



(An Autonomous Institution - AFFILIATED TO ANNA UNIVERSITY, CHENNAI)

S.P.G.Chidambara Nadar - C.Nagammal Campus

S.P.G.C. Nagar, K.Vellakulam – 625 701 (Near VIRUDHUNAGAR).

B.TECH. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
REGULATION – 2021
AUTONOMOUS SYLLABUS
CHOICE BASED CREDIT SYSTEM
VII TO VIII SEMESTER CURRICULUM AND SYLLABI

VISION:

To make the Department of Computer Science and Engineering of this Institution the unique of its kind in the field of Research and Development activities in this part of world.

MISSION:

To impart highly innovative and technical knowledge to the urban and unreachable rural student folks in Computer Science and Engineering through "Total Quality Education".

PROGRAMME EDUCATIONAL OBJECTIVES (PEOs):

- PEO 1:** Apply the basic engineering skills and domain knowledge for developing effective computing solutions to address various social issues.
- PEO 2:** Able to have successful career in technical / managerial roles in multi-disciplinary environment.
- PEO 3:** To confront the evolving technical challenges and problems in the areas of computing.

PROGRAM OUTCOMES:

After going through the four years of study, the Artificial Intelligence and Data Science graduates will have the ability to

POs	Graduate Attribute	Programme Outcome
1	Engineering knowledge	Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2	Problem analysis	Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3	Design/Development of solutions	Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4	Conduct investigations of complex problems	Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions
5	Modern tool usage	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations
6	The engineer and society	Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice

7	Environment and sustainability	Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8	Ethics	Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9	Individual and team work	Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10	Communication	Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11	Project management and finance	Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12	Life-long learning	Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAMME SPECIFIC OUTCOMES (PSOs):

PSO1 : Professional Skills: To apply learned skills to build optimized solutions pertaining to Data Processing, Artificial Intelligence and Machine Learning.

PSO2 : Problem - Solving Skills: To analyze data using domain knowledge to get insights and develop appropriate solutions.

REGULATION - 2021
CHOICE BASED CREDIT SYSTEM
B.TECH. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
CURRICULUM AND SYLLABI FOR SEMESTER VII TO VIII

SEMESTER VII

S.NO.	COURSE CODE	COURSE TITLE	CATE GORY	CONTACT PERIODS	L	T	P	C
THEORY								
1	AI2401	Data Exploration and Visualization	PC	3	3	0	0	3
2	GE2401	Universal Human Values and Ethics	HS	2	2	0	0	2
3		Management Elective	HS	3	3	0	0	3
4		Open Elective - II*	OE	3	3	0	0	3
5		Open Elective - III*	OE	3	3	0	0	3
6		Open Elective - IV*	OE	3	3	0	0	3
PRACTICALS								
7	AI2402	Data Visualization Laboratory	PC	4	0	0	4	2
TOTAL				21	17	0	4	19

* Open Elective Shall be chosen from the list of open electives offered by other Programmes

SEMESTER VIII

S.NO.	COURSE CODE	COURSE TITLE	CATE GORY	CONTACT PERIODS	L	T	P	C
THEORY								
1	AI2451	Project Work	EM	20	0	0	20	10
TOTAL				20	0	0	20	10

MANAGEMENT ELECTIVE

SL. NO.	COURSE CODE	COURSE NAME	CATE GORY	CONTACT PERIODS	L	T	P	C
1	GE2491	Principles of Management	HS	3	3	0	0	3
2	GE2492	Total Quality Management	HS	3	3	0	0	3

OPEN ELECTIVE – II (Offered to EEE, BT, Civil, Mech, MTR)

SL. NO.	COURSE CODE	COURSE NAME	CATE GORY	CONTACT PERIODS	L	T	P	C
1	OAI701	Fundamentals of Big Data	OE	3	3	0	0	3

OPEN ELECTIVE – III (Offered to ECE, EEE, BT, Civil, Mech, MTR)

SL. NO.	COURSE CODE	COURSE NAME	CATE GORY	CONTACT PERIODS	L	T	P	C
1	OAI702	Information Retrieval Basics	OE	3	3	0	0	3

OPEN ELECTIVE – IV (Offered to ECE, EEE, BT, Civil, Mech, MTR)

SL. NO.	COURSE CODE	COURSE NAME	CATE GORY	CONTACT PERIODS	L	T	P	C
1	OAI703	Introduction to Operating Systems	OE	3	3	0	0	3

Course Code	Course Name	L	T	P	C
AI2401	DATA EXPLORATION AND VISUALIZATION	3	0	0	3

Category: Professional Core Course

a. Preamble

This course focuses on the fundamental concepts of Data Visualization and Storytelling through the Tableau interface. This course enables the students to organize and present information intuitively and also get exposure to applying various visualization tools to develop real-time applications.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Illustrate the fundamentals of Data Analysis and transformation techniques used to explore the data	K2
CO2	Demonstrate various File types and Data connections in the Tableau environment	K2
CO3	Build meaningful Dashboards and Stories to simplify and slice the data	K3
CO4	Apply different chart and map forms to visualize and insight into the data	K3
CO5	Identify Statistics tools to perform advanced analysis using Tableau	K3

c. Course Syllabus

Total: 45 Periods

EXPLORATORY DATA ANALYSIS

9

Fundamentals of Exploratory Data Analysis (EDA) - Understanding Data Science - Significance of EDA - Making sense of Data - Comparing EDA with classical and Bayesian analysis - Software tools for EDA - Visual aids for EDA - Data Transformation techniques - Pivot Tables and Cross-tabulations.

FUNDAMENTALS OF DATA VISUALIZATION

9

Introduction to Visualization - Top Data Visualization Tools - Reasons to make a switch to Tableau - Positioning of Tableau - File types in Tableau - Tableau environment - Connection to various data Sources - Custom SQL.

DASHBOARD AND STORIES

9

Filtering - Sorting - Grouping - Number functions - String functions - Logical functions - Date functions - Aggregate functions - Creating a dashboard - Dashboard actions - Creating a Story.

CHART AND MAP FORMS

9

Pie charts – Tree maps – Heat map - Line graph – Gantt chart – Scatter plot – Histogram - Word cloud – Waterfall charts - Bump charts – Bullet graph.

STATISTICS TOOLS

9

Need for Statistics - Descriptive statistics - Inferential statistics - Number summary – Spread of data - Box plot - Statistics tools in Tableau – Reference lines - Trend lines - Forecasting - Integration of tableau with R.

d. Activities

Students shall be exposed to making use of visualization tools using tableau and developing various applications.

e. Learning Resources

Text Books

1. Suresh Kumar Mukhiya and Usman Ahmed, *Hands-On Exploratory Data Analysis with Python*, Packt Publishing, 2020.
2. Seema Acharya and Subashini Chellappan, *Pro Tableau*, Apress, 2017.

Reference Books

1. Claus O Wilke, *Fundamentals of Data Visualization*, 1st Edition, O'Reilly, 2019.
2. Arshad Khan, *Jumpstart Tableau*, Updated edition, Apress, 2016.
3. Ben Jones, *Communicating Data with Tableau*, 1st Edition, O'Reilly, 2014.

Course Code	Course Name	L	T	P	C
GE2401	UNIVERSAL HUMAN VALUES AND ETHICS	2	0	0	2

Category: Science and Humanities

a. Preamble

This course is intended to provide a much-needed orientational input in value education to the young enquiring minds.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Define the essential complementarity between 'VALUES' and 'SKILLS' for ensuring happiness and prosperity.	K2
CO2	Explore Human being as the Co-existence of the Self and the Body.	K3
CO3	Develop holistic perspective towards value-based living in a natural way.	K3
CO4	Explain the interconnectedness of the four orders of Nature and existence.	K2
CO5	Comprehend the ethics of human values, Humanistic education and constitution, strategies of value-based life and profession.	K2

c. Course Syllabus

Total : 30 Periods

INTRODUCTION TO VALUE EDUCATION

6

Right Understanding, Relationship and Physical Facility (Holistic Development and the Role of Education), Understanding Value Education, Self-exploration as the Process for Value Education, Continuous Happiness and Prosperity - the Basic Human Aspirations, Happiness and Prosperity – Current Scenario, Method to Fulfil the Basic Human Aspirations.

HARMONY IN THE HUMAN BEING

6

Understanding Human being as the Co-existence of the Self and the Body, Distinguishing between the Needs of the Self and the Body, The Body as an Instrument of the Self, Understanding Harmony in the Self, Harmony of the Self with the Body, Programme to ensure self-regulation and Health.

HARMONY IN THE FAMILY AND SOCIETY

6

Harmony in the Family - the Basic Unit of Human Interaction, 'Trust' - the Foundational Value in Relationship, 'Respect' - as the Right Evaluation, Other Feelings, Justice in Human-to-Human Relationship, Understanding Harmony in the Society, Vision for the Universal Human Order.

HARMONY IN THE NATURE/EXISTENCE **6**

Understanding Harmony in the Nature, Interconnectedness, self-regulation and Mutual Fulfilment among the Four Orders of Nature, Realizing Existence as Co-existence at All Levels, The Holistic Perception of Harmony in Existence.

IMPLICATIONS OF THE HOLISTIC UNDERSTANDING – A LOOK AT PROFESSIONAL ETHICS **6**

Natural Acceptance of Human Values, Definitiveness of (Ethical) Human Conduct, A Basis for Humanistic Education, Humanistic Constitution and Universal Human Order, Competence in Professional Ethics, Holistic Technologies, Production Systems and Management Models- Typical Case Studies, Strategies for Transition towards Value-based Life and Profession.

d. Activities

Practice Sessions - **Introduction to Value Education**

- 1 Sharing about Oneself
- 2 Exploring Human Consciousness
- 3 Exploring Natural Acceptance

Practice Sessions– Harmony in the Human Being

- 4 Exploring the difference of Needs of Self and Body
- 5 Exploring Sources of Imagination in the Self
- 6 Exploring Harmony of Self with the Body

Practice Sessions– Harmony in the Family and Society

- 7 Exploring the Feeling of Trust
- 8 Exploring the Feeling of Respect
- 9 Exploring Systems to fulfil Human Goal

Practice Sessions– Harmony in the Nature (Existence)

- 10 Exploring the Four Orders of Nature
- 11 Exploring Co-existence in Existence

Practice Sessions– Implications of the Holistic Understanding – a Look at Professional Ethics

12 Exploring Ethical Human Conduct

13 Exploring Humanistic Models in Education

14 Exploring Steps of Transition towards Universal Human Order

e. Learning Resources

https://fdp-si.aicte-india.org/UHV-II_Lectures_PPTs.php

<https://fdp-si.aicte-india.org/UHV-II%20Practice%20Sessions.php>

Text Books

1. R R Gaur, R Asthana, G P Bagaria., *The Textbook A Foundation Course in Human Values and Professional Ethics*, 2nd Revised Edition, Excel Books, New Delhi, 2019.
2. R R Gaur, R Asthana, G P Bagaria., *The Teacher's Manual A Foundation Course in Human Values and Professional Ethics*, 2nd Revised Edition, Excel Books, New Delhi, 2019.

Reference Books

1. EkParichaya, A Nagaraj., *JeevanVidya*, JeevanVidyaPrakashan, Amarkantak, 1999.
2. A.N. Tripathi ., *Human Values*, New Age Intl. Publishers, New Delhi, 2004.
3. Mohandas Karamchand Gandhi, *The Story of My Experiments with Truth*, 2009.
4. J C Kumarappa, *Economy of Permanence*, 2017.
5. Maulana Abdul Kalam Azad, *India Wins Freedom*, 1988.

Course Code	Course Name	L	T	P	C
AI2402	DATA VISUALIZATION LABORATORY	0	0	4	2

Category: Professional Core Course

a. Preamble

The course has been designed to impart practical knowledge on data visualization using tableau. This course enables the students to present information intuitively and also get exposure to applying visualization techniques using tableau to develop various applications.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Build applications using various File types and Data connections in the Tableau environment	K3
CO2	Apply Data transformation functions to preprocess the data	K3
CO3	Construct meaningful Dashboards and Stories for data exploration and visualization	K3
CO4	Choose different forms of the chart and map to visualize and insight into the data	K3
CO5	Identify the Statistics tools to perform advanced data analysis using Tableau	K3

c. Course Syllabus

Total: 60 Periods

1. Build an application by Connecting various single and multiple data sources using the default data source in Tableau “Sample-Superstore”.
2. Apply various filtering, sorting, grouping, and logical functions to analyze the dataset of COVID cases worldwide and give real-time numbers for various regions.
3. Create Interactive Dashboard and Stories to analyze the dataset of different stocks and derive meaningful information.

4. Analyze a data set of marketing campaigns and visualize the performance of various marketing campaigns using different forms of the chart.
5. Analyze a dataset of various crimes happening at a place and visualize the data using maps and word clouds.
6. Build an application in tableau to predict and infer the risk to the patient's health.
7. Develop an application in tableau to detect anomalies and identify fraudulent transactions.
8. Construct a meaningful dashboard in tableau for Twitter sentiment analysis and integrate it with R Studio.
9. Mini Project.

d. Learning Resources

Reference Book

1. Seema Acharya and Subashini Chellappan, *Pro Tableau*, Apress, 2017.

Course Code	Course Name	L	T	P	C
GE2491	PRINCIPLES OF MANAGEMENT	3	0	0	3

Category: Science and Humanities

a. Preamble

This course introduces fundamental principles of management, emphasizing their universal applicability in diverse organizations. It covers core managerial functions, explores organizational structures, and provides insights into effective global leadership qualities and skills.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Discuss the trends and challenges of management in global scenario, the different types of organization and its effectiveness.	K2
CO2	Describe the strategies and policies which are involved in process planning and decision making.	K2
CO3	Illustrate the structure, purpose, selection and recruitment process in organizations.	K2
CO4	Elucidate the various motivational theories and processes of management including its functions.	K2
CO5	Explain the process and control techniques for budgeting and inventory management.	K2

c. Course Syllabus

Total : 45 Periods

INTRODUCTION TO MANAGEMENT AND ORGANIZATIONS 9

Definition of Management – Nature, Scope and Functions of Management – Evolution of Management – Contributions of FW Taylor (14 principles of Management), Henri Fayol, Elton Mayo, Roethlisberger, H.A.Simon and P.F Drucker- Management theories - Science or Art – Manager Vs Entrepreneur- types of managers managerial roles and skills – Evolution of Management –Scientific, human relations , system and contingency approaches –Current trends and issues in Management.

PLANNING 9

Nature and purpose of planning - Planning process - Types of planning - Objectives - Setting objectives - Policies - Planning premises - Strategic Management - Planning Tools and Techniques - Decision making steps and process.

ORGANISING **9**

Nature and purpose - Formal and informal organization - Organization chart - Organization structure - Types - Line and staff authority - Departmentalization - delegation of authority - Centralization and decentralization - Job Design - Human Resource Management - HR Planning, Recruitment, selection, Training and Development, Performance Management , Career planning and management.

DIRECTING **9**

Directing meaning - importance - principles of directing - Motivation - Motivation theories -Motivational techniques - Job satisfaction - Job enrichment - Leadership -14 types and theories of leadership - Communication- Process of communication, types of communication and its uses - Barrier in communication - Effective Communication- Communication and IT.

CONTROLLING **9**

System and process of controlling – Budgetary and non - Budgetary control techniques - Use of computers and IT in Management control - Productivity problems and management - Inventory Management - PERT, CPM - Application - Control and performance - Direct and preventive control.

d. Activities

Students shall be given exposure to various concepts of delegation of authority, centralization, and decentralization within the college premises.

e. Learning Resources

Text Books

1. Harold Koontz and Heinz Weihrich, *Essentials of Management*, Tata McGraw Hill,2020.
2. Stephen P. Robbins and Mary Coulter, *Management*, Pearson ,2019.

Reference Books

1. Robert Kreitner and Mamata Mohapatra, *Management*, Biztantra, 2008.
2. Stephen A. Robbins and David A. Decenzo and Mary Coulter, *Fundamentals of Management*, Pearson Education, 9th Edition,2016.
3. Tripathy PC and Reddy PN, *Principles of Management*, Tata McGraw Hill, 2021.

Course Code	Course Name	L	T	P	C
GE2492	TOTAL QUALITY MANAGEMENT	3	0	0	3

Category: Science and Humanities

a. Preamble

Quality is the mantra for success or even for the survival of any organization in this competitive global market. Total Quality Management (TQM) is an enhancement to the traditional way of doing business. It is a proven technique to guarantee survival in world class competition. It integrates fundamental management techniques, existing improvement efforts, and technical tools under a disciplined approach. At the end of the course the students are expected to recognize the quality issues in an organization and analyze the ways to solve those using TQM techniques, and demonstrate skills in using modern TQM tools.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Describe the concepts of TQM for an enterprise.	K2
CO2	Comprehend the TQM principles and its implementation.	K2
CO3	Discuss the various traditional and new TQM tools.	K2
CO4	Examine the fundamental concepts of QFD and TPM with applications.	K3
CO5	Apply QMS and EMS in business organization.	K3

c. Course Syllabus

Total : 45 Periods

INTRODUCTION

9

Concept of Quality and Quality Management - Determinants of quality of a product and Service - Quality vs. Reliability - Definition of TQM - Basic concepts of TQM – TQM Framework - Barriers to TQM - Benefits of TQM – Gurus of TQM (Brief Introduction) - Quality statements – vision, mission, and policy.

TQM PRINCIPLES

9

Continuous Improvement Process - Deming Philosophy - Juran Trilogy - PDSA cycle - Kaizen - Concepts of Quality circle - Japanese 5S principles and 8D methodology.

TQM TOOLS AND TECHNIQUES I

9

The seven traditional tools of quality - New management tools - Six-sigma Process Capability - Bench marking: Reasons for benchmarking , Benchmarking process, Understanding Current Performance, Planning, Pitfalls and Criticisms of Benchmarking - FMEA: Intent, Documentation, Stages: Design FMEA and Process FMEA.

TQM TOOLS AND TECHNIQUES II

9

Quality circles - Quality Function Deployment: QFD Team - Benefits of QFD - Voice of the customers - Organization of Information - House of Quality - QFD Process - Taguchi quality loss function - TPM: Concepts, improvement needs - Performance measures - Cost of Quality - Applications.

QUALITY AND ENVIRONMENTAL MANAGEMENT SYSTEM

9

Introduction - Benefits of ISO Registration - ISO 9000 Series of Standards - Sector Specific Standards: AS 9100, TS16949 and TL 9000 - ISO 9001 Requirements - Implementation - Documentation - Internal Audits – Registration - ENVIRONMENTAL MANAGEMENT SYSTEM: Introduction - ISO 14000 Series Standards - Concepts of ISO 14001 - Requirements of ISO 14001 – Benefits and applications of EMS.

d. Activities

Students shall be exposed to learn the knowledge and skills necessary to drive organizational excellence through the implementation of effective quality management strategies.

e. Learning Resources

Text Books

1. Dale H.Besterfield, Carol B.Michna,Glen H. Besterfield,Mary B.Sacre, Hemant Urdhwareshe and Rashmi Urdhwareshe, *Total Quality Management*, Revised 3rd Edition, Pearson Education Asia, 2013.
2. Suganthi L & Anand Samuel, *Total Quality Management*, Prentice Hall Publications, 2004.

Reference Books

1. Kiran.D.R, *Total Quality Management: Key concepts and case studies*, Butterworth – Heinemann Limited, 2016.
2. Shridhara Bhat K, *Total Quality Management: Text and Cases*, Himalaya Publishing House India, 2nd Edition, 2016.

Course Code	Course Name	L	T	P	C
AI2451	PROJECT WORK	0	0	20	10

Category: Employability Enhancement Course

a. Preamble

This course develops the ability to solve a specific problem right from its identification and literature review till the successful solution of the same. It is used to train the students in preparing project reports and to face reviews and viva voce examination

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Identify challenging engineering problems/societal needs to propose project-based solutions	K3
CO2	Build critical-thinking and analytical decision-making capabilities to find solution by formulating proper methodology	K3
CO3	Analyze various algorithmic strategies using technological tools to provide software solutions	K4
CO4	Discover solutions to identified problems	K5
CO5	Build an independent project, resulting in at least a publication in reputed journals or conference proceedings	K6

Total : 300 Periods

The students will be working in single or group of 3 to 4 on a scientific problem approved by the Head of the Department under the guidance of the faculty member and prepare a comprehensive report after completing the work to the satisfaction of the supervisor. The progress of the project is evaluated based on a minimum of three reviews. The review committee may be constituted by the Head of the Department. A project report is required at the end of the semester. The project work is evaluated based on project evaluation process as recommended in the respective regulation

The student can also be permitted to work on the project in Industry/Research organization with the due permission from Head of the Department. The Engineer/Scientist from Industry/Research Organization can jointly act as supervisor in addition to the Project Supervisor. The student should undergo project evaluation process as recommended in the respective regulations.

Course Code	Course Name	L	T	P	C
OAI701	FUNDAMENTALS OF BIG DATA	3	0	0	3

Category: Open Elective Course

a. Preamble

The course enables the students to understand Big Data processing used in different business intelligence applications and provide an in-depth coverage of Map Reduce analytics using Hadoop Eco system tools. The students will gain programming knowledge in Pig, Hive, Hbase to handle the Big Data applications and they will get exposure in blooming Big Data technologies.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Explain the fundamental concepts of Big Data tools and techniques	K2
CO2	Apply Map Reduce algorithms in Hadoop framework	K3
CO3	Construct NoSQL data models for the appropriate databases	K3
CO4	Develop scripts using Pig Latin	K3
CO5	Build HiveQL queries for data analytics	K3

c. Course Syllabus

Total : 45 Periods

INTRODUCTION TO BIG DATA

9

Introduction - Types of Digital Data - Characteristics of Big Data - 3Vs of Big Data - Evolution of Big Data - Architecture of Big Data systems - Advantages of Big Data - Challenges with Big Data - Big Data Use Cases -Analytics - Descriptive Analytics - Diagnostic Analytics - Predictive Analytics - Prescriptive Analytics

DATA ANALYTICS USING HADOOP AND MAPREDUCE FRAMEWORK

9

Introduction to Hadoop - RDBMS versus Hadoop - Hadoop Overview - HDFS (Hadoop Distributed File System) - Processing Data with Hadoop - Introduction to MapReduce - Features of MapReduce. Algorithms Using MapReduce - Matrix-Vector Multiplication, Relational Algebra Operations, Grouping and Aggregation - Extensions to MapReduce.

NOSQL DATA MANAGEMENT FOR BIG DATA 9

NoSQL Databases - CAP Theorem - BASE Concept - NoSQL data Models - Increasing Flexibility for Data Manipulation-Key Value Stores - Document Stores - Tabular Stores - Object Data Stores - Graph Databases - Hive - Sharding - Hbase

DATA ANALYTICS USING PIG 9

Introduction - Installation and Execution - Pig Data Model - Pig Latin - Input, Output-Relational Operators - User Defined Functions - Join Implementations - Integrating Pig with Legacy Code and Map Reduce - Developing and Testing Pig Latin Scripts

DATA ANALYTICS USING HIVE 9

Introduction - Data Types and File Formats - Databases in Hive – HiveQL - Data Definition - Data Manipulation - Queries - Views - Indexes - Schema Design

d. Activities

Students shall be exposed to in Hadoop, MapReduce, Pig, Hive to handle the Big Data applications.

e. Learning Resources

Text Books

1. David Dietrich, Barry Heller & Beibei Yang, *Data Science and Big data Analytics*, EMC, 2013.
2. Seema Acharya and Subhasini Chellappan, *Big Data and Analytics*, Wiley Publication, 2015.
3. Raj Kamal and Preeti Saxena, *Big Data Analytics: Introduction to Hadoop, Spark, and Machine-Learning*, McGraw Hill, 2018.
4. Tom White, *Hadoop: The Definitive Guide*, Third Edition, O'reily Media, 2015.
5. EMC Education Services, *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*, 2015.

Reference Books

1. Mark Van Rijmenam, *Think Bigger: Developing a Successful Big Data Strategy for your Business*, 1st Edition, Amazon, 2014.
2. Hurwitz JS, Nugent A, Halper F and Kaufman M, *Big data for dummies*, John Wiley & Sons, 2013.
3. Bill Franks, *Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics*, John Wiley & sons, 2012.

4. Jure Leskovec, Anand Rajaraman, Jeffrey Ullman, *Mining of Massive Datasets*, Cambridge University Press, 2020.
5. Alan Gates and Daniel Dai, *Pig: Data flow Scripting with Hadoop*, O'Reilly Media, 2016.
6. Jason Rutherglen, Dean Wampler and Edward Capriolo, *Programming Hive*, 1st Edition, O'Reilly Media, 2012.
7. Dayong Du, *Apache Hive Essentials*, Packet Publishing, 2015.

Course Code	Course Name	L	T	P	C
OAI702	INFORMATION RETRIEVAL BASICS	3	0	0	3

Category: Open Elective Course

a. Preamble

This course enables the students to understand the basic concepts about information retrieval, web retrieval and recommender systems. The students get familiarized with web crawler, scheduling algorithms and evaluation process. This course focuses on various retrieval algorithms and recommender systems and their techniques useful for a real time user.

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Explain the architectural design and basic concepts of Information Retrieval	K2
CO2	Identify an appropriate information retrieval model to retrieve documents	K3
CO3	Illustrate various search engine architectures and ranking functions for Web Retrieval	K2
CO4	Outline the applications of a web crawler, and architecture implementation for web crawling	K2
CO5	Demonstrate characteristics and potentials of different prediction techniques in Recommender Systems	K2

Total : 45 Periods

c. Course Syllabus

INTRODUCTION

9

Information Retrieval (IR) - The IR Problem - Information versus Data Retrieval - The Software Architecture of the IR System - The Retrieval and Ranking Processes - The Web - How the web changed Search - Practical Issues on the Web - Search Interfaces - Visualization in Search Interface.

MODELING AND RETRIEVAL EVALUATION

9

Basic IR Models - Boolean Model - TF-IDF (Term Frequency/Inverse Document Frequency) Weighting - Vector Model - Probabilistic Model.

WEB RETRIEVAL **9**

The Web - Search Engine Architectures - Cluster based Architecture - Distributed Architectures - Search Engine Ranking - Link based Ranking - Simple Ranking Functions.

WEB CRAWLING **9**

Search Engine User Interaction - Browsing - Applications of a Web Crawler - Architecture and Implementation – Evaluation.

RECOMMENDER SYSTEM **9**

Recommender Systems Functions - Recommendation Techniques - Basics of Content- based Recommender Systems - High Level Architecture - Advantages and Drawbacks of Content - based Filtering - Collaborative Filtering.

d. Activities

Students shall be exposed to the core concepts of information retrieval and its associated algorithms using case study.

e. Learning Resources

Text Books

1. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, *Modern Information Retrieval: The Concepts and Technology behind Search*, 2nd Edition, ACM Press Books, 2011.
2. Iresh A. Dhotre, *Information Retrieval Techniques*, Technical Publications, 2020.
Ricci F, Rokach, Shapira L and Kantor B, *Recommender Systems Handbook*, 1st Edition, 2011.

Reference Book

1. Manning C, Raghavan P and Schütze H, *Introduction to Information Retrieval*, Cambridge University Press, 2008.
2. Stefan Buettcher, Charles L. A. Clarke and Gordon V. Cormack, *Information Retrieval: Implementing and Evaluating Search Engines*, The MIT Press, 2010.

Course Code	Course Name	L	T	P	C
OAI703	INTRODUCTION TO OPERATING SYSTEMS	3	0	0	3

Category: Open Elective Course

a. Preamble

This course enables the students to understand the basic concepts about operating system, processes and threads. The students get familiarized with the scheduling algorithms and deadlock handling mechanisms. This course focuses on various memory management schemes and file systems

b. Course Outcome

After successful completion of the course, the students will be able to

CO. No.	Course Outcome	Knowledge Level
CO1	Understand the basic concepts of operating systems.	K2
CO2	Illustrate the concepts of process, threads and CPU scheduling algorithms.	K2
CO3	Apply the algorithms used for concurrency and deadlock handling.	K3
CO4	Apply various methods and techniques to solve problems relevant to process scheduling, deadlock, Memory management and paging for the different processes.	K3
CO5	Demonstrate the concept of file systems.	K2

Total : 45 Periods

c. Course Syllabus

INTRODUCTION **9**

Operating System Overview - Operating system objectives and functions - Operating System Structures – Operating System Services - System Calls- System Programs .

PROCESS MANAGEMENT **9**

Processes - Process Concept, Process Scheduling, CPU Scheduling - Scheduling criteria, Scheduling algorithms – Inter Process Communication, Threads- Overview, Multi threading models.

CONCURRENCY CONTROL 9

Process Synchronization - The critical-section problem - Semaphores, Deadlock - Methods for handling deadlocks, Deadlock prevention, Deadlock avoidance, Deadlock detection.

MEMORY MANAGEMENT 9

Main Memory - Swapping- Paging - Segmentation, Segmentation with paging, Virtual Memory - Demand Paging, Page Replacement, Thrashing.

FILE SYSTEMS 9

File system Interface - Directory structure - File system mounting - File System implementation - File system structure, Disk scheduling, Device drivers

d. Activities

Students shall be exposed to the core concepts of operating systems using case study.

e. Learning Resources

Text Books

1. Leland L Beck, *System Software: An Introduction to Systems Programming*, 3rd Edition, Pearson Education Asia, 1997.
2. Abraham Silberschatz, Peter Baer Galvin and Greg Gagne, *Operating System Concepts*, 9th Edition, John Wiley and Sons Inc, 2018.

Reference Books

1. Andrew S Tanenbaum, *Modern Operating Systems*, 2nd Edition, Pearson Education, 2004.
2. Elmasri, R, Carrick, A and Levine, D, *Operating Systems – A Spiral Approach*, Tata
3. Gary Nutt, *Operating Systems*, 3rd Edition, Pearson Education, 2004.