Proposed Structure of PG Engineering Program

Master of

Engineering

(M.E.)

Computer Science & Engineering

2020-21

IPS ACADEMY

INSTITUTE OF ENGINEERING & SCIENCE, INDORE

(A UGC Autonomous Institute affiliated to RGPV)

Semester I (First Year)
Course: M.E.

S.No.	Course Code	Course Title	Hrs./ week			Credits
			L	T	P	
1	PSCC-MCS101	Artificial Intelligence: Principal and Technique	3	1	0	4
2	PSCC-MCS102	Advanced Data Structures and Algorithm	3	1	0	4
3	PSMC-MCS101	Mathematics for AI	3	0	0	3
4	PSEC-MCS101	Program Specific Elective Course-I	3	0	0	3
5	LC-MCS101	AI Lab	0	0	4	2
6	LC-MCS102	Advanced Data Structures Lab	0	0	4	2
7	MLC-MCS101	Advanced DBMS	2	0	0	2
8	AUD-MCS101	Audit Course-I	2	0	0	0
					Total credits	**20**

Program Specific Elective-I	**Audit Course-I**
Computer Vision	Disaster Management
Information Retrieval	Stress Management
Big Data Analytics	Economics Policies in India
Natural Language Processing	ICT for development

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

PSCC-MCS101	**Artificial Intelligence: Principal and Technique**	**3L : 1T (4 hrs.)**	**4 credits**

Prerequisite: Computer programming, Mathematics, Analysis and Design of Algorithms

Course Objective:

The objective of this course is to provide a strong foundation of fundamental concepts in Artificial Intelligence.

Course Contents: (40 hrs.)

Module 1: **(08 hrs.)**
Meaning and definition of artificial intelligence, types of production systems, Characteristics of production systems, Study and comparison of breadth first search and depth first search Techniques, Other Search Techniques like hill Climbing, Best first Search, A* algorithm, types of control strategies.

Module 2: **(10 hrs.)**
Knowledge Representation, Problems in representing knowledge, knowledge representation using propositional and predicate logic, comparison of propositional and predicate logic, Resolution, refutation, deduction, theorem proving, inferencing, monotonic and nonmonotonic reasoning.

Module 3: **(6 hrs.)**
Probabilistic reasoning, Baye's theorem, semantic networks, scripts, schemas, frames, conceptual dependency, fuzzy logic, forward and backward reasoning.

Module 4: **(08 hrs.)**
Game playing techniques like minimax procedure, alpha-beta cut-offs etc, planning, Study of the block world problem in robotics, Introduction to understanding and natural languages processing.

Module 5: **(08 hrs.)**
Introduction to learning, Various techniques used in learning, introduction to neural networks, applications of neural networks, common sense, reasoning, some example of expert systems.

Expert Systems: Introduction to expert system and application of expert systems, various expert system shells, vidwan frame work, knowledge acquisition, case studies, MYCIN.

Learning: Rote learning, learning by induction, explanation based learning

Course Outcome:

1. State the core concepts of Artificial Intelligence, It's foundation and principles.
2. Examine the useful search techniques; learn their advantages, disadvantages and be able to develop intelligent systems.
3. Learn the practical applicability of intelligent systems, specifically its applications.
4. Understand AI in different areas like NLP, Pattern Recognition, game planning etc.
5. Understand important concepts like Expert Systems, AI applications.

List of Text / Reference Books:

1. Elaine Rich and Kevin Knight "Artificial Intelligence" - Tata McGraw Hill, Third Edititition,2017.
2. Nelsson N.J., "Principles of Artificial Intelligence", Springer-Verlag Berlin Heidelberg, 1st Edition 1982.
3. Dan W. Patterson "Introduction to Artificial Intelligence and Expert Systems", Prentice-Hall of India Pvt.Ltd, 1990.
4. Clocksin & C.S.Melish "Programming in PROLOG", Narosa Publishing House, Fifth Edition 2008.
5. M.Sasikumar,S.Ramani etc. "Rule based Expert System", Narosa Publishing House, First Edition, 2007.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

PSCC-MCS102	Advanced Data Structures and Algorithm	3L:1T (4 hrs.)	4 credits

Prerequisite: Data Structure, Basic knowledge of C/C++,

Course Objective:

The objective of this course is intended to provide the foundations of the practical implementation and usage of Algorithms and Data Structures.

Course Contents: (40 hrs.)

Module 1: **(08 hrs.)**
Introduction to Data Structure: Concepts of Data and Information, Classification of Data Structures, Abstract Data Types, Implementation Aspects: Memory Representation, Data Structures Operations and its Cost Estimation, Introduction to linear data structures- Arrays, Linked List: Representation of linked list in memory, different implementation of linked list, Circular linked list, doubly linked list, Application of linked list: polynomial manipulation using linked list.

Module 2: **(08 hrs.)**
Stacks: Stacks as ADT, Different implementation of stack, multiple stacks, Application of Stack: Conversion of infix to postfix notation using stack, evaluation of postfix expression, Recursion, Queues: Queues as ADT, Different implementation of queue, Circular queue, Concept of Dqueue and Priority Queue, Queue simulation, Application of queues.

Module 3: **(12 hrs.)**
Algorithms, Designing algorithms, analyzing algorithms, asymptotic notations, heap and heap sort, Introduction to divide and conquer technique, analysis, design and comparison of various algorithms based on this technique, example binary search, merge sort, quick sort, strassen's matrix multiplication,

Study of Greedy Strategy, examples of greedy method like Optimal Merge Patterns, Huffman Coding, minimum spanning trees, knapsack problem, job sequencing with deadlines, Shortest Path Algorithm

Module 4: **(10 hrs.)**
Backtracking concept and its examples like 8 queen's problem, Hamiltonian cycle, Graph coloring problem, Introduction to branch & bound method, Examples of branch and bound method like traveling salesman problem, Meaning of lower bound theory and its use in solving algebraic problem, introduction to parallel algorithms.

Module 5: **(06 hrs.)**

Graphs: Introduction, Classification of graph: Directed and Undirected graphs, Representation, Graph algorithm: Minimum Spanning Tree (MST)- Kruskal, Prim's algorithms, Dijkstra's shortest path algorithm; Comparison between different graph algorithms, Application of graphs, Binary search trees, height balanced trees, 2-3 trees, B-trees, basic search and traversal techniques for trees and graphs (In order, preorder, postorder, DFS, BFS), NP-completeness.

Course Outcome:

1. Design and analyze programming problem statements.
2. Choose appropriate data structures and algorithms, understand the ADT/libraries, and use it to design algorithms for a specific problem.
3. Learn basic concepts of algorithms and sorting & searching techniques.
4. Understand the necessary mathematical abstraction to solve problems and Come up with analysis of efficiency and proofs of correctness
5. Comprehend and select algorithm design approaches in a problem specific manner for tree and graph.

List of Text / Reference Books:

1. AM Tanenbaum, Y Langsam& MJ Augustein, "Data structure using C and C++", Prentice Hall India, Second Addition, 2000.
2. Robert Kruse, Bruse Leung, "Data structures & Program Design in C", Pearson Education, Second Addition, 2000.
3. E. Horowitz, Sahni and D. Mehta, "Analysis & Design of Algorithm", 2002, Galgotia Publication 2002
4. Aho, Hopcroft, Ullman, "Data Structures and Algorithms", Pearson Education.
5. N. Wirth, "Algorithms + Data Structure = Programs", Prentice Hall.
6. Richard, Gilberg Behrouz, Forouzan, "Data structure – A Pseudocode Approach with C", Thomson press.
7. Coremen Thomas, Leiserson CE, Rivest RL, "Introduction to Algorithms" PHI.
8. Ullmann, "Analysis & Design of Algorithm".
9. Michael T Goodrich, Robarto Tamassia, "Algorithm Design", Wiely Publication, India
10. Rajesh K Shukla, "Analysis and Design of Algorithms: A Beginner's Approach", Wiley publications, 2015.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

Branch	Subject Title	Subject Code		
Computer Science and Engineering	Mathematics for AI	PSMC-MCS101	3L:0T:0P	3 Credits

Course Objective: This course introduces the applications of Advanced Computational Mathematics in the Computer Science and Engineering. It covers the concepts of linear algebra, solution of partial differential equations, integral transforms, concepts of probability, stochastic process, queuing theory, fuzzy sets, fuzzy logic and MATLAB which are used in solving problems related with Computer Science and Engineering.

Module 1: Linear Algebra (7 Hours)

Linear transformation, Vector spaces, Hash function, Hermite polynomial, Heaviside's unit function and Error function. Elementary concepts of modular mathematics.

Module 2: Partial Differential Equations and Transforms Calculus (8 Hours)

Solution of partial differential equation (PDE) by separation of variable method, Numerical solution of PDE (Laplace, Poisson's, Parabolic) using finite difference methods, Elementary properties of FT, DFT, WFT, wavelet transform.

Module 3: Concepts of Probability (8 Hours)

Probability, Compound probability and discrete random variable. Binomial, Normal and Poisson's distributions, Sampling distribution, Elementary concept of estimation and Theory of hypothesis, Recurred relations.

Module 4: Stochastic Process and Queuing theory (9 Hours)

Stochastic process, Markov process transition, Probability transition, Probability matrix, First and higher order markov process, Application of eigen value problems in markov process, Markov chain. Queuing system, Transient and steady state, Traffic intensity, Distribution queuing system, Concepts of queuing models (M/M/1: Infinity/ Infinity/ FCFS), (M/M/1: N/ Infinity/ FC FS), (M/M/S: Infinity/ Infinity/ FC FS).

Module 5: Fuzzy Sets, Fuzzy Logic and MATLAB (8 Hours)

Operations of fuzzy sets, Fuzzy arithmetic & relations, Fuzzy relation equations, Fuzzy logics. MATLAB introduction, Programming in MATLAB scripts, Functions and their application.

Course Outcomes:

CO 1. To identify and solve problems using the concepts of linear algebra.
CO 2. To understand and use the concepts of solution of partial differential equations and transform calculus in practical problems appear in Computer Science and Engineering.
CO 3. To identify and solve problems regarding probability and hypothesis.
CO 4. To explain and apply the concept of stochastic process and queuing theory in engineering problems.
CO 5. To understand and apply the concept of fuzzy sets, fuzzy logic and MATLAB in different problems of Computer Science and Engineering.

Textbooks/References:

1. M. K. Jain, Numerical Solution of Differential equations, John Wiley & Sons, 2018.
2. R. N.Bracewell, Fourier Transform & It's Applications, Tata McGraw Hill, 2014.
3. S. C. Chapra, Applied Numerical Methods with MATLAB, Tata McGraw Hill, 2013.
4. T. J. Ross, Fuzzy Logic with Engineering Application, Wiley publisher, 2016.
5. S. Ross, A First Course in Probability, Pearson education India, 9th edition, 2013.
6. Erwin kreyszig, Advanced Engineering Mathematics, John Wiley & Sons, 10th Edition, 2015.
7. R. J. Beerwends, Fourier and Laplace transform, Cambridge university press, 2003.
8. E. J. Watson, Laplace Transform and Applications, Publisher: Van Nostrand Reinhold, 1980.
9. P. G. Hoel, S. C. Port, C. J. Stone, Introduction to Probability Theory, Universal Book Stall, 2012.
10. H. J. Zimmersoms, Fuzzy Sets Theory & its Applications, 4th edition 2007.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering

I-Semester

LC-MCS101	AI Lab	4P (4 hrs.)	2 credits

Prerequisite: Computer programming (C/C++), Analysis and Design of Algorithms, Mathematics

Course Objective:
To provide students with and practical base in Artificial Intelligence.

List of Experiments:

Week 1
1. Write a program to implementation of DFS.
2. Write a program to implementation of BFS.

Week 2
1. Write a Program to find the solution for traveling salesman Problem.

Week 3
1. Write a program to implement Simulated Annealing Algorithm.
2. Write a program to find the solution for wampus world problem.

Week 4
1. Write a program to implement 8 puzzle problems.

Week 5
1. Write a program to implement Tower of Hanoi problem

Week 6
1. Write a program to implement A* Algorithm

Week 7
1. Write a program to implement Hill Climbing Algorithm

Week 8
1. To Study JESS expert system

Week 9
1. To Study RVD expert system

Week 10
1. Write a Program to Perform Fibonacci Series
2. Write a Program to Check Sides of a Triangle

Course Outcomes:

1. Able to understanding of the major areas and challenges of AI.
2. Ability to apply basic AI algorithms to solve problems.
3. Able to describe search strategies and solve problems by applying a suitable search method.
4. Able to describe and apply knowledge representation.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

LC-MCS102	Advanced Data Structures Lab	4P (4 hrs.)	2 credits

Prerequisite: Data Structure, Basic knowledge of C/C++,

Course Objectives:

1. Understand the foundations of the practical implementation and usage of Algorithms and Data Structures.

List of Experiments:

Week 1

1. Write a program for Iterative and Recursive Binary Search.

Week 2

1. Write a program for Merge Sort.
2. Write a program for Quick Sort.

Week 3

1. Write a program for Strassen's Matrix Multiplication.
2. Write a program for optimal merge patterns.
3.

Week 4

1. Write a program for Huffman coding.

Week 5

1. Write a program for minimum spanning trees using Kruskal's and Prims algorithm.
2. Write a program for single sources shortest path algorithm.

Week 6

1. Write a program for Floye-Warshal algorithm.
2. Write a program for traveling salesman problem and Hamiltonian cycle problem.

Week 7

1. Min/Max Heap
2. Leftist Heap

Week 8

1. AVL Trees
2. Red-Black Trees

3. B-Trees
4. Segment Trees

Week 9

1. Line segment intersection

Week 10

1. Convex Hull
2. Voronoi Diagram

Course Outcomes:

1. Design and analyze programming problem statements.
2. Choose appropriate data structures and algorithms, understand the ADT/libraries, and use it to design algorithms for a specific problem.
3. Learn basic concepts of algorithms and sorting & searching techniques.
4. Implement heap and various tree structure like AVL, Red-black, B and Segment trees.
5. Solve the problems such as line segment intersection, convex shell and Voronoi diagram.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

MLC-MCS101	Advanced DBMS	2L (2 hrs.)	2 credits

Prerequisite: Database Management System
.

Course Objective:

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

Course Contents: (40 hrs.)

Module 1: (10 hrs.)
An overview of database, The Extended Entity Relationship Model and Object Model: The ER model revisited, Motivation for complex data types, User defined abstract data types and structured types, Subclasses, Super classes, Inheritance, Specialization and Generalization, Constraints and characteristics of specialization and Generalization, Relationship types.

Module 2: (8 hrs.)
Query Processing, Optimization & Database Tuning: Algorithms for Executing Query Operations, Heuristics for Query Optimizations, Estimations of Query Processing Cost, Join Strategies for Parallel Processors, Database Workloads, Tuning Decisions, DBMS Benchmarks, Clustering & Indexing, Multiple Attribute Search Keys, Query Evaluation Plans, Pipelined Evaluations, System Catalogue in RDBMS.

Module 3: (8 hrs.)
Distributed Database System: Structure of Distributed Database, Data Fragmentation, Data Model, Query Processing, Semi Join, Parallel & Pipeline Join, Distributed Query Processing, Concurrency Control in Distributed Database System, Recovery in Distributed Database System, Distributed Deadlock Detection and Resolution, Commit Protocols.

Module 4: (08 hrs.)
Enhanced Data Model for Advanced Applications: Database Operating System, Introduction to Temporal Database Concepts, Spatial And Multimedia Databases, Data Mining, Active Database System, Deductive Databases, Database Machines, Web Databases, Advanced Transaction Models, Issues in Real Time Database Design.

Module 5: (06 hrs.)
Accessing databases from Web, JavaScript, JDBC, Java Servlets, database technology to Web related areas such as semi-structured databases and data integration, XML, XQuery, XPath, XML Schemas, distributed database design, transactions and query processing.

Course Outcome:

1. Explain and evaluate the fundamental theories and requirements that influence the design of modern database systems
2. Assess and apply database functions and packages suitable for enterprise database development and database management
3. Critically evaluate alternative designs and architectures for databases and discuss and evaluate methods of storing, managing and interrogating complex data.
4. Explain and critically evaluate different database solutions for data exchange.
5. Analyze the background processes involved in queries and transactions, and explain how these impact on database operation and design

List of Text / Reference Books:

1. R. Ramakrishnan, J. Gehrke, "Database Management Systems", McGraw Hill, 2004
2. A. Silberschatz, H. Korth, S. Sudarshan, "Database system concepts", 5 Edition, McGraw Hill, 2008.
3. Martin Kleppmann, "Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems", 1st Ed, O'Reilly 2017.
4. Date C. J., "An Introduction to Database Systems", Addison Wesley Longman (8th Ed), 2003.
5. Silberschatz A., Korth H., and Sudarshan S., "Database System Concepts", McGraw-Hill (6th Ed), 2010.
6. Thomas M. Connolly, Carolyn Begg, "Database Systems: practical approach to design, implementation, and management", Pearson Education Limited, (6th edition), 2015.
7. Wilfried Lemahieu, Seppe vanden Broucke, Bart Baesens, "Principles of Database Management: Practical Guide to Storing, Managing and Analyzing Big and Small Data", Cambridge University press (1st Ed), 2018.
8. Peter Adams : SQL: The Ultimate Guide from Beginner to Expert - Learn and Master SQL in No Time, Addison Wesley, 2016.
9. Majumdar & Bhattacharya, "Database Management System", TMH.
10. Elmasri, Navathe, "Fundamentals of Database Systems", Addison Wesley.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

PSEC-MCS101	Computer Vision	3L (3 hrs.)	3 credits

Prerequisite: Data structures, Computer Graphics, Mathematics

Course Objective:

The main objective of this course is to introduce students the fundamentals of image formation, methods and techniques of computer vision and pattern recognition and to develop an appreciation for various issues in the design of computer vision and object recognition systems.

Course Contents: (40 hrs.)

Module 1: **(10 hrs.)**
Introduction: Image Processing, Computer Vision and Computer Graphics, What is Computer Vision-Low-level, Mid-level, High-level, Overview of Diverse Computer Vision Applications: Document Image Analysis, Biometrics, Object Recognition, Tracking, Medical Image Analysis, Content-Based Image Retrieval, Video Data Processing, Multimedia, Virtual Reality and Augmented Reality

Module 2: **(10 hrs.)**
Image Formation Models: Monocular imaging system, Radiosity: The 'Physics' of Image Formation, Radiance, Irradiance, BRDF, color etc, Orthographic & Perspective Projection, Camera model and Camera calibration, Binocular imaging systems, Multiple views geometry, Structure determination, shape from shading, Photometric Stereo, Depth from Defocus, Construction of 3D model from images.

Module 3: **(8 hrs.)**
Image Processing and Feature Extraction: Image preprocessing, Image representations (continuous and discrete), Edge detection, Motion Estimation: Regularization theory, Optical computation, Stereo Vision, Motion estimation, Structure from motion.

Module 4: **(08 hrs.)**
Shape Representation and Segmentation: Contour based representation, Region based representation, Deformable curves and surfaces, Snakes and active contours, Level set representations, Fourier and wavelet descriptors, Medial representations, Multi resolution analysis

Module 5: **(04 hrs.)**
Object recognition: Hough transforms and other simple object recognition methods, Shape correspondence and shape matching

Course Outcome:

1. To implement fundamental image processing techniques required for computer vision.
2. Understand Image formation process.
3. Extract features form images and do analysis of images.
4. To perform shape analysis and generate 3D model from images.
5. Understand video processing, object recognition, motion computation and 3D vision and geometry

List of Text / Reference Books:

1. R. C. Gonzalez, R. E. Woods. “Digital Image Processing”, Addison Wesley Longman, Inc., 1992.
2. D. H. Ballard, C. M. Brown., “Computer Vision”, Prentice-Hall, Englewood Cliffs, 1982.
3. Richard Szeliski, “Computer Vision: Algorithms and Applications (CVAA)”, Springer, 2010.
4. E. R. Davies, “Computer & Machine Vision”, Fourth Edition, Academic Press, 2012.
5. Simon J. D. Prince, “Computer Vision: Models, Learning, and Inference”, Cambridge University Press, 2012.
6. Mark Nixon and Alberto S. Aquado, “Feature Extraction & Image Processing for Computer Vision”, Third Edition, Academic Press, 2012.
7. Sonka, Hlavac, and Boyle Thomson, “Image Processing, Analysis, and Machine Vision”.
8. D. Forsyth and J. Ponce, “Computer Vision - A modern approach”, Prentice Hall Robot Vision, by B. K. P. Horn, McGraw-Hill.
9. E. Trucco and A. Verri, “Introductory Techniques for 3D Computer Vision”, by, Publisher: Prentice Hall.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

PSEC-MCS101	Information Retrieval	3L (3 hrs.)	3 credits

Prerequisite: Data Mining, Data Structure and Algorithms

Course Objective:

The main objective of this course is to present the scientific support in the field of information search and retrieval. This course explores the fundamental relationship between information retrieval, hypermedia architectures, and semantic models, thus deploying and testing several important retrieval models such as vector space, Boolean and query expansion.

Course Contents: (30 hrs.)

Module 1: **(8 hrs.)**
Introduction to Information Retrieval, Retrieval strategies: vector space model, Probabilistic retrieval strategies: Simple term weights, Non binary independence model, Language models.

Module 2: **(6 hrs.)**
Retrieval Utilities: Relevance feedback, clustering, N-grams, Regression analysis, Thesauri

Module 3: **(6 hrs.)**
Retrieval utilities: Semantic networks, parsing, Cross–Language: Information Retrieval: Introduction, Crossing the Language barrier.

Module 4: **(4 hrs.)**
Efficiency: Inverted Index, Query processing, Signature files, Duplicate document detection.

Module 5: **(06 hrs.)**
Integrating structured data and text: A historical progression, Information retrieval as relational application, Semi Structured search using a relational schema, Distributed Information Retrieval: A theoretical Model of Distributed retrieval, web search

Course Outcome:

1. Understanding information retrieval strategies with examples.
2. Knowledge gathering about Retrieval Utilities and relevance feedback.
3. Applying and examine case studies
4. Understanding concept of query processing.
5. Knowledge gathering about integrated structure data, text and models.

List of Text / Reference Books:

1. David A. Grossman, OphirFrieder, “Information Retrieval – Algorithms and Heuristics”, Springer, 2nd Edition(Distributed by Universal Press), 2004
2. Gerald J Kowalski, Mark T Maybury, “Information Storage and Retrieval Systems: Theory and Implementation”, Springer, 2004.
3. Soumen Chakrabarti, “Mining the Web: Discovering Knowledge from Hypertext Data, Morgan”, Kaufmann Publishers, 2002.
4. Christopher D Manning, Prabhakar Raghavan, HinrichSchutze, “An Introduction to Information Retrieval”, Cambridge University Press, England, 2009.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering

I-Semester

PSEC-MCS101	**Big Data Analytics**	**3L (3 hrs.)**	**3 credits**

Prerequisite: Data structures and algorithms, Data Base Management System

Course Objective: The main goal of this course is to help students learn, understand and practice big data analytics and machine learning approaches, which include the study of modern computing big data technologies and scaling up machine learning techniques focusing on industry applications.

Course Contents: (40 hrs.)

Module 1: **(06 hrs.)**

Introduction to Big data, Big data characteristics, Types of big data, Traditional versus Big data, Evolution of Big data, challenges with Big Data, Technologies available for Big Data, Infrastructure for Big data, Use of Data Analytics, Desired properties of Big Data system.

Module 2: **(10 hrs.)**

Introduction to Hadoop, Core Hadoop components, Hadoop Eco system, Hive Physical Architecture, Hadoop limitations, RDBMS Versus Hadoop, Hadoop Distributed File system, Processing Data with Hadoop, Managing Resources and Application with Hadoop YARN, MapReduce programming.

Module 3: **(10 hrs.)**

Introduction to Hive: Hive Architecture, Hive Data types, Hive Query Language, Introduction to Pig, Anatomy of Pig, Pig on Hadoop, Use Case for Pig, ETL Processing, Data types in Pig, running Pig, Execution model of Pig, Operators, functions.

Module 4: **(08 hrs.)**

Introduction to NoSQL, NoSQL Business Drivers, NoSQL Data architectural patterns, Variations of NOSQL, Architectural patterns using NoSQL, Introduction to MangoDB

Module 5: **(06 hrs.)**

Mining social Network Graphs: Introduction Applications of social Network mining, Social Networks as a Graph, Types of social Networks, Clustering of social Graphs, Direct Discovery of communities in a social graph, Introduction to recommender system.

Course Outcome:

1. Students should be able to understand the concept and challenges of Big data.
2. Students should be able to demonstrate knowledge of big data analytics.
3. Students should be able to develop Big Data Solutions using Hadoop Eco System.
4. Students should be able to gain hands-on experience on large-scale analytics tools.
5. Students should be able to analyze the social network graphs.

List of Text / Reference Books:

1. Radha Shankarmani, M. Vijaylakshmi, " Big Data Analytics", Wiley, Second edition, Januray 2016.
2. Seema Acharya, Subhashini Chellappan, " Big Data and Analytics", Wiley, Second edition, 2015.
3. Michael Minelli, Michele Chambers, AmbigaDhiraj, “Big Data Big Analytics”,Wiley, First edition, February 2013.
4. Tom White, “HADOOP: The definitive Guide”, O Reilly 2012
5. Chris Eaton, Dirk deroos, “Understanding Big Data”, McGraw Hill, 2012.

IPS Academy
Institute of Engineering & Science
Department of Computer Science & Engineering
I-Semester

PSEC-MCS101	**Natural Language Processing**	**3L (3 hrs.)**	**3 credits**

Prerequisite: Artificial Intelligence, Compiler Design, Theory of Computation

Course Objective:

This course introduces the fundamental concepts and techniques of natural language processing (NLP). Students will gain an in-depth understanding of the computational properties of natural languages and the commonly used algorithms for processing linguistic information.

Course Contents: (40 hrs.)

Module 1: **(06 hrs.)**
Introduction to AI and NLP, History of NLP, Study of Human Languages, Ambiguity and Uncertainty in language, NLP Phases, Introduction to Linguistic Resources, Corpus, Elements of Corpus Design, Types of Treebank Corpus, Applications of Treebank Corpus

Module 2: **(08 hrs.)**
Parsing techniques, context free grammer, recursive transitions nets (RNT), augmented transition nets (ATN), case and logic grammers, semantic analysis, Regular Expression, Properties and Examples, Finite State Autometa, Concept of derivation, types of derivations

Module 3: **(10 hrs.)**
Game playing, Minimax search procedure, alpha-beta cutoffs, additional refinements, Planning, Overview an example domain the block word, component of planning systems, goal stack planning, non linear planning.

Module 4: **(10 hrs.)**
Approaches and Methods to Word Sense Disambiguation (WSD), Applications of Word Sense Disambiguation (WSD), Difficulties in WSD, Concept of Coherence, Discourse structure, Algorithms for Discourse Segmentation, Text Coherence, Building Hierarchical Discourse Structure, Reference Resolution, Terminology Used in Reference Resolution, Types of Referring Expressions, Reference Resolution Tasks.

Module 5: **(06 hrs.)**
Natural Language Processing — Part of Speech (PoS) Tagging, Natural Language Processing: Natural Language Inception, Information Retrieval, Applications of NLP, Types of Machine Translation Systems, Approaches to Machine Translation (MT), Fighting Spam, Sentiment Analysis, Introduction to Python.

Course Outcome:

1. Understand approaches to syntax and semantics in NLP.
2. Understand approaches to discourse, generation, dialogue and summarization within NLP.
3. Understand current methods for statistical approaches to machine translation.
4. Acquiring knowledge about Word Sense Disambiguation and its applications.
5. Understand machine learning techniques in NLP and various algorithms applied within NLP.

List of Text / Reference Books:

1. Jurafsky and Martin, "Speech and Language Processing", Prentice Hall, 2000.
2. Lutz and Ascher, "Learning Python", O'Reilly, ISBN: 0596002815.
3. Manning and Schutze, "Statistical Natural Language Processing", MIT Press; 1st edition (June 18, 1999).
4. Tom Mitchell, "Machine Learning", McGraw Hill, 1997, ISBN 0070428077.
5. Cover, T. M. and J. A. Thomas, " Elements of Information Theory" Wiley Publications, 1991.
6. Charniak, E, "Statistical Language Learning", The MIT Press. 1996.
7. Jelinek F, "Statistical Methods for Speech Recognition", The MIT Press, 1998.