models Case Study - Herd Behavior - Information Cascades Methods - Social Similarity - assortativity - Social Forces - Influence homophily - Confounding – Assortativity measures - Influence measures - Predictive Models

UNIT–V

Recommendation Vs Search - Recommendation Challenges - Recommender algorithms - Content Based Methods– Collaborative Filtering - Memory Based - Model Based - Social Media

Recommendation - User friendship - Recommendation Evaluation - Precision - Recall -Behavioral- User Behavior - User - Community behavior - User Entity behavior - Behavioral Analytics – Methodology

TEXT BOOK:

1. Social Media Mining: An Introduction - Reza Zafarani , MohhammadAbiElasi - Published by Cambridge press, 2014 - (Free Ebook available http://dmml.asu.edu/smm/chapter)

REFERENCE:

1.. Data Mining for Social Network Data- Memon, N., Xu, J.J., Hicks, D.L., Chen, H. (Eds.), Springer - Annals of Information Systems ,ISBN 978-1-4419-6287-4

MDS18R5036	SENTIMENT ANALYSIS	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES:

The main objective of this course is making the students to gain knowledge about sentiment analysis and its applications

COURSE OUTCOMES:

CO1: Understand the basics of sentiment analysis and its applications
CO2: Demonstrate about sentiment classification
CO3: Understand about aspect based sentiment analysis
CO4: Understand about sentiment lexicon generation an aspect based summarization
CO5: Understand about opinion search and retrieval

UNIT- I

Sentiment Analysis Applications: Sentiment Analysis Research – Opinion Spam Detection. The Problem of Sentiment Analysis: Problem Definitions – Opinion Summarization - Different Types of Opinions – Subjectivity and Emotion – Author and Reader Standing Point. Document Sentiment Classification: Sentiment Classification Using Supervised Learning - Sentiment Classification Using Unsupervised Learning – Sentiment Rating Prediction - Cross-Domain Sentiment Classification – Cross-Language Sentiment Classification.

UNIT-II

Sentence Subjectivity and Sentiment Classification: Subjectivity Classification - Sentence Sentiment Classification – Dealing with Conditional Sentences – Dealing with Sarcastic Sentences - Cross language Subjectivity and Sentiment Classification – Using Discourse Information for Sentiment Classification.

UNIT-III

Aspect-based Sentiment Analysis: Aspect Sentiment Classification – Basic Rules of Opinions and Compositional Semantics – Aspect Extraction – Identifying Resource Usage Aspect –Simultaneous Opinion Lexicon Expansion and Aspect. Extraction: Grouping Aspects into Categories – Entity, Opinion Holder and Time Extraction – Co reference Resolution and Word Sense Disambiguation.

UNIT- IV

Sentiment Lexicon Generation: Dictionary-based Approach – Corpus-based Approach –Desirable and Undesirable Facts. Opinion Summarization: Aspect-based Opinion Summarization – Improvements to Aspect-based Opinion Summarization – Contrastive View Summarization – Traditional Summarization Analysis of Comparative Opinions: Problem Definitions – Identify Comparative Sentences – Identifying Preferred Entities .

UNIT- V

Opinion Search and Retrieval: Web Search vs. Opinion Search – Existing Opinion Retrieval Techniques Opinion Spam Detection: Types of Spam and Spamming – Supervised Spam Detection – Unsupervised Spam Detection – Group Spam Detection.

TEXT BOOK:

1. Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies), Bing Liu, Morgan & Claypool Publishers (2012)

REFERENCES

1. Sentiment Analysis: Mining Opinions, Sentiments, and Emotions, Bing Liu, Cambridge University Press (2015)
2. http://nptel.ac.in/courses/106105158/61
3. Sentiment Analysis: Second Edition, GerardusBlokdyk ,Createspace Independent Publishing Platform (2018)

MDS18R4037	HEALTH CARE ANALYTICS	L	T	P	C
		3	0	2	4

COURSE OBJECTIVES:

The main objective of this course is making the students to gain knowlwdge about health care analytics and its applications

COURSE OUTCOMES:

CO1: Understand about health care analytics and benefits of Electronic health records.
CO2: Understand about Bio medical image analysis
CO3: Understand about Natural language processing and biomedical miniing
CO4: Understand about information reteival for health care.

CO5: Demonstrate about applications and practical systems for health care.

UNIT-I

Introduction: Introduction to Healthcare Data Analytics- Electronic Health Records-Components of EHR- Coding Systems- Benefits of EHR- Barrier to Adopting EHR-Challenges- Phenotyping Algorithms.

UNIT-II

Analysis: Biomedical Image Analysis- Mining of Sensor Data in Healthcare- Biomedical Signal Analysis- Genomic Data Analysis for Personalized Medicine.

UNIT-III

Analytics: Natural Language Processing and Data Mining for Clinical Text- Mining the Biomedical- Social Media Analytics for Healthcare.

UNIT-IV

Advanced Data Analytics: Advanced Data Analytics for Healthcare- Review of Clinical Prediction Models- Temporal Data Mining for Healthcare Data- Visual Analytics for Healthcare- Predictive Models for Integrating Clinical and Genomic Data- Information Retrieval for Healthcare- Privacy-Preserving Data Publishing Methods in Healthcare.

UNIT-V

Applications: Applications and Practical Systems for Healthcare- Data Analytics for Pervasive Health- Fraud Detection in Healthcare- Data Analytics for Pharmaceutical Discoveries- Clinical Decision Support Systems- Computer-Assisted Medical Image Analysis Systems- Mobile Imaging and Analytics for Biomedical Data.

TEXT BOOK:

1. Chandan K. Reddy and Charu C Aggarwal, "Healthcare data analytics", Taylor & Francis, 2015

REFERNCE:

1. Hui Yang and Eva K. Lee, "Healthcare Analytics: From Data to Knowledge to Healthcare Improvement, Wiley, 2016.