B.Tech Computer Science and Engineering (Artificial Intelligence and Data Science)
Scheme of Studies/Examination (w.e.f. Session 2022-23)
Semester VI

S. No.	Course No.	Subject	L:T:P	Hours/ Week	Credits	Examination Schedule				Duration of Exam (Hrs.)
						Major Test	Minor Test	Practical	Total	
1	PC-CS-AIDS-302A	Compiler Design	3:0:0	3	3	75	25	0	100	3
2	ES-CS-AIDS-304A	Applied Statistical Analysis for AI	3:0:0	3	3	75	25	0	100	3
3	PC-CS-AIDS-306A	Big Data Analytics	3:0:0	3	3	75	25	0	100	3
4	PC-CS-AIDS-308A	Applied Machine Learning	3:0:0	3	3	75	25	0	100	3
5	OEC	OEC Elective-I	3:0:0	3	3	75	25	0	100	3
6	PC-CS-AIDS-310A	Soft Computing	3:0:0	3	3	75	25	0	100	3
7	PC-CS-AIDS-312LA	Applied Machine Learning Lab	0:0:2	2	1	0	40	60	100	3
8	PC-CS-AIDS-314LA	Big Data Analytics Lab	0:0:2	2	1	0	40	60	100	3
9	ES-CS-AIDS-316LA	Applied Statistical Analysis for AI Lab	0:0:2	2	1	0	40	60	100	3
		Total		**24**	**21**	**450**	**270**	**180**	**900**	

OEC Elective-I
Soft Skills and Interpersonal Communication: OE-CS-AIDS-302
Management Information System: OE-CS-AIDS-304
Enterprise Resource Planning: OE-CS-AIDS-306

*The students will choose any One Open Elective course out of the given elective list in VI Semester.

PC-CS-AIDS-302A	Compiler Design						
Lecture	Tutorial	Practical	Credit	Major Test	Minor Test	Total	Time
3	0	0	3	75	25	100	3 Hrs.
Purpose	To introduce complier design concepts and their implementation.						
Course Outcomes (CO)							
CO1	To understand the role and designing of a lexical analyzer.						
CO2	To analyze the role and designing of syntax analyzer or parser.						
CO3	To identify the role of semantic analyzer and intermediate code generation.						
CO4	To explore the design importance of optimization of codes and error detection.						

UNIT-I

Introduction to Language Processing System, Compiling Analysis of the Source Program, Phases of a Compiler, Compiler Construction Tools. Lexical Analysis –Regular Expression, Introduction to Finite Automata and Regular Expression, Conversion of Regular Expression to NFA, Role of Lexical Analyzer, Specification of Tokens.

UNIT-II

Syntax Analysis: Role of the Parser, Abstract Syntax Trees, Ambiguity in Context-Free Grammars, Types of Parsing: Top-Down Parsing, Recursive Descent Parsing, LL Parser, Back Tracking, Bottom-Up Parsing, SLR Parser, Canonical LR Parser, LALR Parser.

UNIT-III

Semantic Analysis: Semantic Errors, Attribute Grammar, Synthesized attributes, Static Allocation, Stack Allocation, Heap Allocation, Activation Trees, Symbol Table, Intermediate Code Generation and Code Intermediate languages, Declarations, Assignment Statements, Boolean Expressions, Case Statements, DAG representation of Basic Blocks, A simple Code generator from DAG, Issues in the Design of Code Generator.

UNIT-1V

Code Optimization and Run Time Environments, Principal Sources of Optimization, Machine-independent Optimization, Machine-dependent Optimization, Optimization of Basic Blocks, Loop Optimization, Peephole Optimization, Introduction to Global Data Flow Analysis, Storage Organization, Static Storage Management, Heap Storage management, Parameter Passing. Error Recovery, Panic mode, Statement mode, Global correction.

Suggested Books:

1. Alfred Aho, Ravi Sethi, Jeffrey DUllman,"Compilers Principles, Techniques and Tools", Pearson Education Asia, 2018.
2. Allen I. Holub" Compiler Design in C",Prentice Hall of India,2003.
3. C.N.Fischer and R.J.LeBlanc,"Crafting a compiler with C",Benjamin Cummings,2003.
4. V Raghavan,"Principles of Compiler Design",Second Edition,Tata Mc Graw-Hill,2018.
5. Henk Alblas and Albert Nymeyer,"Practice and Principles of Compiler Building with C", PHI, 2001.
6. Kenneth C. Louden, "Compiler Construction: Principles and Practice", Thompson Learning, 2003.

ES-CS-AIDS-304A	Applied Statistical Analysis for AI						
Lecture	Tutorial	Practical	Credit	Major Test	Minor Test	Total	Time
3	0	0	3	75	25	100	3 Hrs.
Purpose	To gain a broad understanding of the statistical analysis in Artificial Intelligence.						
Course Outcomes -At the end of this course students will be able to:							
CO1	Explore the Statistical Analysis concepts with the irrelationships and process.						
CO2	Explain the concept of describing, transforming and summarizing data using various Statistical methods and apply them to solve real world problems.						
CO3	Understand and apply testing hypothesis with real life datasets.						
CO4	Examine and analyze the relationships to find the correlation and regression and their Applications in real life.						
CO5	Explore the advanced techniques with applications of decision trees, neural networks.						

UNIT – I

Introduction to basic concepts of Statistics, The Scientific Method, Basic Steps of the Research Process, Experimental Data and Survey Data, Populations and Samples, Census and Sampling Method, Parameter and Statistic, Independent and Dependent Variables, Examining Relationships, Introduction to SPSS Statistics.

UNIT – II

Introduction, Types of Data, Data Transformation, Summarizing Data: Graphical Methods, Summarizing Data: Measures of Central Tendency, Summarizing Data: Measures of Dispersion, Levels of Measurement, Random Variables and Probability Distributions, Discrete and Continuous Random Variable, Making Inferences about Populations from samples, Estimator and Estimate, Confidence Interval for Population Mean (Large Sample).

UNIT– III

Introduction, Null and Alternative Hypothesis, Type I and Type II Error, The Procedure of Hypothesis Testing; Hypothesis Testing of a Population Mean: Sample, a proportion (One Sample), Population Variance, Population Mean: Two Independent Samples (), Dependent Samples (Paired Samples), Two Population Proportion, Two Population Variances; Analysis of Variance(ANOVA).

UNIT– IV

Introduction, Types of Correlation, Karl Pearson Coefficient Correlation, Spearman's Rank Order Correlation, Partial Correlation, Residuals and Plots, Simple Linear Regression, Multiple Regression Model, Repeated Measures, Non-linear Regression, Polynomial Regression Models, Decision Trees, Neural Networks, Cluster Analysis, Factor Analysis.

Suggested Books:
1. Probability for Statistics and Machine Learning: Anirban Das Gupta –2011.
2. An Introduction to Statistics with Python With Applications in the Life Sciences By Thomas Hasl wanter,2016.
3. Applied Statistics: A hand book of techniques-Zenon Reynaro wych,springer verlag.
4. Joel Grus,"Data Science from Scratch: First Principles with Python",O' Reilly Media Media.
5. Jain V.K.,"Data Sciences",Khanna Publishing House, Delhi.
6. Applied Statistics-principles and Examples-D.R. Coxand E.J. Snell.
7. Appliedstatisticalmethods,IrvingW.Burr,Academicpress.
8. Probability, Statistics and Random process, Dr.K. Murugesan & P.Gurusamy by Anuradha

Agencies, Deepthi publications.
9. Jain V.K.,"Big Data and Hadoop",Khanna Publishing House, Delhi.
10. Jiawei Han and Jian Pei, "Data Mining Concepts and Techniques", Third Edition, Morgan Kaufmann Publishers.

PC-CS-AIDS-306A	Big Data Analytics						
L	T	P	Credit	Major Test	Minor Test	Total	Time
3	0	0	3	75	25	100	3 Hour
Purpose	The course provides grounding in basic and advanced methods to big data technology and tools.						
Course Outcomes -At the end of this course students will be able to:							
CO1	Understand Big Data and its analytics in the real world.						
CO2	Analyze the Big Data framework like Hadoop and NO SQL to efficiently store and process Big Data to generate analytics.						
CO3	Design of Algorithms to solve Data Intensive Problems using Map Reduce Paradigm 3 4						
CO4	Design and Implementation of Big Data Analytics using pig and spark to solve data intensive problems and to generate analytics.						
CO5	Implement Big Data Activities using Hive.						

UNIT-I

Introduction To Big Data - Distributed file system, Big Data and its importance, Four Vs, Drivers for Big data, big data analytics, big data applications. Algorithms using map reduce, Matrix-Vector Multiplication by Map Reduce.

UNIT-II

Introduction To Hadoop- Big Data – Apache Hadoop & Hadoop Eco System – Moving Data in and out of Hadoop– Understanding inputs and outputs of Map Reduce-Data Serialization.

UNIT-III

Hadoop Architecture - Hadoop Architecture, Hadoop Storage: HDFS, Common Hadoop Shell commands, Anatomy of File Write and Read., Name Node, Secondary Name Node, and Data Node, Hadoop Map Reduce paradigm, Map and Reduce tasks, Job, Task trackers - Cluster Setup – SSH & Hadoop Configuration– HDFS Administering–Monitoring &Maintenance.

UNIT-IV

Hadoop Ecosystem And Yarn -Hadoop ecosystem components - Schedulers - Fair and Capacity, Hadoop2.0NewFeatures-Name Node High Availability, HDFS Federation, MRv2, YARN, Running MRv1 in YARN.

Suggested Books:

1. Boris lublinsky, Kevint. Smith, Alexey Yakubovich, "Professional Hadoop Solutions" ,Wiley, ISBN:9788126551071,2015.
2. Chris Eaton,Dirk deroos, etal.,"Understanding Big data",Mc Graw Hill,2012.
3. Tom White,"HADOOP:The definitive Guide",Oreilly 2012.
4. Vignesh Prajapati,"Big Data Analytics with R and Hadoop",Packet Publishing 2013.
5. Tom Plunkett, Brian Mac Donald etal, "Oracle Big Data Handbook",Oracle Press,2014.

PC-CS-AIDS-308A	Applied Machine Learning						
Lecture	Tutorial	Practical	Credit	Major Test	Minor Test	Total	Time
3	0	0	3	75	25	100	3 Hrs.
Purpose	Objective of this course is to learn conceptually how machine learning algorithms work and interact with data; the emphasis will be on effective methodology for using machine learning to solve practical problems.						
Course Outcomes-At the end of this course students will be able to:							
CO1	Identify overfit regression models.						
CO2	Compare different regularized regression algorithms and decision tree ensemble algorithms.						
CO3	Explain the confusion matrix and its relation to the ROC curve.						
CO4	Construct training datasets, testing datasets, and model pipelines.						

Unit-I

Introduction- towards intelligent machines, well-posed machine learning problems, examples of applications in diverse fields, data representation, domain knowledge for productive use of machine learning, diversity of data: structured/unstructured, forms of learning, machine learning and data mining, basic linear algebra in machine learning techniques, relevant resources for machine learning.
Statistical learning-machine learning and inferential statistical analysis, descriptive statistics in learning techniques, Bayesian reasoning: a probabilistic approach to inference.

Unit-II

Supervised learning: rationale and basics, learning from observations, bias and variance, why learning works: computational learning theory, occam's razor principle and overfitting avoidance, heuristic search in inductive learning, estimating generalization errors, metrics for assessing regression (numeric prediction) accuracy, metrics for assessing classification (pattern recognition) accuracy, an overview of the design cycle and issues in machine learning.
Learning with support vector machines (SVM) and Random Forests-introduction, linear discriminant functions for binary classification, perceptron algorithm, linear maximal margin classifier for linearly separable data, linear soft margin classifier for overlapping classes, nonlinear classifier, regression by support vector machines, Decision tree learning, Building a decision tree, combining weak to strong learners via random forest, choosing a split with information gain.

Unit-III

Unsupervised learning-Data clustering and data transformations, engineering the data, overview of basic clustering methods, k-means clustering, fuzzy k-means clustering, expectation-maximization (EM) algorithm and gaussian mixtures clustering, some useful data transformations, entropy–based method for attribute discretization, principal components analysis (PCA) for attribute reduction, rough sets-based methods for attribute reduction. k-nearest neighbor (k-nn) classifier, discriminant functions and regression functions, linear regression with least square error criterion, logistic regression for classification tasks, fisher's linear discriminant and thresholding for classification, minimum description length principle.

Unit-IV

Learning with neural networks-towards cognitive machine, neuron models, network architectures, perceptron, linear neuron and the widrow-h off learning rule, the error-correction delta rule, multi-layer perceptron (MLP) networks and the error-back propagation algorithm, multi-class discrimination with MLP networks, radial basis functions (RBF)networks, genetic-neural systems
Fuzzy inference systems-introduction, cognitive uncertainty and fuzzy rule-base, fuzzy quantification of knowledge, fuzzy rule-base and approximate reasoning, Mamdani model for fuzzy inference systems, takagi-sugeno fuzzy model, neuro-fuzzy inference systems, genetic-fuzzy systems

Suggested Books:

1. M.Gopal, Applied Machine learning, McGraw-Hill Education
2. David Forsyth, Applied Machine learning, Springer
3. Pascal Bugnion, Patrick R.Nicolas, Alex Kozlov, Scala: Applied Machine Learning, Packt.

OE-CS-AIDS-302	Soft Skills and Inter personal Communication						
Lecture	Tutorial	Practical	Credit	Major Test	Minor Test	Total	Time
3	0	0	3	75	25	100	3 Hrs.
Purpose	To Develop broad career plans, evaluate the employment market, identify the organizations to get good placement, match the job requirements and skill sets.						
Course Outcomes (CO):By the end of the course, the students should be able to:							
CO1	Develop effective communication skills (spoken and written).						
CO2	Develop effective presentation skills.						
CO3	Conduct effective business correspondence and prepare business reports which produce results.						
CO4	Become self-confident individuals by mastering inter-personal skills, team management skills, and leadership skills.						

Unit-I

Introduction, Need for Communication, Process of Communication-Written and Verbal Communication, Visual communication, Signs, Signals and Symbols, Silence as a Mode of Communication - Inter-cultural, Intra-cultural, Cross-cultural and International communication-Communications skills, Communication through Questionnaires, Business Letter Writing, Electronic Communication. Barriers to Communication Improving Communication Skills-Preparation of Promotional Material -Non-verbal communication -Body language -Postures and gestures -Value of time-Organizational body language – Importance of Listening-Emotional Intelligence.

Unit-II

Business Cases and Presentations, Letters within the Organizations, Letters from Top Management, Circulars and Memos - Business Presentations to Customers and other stakeholders, presenting a Positive Image through Verbal and Non-verbal Cues, Preparing and Delivering the Presentations, Use of Audio-visual Aids -Report Writing.

Unit-III

Individual Interaction and skills Basic Interaction Skills –Within family, Society Personal and interpersonal intrapersonal skills Types of skills; conceptual, supervisory, technical, managerial and decision-making skills. Problem Solving, Lateral Thinking Self Awareness and Self Esteem Group Influence on Interaction Skills Human relations examples through role–play and cases.

Unit-IV:

Leadership Skills Working individually and in a team Leadership skills 15 Lectures Leadership Lessons through Literature Team work & Team building Interpersonal skills–Conversation, Feedback, Feedforward Interpersonal skills–Delegation, Humor, Trust, Expectations, Values, Status, Compatibility and their role in building team– work Conflict Management – Types of conflicts, how to cope with them, small cases including role – plays will be used as teaching methodology. **Negotiation Skills** (To be Taught through Role Plays and Cases) Types of Negotiation Strategies Selling skills – Selling to customers Selling to Superiors Selling to peer groups, team mates & sub ordinates Conceptual selling, Strategic selling skills–Body language.

Suggested Books:

1. A Practical Guide to Soft Skills Communication, Psychology, and Ethics for Your Professional Life by Richard Almonte, Taylor & Francis.
2. Soft Skills for Interpersonal Communication, By Jeypaul Jesudoss, T.Ravindran, OXFORD University Press.

OE-CS- AIDS-304	Management Information System						
L	T	P	Credit	Major Test	Minor Test	Total	Time
3	0	0	3	75	25	100	3 Hour
Purpose	To familiarize the students with Management Information System.						
	Course Outcomes						
CO1	To provide introduction to relational model.						
CO2	To learn about ER diagrams and SQL.						
CO3	To understand about the concept of functional dependencies.						
CO4	To understand about Query Processing and Transaction Processing.						

UNIT-I

Introduction: Definition information system, role and impact of MIS, the challenges of Information system, Nature of MIS, Characteristics of MIS, Myths regarding MIS, Requirements of MIS, Problems & Solutions in implementing MIS, Benefits of MIS, Limitations of MIS, Significance of MIS, Components of MIS. Role of MIS, Major Management challenge to building and using information system in Organization, functions of management.

UNIT-II

Information system and Organizations: The relationship between Organization and Information System, Information needs of different organization levels: Information concept as quality product, classification and value of information, methods of data and information collection. Strategic role of information system, Salient features of Organization, Information, management and decision making, How Organization affect Information Systems, How Information system affect Organization, Ethical and Social impact of information system.

UNIT-III

Business application of Information System: Foundation Concepts Information systems in Business: Information system and technology, Business Applications, Development and Management. The inter networked E-business Enterprise: Internet, and Extranet in business. Electronic Commerce System: Electronics commerce Fundamentals, Commerce Application and issues. E-business Decision Support: Decision support in E-Business, Artificial Intelligence Technologies in business.

UNIT-IV

Technical Foundation of Information System: Computers and information processing, Computer Hardware, Computer software, Managing data resources, Telecommunication, Enterprise: wide computing and networking.
Strategic and Managerial Implications of Information Systems: Strategic Information System: Introduction, Characteristics of Strategic Information Systems, Strategic Information Systems (SISP), Strategies for developing an SIS, Potential Barriers to developing a Strategic Information System (SIS), Decision Support System (DSS): Decision making concepts, methods, tools and procedures. Managing Information Resources: Introduction, IRM, Principal of Managing Information Resources, IRM functions, Computer Security: Introduction, Computer Security, Types of Computer Security, Disaster Recovery Plan.

Suggested Books:
1. W.S.Jawadakar,"Management Information System",McGraw Hill.
2. J.O.Brien,"Management Information System",TMH, New Delhi.

3. Uma G.Gupta,"Management Information System"Fifth Edition TMH.
4. Kenneth C. Laudon,"Management Information System Organization and Technology" TMH.

OE-CS-AIDS-306	Enterprise Resource Planning						
L	T	P	Credit	Major Test	Minor Test	Total	Time
3	-	-	3	75	25	100	3hrs
Purpose	To describe the concept of ERP and the ERP model; define key terms; explain the transition from MRP to ERP; identify the levels of ERP maturity and is to explain how ERP is used to integrate business processes; define and analyze a process; create a process map and improve and/or simplify the process; apply the result to an ERP implementation.						
Course Outcomes (CO)							
CO1	Design model for ERP for large projects and to design model for E-commerce architecture for any application						
CO2	Describe the advantages, strategic value, and organizational impact of utilizing an ERP system for the management of information across the functional areas of a business: sales and marketing, accounting and finance, human resource management, and supply chain.						
CO3	Demonstrate a working knowledge of how data and transactions are integrated in an ERP system to manage the sales order process, production process, and procurement process.						
CO4	Evaluate organizational opportunities and challenges in the design system with in a business scenario.						

UNIT-I

Introduction: ERP, Origin, Benefits, Evolution and Structure: Conceptual Model of ERP, the Evolution of ERP, the Structure of ERP, Business Process Reengineering, Dataware Housing, Data Mining, Online Analytic Processing (OLAP), Product Life Cycle Management (PLM), LAP, Supply chain Management.

UNIT-II

ERP Marketplace and Marketplace Dynamics: Market Overview, Marketplace Dynamics, the Changing ERP Market. ERP- Functional Modules: Introduction, Functional Modules of ERP Software, Integration of ERP, Supply chain and Customer Relationship Applications.

UNIT-III

ERP Implementation Basics, ERP Implementation Life Cycle, Role of SDLC/SSAD, Object Oriented Architecture, Consultants, Vendors and Employees.

UNIT-IV

ERP & E-Commerce, Future Directives-in ERP, ERP and Internet, Critical success and failure factors, Integrating ERP into organizational culture. Using ERP tool: either SAP or ORACLE format to case study.

Suggested Books:
1. Vinod Kumar Garg and Venkita Krishnan N K,"Enterprise Resource Planning Concepts and Practice", PHI.
2. Joseph A Brady, Ellen FMonk,Bret Wagner,"Concepts in Enterprise Resource Planning",Thomps on Course Technology.
3. Alexis Leon,"ERP Demystified",Tata McGraw Hill.
4. Rahul V.Altekar"Enterprise Resource Planning",Tata McGraw Hill.
5. Vinod Kumar Garg and Venkita Krishnan NK,"Enterprise Resource Planning–A Concepts and Practice", PHI.
6. Mary Summer,"Enterprise Resource Planning"-Pearson Education.

PC-CS-AIDS-310A	Soft Computing						
Lecture	**Tutorial**	**Practical**	**Credit**	**Major Test**	**MinorTest**	**Total**	**Time**
3	**0**	**0**	**3**	**75**	**25**	**100**	**3 Hrs.**
Purpose	Soft Computing is a consortia of methodologies which collectively provide a body of concepts and techniques for designing intelligent systems.						
Course Outcomes (CO)							
CO1	The main objective of the Soft Computing Techniques to Improve Data Analysis						
CO2	To strengthen the dialogue between the statistics and soft computing research communities in order to cross-pollinate both fields						
CO3	To develop Solutions and generate mutual improvement activities						
CO4	To develop practical data analysis skills, which can be applied to practical problems						

Unit-I

Introduction: What is Soft Computing. Difference between Hard and Soft computing, Requirement of Soft computing, Major Areas of Soft Computing, Applications of Soft Computing.

Unit-II

Neural Network, Learning rules and various activation functions, Single layer Perceptron, Back Propagation networks, Architecture of Back propagation (BP) Networks, Backpropagation Learning, Variation of Standard Back propagation Neural Network, Introduction to Associative Memory, Adaptive Resonance theory and Self Organizing Map, Recent Applications.

Unit-III

Fuzzy Systems: Fuzzy Set theory, Fuzzy versus Crisp set, Fuzzy Relation, Fuzzification, Minmax Composition, Defuzzification Method, Fuzzy Logic, Fuzzy Rule based systems, Predicate logic, Fuzzy Decision Making, Fuzzy Control Systems, Fuzzy Classification. Hybrid Systems: Sequential Hybrid Systems, Auxiliary Hybrid Systems, Fuzzy Backpropagation Networks: LR type Fuzzy numbers, Fuzzy Neuron, Fuzzy BP Architecture, Learning in Fuzzy BP, Application of Fuzzy BP Networks, Embedded Hybrid Systems, Neuro-Fuzzy Hybrid Systems, Neuro-Genetic Hybrid Systems, Fuzzy-Genetic Hybrid Systems.

Unit-IV

Genetic Algorithm: History of Genetic Algorithms (GA), Working Principle, Various Encoding methods, Fitness function, GA Operators-Reproduction, Crossover, Mutation, Convergence of GA, Bitwise operation in GA, Multi-level Optimization. GA based Backpropagation Networks: GA based Weight Determination, K-factor determination in Columns.

Suggested Books:

1. Principles of Soft Computing by S.N.Sivanandam & S.N.Deepaby Wiley,India edition.
2. Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis & Applications, S. Rajasekaran, G. A. Vijaya lakshami, PHI.
3. Genetic Algorithms: Search and Optimization, E.Goldberg.
4. Neuro-Fuzzy Systems, ChinTengL in,C.S.George Lee,PHI.
5. Build_Neural_Network_With_MS_Excel_sample by Joechoong.

PC-CS-AIDS-312LA	Applied Machine Learning Lab						
Lecture	Tutorial	Practical	Credit	Minor Test	Practical	Total	Time
0	0	2	1	40	60	100	3 Hrs.
Purpose	To understand and implement advanced Machine Learning operations in Python.						
Course Outcomes-At the end of this course students will be able to:							
CO1	Perform advanced data cleaning, exploration, and visualization						
CO2	Engineer features based on conditional relationships between existing features						
CO3	Build and finalize a machine learning classifier						
CO4	Build machine learning applications in different domains						

List of Practicals

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.CSVfile.

2. 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.

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

4. Build an Artificial Neural Network by implementing the Back propagation algorithm and test the same using appropriate data sets.

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

6. 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.

7. 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.

8. 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.

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

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

PC-CS-AIDS-314LA	Big Data Analytics Lab						
Lecture	Tutorial	Practical	Credit	Minor Test	Practical	Total	Time
0	0	2	1	40	60	100	3 Hrs.
Purpose	To understand and implement advanced Big Data operations in Hadoop architecture.						
Course Outcomes-At the end of this course students will be able to:							
CO1	**Demonstrate** the knowledge of big data analytics and implement different file management task in Hadoop.						
CO2	**Understand** Map Reduce Paradigm and develop data applications using variety of systems.						
CO3	**Analyze** and perform different operations on data using Pig Latin scripts.						
CO4	**Illustrate** and apply different operations on relations and databases using Hive.						

List of Practicals

1. To Study of Big Data Analytics and Hadoop Architecture.

2. Installation of Single Node Hadoop Cluster on Ubuntu

3. Hadoop Programming: Word Count Map Reduce Program Using Eclipse

4. Implementing Matrix Multiplication Using One Map-Reduce Step.

5. Implementing Relational Algorithm on Pig.

6. Implementing database operations on Hive.

7. Implementing Bloom Filter using Map-Reduce

8. Implementing Frequent Item set algorithm using Map-Reduce.

9. Implementing Clustering algorithm using Map-Reduce

10. Implementing Page Rank algorithm using Map-Reduce

11. Mini Project:

    Few topics for Projects:

        a. Twitter data analysis

        b. Fraud Detection

        c. Text Miningd. Equity Analysis etc.

Few websites for sample data: www.data.gov.in, www.nseindia.in, www.censusindia.gov.in,

www.importexportdata.in.

ES-CS-AIDS-316LA	Applied Statistical Analysis for AI Lab						
Lecture	Tutorial	Practical	Credit	Minor Test	Practical	Total	Time
0	0	2	1	40	60	100	3 Hrs.
Purpose	To implement statistical analysis functions in R language.						
Course Outcomes -At the end of this course students will be able to:							
CO1	Implement basic Statistical operations in R language.						
CO2	Implement regression techniques.						
CO3	Implement hypothesis testing with real time applications.						
CO4	Implement and evaluate various probability distributions for real world problems.						

List of Practicals

1. Calculation of coefficient of correlation.

2. Calculation spearman rank correlation.

3. Simple linear regression, residuals,estimate of intercept, regression coefficients.

4. Residual plots, regression diagnostics.

5. Multiple linear regression and regression estimates.

6. Calculation of multiple correlation and partial correlation.

8. Polynomial regression and regression estimates.

9. Calculation of Type I and Type II error probabilities.

10. Calculation of size of critical region, power of the test for the mean of a normal distribution with known and unknown variance and plotting graph of the power function.

11. Calculation of size of most powerful critical region (NP lemma).

12. Evaluating shortest confidence interval for mean of normal distribution when variance is known / unknown.

13. Evaluating shortest confidence interval for variance of normal distribution when mean is known / unknown.

14. Calculation of power of the test for the Bernoulli distribution with probability(p)in case of simple hypothesis and power curves and plotting the graph of the power function.

15. Calculation of likelihood ratio test(LRT)for simple hypothesis and composite hypothesis.