**MCA-41 PYTHON PROGRAMMING**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

**Note:-** Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives**

1. Basics of Python programming
2. Decision Making and Functions in Python
3. Object Oriented Programming using Python
4. Files Handling in Python
5. GUI Programming and Databases operations in Python
6. Network Programming in Python

**Courses Learning Outcomes:**

1. Describe the Numbers, Math functions, Strings, List, Tuples and Dictionaries in Python
2. Express different Decision Making statements and Functions
3. Interpret Object oriented programming in Python
4. Understand and summarize different File handling operations
5. Design and develop Client Server network applications using Python

**UNIT-I**

Introduction Installing And Working With Python, Tokens, Operators, Data Types, Sequence Types, Mapping Types: Dictionaries, Tuples. Control Structures: Conditional Branching, Looping, Custom Functions: Names And Doc string, Argument And Parameter Unpacking.

**UNIT-II**

Sequences, iteration and recursion. Modules And Packages Modules And Packages, Overview Of Python Standard Library. OOPS Concepts And Classes In Python,.

**UNIT-III**

Exception handling: catching and raising exceptions, custom exceptions. File Handling and Writing and Reading, Binary Data, Writing and Parsing Text Files, Writing and Parsing XML Files, Random Access Binary Files.

**UNIT-IV**

Database programming. Numpy basics: Introduction, data types, arrays: indexing, slicing, shape, iteration, join, split, search, sort, filter.

**Text/Reference Books:**

1. Mark Summerfield, “Programming in python
2. A Complete Introduction to Python Programming”.
3. Mark Lutz, “Learning Python”, O Reily, 4th Edition, 2009.
4. Brian K. Jones, “Python Cookbook”.
5. Alex Martelli, “Python in a nutshell”.
6. Tim Hall and J-P Stacey, “Python 3 for Absolute Beginners”, 2009.

**MCA-42 R PROGRAMMING**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

**Note:-** Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives**

1. Understanding of R System and installation and configuration of R-Environment and R-Studio.
2. Understanding R Packages, their installation and management.
3. Understanding of nuts and bolts of R.
4. Application of R Programming in Daily life problems
5. Visualizing data using R with different type of graphs and charts
6. Applying R Advance features to solve complex problems and fine tuning R Processes

**Course Learning Outcomes:**

1. Understand the basics in R programming in terms of constructs, control statements, string functions
2. Understand the use of R for Big Data analytics
3. Learn to apply R programming for Text processing
4. Able to appreciate and apply the R programming from a statistical perspective

**Unit I**

Introduction to R, variables, data types, comments, operators, decision making, loops.

**Unit II**

Functions, Mathematical functions, R Data Structures: String, functions on string, Vectors, working with vectors, Lists, Arrays, R Matrices, Data Frames, Factors.

**Unit III**

R Graphics: R Plot, R Line, R Scatterplot, R Pie Charts, R Bars

**Unit IV**

R Statistics: Introduction, R Data Set, R Max and Min, R Mean Median Mode, R Percentiles.

**Text/References Books:**

1. Introduction to Statistics and Data Analysis  - With Exercises, Solutions and Applications in R  By Christian Heumann, Michael Schomaker and Shalabh, Springer, 2016
2. The R Software-Fundamentals of Programming and Statistical Analysis -Pierre Lafaye de Micheaux, Rémy Drouilhet, Benoit Liquet, Springer 2013
3. A Beginner's Guide to R (Use R) By Alain F. Zuur, Elena N. Ieno, Erik H.W.G. Meesters, Springer 2009

**MCA-43(i) SOFT COMPUTING**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

**Note:-** Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives:**

1. Study the needs and applications of soft computing.
2. To learn how natural and biological systems influence the computational field.
3. Provide knowledge and applications of Neural Network, Fuzzy Logic , Genetic Algorithms, Natural Computing.

**Course Learning Outcomes:**

1. Have an in-depth understanding of some of the soft computing techniques.
2. Identify the situations for which it is beneficial to apply soft computing techniques.
3. Describe and Apply suitable soft computing techniques for the problems which could not be otherwise solved efficiently.
4. Be able to understand how large numbers of agents can self-organize and adapt.

**UNIT I**

Introduction to Soft Computing- Introduction, Artificial Intelligence, Artificial Neural Network, Fuzzy System, Genetic Algorithms, Swarm Intelligent Systems, Expert System.

**UNIT II**

Artificial Neural Network –First Generation – Biological Inspiration, ANNs and Classification of ANNs. Perceptron Network, Adaline Network, Madaline Network.

Artificial Neural Network –Second Generation- Introduction, Back propagation Neural Network, Hamming Neural Network, Bi-directional Associative memory.

**UNIT III**

Artificial Neural Network –Third Generation**-** Introduction**,** Spikes Neuron Models

Fuzzy Logic- Probability Theory vs Possibility Theory, Fuzzy Set and Operations, Applications

**UNIT IV**

Genetic Algorithm- Procedures, Applications.

Swarm Intelligence- Background of SIS, Ant Colony System and its working, Particle swarm intelligence system, Artificial Bee Colony System, Cuckoo search algorithm.

**Text/Reference Books:**

1. Soft Computing with Matlab Programming- N.P. Padhy, S.P. Simon, Oxford Publications
2. Zbigniew Michalewicz, Genetic algorithms +Data Structures = Evolution Programs, Springers-Verlag.
3. M. Mitchell, An Introduction to Genetic Algorithms, Prentice-Hall.
4. S. Rajasekaran & G. A. Vijayalakshmi Pai, Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis & Applications, PHI.
5. Simon O. Haykin, Neural Networks, A Comprehensive Foundation, PHI.

**MCA-43(ii) MACHINE LEARNING**

**L/T - 4 Total Credits – 4 Internal Marks 30**

**External Marks 70**

**Notes: -** Total 09 questions are to bet set by the examiner. First question will be compulsory consisting of 5 short answer type question (each carry 2 marks) covering the entire syllabus uniformly. In addition, 08 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A Candidate is required to attempt five questions in all selecting one from each unit including the compulsory question.

**Course Learning Objectives:**

1. Be able to understand and apply supervised and unsupervised learning algorithms.
2. Understand the fundamental concepts in machine learning and popular machine learning algorithms.
3. Understand the basic concept of Deep Learning.
4. Be able to solve the problems related to the application of machine learning algorithms with programming.

**Course Learning Outcomes:**

1. Recognize major programming languages.
2. Identify potential applications of machine learning in practice.
3. Select the suitable machine learning tasks for given application.
4. Implement feature extraction and selection to represent data as features to serve as input to machine learning models.

**Unit 1**

Introduction to Machine Learning, Artificial Intelligence and Machine Learning, types of Machine Learning, key Elements, applications, model and cost functions, Python basics, vector and matrices.

**Unit II**

NumPy, Pandas, Data Processing, Supervised Learning, Regression: Linear Regression, Multivariable Regression, Logistic Regression. Classification: Introduction, classification methods.

**Unit III**

Support Vector Machine. Unsupervised Learning, Clustering: K-Means, forecasting. Optimization: Introduction, Gradient Based methods, derivative-free optimization.

**Unit IV**

Neural Network: Introduction, Deep Neural Network, Convolutional Neural Network, Recurrent Neural Network, Graph Neural Network. Ensemble Learning, Reinforcement Learning, Deep Learning.

**Text/Reference Books:**

1. Machine Learning using Python. Manaranjan Pradhan, U Dinesh Kumar. Wiley. 2019
2. Machine Learning. Tom Mitchell. First Edition. McGraw- Hill, 1997
3. Machine Learning. Anuradha Srinivasaraghavan, Vincy JOdeph. Wiley. 2019
4. Deep Learning Ian Goodfellow, Yoshua Bengio, Aaron Courville. MIT Press. 2016.
5. Introduction to machine learning, E. Alpaydin. MIT Press, 2e, 2009
6. Machine Learning in Action, P. Harrington. Manning Publication. 2012.
7. Machine Learning and Pattern Recognition. C.M. Bishop. Springer.

**MCA-43(iii) GENETIC ALGORITHM**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

**Note:-**Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives**

1. A gentile introduction to evolutionary algorithm, its working principles.
2. Mathematical underpinnings, its extensions to solve different kinds of optimization problems including nonlinear constraint handling.
3. Multi-objective optimization and scheduling problems will be introduced.

**Course Learning Outcomes:**

1. Explain the principles underlying Evolutionary Computation in general and Genetic Algorithms in particular.
2. Apply Evolutionary Computation Methods to find solutions to complex problems.
3. Analyze and experiment with parameter choices in the use of Evolutionary Computation.
4. Summarize current research in Genetic Algorithms and Evolutionary Computing.

**Unit I**

**INTRODUCTION:** Introduction to Evolutionary Algorithm, History of Evolutionary Algorithm, Advantage of evolutionary algorithm, Application of evolutionary algorithm, Biological and AI , Introduction of Genetic Algorithm, Difference between traditional approach and Evolutionary algorithm.

**Unit II**

**Genetic modelling**: Basic Terminologies and operators like individual, gene, alleles, phenotype and fitness function. Simple genetic algorithm, its representation, Advantage and Limitation of genetic algorithm.

**Unit III**

**Operators of GA:** Selection: Roullete wheel selection, random, rank, tournament, Boltzmann selection, Crossover: Singlepoint, twopoint, multipoint, ordered, uniform, Mutation: Flipping, Interchanging, reversing, replacement and crossover rate, mutation rate, convergence criteria.

**Unit IV**

**Theoretical Analysis of Evolutionary Algorithms**: Diploid, Dominance and abeyance, inversion and reordering operator, fitness scaling, Niching and Speciation.

**Text/Reference Books:**

1. D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Pearson Education Asia.
2. M. Mitchell, *An introduction to genetic algorithms*, MIT Press.
3. L. D. Davis, *Evolutionary algorithms*, Springer-Verlag.
4. K. Srinivasa Raju and D. Nagesh Kumar, ***M****ulti-criterion Analysis in Engineering and Managemet,*PHI Learning Pvt. Ltd., New Delhi, India.
5. S.N. Sivanandam, S.N. Deepa, *Introduction to Genetic Algorithms*, Springer.
6. Eiben And Smith, *Introduction To Evolutionary Computing*, Springer

**MCA-44(i) DATA WAREHOUSING AND DATA MINING**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

**Note:-** Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives**

1. To identify the scope and essentiality of Data Warehousing and Mining.
2. To analyze data, choose relevant models and algorithms for respective applications.
3. To study spatial and web data mining.
4. To develop research interest towards advances in data mining.

**Course Learning Outcomes:**

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

**Unit I**

Data Mining: Introduction: Motivation, Importance, data mining, kind of data, Functionalities, interesting patterns, classification of data mining system, Major issues, Data Mining Primitives. Data Preparation: Preprocess, Data cleaning, Data Integration and transformation, Data reduction, Discritization and concept hierarchy generation.

**Unit II**

Data warehouse and OLAP Technology for data mining: data warehouse, difference between operational data base systems and data warehouse, A Multidimensional Data Model, Architecture, Implementation, data warehousing to data mining, Data warehouse usage.

**Unit III**

Association Rule Mining: Mining single-dimensional Boolean association rules from transactional databases, mining multilevel association rules from transaction databases, Mining multidimensional association rules from relational databases and data warehouses, From association mining to correlation analysis, constraint-based association Mining.

**Unit IV**

Classification and prediction, issues, classification by decision induction, Bayesian classification, classification by back propagation, classification based on concepts from association rule mining other classification methods .Cluster Analysis: What is Cluster Analysis, Types of Data in Cluster Analysis, Applications and Trends in Data Mining.

**Text/Reference Books:**

1. Ale Berson, Stephen Smith, Korth Theorling, *Data Mining*,TMH.
2. Adruaans, Longman, Addison-wesley *Data Mining,*
3. Addison-wesley Longman, *Data Warehousing in the Real World.*
4. Chanchal Singh, *Data Mining and Warehousing*, Wiley.
5. John E, Herbert P, *Data Mining*.

**MCA-44(ii) BIG DATA ANALYTICS**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

Note:-Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives:**

1. To provide an overview of an exciting growing field of big data analytics.
2. To introduce the tools required to manage and analyze big data like Hadoop, NoSql MapReduce.
3. To teach the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.
4. To enable students to have skills that will help them to solve complex real-world problems in for decision support.

**Course Learning Outcomes:**

1. Understand the key issues in big data management and its associated applications in intelligent business and scientific computing.
2. Acquire fundamental enabling techniques and scalable algorithms like Hadoop, Map Reduce and NO SQL in big data analytics.
3. Interpret business models and scientific computing paradigms, and apply software tools for big data analytics.
4. Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc.

**UNIT I**

Introduction, Wholeness of Big Data, Big Data Sources and Applications, Big Data Architecture.

**UNIT II**

Distributed Computing Using Hadoop, Parallel Processing with Map Reduce, Application and Programming. NoSQL Databases, Big Data Programming Languages – Apache Hive, Apache Pig.

**UNIT III**

BIG DATA PRIVACY, ETHICS AND SECURITY - Privacy – Reidentification of Anonymous People – Why Big Data Privacy is self regulating? , Ethics – Ownership – Ethical Guidelines – Big Data Security – Organizational Security

**UNIT IV**

SECURITY, COMPLIANCE, AUDITING, AND PROTECTION- Steps to secure big data – Classifying Data – Protecting – Big Data Compliance – Intellectual Property Challenge – Research Questions in Cloud Security – Open Problems, HADOOP SECURITY DESIGN Kerberos – Default Hadoop Model without security - Hadoop Kerberos Security Implementation & Configuration, DATA SECURITY & EVENT LOGGING Integrating Hadoop with Enterprise Security Systems - Securing Sensitive Data in Hadoop – SIEM system – Setting up audit logging in Hadoop cluster.

**Text/Reference Books:**

1. Big Data- Anil Maheshwari, McGraw hill.
2. Frank Ohlhorst John Wiley & Sons, “Big Data Analytics: Turning Big Data into Big Money”, John Wiley & Sons, 2013.
3. Sherif Sakr, “Large Scale and Big Data: Processing and Management”, CRC Press, 2014.
4. Sudeesh Narayanan, “Securing Hadoop”, Packt Publishing, 2013.
5. Ben Spivey, Joey Echeverria, “Hadoop Security Protecting Your Big Data Problem”, O’Reilly Media, 2015.
6. Boris Lublinsky, Kevin t. Smith, Alexey Yakubovich, “Professional Hadoop Solutions”, Wiley, ISBN: 9788126551071, 2015.
7. Chris Eaton, Dirk Deroos et al. , “Understanding Big data ”, McGraw Hill, 2012. 3. Tom White, “HADOOP: The definitive Guide” , O Reilly 2012

**MCA-45(iii) DATA SCIENCE**

**L/T - 4 Total Credits – 4 Internal - 30 Marks External – 70 Marks**

Note:-Total 09 Questions are to be set by the examiner. First question will be compulsory consisting of 5 short (each 2 marks) questions covering entire syllabus uniformly. In addition 8 more questions will be set unit wise comprising 2 questions from each unit of the given syllabus. A candidate is required to attempt five questions in all selecting one question from each unit including the compulsory question.

**Course Learning Objectives**:

1. To know the fundamental concepts of data science and analytics
2. To learn various techniques for mining data streams
3. To learn event modelling for different applications.
4. To know about Hadoop and Map Reduce procedure

**Course Learning Outcomes**:

1. Work with big data platform and its analysis techniques.
2. Design efficient algorithms for mining the data from large volumes.
3. Model a framework for Human Activity Recognition.
4. Development with cloud databases

**UNIT I**

INTRODUCTION TO DATA SCIENCE – Applications - Data Science Process – Exploratory Data analysis – Collection of data – Graphical presentation of data – Classification of data – Storage and retrieval of data – Big data – Challenges of Conventional Systems - Web Data – Evolution Of Analytic Scalability - Analytic Processes and Tools - Analysis vs Reporting - Modern Data Analytic Tools - Statistical Concepts: Sampling Distributions - Re-Sampling - Statistical Inference - Prediction Error.

**UNIT II**

DATA ANALYSIS: Correlation – Regression – Probability – Conditional Probability – Random Variables – Analysis using Mean, Median, Mode, Standard Deviation, Skewness, Kurtosis- Regression Modeling - Multivariate Analysis - Bayesian Modeling - Inference and Bayesian Networks - Support Vector and Kernel Methods - Analysis of Time Series: Linear Systems Analysis - Nonlinear Dynamics –

**UNIT III**

DATA MINING TECHNIQUES: Rule Induction - Neural Networks: Learning and Generalization - Competitive Learning - Principal Component Analysis and Neural Networks - Fuzzy Logic: Extracting Fuzzy Models from Data - Fuzzy Decision Trees - Stochastic Search Methods- Neuro-Fuzzy Modelling – Association rule mining – Clustering – Outlier Analysis – Sequential Pattern Mining – Temporal mining – Spatial mining – Web mining.

**UNIT IV**

FRAMEWORKS AND VISUALIZATION: Map Reduce – Hadoop, Hive, MapR – Sharding – NoSQL Databases – Cloud databases - S3 - Hadoop Distributed File Systems – Visualizations - Visual Data Analysis Techniques - Interaction Techniques – Social Network Analysis – Collective Inferencing – Egonets - Systems and Applications.

**Text/Reference Books**

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. Bill Franks, “Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics”, John Wiley & sons, 2012.
4. Jiawei Han, Micheline Kamber “Data Mining Concepts and Techniques”, Second Edition, Elsevier, Reprinted 2008.
5. Rachel Schutt, Cathy O'Neil, “Doing Data Science”, O'Reilly Publishers, 2013.
6. Foster Provost, Tom Fawcet, “Data Science for Business”, O'Reilly Publishers, 2013.
7. Bart Baesens, “Analytics in a Big Data World: The Essential Guide to Data Science and its Applications“, Wiley Publishers, 2014.
8. S. N. Sivanandam, S. N Deepa, “Introduction to Neural Networks Using Matlab 6.0”, Tata McGraw- Hill Education, 2006.