(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# **Department of Computer Science & Engineering**

Master of Engineering (M.E.)

# (Data Science)

# **II** Semester

					Maximum Marks Allotted					Contact			
S.No. Subject Code		ory	Subject Name	Theory		Practical		Total Marks	Ho	ours per eek		Total Credits	
		Categ		End Sem	Mid Sem. Exam.	Quiz/ Assignment	End Sem	Term work Lab Work & Sessional		L	Т	Р	
1.	PSCC- MCS201	PSC C	Machine Learning	60	25	15	—	_	100	3	1	_	4
2.	PSCC- MCS202	PSC C	Data Visualization	60	25	15	_	_	100	3	1	_	4
3.	PSEC- MCS201	PSE C	Program Specific Elective Course-II	60	25	15	—	_	100	3	_	_	3
4.	OEC-MCS201	OEC	<b>Open Elective Course</b>	60	25	15	_	—	100	3	_	_	3
5.	LC-MCS201	LC	Machine Learning Lab	_	—	_	60	40	100	—	-	4	2
6.	LC-MCS202	LC	Data Visualization Lab	_	—	_	60	40	100	—	_	4	2
7.	MLC-MCS201	MLC	Research Methodology & IPR	60	25	15	_	_	100	2	_	—	2
8.	AUD-MCS201	AUD	Audit Course-II	—	—	_	_	_		2	—	_	0
			Total	300	125	75	120	80	700	16	2	8	20

Program Specific Elective Course-II	Open Elective Course	Audit Course-II
PSEC-MCS201(A) Introduction to Statistical	OEC-MCS201(A) Business Analytics	AUD-MCS201(A) Soft Skills and Interpersonal
Methods		Communication
PSEC-MCS201(B) Fuzzy Logic and Application	OEC-MCS201(B) Operation Research	AUD-MCS201(B) Business Communication
PSEC-MCS201(C) Data Warehousing and Mining	OEC-MCS201(C) Industrial Safety	AUD-MCS201(C) Communication Skills
PSEC-MCS201(D) Data Science	OEC-MCS201(D) Voice & Data Network	AUD-MCS201(D) Life Science
		AUD-MCS201(E) Stress Management
		1 Hr Lecture 1 Hr Tutorial 2 Hr Practical

1 Credit 1 Credit

1 Credit

# **IPS Academy, Institute of Engineering & Science** (A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) **Scheme Based on AICTE Flexible Curriculum Department of Computer Science & Engineering** Master of Engineering (M.E.) (Data Science)

**II-Semester** 

PSCC-MCS201	Machine Learning	3L: 1T (4 hrs.)	4 credits
-------------	------------------	-----------------	-----------

### Prerequisite: Artificial Intelligence

## **Course Objective:**

The objective of this course is to introduce the basic concepts and techniques of Machine Learning and to become familiar with various methods used in ML like regression methods, classification methods, clustering methods.

### **Course Contents:**

#### Module 1:

Introduction to machine learning, scope and limitations, regression, probability, statistics and linear algebra for machine learning, convex optimization, data visualization, hypothesis function and testing, data distributions, data preprocessing, data augmentation, normalizing data sets, machine learning models, supervised and unsupervised learning.

#### Module 2:

Linearity v/s non Linearity, activation functions like sigmoid, ReLU, etc., weights and bias, loss function, gradient descent, multilayer network, Back propagation, weight initialization, training, testing, unstable gradient problem, auto encoders, batch normalization, dropout, L1 and L2 regularization, momentum, tuning hyper parameters.

#### Module 3:

Convolution neural network, flattening, sub sampling, padding, stride, convolution layer, pooling layer, loss layer, dance layer 1x1 convolution, inception network, input channels, transfer learning, one shot learning, dimension reductions, implementation of CNN like tensor flow, keras etc.

#### Module 4:

Recurrent neural network, Long short-term memory, gated recurrent unit, translation, beam search and width, Bleu score, attention model, Reinforcement Learning, RL-framework, MDP, Bellman equations, Value Iteration and Policy Iteration, Actor-critic model, Q-learning, SARSA

### Module 5:

Support Vector Machines, Bayesian learning, application of machine learning in computer vision, speech processing, natural language processing etc, Case Study: ImageNet Competition

# (8 hrs.)

(8 hrs.)

# (8 hrs.)

#### (08 hrs.)

# (08 hrs.)

- 1. Apply knowledgeofcomputing andmathematicsto machinelearningproblems, models and algorithms;
- 2. Analyzeaproblemandidentify the computing requirements appropriate for its solution;
- 3. Design, implement, and evaluate an algorithm to meet desired needs
- 4. Apply mathematical foundations, algorithmic principles and computer science theory to the modeling and design of computer-based systems.
- 5. Demonstrates comprehension of the trade-offs involved in designchoices.

- 1. Christopher M. Bishop, "Pattern Recognition and Machine Learning", Springer Verlag New York Inc., 2nd Edition, 2011.
- 2. TomM. Mitchell, "Machine Learning", McGraw Hill Education, First edition, 2017.
- IanGoodfellowand Yoshua Bengioand AaronCourville, "Deep Learning", MITPress, 2016
- 4. Aurelien Geon, "Hands-On Machine Learning with Scikit -Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems", Shroff/O'Reilly; First edition (2017).
- 5. Francois Chollet, "Deep Learning with Python", Manning Publications, 1 edition (10 January 2018).
- 6. Andreas Muller, "Introduction to Machine Learning with Python: A Guide for Data Scientists", Shroff/O 'Reilly; First edition (2016).
- 7. Russell, S. and Norvig, N., "Artificial Intelligence: AModern Approach", Prentice Hall Series in Artificial Intelligence. 2003.

# **IPS Academy, Institute of Engineering & Science** (A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum **Department of Computer Science & Engineering** Master of Engineering (M.E.) (Data Science) **II-Semester**

PSCC-MCS202	Data Visualization	3L : 1T (4 hrs)	4 credits
-------------	--------------------	-----------------	-----------

**Prerequisite:** Python, AI, Mathematics: basic linear algebra, calculus, probability

### **Course Objective:**

The objective of this course is to focus on how accurately represent voluminous complex data set in web and from other data sources and understand the methodologies used to visualize large data sets and to know the process involved in data visualization with security aspects

### **Course Contents:**

#### Module 1:

Introduction to Data Visualization, The Importance of Data Visualization, Data Wrangling, Tools and Libraries for Visualization, Overview of Statistics, Measures of Central Tendency, Measures of Dispersion, Correlation, Types of Data.

#### Module 2:

Creating visual representations, visualization reference model, visual mapping, visual analytics, Design of visualization applications.

#### Module 3:

Deep Dive into Matplotlib: Overview of Plots in Matplotlib, Simplifying Visualizations Using Seaborn, Plots, Charts, Maps.

#### Module 4:

Plotting Geospatial Data: Introduction, The Design Principles of Geoplotlib, Geospatial Visualizations.

#### Module 5:

Interactive Data Visualization for the Web: Introduction, Data-Driven Documents(D3), Technology Fundamentals: HTML, CSS, Java Script, Drawing with Data,

### **Course Outcome:**

- 1. State the core concepts of Data Visualization.
- 2. Illustrate various concepts of visual representations.
- 3. Illustrate various techniques and classification of visualization systems.
- 4. State the core concepts of network, distributed and multiprocessor operating system.
- 5. Describe the concepts of process, simulations and visualization of GIS.

### List of Text / Reference Books:

- 1. Scott Murray, "Interactive data visualization for the web", O"ReillyMedia, Inc., 2013.
- 2. Ben Fry, "Visualizing Data", O"Reilly Media, Inc., 2007.
- 3. Greg Conti, "Security Data Visualization: Graphical Techniques for Network Analysis", No Starch Press Inc, 2007.

# (06 hrs.)

(06 hrs.)

#### (12 hrs.)

#### (08 hrs.)

#### (08 hrs.)

# **IPS Academy, Institute of Engineering & Science** (A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum **Department of Computer Science & Engineering** Master of Engineering (M.E.) (Data Science) **II-Semester**

MLC-MCS201	<b>Research Methodology and IPR</b>	2L (2 hrs.)	2 credits
------------	-------------------------------------	-------------	-----------

**Prerequisite:** Basic knowledge of research, Simple mathematics

#### **Course Objective:**

The course has been developed with orientation towards research related activities and recognizing the ensuing knowledge as property. It will create consciousness for Intellectual Property Rights and its constituents. Learners will be able to perform documentation and administrative procedures relating to IPR in India as well as abroad.

#### **Course Contents:**

#### Module 1:

Meaning of research problem, Sources of research problem, Criteria Characteristics of a good research problem, Errors in selecting a research problem, Scope and objectives of research problem, Approaches of investigation of solutions for research problem, data collection, analysis, interpretation, Necessary instrumentations

#### Module 2:

Effective literature studies approaches, analysis Plagiarism, Research Ethics.

#### Module 3:

Effective technical writing, how to write report, Developing a Research Proposal, Format of research proposal, presentation and assessment by a review committee

#### Module 4:

Nature of Intellectual Property: Patents, Designs, Trade and Copyright, Process of Patenting and Development: technological research, innovation, patenting, development, International Scenario: International cooperation on Intellectual Property, Procedure for grants of patents, Patenting under PCT

#### Module 5:

Patent Rights: Scope of Patent Rights, Licensing and transfer of technology, Patent information, and databases, Geographical Indications, New Developments in IPR: Administration of Patent System, IPR of Biological Systems, Computer Software, Traditional knowledge, Case Studies.

# (10 hrs.)

(12 hrs.)

#### (08 hrs.)

### (04 hrs.)

### (06 hrs.)

- 1. Understanding and formulation of research problem.
- 2. Analyze research related information.
- 3. Understand plagiarism and follow research ethics
- 4. Understand that today's world is controlled by Computer, Information Technology, but tomorrow world will be ruled by ideas, concept, and creativity.
- 5. Understand that IPR protection provides an incentive to inventors for further research work and investment in R & D, which leads to creation of new and better products, and in turn brings about, economic growth and social benefits.

- 1. Wayne Goddard and Stuart Melville, "Research Methodology: An Introduction".
- 2. Ranjit Kumar, "ResearchMethodology: A StepbyStepGuideforbeginners", 2nd Edition.
- 3. Halbert, "Resisting Intellectual Property", Taylor & Francis Ltd, 2007.
- 4. Mayall, "Industrial Design", McGrawHill, 1992.
- 5. Niebel, "Product Design", McGraw Hill, 1974.
- 6. Asimov, "Introduction to Design", PrenticeHall, 1962.
- 7. Robert P. Merges, Peter S. Menell, Mark A. Lemley, "Intellectual Property in New Technological Age", 2016.
- 8. T. Ramappa, "Intellectual Property Rights Under WTO".
- 9. Stuart Melvilleand Wayne Goddard, "Research methodology: Anintroduction for science & engineering students", Pubished By- Kenwyn, South Africa : Juta & Co. Ltd., 1996.

# IPS Academy, Institute of Engineering & Science (A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum Department of Computer Science & Engineering Master of Engineering (M.E.) (Data Science) II-Semester

PSEC- MCS201(A)	Introduction to Statistical Methods	3L (3 hrs.)	3 credits
-----------------	--	-------------	-----------

**Prerequisite:** No prior preparation in probability in statistics is required, but familiarity with basic algebra and calculus (single and multi-variable) is assumed.

#### **Course Objective:**

1. Learn to understand the main features of traditional and modern statistics.

- 2. Learn how to analyze statistical data properly.
- 3. Understand the role of formal statistical theory and informal data analytic methods.
- 4. Gain an understanding of statistical methods relevant to upper division interdisciplinary courses.

5. Sharpen students statistical intuition and abstract reasoning as well as their reasoning from numerical data through community-based and other research.

#### **Course Contents:**

#### Module 1:

Statistical Methods: Definition and scope of Statistics, concepts of statistical population and sample. Data: quantitative and qualitative, attributes, variables, scales of measurement nominal, ordinal, interval and ratio. Presentation: tabular and graphical, including histogram and ogives, consistency and independence of data with special reference to attributes.

#### Module 2:

Measures of Central Tendency: mathematical and positional. Measures of Dispersion: range, quartile deviation, mean deviation, standard deviation, coefficient of variation, Moments, absolute moments, factorial moments, skewness and kurtosis, Sheppard's corrections.

#### Module 3:

Bivariate data: Definition, scatter diagram, simple, partial and multiple correlation (3 variables only), rank correlation. Simple linear regression, principle of least squares and fitting of polynomials and exponential curves

#### (08 hrs.)

# (08 hrs.)

# (08 hrs.)

#### Module 4:

#### Index Numbers: Definition, construction of index numbers and problems thereof for weighted and unweighted index numbers including Laspeyre's, Paasche's, Edgeworth- Marshall and Fisher's Ideal Index numbers. Errors in Index numbers. Chain index numbers, conversion of fixed based to chain based index numbers and vice-versa. Consumer price index numbers. Uses and limitations of index numbers.

#### Module 5:

Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability – classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications.

# List of Text / Reference Books:

1. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edn. The World Press, Kolkata.

2. Gupta, S. C. and Kapoor, V.K. (2008): Fundamentals Of Mathematical Statistics, 4thEdition (Reprint), Sultan Chand &Sons SUGGESTED READINGS:

3. Miller, Irwin and Miller, Marylees(2006): John E.Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.

4. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co.Ltd.

5. Mukhopadhyay, P. (1999): Applied Statistics, New Central Book Agency.

#### (08 hrs.)

#### (08 hrs.)

# IPS Academy, Institute of Engineering & Science (A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum Department of Computer Science & Engineering Master of Engineering (M.E.) (Data Science) II-Semester

PSEC- MCS201(B)Fuzzy Logic and Applications3L	L (3 hrs.) 3 c	redits
---	----------------	--------

#### Prerequisite: Discrete Structure

**Course Objective:** The main objective of this course is to introduce the ideas of fuzzy sets, fuzzy logic and use of heuristics based on human experience.

#### **Course Contents:**

#### Module 1:

Basic concepts of fuzzy set theory, operations of fuzzy sets, properties of fuzzy sets: Crisp relations, Fuzzy relational equations, operations on fuzzy relations, fuzzy systems – propositional logic, Inference, Predicate Logic, Inference in predicate logic, fuzzy logic principles, fuzzy quantifiers, fuzzy inference, fuzzy rule based systems, fuzzification and defuzzification: Types.

#### Module 2:

Fuzzy logic controllers, principles, review of control systems theory, various industrial applications of FLC, adaptive fuzzy systems, fuzzy decision making, Multi objective decision making – fuzzy classification, means clustering, fuzzy pattern recognition, image processing applications, systactic recognition – fuzzy optimization.

#### Module 3:

Fundamentals of neural networks: model of an artificial neuron, neural network architectures, Learning methods – Taxonomy of Neural network architectures, Standard back propagation algorithms, selection of various parameters, variations Applications of back propagation algorithms.

#### Module 4:

Fundamentals of genetic algorithms – genetic modeling, hybrid systems, integration of fuzzy logic, neural networks and genetic algorithms, nontraditional optimization techniques like ant colony optimization, Particle swarm optimization and artificial immune systems: applications in design and manufacturing.

#### Module 5:

Rough Set Theory - Introduction, Fundamental Concepts, Set approximation, Rough membership, Attributes, Optimization, Hidden Markov Models, Decision tree model, Introduction to Swarm Intelligence, Swarm Intelligence Techniques: Ant Colony Optimization, Particle Swarm Optimization, Bee Colony Optimization and Applications of Computational Intelligence.

### (12 hrs.)

(06 hrs.)

#### (10 hrs.)

#### (08 hrs.)

(04 hrs.)

- 1. Develop the skill in basic understanding on fuzzy logic.
- 2. Develop the skill in basic understanding on neural network.
- 3. Explore the functional components of neural classification conducer and the functional components of fuzzy logic classification on controller.
- 4. Developand implement abasictrainableneural network(or) a fuzzy logic system to design and manufacturing.
- 5. Understand the recent advances in fundamentals of genetic algorithm.

- 1. Rajasekaran.S., Vijayalakshmi Pai.G.A. "Neural Networks, Fuzzy Logic and Genetic Algorithms", Prentice Hall of India Private Limited, 2003
- 2. TimothyJ.Ross, "Fuzzylogicwith Engineering Applications", McGraw Hill, 1995
- 3. ZuradaJ.M."Introduction to Artificial Neural Systems", Jaicopublishin ghouse, 1994.
- 4. Klir.G, Yuan B.B. "Fuzzy sets and Fuzzy Logic Prentice Hall of India private limited, 1997.
- 5. Laurance Fausett, "Fundamentals of Neural Networks", Prentice hall, 1992
- 6. Gen, M. and Cheng R. "Genetic Algorithmand Engineering Design", johnwiley 1997.

# **IPS Academy, Institute of Engineering & Science** (A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum **Department of Computer Science & Engineering** Master of Engineering (M.E.) (Data Science) **II-Semester**

PSEC- MCS201(C)	Data Warehousingand Mining	3L (3 hrs.)	3 credits
--------------------	----------------------------	-------------	-----------

Prerequisite: Database Management System (Relational & Non-Relational Databases), SQL

**Course Objective:** The objective of this course is to identify the scope and essentiality of Data Warehousing and Mining and to analyze data, choose relevant models and algorithms for respective applications.

#### **Course Contents:**

#### Module 1:

Introduction to Data Mining: Definitions, KDD v/s Data Mining, DBMS v/s Data Mining, DM techniques, Mining problems, Issues and Challenges in DM, DM Application areas.

#### Module 2:

Association Rules & Clustering Techniques: Introduction, Various association algorithms like APriori, Partition, Pincer search etc., Generalized association rules, Clustering paradigms, Partitioning algorithms like K-Medioid, CLARA, CLARANS, Hierarchical clustering. DBSCAN, BIRCH, CURE, categorical clustering algorithms, STIRR, ROCK, CACTUS

#### Module 3:

Other DM techniques & Web Mining: Application of Neural Network, AI, Fuzzy logic and Genetic algorithm, Decision tree in DM, Web Mining, Web content mining, Web structure Mining, Web Usage Mining.

#### Module 4:

Temporal and spatial DM: Temporal association rules, Sequence Mining, GSP, SPADE, SPIRIT and WUM algorithms, Episode Discovery, Event prediction, Time series analysis, Spatial Mining, Spatial Mining tasks, Spatial clustering, Spatial Trends.

#### Module 5:

Data Mining of Image and Video: A case study, Image and Video representation techniques, feature extraction, motion analysis, content based image and video retrieval, clustering and association paradigm, knowledge discovery.

#### (10 hrs.)

#### (08 hrs.)

#### (06 hrs.)

# (12 hrs.)

(06 hrs.)

- 1. Understand Data Warehouse fundamentals, Data Mining Principles.
- 2. Identify appropriate data mining algorithms to solve real world problems.
- 3. Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining.
- 4. Describe complex data types with respect to spatial and web mining.
- 5. Benefit the user experiences towards research and innovation Integration.

- 1. Arun K.Pujari, "Data Mining Techniques", UniversityPress, First Edition, 2001.
- 2. Adriaans & Zantinge, "Data Mining", Pearson education, First Edition, 1996.
- 3. Berry, Linoff, "Mastering Data Mining", Wiley Publications, 2011.
- 4. Dunham, "Data Mining", Pearson education, 2003.
- 5. Konchandy Manu, "Text Mining Applications" "Cengage Learning, 2006.

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.) (Data Science)

### **II-Semester**

PSEC- MCS201(D)	Data Science	3L (3 hrs.)	3 credits
-----------------	--------------	-------------	-----------

**Prerequisite:** Data Structure, Database management System, Data Visualization

### **Course Objective:**

The objective of this course is to focus on insights about the roles of a Data Scientist.

### **Course Contents:**

#### Module 1:

Introduction to core concepts and technologies: Introduction, Terminology, data science, process, data science toolkit, Types of data, Example applications

#### Module 2:

Data collection and management: Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, using multiple data sources

#### Module 3:

Data analysis: Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT, Basic machine learning algorithms, Linear regression, SVM, Naive Bayes.

#### Module 4:

Data visualization: Introduction, Types of data visualization, Data for visualization: Data types, Data encodings, Retinal variables, mapping variables to encodings, Visual encodings.

#### Module 5:

Applications of Data Science, Technologies for visualization, Bokeh (Python), recent trends in various data collection and analysis techniques, various visualization techniques, application development methods of used in data science.

### **Course Outcome:**

- 1. Apply quantitative modeling and data analysis techniques to the solution of real world business problems
- 2. Communicate findings, and effectively present results using data visualization techniques.
- 3. Demonstrate knowledge of statistical data analysis techniques utilized in business decision making.
- 4. Illustrate data visualization types and techniques.
- 5. Describe various applications and recent trends of data science.

### (12 hrs.)

(10 hrs.)

# (06 hrs.)

#### (06 hrs.)

# (**06 hrs.**)

- 1. Field Cady, "The Data Science Handbook", 1/e, 2018, Publisher: Wiley
- 2. SinanOzdemir, "PrinciplesofDataScience", 1/e, 2016Packt Publishing Limited.
- 3. Peter Bruce, "Practical Statistics for Data Scientists: 50 Essential Concepts"
- 4. Shroff/O'Reilly; "First edition", 2017
- 5. Pang-Ning Tan, "Introduction to Data Mining", Pearson Edu.
- 6. Ricardo Baeza-YatesandBerthierRibeiro-Neto,"ModernInformation Retrieval", Pearson Education.

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.) (Data Science)

**II-Semester** 

OEC- MCS201(A)	<b>Business Analytics</b>	3L (3 hrs.)	3 credits

**Prerequisite:** Basic Computer Engineering, Microsoft Office, Business/strategy Acumen, Mathematics, Data Mining

### **Course Objective:**

The objective of this course is to focus on enable all participants to recognize, understand and apply the language, theory and models of the field of business analytics and foster an ability to critically analyze, synthesis and solve complex unstructured business problems

### **Course Contents:**

#### Module 1:

Business Analytics: Introduction to Business Analytics (BA) – Need, Components (Business Context, Technology, Data Science), Types (Descriptive, Predictive and Prescriptive), Business Intelligence versus Business Analytics, Transaction Processing v/s Analytic Processing, OLTP v/s OLAP, OLAP Operations, Data models for OLTP (ER model) and OLAP, (Star & Snowflake Schema)

#### Module 2:

Types of Digital Data: Definition, Sources, Storage and Characteristics of Structured, Unstructured and Semi Structured Data, Data Warehouse: Definition, characteristics, framework, Data lake, Business Reporting, Visual Analytics: Definition, concepts, Different types of charts and graphs, Emergence of data visualization and visual analytics

#### Module 3:

Data Mining: Concepts and applications, Data mining process, Text & Web Analytics: Text analytics and text mining overview, Text mining applications, Web mining overview, Social media analytics, Sentiment analysis overview, Big Data Analytics: Definition and characteristics of big data, Fundamentals of big data analytics

#### Module 4:

Business Performance Management: Business performance management cycle, KPI, Dashboard Analytics in Business Support Functions: Sales & Marketing Analytics, HR Analytics, Financial Analytics, Production and operations analytics

#### Module 5:

Analytics in Industries: Telecom, Retail, Healthcare, Financial Services

# (08 hrs.)

(10 hrs.)

(10 hrs.)

#### (06 hrs.)

(06 hrs.)

- 1. Understand the key issues in big data management and its associated applications in intelligent business and scientific computing and i dentify and describe complex business problems in terms of analytical models.
- 2. Acquire fundamental knowledge of data warehousing and enabling techniques and scalable algorithms like Hadoop, Map Reduce and NO SQL in big data analytics.
- 3. Understand basic concepts of data mining, web mining and big data analytics.
- 4. Translate results of business analytic projects into effective courses of action and Demonstrate ethical decision-making in structured or unstructured and ambiguous situations.
- 5. Exhibit effective collaboration and leadership skills and Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc.

- 1. RameshSharda, DursunDelen, EfraimTurban, "Business Intelligence: AManagerial Perspective on Analytics", Pearson 3 rd edition.
- 2. R.N.PrasadandSeema Acharya, "FundamentalsofBusiness Analytics", Wiley, 2016.
- 3. U.DineshKumar, "BusinessAnalytics-TheScienceofDataDrivenDecisionMaking", Wiley 2017.
- 4. Anil Maheshwari, "Data Analytics", McGraw Hill, 2017.
- 5. Jesper Thorlund & Gert H.N. Laursen, "Business Analytics for Managers: Taking Business Intelligence Beyond", Wiley Latest.
- 6. Sahil Raj, "Business Analytics", Cengage, Latest.
- 7. James R. Evans, "Business Analytics", Pearson, Latest

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# **Department of Computer Science & Engineering**

Master of Engineering (M.E.)

(Data Science)

**II-Semester** 

OEC- MCS201(B)	<b>Operation Research</b>	3L (3 hrs.)	3 credits
----------------	---------------------------	-------------	-----------

Prerequisite: Mathematics, Computer Programming

#### **Course Objective:**

To impart knowledge in concepts and tools of Operations Research and to understand mathematical models used in Operations Research and to apply these techniques constructively to make effective business decisions.

### **Course Contents:**

#### Module 1:

Introduction to Operations Research: Introduction, Historical Background, Scope of Operations Research, Features of Operations Research, Phases of Operations Research, Types of Operations Research Models, Operations Research Methodology, Operations Research Techniques and Tools, Structure of the Mathematical Model, Limitations of Operations Research

#### Module 2:

Linear Programming: Introduction, Linear Programming Problem, Requirements of LPP, Mathematical Formulation of LPP, Case Studies of LPP, Graphical Methods to Solve Linear Programming Problems, Applications, Advantages, Limitations

#### Module 3:

Graphical Analysis of Linear Programming Problems: Introduction, Graphical Ana lysis, Some Basic Definitions, Graphical Methods to Solve LPP, Some Exceptional Cases, Important Geometric Properties of LPP

#### Module 4:

Simplex Method: Introduction, Standard Form of LPP, Fundamental theorem of LPP, Solution of LPP – Simplex Method, The Simplex Algorithm, Penalty C ost Method or Big M-method, Two-Phase Method, Solved Problems on Minimization, Duality in Linear Programming Problem: Introduction, Importance of Duality Concepts, Formulation of Dual Problem, Economic Interpretation of Duality, Sensitivity Analysis

#### Module 5:

Study of various problems: Transportation Problems: Introduction, Formulation of Transportation Problem (TP), Transportation Algorithm (MODI Method), The initial basic feasible solution, moving towards Optimality, Assignment Problem, Hungarian methods, Routing problem, Travelling salesman problem, Integer Programming Problem., All IPP Algorithm, Branch and Bound Technique.

#### (**12 hrs.**) s of LP

(06 hrs.)

#### (10 hrs.)

(08 hrs.)

#### (04 hrs.)

- 1. Solve Linear ProgrammingProblems
- 2. Solve Transportation and Assignment Problems
- 3. Understandtheusageofgametheoryand Simulationfor Solving Business Problems
- 4. Learn various concepts for research and operations.
- 5. Study various realworld problems.

- 1. Taha H.A., "Operations Research", Sixth Edition, Prentice Hall of India, 2003.
- 2. Shennoy G.V. and Srivastava U.K., "Operation Research for Management", Wiley Eastern, 1994.
- 3. Bazara M.J., Jarvis and Sherali H., "Linear Programming and Network Flows", John Wiley, 1990.
- 4. Philip D.T. and Ravindran A., "Operations Research", John Wiley, 1992.
- 5. Hillier and Libeberman, "Operations Research", Holden Day, 1986
- 6. Budnick F.S., "Principles of Operations Research for Management", Richard D Irwin, 1990.
- 7. Tulsian and Pasdey V., "Quantitative Techniques", Pearson Asia, 2002.

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# **Department of Computer Science & Engineering**

# Master of Engineering (M.E.)

(Data Science)

**II-Semester** 

OEC-MCS201(D)Voice and Data Network3L (3 hrs.)3 credits	OEC-MCS201(D)	Voice and Data Network	3L (3 hrs.)	3 credits
---	---------------	------------------------	-------------	-----------

**Prerequisite:** Computer Network and Data Communication

## **Course Objective:**

The objective of this course is to build an understanding of the fundamental concepts of computer networking and familiarize the student with the basic taxonomy and terminology of the computer networking area.

## **Course Contents:**

#### Module 1:

Network Design Issues, Network Performance Issues, Network Terminology, centralized and distributed approaches for networks design, Issues in design of voice and data networks.

#### Module 2:

Layered and Layer less Communication, Cross layer design of Networks, Voice Networks (wired and wireless) and Switching, Circuit Switching and Packet Switching, Statistical Multiplexing.

### Module 3:

Data Networks and their Design, Link layer design- Link adaptation, Link Layer Protocols, Retransmission. Mechanisms (ARQ), Hybrid ARQ (HARQ), Go Back N, Selective Repeat protocols and their analysis.

### Module 4:

Queuing Models of Networks, Traffic Models, Little's Theorem, Markov chains, M/M/1 and other Markov systems, Multiple Access Protocols, Aloha System, Carrier Sensing, Examples of Local area networks.

### Module 5:

Inter-networking, Bridging, Global Internet, IP protocol and addressing, Sub netting, Classless Inter domain Routing (CIDR), IP address lookup, Routing in Internet, End to End Protocols, TCP and UDP, Congestion Control, Additive Increase/Multiplicative Decrease, Slow Start, Fast Retransmit/ Fast Recovery, Congestion avoidance, RED TCP Throughput Analysis, Quality of Service in Packet Networks. Network Calculus, Packet Scheduling Algorithms.

## **Course Outcome:**

- 1. Protocol, algorithms, trade-offs rationale.
- 2. Routing, transport, DNSresolutions
- 3. Network extensions and next generation architectures.
- 4. Independently understand basic computer network technology.
- 5. Understandand explain Data Communications Systemand its components.

### (06 hrs.)

#### (06 hrs.)

(10 hrs.)

#### (08 hrs.)

# (10 hrs.)

- 1. D. Bertsekas and R. Gallager, "Data Networks", 2nd Edition, Prentice Hall, 1992.
- 2. L. Peterson and B. S. Davie, "Computer Networks: A Systems Approach", 5th Edition, Morgan Kaufman, 2011.
- 3. Kumar, D. Manjunath and J. Kuri, "Communication Networking: An analytical approach", 1st Edition, Morgan Kaufman, 2004.
- 4. Walrand, "Communications Network: A First Course", 2nd Edition, McGraw Hill, 2002.
- 5. Leonard Kleinrock, "Queueing Systems", Volume I: Theory", 1st Edition, John Wiley and Sons, 1975.
- 6. Aaron Kershenbaum, "Telecommunication Network Design Algorithms", McGraw Hill, 1993.
- 7. Vijay Ahuja, "Design and Analysis of Computer Communication Networks", McGraw Hill, 1987

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.) (Data Science) II-Semester

LCMCS201Machine Learning Lab4P (4 hrs.)2 credits	
--	--

#### Prerequisite: Artificial Intelligence

#### **Course Objective:**

Implement the basic concepts and techniques of Machine Learning algorithms.

#### **List of Experiments:**

#### Week 1:

1. Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a

.CSV file.

#### Week 2:

1. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.

#### Week 3:

1. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.

#### Week 4:

1. Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.

#### Week 5:

1. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.

#### Week 6:

1. Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set.

#### Week 7:

1. Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Java/Python ML library classes/API.

### Week 8:

1. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program.

#### Week 9:

1. Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions. Java/Python ML library classes can be used for this problem.

#### Week 10:

1. Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs.

### **Course Outcome:**

- 1. Apply knowledge of computing and mathematics to machine learning problems, models and algorithms;
- 2. Analyzeaproblemandidentifythecomputingrequirementsappropriateforits solution;
- 3. Design, implement, and evaluate an algorithm to meet desired needs
- 4. Apply mathematical foundations, algorithmic principles and computer science theory to the modeling and design of computer-based systems.
- 5. Demonstrates comprehension of the trade -offs involved in design choices.

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.) (Data Science)

# II-Semester

LC-MCS202 Data Visualization Lab	4P (4hrs)	2 credits
----------------------------------	-----------	-----------

Prerequisite: Computer Graphics, AI, Mathematics: basic linear algebra, calculus, probability

### **Course Objective:**

To learn the process involved in data visualization with security aspects.

## List of Experiments:

#### Week 1:

- 1. Introduction to data Visualizationtools.
- 2. Data preprocessing methods on student and labor datasets: Data preprocessing methods on student and labor datasets implement data cube for data warehouse on 3 dimensional data.

#### Week 2:

- 1. Basic and Specialized Visualization Tools
- 2. Data Cleaning: Implement various missing handling mechanisms ,Implement various noisy handling mechanisms

#### Week 3:

**1. Exploratory Analysis:** Develop k-means and MST based clustering techniques: Develop kmeans and MST based clustering techniques, Develop the methodology for assessment of clusters for given dataset.

#### Week 4:

1. Association Analysis: Design algorithms for association rule mining algorithms.

#### Week 5:

**1. Hypothesis Generation:** Derive the hypothesis for association rules to discovery of strong association rules; Use confidence and support thresholds.

#### Week 6:

1. **Transformation Techniques:** Construct Haar wavelet transformation for numerical data, Construct principal component analysis (PCA) for 5-dimensional data.

#### Week 7:

**1. Data Visualization:** Implement binning visualizations for any real time dataset, Implement linear regression techniques.

#### Week 8:

**1. Clusters Assessment:** Visualize the clusters for any synthetic dataset, Implement the program for converting the clusters into histograms

#### Week 9:

**1. Hierarchical Clustering:** Write a program to implement agglomerative clustering technique, Write a program to implement divisive hierarchical clustering technique

#### Week 10:

**1. Scalability Algorithms**: Develop scalable clustering algorithms, Develop scalable apriori algorithm.

#### **Course Outcome:**

- 1. State the core concepts of Data Visualization.
- 2. Illustrate various concepts of visual representations.
- 3. Illustrate various techniques and classification of visualization systems.
- 4. Statethecoreconceptsofnetwork, distributed and multiprocessor operating system.

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.)

(Data Science)

**II-Semester** 

AUD-MCS201(B)	<b>Business Communication</b>	2P (2hrs)	0 credits
---------------	-------------------------------	-----------	-----------

# **Prerequisite:**

# **Course Objectives:**

The primary objective of this course is to exchange ideas, opinions, thoughts, beliefs and information between human beings. In the fields of business, the main objective of communication is to help managers in performing all managerial functions smoothly.

## **Course Contents:**

#### Module-I

#### **Communication:** It's Interpretation

Basics of Communication: Process, Components and Factors of Communication. Types of Communication, Global Aspects, Ethical Aspects, Legal Aspects, Gender Communication, Communication in Organization, Communication during crisis

#### **Module-II**

#### **Communication Core:**

Difficult Communication, Communication Skills for Team and Leadership Effectiveness, Managerial Communication,

Organizational Communication, Persuasive Communication, Strategic Negotiation Skills for Leaders, Intercultural Communication Competence, Cultural studies

#### Module- III

#### **Communication and Media**

Media and Society: The Economics, Politics, Ethics, and Technologies of Mass Communications, Managerial Communication, Strategic Communication in the Digital Era, Technology enabled change

### Module- IV

#### **Business Correspondence**

Writing Skills in business and public administration in India: Effective writing, Job Application, Bio Data, Curriculum Vitae, Personal Resume, Notice, Agenda & Minutes of meetings, Memorandum, Technical Reports: Types & Structure

#### Module-V

#### **Presentation Skills and Public Speaking**

Elements of Professional Presentation, Glossophobia and Lack of Confidence, Tips for Effective Presentation, Factors affecting Public Speaking, Plan and Prepare, Practice and Perform, Presentation of student with the help of audio-visual aids (Presentation prepared by the students will be evaluated)

### **Course Outcomes:**

Program learning outcomes define the knowledge, skills, and abilities students are expected to demonstrate upon completion of an academic program. These learning outcomes are regularly assessed to determine student learning and to evaluate overall program effectiveness.

# **Text Books:**

- 1. Professional Communication by Aruna Koneru, Tata McGraw-Hill Publishing Company Limited, New Delhi,2005
- 2. Effective Technical Communication by M. Ashraf Rizvi, Tata McGraw-Hill Publishing Company Limited, New Delhi,2005
- 3. Communication Skill for Engineers and Scientist by Sangeeta Sharma and Vinod Mishra, PHI Learning, New Delhi, 2015
- 4. Business Communication by Dr. V.G. Sadh, Thakur Publications, Lucknow, 2013
- 5. Business Correspondence and Report Writing by R.C. Sharma and Krishna Mohan, Tata McGraw-Hill Publishing Company Limited, New Delhi,2008

# **Reference Books:**

- 1. Bonet, Diana. The Business of Listening: Third Edition. New Delhi: Viva Books, 2004.
- 2. Bovee, Courtland L, John V. Thill & Barbara E. Schatzman. Business Communication Today: Tenth Edition. New Jersey: Prentice Hall, 2010
- 3. Guffey, Mary Ellen. Essentials of Business Writing. Ohio: SouthWestern College Pubg., 2000.
- 4. Hughes, Shirley. Professional Presentations: A Practical Guide to the Preparation and Performance of Successful Business Presentations. Sydney: McGraw-Hill, 1990
- 5. Monippally, Matthukutty, M. Business Communication Strategies. New Delhi: Tata McGraw-Hill Publishing Company Ltd., 2001.

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.)

(Data Science)

# **II-Semester**

AUD-MCS201(A) Soft Skills & Interperse Communication	onal 2P (2hrs)	0 credits
---	----------------	-----------

# **Prerequisite:**

# **Course Objectives:**

The course objective is to create awareness among the students for the professional world in which the role of individuals as team players and also as responsible leaders materializes to a great extent.

# **Course Contents:**

### Module-I

Introduction to Soft Skills, Aspects of Soft Skills, Effective Communication Skills, Classification of Communication, Personality Development, Positive Thinking, Telephonic Communication Skills

#### Module II

Communicating without Words, Paralanguage Proxemics, Haptics: The Language of Touch, Metacommunication, Listening Skills, Types of Listening, Communication Breakdown

#### Module - III

Negotiation Skills, Culture as Communication, Organizational Communication, Leadership skills, Team Management, Group Dynamics, Negotiation, Assertiveness, Emotional Intelligence, Persuasion, Meeting Management, Adaptability & Work Ethics

#### Module - IV

Advanced Writing Skills, Principles of Business Writing, Business Letters, Business Letters: Format and Style, Types of Business Letter, Business Letters, Business Letters: Format and Style, Types of Business Letter, Writing Reports, Types of Report, Strategies for Report Writing, Evaluation and Organization of Data

### Module –V

Advanced Speaking Skills, Oral Presentation (Analysis of the Recorded Video of the Presentation), Speeches & Debates, Combating Nervousness, Patterns & Methods of Presentation, Oral Presentation: Planning & Preparation Making Effective Presentations, Speeches: Listening speech audio of Eminent Personalities (Martin Luthar King :I Have a Dream, Steve Jobs' Standford's Commencement Speech: Stay Hungry Stay Foolish, Swami Vivekand : Chicago Various Occasions), Video Interviews & Mock Interviews Planning & Preparation

## **Course Outcomes:**

The course, with its interactive and need based modules, will develop an ability to face various challenges of communication as well as behavioral skills faced by individuals at workplace and organizations in bridging the gaps through effective skills of interviews, group discussions, meeting management, presentations and nuances of drafting various business documents for sustainability in today's global world.

# **Text Books:**

- 1. Butterfield, Jeff. Soft Skills for Everyone. New Delhi: Cengage Learning. 2010.
- 2. Soft Skills by G.S. Chouhan and Sangeeta Sharma, Wiley, New Delhi, 2016 Communication Skills by Sanjay Kumar and Pushplata, OUP, New Delhi, 2011
- 3. Soft Skill for Everyone by Jeff Butterfield, Cengage Leaning, New Delhi, 2010
- 4. Communication Skill for Engineers and Scientist by Sangeeta Sharma and Vinod Mishra, PHI Learning, New Delhi, 2015
- 5. Developing Communication Skill by Krishna Mohan, Meera Banerji, McMillan India Limited,2018
- 6. Thorpe, Edgar and Showick Thorpe. Winning at Interviews. Pearson Education. 2004.

## **Reference Books:**

- 1. Theories of Personality by Hall, Calvin S. et al. . NewDelhi: Wiley. rpt. 2011.
- 2. Corporate Conversations by Holtz, Shel. New Delhi: PHI. 2007.
- 3. The Art of Public Speaking by Lucas, Stephen E. McGraw-Hill Book Co. International Edition, 11th Ed. 2014.
- 4. Winning at Interviews by Thorpe, Edgar and Showick Thorpe. Pearson Education. 2004.
- 5. Business Communication for Managers by Penrose, John M., et al. New Delhi: Thomson South Western. 2007

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme Based on AICTE Flexible Curriculum

# Department of Computer Science & Engineering Master of Engineering (M.E.)

(Data Science)

**II-Semester** 

|--|

## **Prerequisite:**

# **Course Objectives:**

To enable participants to communicate clearly and with impact, by improving their verbal and non-verbal communication style, as well as enhancing interpersonal skills

# **Course Contents:**

### Module-I

### Nature of Technical Communication

Communication as sharing, stages of Communication, Channels of Communication, Process of Communication, Technical communication skill- LSRW skills, Organization in Technical Communication: Introduction, Spatial Organization, Chronological organization, General to specific organization, Empathetic organization

#### Module-II

### Listening Process

Listening process, Difference between hearing and listening, Types of listening: Superficial, Appreciative, Focused, Evaluative, Attentive and Empathetic. Barriers to Listening and ways to overcome, Active Listening practices, Effective Listening Strategies

### Module-III

### **Speaking Strategies**

The speech process, Art of conversations and oral skills, Types of Conversation: Formal and Informal, Strategies for good conversation- Improving fluency, articulation, voice quality, pronunciation, accent and intonation, body language

### Module-IV

#### **Reading Skills**

Introduction of Reading, Reading Process, Types of Reading: Extensive & Intensive: Analytical, Study, General, Idea & Exploratory, Reading Techniques: Skimming and Scanning, Benefits of Effective Reading Barriers to Reading: Sub-Vocalization, Reading Word by Word, Hard Gazing, Regression, Reading Speed: Slow, Average, Fast & Very Fast. Reading Comprehension, Reading different types of text by the students (Read translating text: Hindi to English, English to Hindi)

#### Module-V

### Writing Strategies

Mechanics of Writing, Planning, Drafting & Editing, Right Word Selection, Effective Sentence Structure, Developing Logical Paragraph with Analysis, Aristotle's Technique of Analysis, Art of Condensation, Précis Writing, Practice through selected text

## **Course Outcomes:**

Students will develop knowledge, skills, and judgment around human communication that facilitate their ability to work collaboratively with others. Such skills could include communication competencies such as managing conflict, understanding small group processes, active listening, appropriate self-

# **Text Books**

- 1. Developing Communication Skill by Krishna Mohan, Meera Banerji, McMillan India Limited,2018
- 2. Dale Jungk, Applied Writing for Technicians, McGraw Hill, New York, 2004. (ISBN: 07828357-4)
- 3. Reading Comprehension: How to Drastically Improve your Reading Comprehension Fast by Clarence T. Rivers, J. R. Kindle Publishing, 2014.
- 4. English for Effective Communication by Sanjay Kumar and Pushp Lata: Oxford University Press, New Delhi 2015.
- 5. English Language Skills by Aruna Koneru: McGraw Hills Education (India) Pvt. Ltd., New Delhi, 2015.

# **Reference Books**

- 1. Adair, John. Effective Communication. London: Pan Macmillan Ltd., 2003.
- 2. Collins, Patrick. Speak with Power and Confidence. New York: Sterling, 2009
- 3. Hasson, Gill. Brilliant Communication Skills. Great Britain: Pearsonθ Education, 2012.
- 4. Lesikar, Raymond V and Marie E. Flatley. Basic Business Communication: Skills for Empowering the Internet Generation: Ninth Edition. New Delhi: Tata McGraw-Hill, 2002.
- 5. Rutherford, Andrea J. Basic Communication Skills for Technology: Second Edition. Delhi: Pearson Education, 2007