ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV
MCACC 401- Artificial Intelligence

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To learn and understand an overview of artificial intelligence (AI) principles and approaches.
2. To get an overview of different domains of AI as in Computer Vision, Machine Learning, Natural Language Possessing, etc.
3. To develop a basic understanding of the building blocks of AI as presented in terms of intelligent agents: Search, Knowledge representation, inference, logic, and learning.
4. To learn about different Machine Learning models and techniques.
5. To understand basics of Pattern Recognition approaches and techniques.

UNIT-I **Lectures: 09**

Introduction- Introduction to Artificial Intelligence, Foundations and History of Artificial Intelligence, Applications of Artificial Intelligence, Intelligent Agents, Structure of Intelligent Agents. Computer vision, Natural Language Possessing.

UNIT-II **Lectures: 09**

Introduction to Search- Searching for solutions, Uniformed search strategies, Informed search strategies, Local search algorithms and optimistic problems, Adversarial Search, Search for games, Alpha - Beta pruning.

UNIT-III **Lectures: 09**

Knowledge Representation &Reasoning- Propositional logic, Theory of first order logic, Inference in First order logic, Forward & Backward chaining, Resolution, Probabilistic reasoning, Utility theory, Hidden Markov Models (HMM), Bayesian Networks.

UNIT-IV **Lectures: 09**

Machine Learning- Supervised and unsupervised learning, Decision trees, Statistical learning models, Learning with complete data - Naive Bayes models, Learning with hidden data - EM algorithm, Reinforcement learning.

UNIT-V **Lectures: 09**

Pattern Recognition- Introduction, Design principles of pattern recognition system, Statistical Pattern recognition, Parameter estimation methods - Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA), Classification Techniques – Nearest Neighbour (NN) Rule, Bayes Classifier, Support Vector Machine (SVM), K – means clustering.

Course Outcome: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand and become familiar with basic principles	K1, K2

	of AI toward problem solving, inference, perception, knowledge representation, and learning.	
2	CO2. Demonstrate knowledge of the building blocks of AI as presented in terms of intelligent agents.	K1, K2
3	CO3. Analyze and formalize the problem and select amongst different search or game based techniques to solve them.	K2, K4
4	CO4. Experiment with machine learning models, explore the current scope, potential, limitations, and implications.	K2, K3
5	CO5. Attain the capability to understand pattern-recognition based techniques, solve problems with uncertain information using Bayesian approaches and some other types of classifiers.	K1, K2

Suggested Readings:

1. Stuart Russell, Peter Norvig, "Artificial Intelligence – A Modern Approach", Pearson Education
2. Elaine Rich and Kevin Knight, "Artificial Intelligence", Tata McGraw-hill Education Pvt. Ltd.
3. E.Charniak and D McDermott, "Introduction to Artificial Intelligence", Pearson Education
4. Dan W. Patterson, "Artificial Intelligence and Expert Systems", Prentice Hall of India.
5. T. M. Mitchell, Machine Learning, McGraw-Hill, 1997.
6. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2006.
7. Richard O. Duda, Peter E. Hart and David G. Stork, "Pattern Classification", 2nd Edition, John Wiley, 2006.

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-III)
MCAE31: Machine Learning

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To understand the basics of machine learning and need for it, understand human learning aspects and relate it with machine learning concepts.
2. To understand implementation aspects of machine learning through Scikit learn and understand the concepts and importance of feature selection and filtering.
3. To learn about linear and logistic regression models, their concepts and methodologies.
4. To understand working of SVM, Bayes classifier and their methodologies.
5. To understand, evaluate and analyze different machine learning models such as decision trees, perceptron, clustering basics, etc.

UNIT-I **Lectures: 09**

Introduction- Introduction to different types of learning, Supervised and Unsupervised learning – Reinforcement learning- Basics of Neural network models, beyond machine learning-deep learning and bio inspired adaptive systems, Machine learning and big data.
Important Elements of Machine Learning- Data Representation, Diversity of Data: Structured/Unstructured, Statistical learning approaches, Basic Linear Algebra in Machine Learning Techniques.

UNIT-II **Lectures: 08**

Scikit- learn Dataset, Creating training and test sets, managing categorical data, Managing missing features, Data scaling and normalization, Feature selection and Filtering, Principle Component Analysis (PCA)-non negative matrix factorization, Sparse PCA, Kernel PCA. Atom Extraction and Dictionary Learning.

UNIT-III **Lectures: 09**

Linear regression- Linear models, A bi-dimensional example, Linear Regression and higher Dimensionality, Ridge, Lasso and Elastic Net, Robust regression with random sample consensus, Polynomial regression, **Logistic Regression**-Linear classification, Implementation and Optimizations, Stochastic gradient descendent algorithms.

UNIT-IV **Lectures: 10**

Introduction to Bayes Decision Theory, Bayes Theorem, Maximum Likelihood hypotheses for predicting probabilities, class conditional probability distributions, Naïve Bayes Classifiers.
Support Vector Machines (SVM)- Introduction, Hyperplanes and Support Vectors, Linear Discriminant Functions for Binary Classification, Perceptron Algorithm, Large Margin Classifier for linearly separable data, Linear Soft Margin Classifier for Overlapping Classes, Kernel Induced Feature Spaces, Nonlinear Classifier, Regression by Support vector Machines.

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

UNIT-V **Lectures: 09**

Decision Tree Learning: Introduction, Example of classification decision tree, measures of impurity for evaluating splits in decision trees, ID3, C4.5, and CART decision trees, pruning the tree, strengths and weakness of decision tree approach.
Clustering Fundamentals- Basics, Types of clustering methods, Clustering algorithms as K-means.
Multilayer Perceptron Networks and error back propagation algorithm, Radial Basis Functions Networks.

Course Outcome: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand and comprehend basics of machine learning techniques and design issues.	K1, K2
2	CO2. Learn and study about dataset for training and testing, managing categorical data, feature selection and filtering, etc. and understand their implementations.	K2, K3
3	CO3. Comprehend and understand linear and logistic regression models and their methodologies.	K2, K4
4	CO4. Understand basics of Bayes Decision Model and Support Vector Machines concepts and techniques.	K1, K2
5	CO5. Comprehend and understand working of Decision trees, fundamental concept of clustering and multilayer perceptron networks overview.	K1, K2

Suggested Readings:

1. T. M. Mitchell, Machine Learning, McGraw-Hill, 1997.
2. E. Alpaydin, Introduction to Machine Learning, Prentice Hall of India, 2006.
3. Applied Machine Learning, M. Gopal, McGraw Hill Education.
4. Machine Learning: A Probabilistic Perspective, Kevin Murphy, MIT Press,2012.
5. Peter Harrington, Machine Learning in action, Wiley.
6. The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani, Jerome Friedman, Springer 2009.
7. Pattern Recognition and Machine Learning, Christopher Bishop, Springer,2007.
8. Vladimir N. Vapnik, Statistical Learning Theory, John Wiley and Sons, 1998.
9. Shawe-Taylor J. and Cristianini N., Cambridge, Introduction to Support Vector Machines, University Press, 2000.
10. Peter Flach, "Machine Learning: The Art and Science of Algorithms that Make Sense of Data", Cambridge University Press, Edition 2012, ISBN-10: 1107422221; ISBN-13: 978-1107422223.
11. Giuseppe Bonaccorso, "Machine Learning Algorithms", Packt Publishing Limited, ISBN10: 1785889621, ISBN-13: 978-1785889622.

12. Stephan Marsland, "Machine Learning - An Algorithmic Perspective", Chapman and Hall, 1st Edition, 2009.
13. Nils Nilsson, "Introduction to Machine Learning", MIT Press, 1997.
14. Building Machine Learning Systems with Python – WilliRichert, Luis Pedro Coelho

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-III)
MCAE32: Neural Network

Credit: 03, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To understand the fundamental theory and concepts of computational intelligence methods, in particular neural networks, fuzzy systems, genetic algorithms and their applications in the area of machine intelligence.
2. To understand human learning aspects and relate it with machine intelligence.
3. To understand the concepts of Multi-layered network architecture, Recurrent Neural Network and their working.
4. To understand the basics of an evolutionary computing algorithms and its application to engineering optimization problems.

UNIT-I **Lectures: 09**

Neuro computing and Neuroscience: Historical notes, human Brain, neuron Model, Knowledge representation, Al and NN. Learning process: Supervised and unsupervised learning, Error correction learning, competitive learning, adaptation, statistical nature of the learning process.

UNIT-II **Lectures: 09**

Data processing: Scaling, normalization, Transformation (FT/FFT), principal component analysis, regression, covariance matrix, eigen values & eigen vectors. Basic Models of Artificial neurons, activation. Functions, aggregation function, single neuron computation, multilayer perceptron, least mean square algorithm, gradient descent rule, nonlinearly separable problems and bench mark problems in NN.

UNIT-III **Lectures: 09**

Multi-layered network architecture, back propagation algorithm, heuristics for making BP-algorithm performs better. Accelerated learning BP (like recursive least square, quick prop, RPROP algorithm), approximation properties of RBF networks and comparison with multilayer perceptron.

UNIT-IV **Lectures: 09**

Recurrent network and temporal feed-forward network, implementation with BP, self-organizing map and SOM algorithm, properties of feature map and computer simulation. Principal component and independent component analysis, application to image and signal processing.

UNIT-V **Lectures: 09**

Complex valued NN and complex valued BP, analyticity of activation function, application in 2D information processing. Complexity analysis of network models. Soft computing. Neuro-Fuzzy-genetic algorithm Integration.

Course Outcome: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand and demonstrate knowledge of Neuro computing and Neuroscience for understanding human learning aspects and relate it to machine intelligence.	K1, K2
2	CO2. Explain the concepts of data processing and attain the capability to apply them in various real life problem domains.	K2, K3
3	CO3. Describe and analyze the capability of Multi-layered network architecture, back propagation algorithm, multilayer perceptron, etc.	K2
4	CO4. Understand, describe and analyze Recurrent network, temporal feed-forward network, self-organizing map, Principal component, etc. and their applications.	K2, K4
5	CO5. Comprehend and attain capability to understand and analyze complex neural network models, fuzzy-systems, genetic algorithm, etc.	K1, K2

Suggested Readings:

1. J.A. Anderson, An Intoduction to Neural Networks, MIT
2. Hagen Demuth Beale, Neural Network Design, Cengage Learning
3. R.L. Harvey, Neural Network Principles, PHI
4. Kosko, Neural Network and Fuzzy Sets, PHI

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-III)
MCAE33: Natural Language Processing

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorial: 15 Hours

OBJECTIVES OF THE COURSE:

1. To provide the student with knowledge of various levels of analysis involved in NLP
2. To understand language modeling
3. To gain knowledge in automated natural language generation and machine Translation

UNIT-I **Lectures: 09**

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

UNIT-II **Lectures: 09**

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

UNIT-III **Lectures: 09**

Semantic Analysis and Discourse Processing: Meaning Representation-Lexical Semantics-Ambiguity-Word SenseDisambiguation DISCOURSE PROCESSING:Introduction- cohesion-Reference Resolution-Discourse Coherence and Structure

UNIT-IV **Lectures: 09**

Natural Language Generation and Machine Translation: Introduction-Architecture of NLG Systems- Generation Tasks and Representations-Application ofNLG MACHINE
Translation: Introduction-Problems in Machine Translation-Characteristics of Indian Languages- Machine Translation Approaches-Translation involvingIndian Languages

UNIT-V **Lectures: 09**

Information Retrieval and Lexical Resources: -Design features of Information Retrieval Systems-Classical, Non-classical, Alternative Modelsof Information Retrieval – Evaluation
Lexical Resources: Introduction-WordNet-FrameNet-Stemmers-POS Tagger- Research Corpora

Course Outcome: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Study and understand basic concepts, background and representation of natural languages.	K1, K2
2	CO2. Analyze various real-world applications of NLP	K3

3	CO3. Apply different parsing techniques in NLP.	K3, K4
4	CO4. Understand grammatical concepts and apply them in NLP	K2, K3
5	CO5. Apply various statistical probabilities grammar methods to handle and evaluate ambiguity.	K3, K5

Suggested Readings:

1. TanveerSiddiqui, U.S. Tiwary, "Natural Language Processing and Information Retrieval", Oxford University Press, 2008
2. Daniel Jurafsky and James H Martin," Speech and Language Processing: An introduction toNatural Language Processing, Computational Linguistics and Speech Recognition", Prentice Hall, 2nd Edition, 2008.
3. James Allen, Bejamin/cummings, "Natural Language Understanding", 2 nd edition, 1995

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-III)
MCAE34: Pattern Recognition

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To learn the fundamentals of pattern recognition and its relevance to classical and modern problems.
2. To develop a basic understanding of the scientific discipline whose goal is the classification of objects into a number of categories or classes.
3. To introduce the student to the basic concepts and methods for the recognition of patterns in data. This is accomplished via the presentation of the underlying theory and algorithmic approaches for the detection and characterization of patterns in multidimensional data.

UNIT-I **Lectures: 09**

Introduction: Basics of pattern recognition, Design principles of pattern recognition system, Learning and adaptation, Pattern recognition approaches, Mathematical foundations – Linear algebra, Probability Theory, Expectation, mean and covariance, Normal distribution, multivariate normal densities, Chi squared test.

UNIT-II **Lectures: 09**

Statistical Patten Recognition: Bayesian Decision Theory, Classifiers, Normal density and discriminant functions.

UNIT-III **Lectures: 09**

Parameter estimation methods: Maximum- Likelihood estimation, Bayesian Parameter estimation, Dimension reduction methods-Principal Component Analysis (PCA),Fisher Linear discriminate analysis, Expectation-maximization(EM),Hidden Markov Models(HMM), Gaussian mixture models.

UNIT-IV **Lectures: 09**

Nonparametric Techniques: Density Estimation, Parzen Windows, K-Nearest Neighbor Estimation, Nearest Neighbor Rule, Fuzzy classification.

UNIT-V **Lectures: 09**

Unsupervised Learning & Clustering: Criterion functions for clustering, Clustering Techniques:
Iterative square - error partitional clustering – K means, agglomerative hierarchical clustering, Cluster validation.

Course Outcome: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand and become familiar with concepts of pattern recognition	K1, K2
2	CO2. Explain and distinguish procedures, methods and	K1, K2

	algorithms related to pattern recognition	
3	CO3. Summarize, analyze, relate and compare simple pattern classifiers, statistical pattern recognition techniques, non-parametric techniques, etc.	K2, K4
4	CO4. Understand and experiment with pattern recognition techniques using unsupervised learning and clustering.	K1, K2
5	CO5. Attain the capability to represent, design and develop pattern recognition systems for real-world problems.	K2, K3

Suggested Readings:

1. Richard O. Duda, Peter E. Hart and David G. Stork, "Pattern Classification", 2nd Edition, John Wiley, 2006.
2. C. M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2009.
3. S. Theodoridis and K. Koutroumbas, "Pattern Recognition", 4th Edition, Academic Press,2009.

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-IV)
MCAE41- Data Warehousing & Data Mining

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To understand data warehouse concepts, architecture, business analysis and tools.
2. To understand data pre-processing and data visualization techniques.
3. To study algorithms for finding hidden and interesting patterns in data.
4. To understand and apply various classification and clustering techniques using tools.

UNIT-I **Lectures: 09**

Data Warehousing and On-Line Analytical Processing: Data Warehouse basic concepts, Data Warehouse Modeling, Data Cube and OLAP, Data Warehouse Design and Usage, Datawarehouse Implementation, Data Generalization by Attribute-Oriented Induction.
Data Cube Technology: Efficient Methods for Data Cube Computation, Exploration and Discovery in Multidimensional Databases.

UNIT-II **Lectures: 09**

Introduction to Data Mining: Motivation, Importance, Definition of Data Mining, Kind of Data, Data Mining Functionalities, Kinds of Patterns, Classification of Data Mining Systems, Data Mining Task Primitives, Integration of A Data Mining System With A Database or Data Warehouse System, Major Issues In Data Mining, Types of Data Sets and Attribute Values, Basic Statistical Descriptions of Data, Data Visualization, Measuring Data Similarity. **Preprocessing:** Data Quality, Major Tasks in Data Preprocessing, Data Reduction, Data Transformation and Data Discretization, Data Cleaning and Data Integration

UNIT-III **Lectures: 09**

Mining Frequent Patterns, Associations and Correlations: Basic Concepts, Efficient and Scalable Frequent Item set Mining Methods, Are All the Pattern Interesting, Pattern Evaluation Methods, Applications of frequent pattern and associations.
Frequent Pattern and Association Mining: A Road Map, Mining Various Kinds of Association Rules, Constraint-Based Frequent Pattern Mining, Extended Applications of Frequent Patterns

UNIT-IV **Lectures: 09**

Classification: Decision Tree Induction, Bayesian Classification Methods, Rule-Based Classification, Model Evaluation and Selection, Bayesian Belief Networks, Classification by Neural Networks, Support Vector Machines, Pattern-Based Classification, Lazy Learners (or Learning from Your Neighbors).

UNIT-V **Lectures: 09**

Clustering Analysis: Basic Concepts of Cluster Analysis, Clustering structures, Major ClusteringApproaches, Partitioning Methods, Hierarchical Methods, Density-Based Methods,

Model Based Clustering - The Expectation-Maximization Method, Other Clustering Techniques, Clustering High-Dimensional Data, Constraint-Based and User-Guided Cluster Analysis, Link-Based Cluster Analysis, Semi-Supervised Clustering and Classification, Bi-Clustering, Collaborative Clustering.

Course Outcomes: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Demonstrate the concept of Data warehouse system and perform business analysis with OLAP tools	K1, K2
2	CO2. Understand and Apply suitable pre-processing and visualization techniques for data analysis	K2, K3
3	CO3. Apply frequent pattern and association rule mining techniques for data analysis	K3
4	CO4. Understand and Apply appropriate classification techniques for data analysis	K2, K3
5	CO5. Understand and apply various clustering techniques for data analysis	K2, K3

Suggested Readings:

1. Jiawei Han, MichelineKamber, Jian Pei (2012), Data Mining: Concepts and Techniques, 3rdedition, Elsevier, United States of America.
2. Margaret H Dunham (2006), Data Mining Introductory and Advanced Topics, 2ndedition,PearsonEducation, New Delhi, India.
3. AmiteshSinha (2007), Data Warehousing, Thomson Learning, India.
4. Xingdong Wu, Vipin Kumar (2009), the Top Ten Algorithms in Data Mining, CRC Press, UK.

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-IV)
MCAE42- Big Data Analytics

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To Understand the concepts of Big Data fundamentals.
2. To Learn the importance of Hadoop and SQL Languages.
3. To use Hive and NoSQL

UNIT-I **Lectures: 09**

Introduction to Big Data: Types of digital data, history of Big Data innovation, introduction to Big Data platform, drivers for Big Data, Big Data architecture and characteristics, 5 Vs of Big Data, Big Data technology components, Big Data importance and applications, Big Data features – security, compliance, auditing and protection, Big Data privacy and ethics, Big Data Analytics, Challenges of conventional systems, intelligent data analysis, nature of data, analytic processes and tools, analysis vs reporting, modern data analytic tools.

UNIT-II **Lectures: 09**

Hadoop: History of Hadoop, Apache Hadoop, the Hadoop Distributed File System, components of Hadoop, data format, analyzing data with Hadoop, scaling out, Hadoop streaming, Hadoop pipes, Hadoop Echo System.
Map Reduce: Map Reduce framework and basics, how Map Reduce works, developing a Map Reduce application, unit tests with MR unit, test data and local tests, anatomy of a Map Reduce job run, failures, job scheduling, shuffle and sort, task execution, Map Reduce types, input formats, output formats, Map Reduce features, Real-world Map Reduce

UNIT-III **Lectures: 09**

HDFS (Hadoop Distributed File System): Design of HDFS, HDFS concepts, benefits and challenges, file sizes, block sizes and block abstraction in HDFS, data replication, how does HDFS store, read, and write files, Java interfaces to HDFS, command line interface, Hadoop file system interfaces, data flow, data ingest with Flume and Scoop, Hadoop archives, Hadoop I/O: compression, serialization, Avro and file-based data structures. Hadoop Environment: Setting up a Hadoop cluster, cluster specification, cluster setup and installation, Hadoop configuration, security in Hadoop, administering Hadoop, HDFS monitoring & maintenance, Hadoop benchmarks, Hadoop in the cloud

UNIT-IV **Lectures: 09**

Hadoop Eco System and YARN: Hadoop ecosystem components, schedulers, fair and capacity, Hadoop 2.0 New Features - NameNode high availability, HDFS federation, MRv2, YARN, Running MRv1 in YARN.
NoSQL Databases: Introduction to NoSQL
MongoDB: Introduction, data types, creating, updating and deleing documents, querying, introduction to indexing, capped collections

Spark: Installing spark, spark applications, jobs, stages and tasks, Resilient Distributed 09 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 **Lectures: 09**

Hadoop Eco System Frameworks: Applications on Big Data using Pig, Hive and HBase
Pig: Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators,
Hive: Apache Hive architecture and installation, Hive shell, Hive services, Hive meta store, comparison with traditional databases, HiveQL, tables, querying data and user defined functions, sorting and aggregating, Map Reduce scripts, joins & sub queries.
HBase: Hbase concepts, clients, example, Hbasevs RDBMS, advanced usage, schema design, advance indexing, Zookeeper – how it helps in monitoring a cluster, how to build applications with Zookeeper. IBM Big Data strategy, introduction to Info sphere, Big Insights and Big Sheets, introduction to Big SQL.

Course Outcomes: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand and Identify Big Data and its business implications.	K2, K3
2	CO2. Make use various techniques for mining data stream.	K3
3	CO3. List the components of Hadoop and Hadoop Eco-System	K1
4	CO4. Apply Map Reduce programming model to access and process data on Distributed File System	K3
5	CO5. Manage job execution in Hadoop environment and develop Big Data solutions by applying Hadoop Eco System components	K3

Suggested Readings:

1. Michael Minelli, Michelle Chambers, and AmbigaDhiraj, "Big Data, Big Analytics: Emerging Business, Intelligence and Analytic Trends for Today's Businesses", Wiley
2. Big-Data Black Book, DT Editorial Services, Wiley
3. Dirk deRoos, Chris Eaton, George Lapis, Paul Zikopoulos, Tom Deutsch, "Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data", McGrawHill.
4. Thomas Erl, WajidKhattak, Paul Buhler, "Big Data Fundamentals: Concepts, Drivers and Techniques", Prentice Hall.
5. Bart Baesens "Analytics in a Big Data World: The Essential Guide to Data Science and its Applications (WILEY Big Data Series)", John Wiley & Sons
6. ArshdeepBahga, Vijay Madisetti, "Big Data Science & Analytics: A HandsOnApproach", VPT
7. AnandRajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", CUP

8. Tom White, "Hadoop: The Definitive Guide", O'Reilly.
9. Eric Sammer, "Hadoop Operations", O'Reilly.
10. Chuck Lam, “Hadoop in Action”, MANNING Publishers
11. Deepak Vohra, “Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-Related Frameworks and Tools”, Apress
12. E. Capriolo, D. Wampler, and J. Rutherglen, "Programming Hive", O'Reilly
13. Lars George, "HBase: The Definitive Guide", O'Reilly.
14. Alan Gates, "Programming Pig", O'Reilly.
15. Michael Berthold, David J. Hand, “Intelligent Data Analysis”, Springer
16. Bill Franks, “Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics”, John Wiley & sons
17. Glenn J. Myatt, “Making Sense of Data”, John Wiley & Sons
18. Pete Warden, “Big Data Glossary”, O’Reilly

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-IV)
MCAE43- Advanced Database Management System

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To Understand the principles of distributed databases and how they differ from centralized databases.
2. To enhance the previous knowledge of database systems by deepening the understanding of the theoretical and practical aspects of the database technologies and showing the need for distributed database technology to tackle deficiencies of the centralized database systems.
3. Focuses on understanding the concepts of designing and managing distributed databases.
4. To Provide a cohesive overview regarding the importance of data management and data analytics in the era we are living

UNIT-I **Lectures: 09**

Introduction
Distributed Data Processing, Distributed Database System, Promises of DDBSs, Problem areas, Architectural Models for Distributed DBMS, DDMBS Architecture, Alternative Design Strategies, Distribution Design Issues, Fragmentation, and Allocation.

UNIT–II **Lectures: 09**

Query Processing and Decomposition
Query processing objectives, characterization of query processors, layers of query processing, query decomposition, localization of distributed data, Query optimization, centralized query optimization, distributed query optimization algorithms.

UNIT-III **Lectures: 09**

Transaction Management
Definition, properties of transaction, types of transactions, distributed concurrency control: serializability, concurrency control mechanisms & algorithms, time - stamped & optimistic concurrency control Algorithms, deadlock Management.

UNIT–IV **Lectures: 09**

Distributed DBMS Reliability
Reliability concepts and measures, fault-tolerance in distributed systems, failures in Distributed DBMS, local & distributed reliability protocols, site failures and network partitioning. Parallel database system architectures, parallel data placement, parallel query processing, load balancing, database clusters.

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

UNIT-V **Lectures: 09**

Distributed Object Database Management Systems

Fundamental object concepts and models, object distributed design, architectural issues, object management, distributed object storage, object query Processing, Inheritance, object identity, persistent programming languages, persistence of objects, comparison OODBMS and ORDBMS

Course Outcomes: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand theoretical and practical aspects of distributed database systems.	K1
2	CO2. Study and identify various issues related to the development of Query processing and decomposition in distributed database system.	K2
3	CO3. Illustrate the concept and significance of transactions in distributed database system.	K2
4	CO4. Identify the concept of Reliability, fault-tolerance and Parallel database system architectures.	K3
5	CO5. Understand the design aspects of object-oriented database system and related development.	K1

Suggested Readings:

1. M. Tamer OZSU and Patuck Valduriez: Principles of Distributed Database Systems, Pearson Edn. Asia, 2001.
2. Stefano Ceri and Giuseppe Pelagatti: Distributed Databases, McGraw Hill.
3. Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom: "Database Systems: The Complete Book", Second Edition, Pearson International Edition
4. Garcia-Molina, Ullman, Widom,' Database System Implementation' Pearson Education
5. Silberschatz, Korth and Sudershan, Database System Concept',McGrawHill

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV (Elective-IV)
MCAE44- Compiler Design

Credit: 04, IA Marks: 30, ESE Marks: 70
Lectures: 45 Hours, Tutorials: 15 Hours

OBJECTIVES OF THE COURSE:

1. To Explore the principles, algorithms, and data structures involved in the design and construction of compilers.
2. The course is intended to teach the students the basic techniques that underlie the practice of Compiler Construction.
3. To Introduce the major concept areas of language translation and compiler design and to develop an awareness of the functions.

UNIT-I **Lectures: 09**

Introduction to Compiler: Phases and passes, Bootstrapping, Finite state machines and regular expressions and their applications to lexical analysis, Optimization of DFA-Based Pattern Matchers implementation of lexical analyzers, lexical-analyzer generator, LEX compiler, Formal grammars and their application to syntax analysis, BNF notation, ambiguity, YACC. The syntactic specification of programming languages: Context free grammars, derivation and parse trees, capabilities of CFG.

UNIT–II **Lectures: 09**

Basic Parsing Techniques: Parsers, Shift reduce parsing, operator precedence parsing, top down parsing, predictive parsers Automatic Construction of efficient Parsers: LR parsers, the canonical Collection of LR(0) items, constructing SLR parsing tables, constructing Canonical LR parsing tables, Constructing LALR parsing tables, using ambiguous grammars, an automatic parser generator, implementation of LR parsing tables

UNIT-III **Lectures: 09**

Syntax-directed Translation: Syntax-directed Translation schemes, Implementation of Syntax directed Translators, Intermediate code, postfix notation, Parse trees & syntax trees, three address code, quadruple & triples, translation of assignment statements, Boolean expressions, statements that alter the flow of control, postfix translation, translation with a top-down parser. More about translation: Array references in arithmetic expressions, procedures call, declarations and case statements.

UNIT–IV **Lectures: 09**

Symbol Tables: Data structure for symbols tables, representing scope information. Run-Time Administration: Implementation of simple stack allocation scheme, storage allocation in block structured language. Error Detection & Recovery: Lexical Phase errors, syntactic phase errors semantic errors

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

UNIT-V **Lectures: 09**

Code Generation: Design Issues, the Target Language. Addresses in the Target Code, Basic Blocks and Flow Graphs, Optimization of Basic Blocks, Code Generator. Code optimization: Machine-Independent Optimizations, Loop optimization, DAG representation of basic blocks, value numbers and algebraic laws, Global Data-Flow analysis.

Course Outcomes: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Acquire knowledge of different phases and passes of the compiler and also able to use the compiler tools like LEX, YACC, etc.	K3
2	CO2. Understand the parser and its types i.e. Top-Down and Bottom-up parsers and construction of LL, SLR, CLR, and LALR parsing table.	K2, K3
3	CO3. Analyze the compiler using syntax-directed translation method and get knowledge about the synthesized and inherited attributes.	K4,
4	CO4. Acquire knowledge about run time data structure like symbol table organization and different techniques used in that.	K3
5	CO5. Understand the target machine's run time environment, its instruction set for code generation and techniques used for code optimization.	K12, K4

Suggested Readings:

1. K. Muneeswaran,CompilerDesign,FirstEdition,Oxford University Press.
2. J.P. Bennet, "Introduction to Compiler Techniques", Second Edition, Tata McGraw-Hill,2003.
3. HenkAlblas and Albert Nymeyer, "Practice and Principles of Compiler Building with C", PHI, 2001.
4. Aho, Sethi& Ullman, "Compilers: Principles, Techniques and Tools", Pearson Education
5. V Raghvan, "Principles of Compiler Design", TMH
6. Kenneth Louden," Compiler Construction", Cengage Learning.
7. Charles Fischer and Ricard LeBlanc," Crafting a Compiler with C", Pearson Education

ख़्वाजा मुईनुद्दीन चिश्ती भाषा विश्वविद्यालय, लखनऊ, उत्तर प्रदेश (भारत)
Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh (India)
U.P. State Government University
(Recognised Under Section 2(F) & 12(B) of the UGC Act 1956 & B.Tech Approved by AICTE)

MCA Semester IV
MCAP402- Project- Dissertation (Software/Research) **Credit: 12, IA Marks: 150, ESE Marks: 350**

OBJECTIVES OF THE COURSE:

To help students develop openness to new ideas in computer science, develop the ability to draw reasonable inferences from observations and learn to formulate and solve new computer science problems using analytical and problem-solving skills; To help students develop the ability to synthesize and integrate information and ideas, develop the ability to think creatively, develop the ability to think holistically and develop the ability to distinguish between facts and opinion; To help students acquire the necessary competences to build a real-life software system by completing different software life cycle phases (like, specification, architecture, design, implementation, validation, documentation, etc. To help students acquire the necessary competences to build a Research Project.

Major-Project using Web Engineering Tools: Design and Implementation of Web Applications, Web Services, Mobile Applications etc. Students are required to incorporate the followings: Dynamic Pages, Adding Dynamic Functionality Interactive User Interface Database in the back-end XML and Databases Provision for EDIs Adding Security Features, etc.

For developing the project, students may use the followings:

J2EE Platform, .NET Platform, Eclipse JAVA, C#, VC++, Python, etc Microsoft Frontpage/Flash/PHP/Dreamweaver etc. XML, DHTML, CGI, Scripting Languages (JSP, ASP, PHP), PYTHON or Suitable Technologies.

Course Outcomes: After successful completion of this course students will be able to:

S. No.	Course Outcome	Bloom's Taxonomy
1	CO1. Understand project characteristics and various stages of a project.	K2
2	CO2. Understand the conceptual clarity about project organization and feasibility analyses – Market, Technical, Financial and Economic	K2, K3
3	CO3. Demonstrate a sound technical knowledge of the project.	K2, K4
4	CO4. Illustrate &Apply different software package for project	K4, K5