M.Tech. DATA ANALYTICS

Credit Based Flexible Curriculum
(Applicable from 2022-23)

Department of Computer Applications
National Institute of Technology
Tiruchirappalli - 620 015
Tamilnadu, India
About us:

The Department of Computer Applications is one of the pioneering departments of the institution that offers Information Technology courses such as MCA, M.Sc Computer Science and M.Tech (Data Analytics) and is one among the top five offering MCA courses in the country. It is committed to impart quality education in the sub-fields of IT, a field growing in leaps and bounds.

Vision:

Towards a school of Information Science and Technology conforming to international standards

Mission:

• To offer state-of-art education in Information Science and Technology
• To provide strong theoretical foundation complemented with extensive practical training
• To inculcate value-based, socially committed professionalism to the cause of overall development of students and society

M. Tech. Data Analytics

Objectives of the Programme:

• To provide students with a comprehensive foundation for applying statistical methods to resolve real-world problems in data analytics.
• To develop the ability to apply computational techniques to draw insight from big data in a variety of application domains.
• To prepare students for careers in data analytics with the ability to conduct in-depth data analysis, synthesis and evaluation.
• To manage real life complex data analytics projects.
• To prepare students for research in data analytics with the ability to provide solutions to recent research issues.

Qualification for Admission:

Bachelor’s degree in Engineering / Technology in CSE / IT or MSc. (Computer Science / Information Technology / Software Engineering) or MCA from a recognized institution and a valid GATE score in Computer Science. The percentage and CGPA requirement for admission may be fixed as per NIT Trichy norms.
Head of the Department:

Dr. P. J. A. Alphonse

Members:

1. Dr. S. R. Balasundaram, Professor
2. Dr. S. Nickolas, Professor
3. Dr. Michael Arock, Professor
4. Dr. P. J. A. Alphonse, Professor
5. Dr. S. Domnic, Associate Professor
6. Dr. G. R. Gangadharan, Associate Professor
7. Dr. B. Janet, Assistant Professor
8. Dr. S. Sangeetha, Assistant Professor
9. Dr. R. Eswari, Assistant Professor
10. Dr. U. Srinivasulu Reddy, Assistant Professor
11. Dr. B. Balaji, Assistant Professor
12. Dr. K. Selvakumar, Assistant Professor
13. Dr. Jitendra Kumar, Assistant Professor
14. Dr. Ghanshyam S. Bopche, Assistant Professor
# SYLLABUS

## Semester I

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course name</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA611</td>
<td>Mathematical Foundations for Data Analytics</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA613</td>
<td>Advanced Data Structures and Algorithms</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>CA615</td>
<td>Machine Learning Techniques</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA617</td>
<td>Big Data Analytics</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6AX</td>
<td>Elective-1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6BX</td>
<td>Elective-2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA601</td>
<td>Big Data Analytics Lab</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>CA603</td>
<td>Machine Learning Lab</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

## Semester II

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course name</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA610</td>
<td>Distributed and Cloud Computing</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA612</td>
<td>Deep Learning and its Applications</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6CX</td>
<td>Elective-3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6DX</td>
<td>Elective-4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6EX</td>
<td>Elective-5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6FX</td>
<td>Elective-6</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA602</td>
<td>Distributed and Cloud Computing Lab</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>CA604</td>
<td>Deep Learning Lab</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
</tbody>
</table>

## Semester III

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course name</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA649</td>
<td>Project Work Phase I</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

## Semester IV

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course name</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA650</td>
<td>Project Work Phase II</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>
# LIST OF ELECTIVES

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course name</th>
<th>L</th>
<th>T</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEMESTER I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA6A1</td>
<td>Data Mining and Warehousing</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6A2</td>
<td>Soft Computing Techniques</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6A3</td>
<td>Next Generation Data Base Systems</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6B1</td>
<td>DevOps</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6B2</td>
<td>Analytics for Strategic Market Planning</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6B3</td>
<td>Financial Risk Analytics and Management</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>SEMESTER II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA6C1</td>
<td>Intelligent Systems and Process Automation</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6C2</td>
<td>Cognitive Sciences</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6C3</td>
<td>Blockchain Technology</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6D1</td>
<td>Natural Language Processing</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6D2</td>
<td>Social Network Analysis</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6D3</td>
<td>Supply Chain Analytics</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6E1</td>
<td>Web Analytics</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6E2</td>
<td>Federated Learning</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6E3</td>
<td>Human Resource Analytics</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6F1</td>
<td>Data Visualization</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6F2</td>
<td>Pattern Recognition</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>CA6F3</td>
<td>Image and Video Analytics</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Semester – I

CA611 – MATHEMATICAL FOUNDATIONS OF DATA ANALYTICS

Objectives:

• To learn the mathematical foundations applicable to data analytics.


Sampling Distributions and Descriptive Statistics - The Central Limit Theorem – Sampling Distributions (Chi-Square, t, F, z) – Test of Hypothesis – Testing for Attributes – Mean of Normal Population – One-tailed and two-tailed tests – F-test and Chi-Square test – Analysis of variance ANOVA – One way and two-way classifications

REFERENCES:

COURSE OUTCOMES:
The students will be able to:
- Deduce complex tasks by various Mathematical logic.
- Solve the problems using the concepts of Linear Equations, Discrete and Continuous Probabilities.
CA613 – ADVANCED DATA STRUCTURES AND ALGORITHMS

Objectives:

- To introduce different advanced data structures and algorithms.


Backtracking and Branch-and-Bound strategies with applications – Randomized algorithms – Examples.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:

- Write structured pseudo code for the given problem.
- Use advanced data structures to solve real-time problems.
CA615 - MACHINE LEARNING TECHNIQUES

Objectives:
- To introduce the basic concepts and techniques of Machine Learning.
- To develop the skills in using recent machine learning software for solving practical problems.
- To be familiar with a set of well-known supervised, semi-supervised and unsupervised learning algorithms.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:
- Select real-world applications that needs machine learning based solutions.
- Implement and apply machine learning algorithms.
- Select appropriate algorithms for solving a particular group of real-world problems.
- Recognize the characteristics of machine learning techniques that are useful to solve real-world problems.
CA617 - BIG DATA ANALYTICS

Objectives:
- To understand the Big Data Platform and its Use cases.
- To provide an overview of Apache Hadoop Ecosystem.
- To provide HDFS Concepts and Interfacing with HDFS.
- To understand Map Reduce concepts.


Map Reduce: Anatomy of a Map Reduce - Job Run – Failures – Job Scheduling – Shuffle and Sort – Task Execution – Map Reduce Types and Formats – Map Reduce Features - Composing map reduce calculations.


REFERENCES:
COURSE OUTCOMES:
The students will be able to:
  • Describe big data and use cases from selected business domains
  • List the components of Hadoop and Hadoop Eco-System
  • Access and Process Data on Distributed File System
  • Manage Job Execution in Hadoop Environment
  • Develop Big Data Solutions using Hadoop Eco System
CA601 - BIG DATA ANALYTICS LAB

Objectives:

- To optimize business decisions and create competitive advantage with Big Data analytics.
- To impart the architectural concepts of Hadoop and introduce map reduce paradigm.
- To introduce PIG & HIVE in Hadoop ecosystem.
- To develop Big Data applications for streaming data using Apache Spark.

Exercises:

Problems to
- Perform setting up and installing Hadoop
- Implement file management tasks
- Run different MapReduce programs
- Install and run Pig and write Pig Latin scripts
- Install and run Hive and use Hive
- Install and deploy Apache Spark cluster and run apache spark applications
- Develop Big Data applications for streaming data using Apache Spark

COURSE OUTCOMES:

The students will be able to:
- Preparing for data summarization, query, and modeling
- Applying data modeling techniques to large data sets
- Creating applications for Big Data analytics
- Building a complete business data analytic solution
CA603 - MACHINE LEARNING LAB

Objectives:

- To introduce basic machine learning techniques.
- To develop the skills in using recent machine learning software for solving practical problems in high-performance computing environment.
- To develop the skills in applying appropriate supervised, semi-supervised or unsupervised learning algorithms for solving practical problems.

COURSE OUTCOMES:
The students will be able to:

- Implement and apply machine learning algorithms to solve problems.
- Select appropriate algorithms for solving real-world problems.
- Use machine learning techniques in high-performance computing environment to solve real-world problems.
Semester – II

CA610 – DISTRIBUTED AND CLOUD COMPUTING

Objectives:
- To explore basic concepts and practices of distributed computing.
- To understand Cloud Computing concepts, technologies, architecture and applications.
- To understand different cloud programming platforms and tools to develop and deploy applications on cloud.

Fundamentals of Distributed Computing – Inter Process Communications – Distributed Computing Paradigms - Distributed Objects


REFERENCES:
COURSE OUTCOMES:
The students will be able to:
  • Acquire knowledge of the concepts and technologies of distributed and cloud Computing.
  • Demonstrate where to apply parallel and distributed techniques.
  • Use various performance criteria to evaluate the quality of the cloud architecture.
CA612 – DEEP LEARNING AND ITS APPLICATIONS

Objectives:
- To introduce the techniques of deep learning.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:
- Explore the essentials of Deep Learning and Deep Network architectures
- Define, train and use Deep Neural Networks for solving real world problems that require artificial intelligence based solutions
CA602 – DISTRIBUTED AND CLOUD COMPUTING LAB

Objectives:

- To introduce basic distributed and cloud computing techniques.
- To develop the skills in applying appropriate distributed and cloud computing techniques for solving practical problems.

- MPI-code parallel program to multiply two large scale N X N matrices
- Problem on Load Sharing Facility (LSF) for cluster Computing.
- Problem related to analysis of the performance of Xen live migration for I/O read-intensive applications.
- Problem related to analysis of the performance of Xen live migration for I/O write-intensive applications.
- Problem on fat-tree concept. (scalability related)
- Program that uses publish-subscribe to communicate between entities developed in different languages.
- Program to run the GAE platform for backup storage of a large amount of different type of data (text, images, video, audio etc.).
- MPI code on any GT$-enable grid platform (set up one with a research team). Write a report to explain the grid experimental setting using the Globus tools.
- Problem related to issues in P2P computing.
- Problem related to WSN.

COURSE OUTCOMES:
The students will be able to:

- Implement and apply distributed and cloud computing techniques to solve problems.
- Select appropriate algorithms for solving real-world problems.
CA604 – DEEP LEARNING LAB

Objectives:

- To introduce basic deep learning techniques.
- To develop the skills in using recent deep learning software for solving practical problems in high-performance computing environment.
- To develop the skills in applying appropriate deep learning algorithms for solving practical problems.

- Problem Selection
- Literature Review and Data Collection
- Data Preprocessing
- Model Selection (Deep Learning Models)
- Model Evaluation
- Model Deployment
- Results Analysis

COURSE OUTCOMES:
The students will be able to:

- Implement and apply deep learning algorithms to solve problems.
- Use deep learning techniques in a high-performance computing environment to solve real-world problems
Semester – III

CA649 PROJECT WORK – PHASE I

• Internal project work of 6 Months duration to be extended in phase II.

Semester – IV

CA650 PROJECT WORK – PHASE II

• Internal project work of 6 Months duration with submission of thesis and viva-voce examination
CA6A1: DATA MINING AND WAREHOUSING

Objectives:

- To develop and understand the necessity of Data Mining, the strengths and limitations of popular data mining techniques and how they can be applied in business applications.
- To understand the overall architecture of a data warehouse and methods for data gathering and data pre-processing.
- To explore data mining and data warehousing applications in various real-world datasets.

Fundamentals of data mining and Data Preprocessing: Motivation – Importance – Definition of Data Mining - Data Mining Functionalities - Classification of Data Mining systems - Data Mining Task Primitives - Integration of a Data Mining System with a Database or a Data Warehouse System - Major issues in Data Mining. Types of Data Sets and Attribute Values - Basic Statistical Descriptions of Data - Data Visualization - Measuring Data Similarity Data – Pre-processing: Need for Pre-processing the Data - Data Cleaning - Data Integration and Transformation - Data Reduction - Discretization and Concept Hierarchy Generation

Data Warehouse and OLAP Technology for Data Mining: Data Warehouse - Multidimensional Data Model - Data Warehouse Architecture - Data Warehouse Implementation - Further Development of Data Cube Technology - From Data Warehousing to Data Mining Data Cube Computation and Data Generalization: Efficient Methods for Data Cube Computation – Further Development of Data Cube and OLAP Technology

Mining Frequent Patterns – Associations and Correlations: Basic Concepts on Frequent Itemsets - Efficient and Scalable Frequent Item set Mining Methods - Mining various kinds of Association Rules – Apriori Algorithm – FP-tree algorithm - From Association Mining to Correlation Analysis – Constraint-Based Association Mining

Classification and Prediction: Issues Regarding Classification and Prediction - Classification by Decision Tree Induction - Bayesian Classification - Rule-Based Classification - Classification by Back propagation - Support Vector Machines – Prediction - Accuracy and Error measures - Evaluating the accuracy of a Classifier or a Predictor - Ensemble Methods

Clustering Methods: Cluster Analysis Introduction - Types of Data in Cluster Analysis - A Categorization of Major Clustering Methods- Partitioning Methods – Hierarchical Methods – Density-Based Methods – GridBased Methods – Model-Based Clustering Methods - Clustering
High-Dimensional Data – Constraint Based Cluster Analysis; Outlier Analysis; Mining Stream – Time-Series and Sequence Data – Graph Mining – Mining the Web Data

REFERENCES:
3. Arun K Pujari, Data Mining Techniques, Universities Press, 2010

COURSE OUTCOMES:
The students will be able to:
• Learn the task of data mining as an important phase of the knowledge discovery process
• Understand the Data warehouse and its implementation
• Understand the mining methodologies
• Study and Implement machine learning algorithms
• Apply Data Mining for Knowledge extraction from real-world databases.
CA6A2 – SOFT COMPUTING TECHNIQUES

Objectives:

- To learn the techniques of soft computing.
- To know artificial neural network concepts.
- To learn the applications of evolutionary and genetic algorithms.
- To know the design of fuzzy controller and rough sets.


REFERENCES:

5. E. Bonabeau, M. Dorigo, G. Theraulaz, Swarm Intelligence: From Natural to Artificial Systems, Oxford Press, 1999

**COURSE OUTCOMES:**
The students will be able to:
- Explain the basics of soft computing and their suitable industry related applications
- Apply neural network principles and algorithms for given problems
- Apply the principles of fuzzy and hybrid algorithms for real time applications
- Solve problems using evolutionary algorithms
CA6A3 – NEXT GENERATION DATABASE SYSTEMS

Objectives:

- To explore the fundamental concepts of next generation database systems.
- To design and develop NoSQL database management systems.

Database Revolutions – Relational Databases – Advanced SQL queries - Transaction Management – Data warehouse and Data Mining – Information Retrieval.


Distributed Database Patterns – Consistency Models – Data Models and Storage – Languages and Programming Interfaces – Emerging Trends in Databases.

REFERENCES:

COURSE OUTCOMES:
The students will be able to:

- Understand how NoSQL databases differ from relational databases from a theoretical perspective.
- Design NoSQL database management systems.
- Select a particular NoSQL database for specific use cases.
CA6B1 – DEVOPS

Objectives:
- To learn the rise of agile methodologies
- To define and design purpose and applications of DevOps


Basics of DevOps: Introduction to DevOps - DevOps and Agile - Minimum Viable Product - Application Deployment - Continuous Integration - Continuous Delivery


Continuous Delivery and Continuous Integration: Rebase frequently from the mainline - Check-in frequently - Frequent build – Automate the testing as much as possible - Automate the deployment - Development operations - Use a version control system - Use repository tools - Use a Continuous Integration tool - Automate the packaging - Automating the deployments - Automating the testing


REFERENCES:
2. Paul Swartout, Continuous Delivery and DevOps A Quickstart Guide, Packt Publishing Limited 2018

COURSE OUTCOMES:
The students will be able to:
- Understand the basic concepts DevOps.
- Understand the real world applications of DevOps.
CA6B2 – ANALYTICS FOR STRATEGIC MARKET PLANNING

Objectives:

- To build marketing response models for strategic marketing decisions.

Marketing Analytics: Basics of marketing analytics - marketing decisions models - characteristics - types and benefits of marketing decisions models - Response models - types - calibration - objectives - interactions effects - dynamic effects - competitive effects - models in individual levels - shared experience and qualitative models.


Positioning: Differentiation and positioning - perceptual maps: developing perceptual map – multi dimensional scaling - techniques – attribute based and similarity based - joint space mapping.

Strategic Market Analysis: Strategic marketing decisions - market demand and trend analysis - product life cycle - cost dynamics: scale and experience effects.


REFERENCES:


COURSE OUTCOMES:

The students will be able to:

- Take data empowered strategic marketing decisions by using analytical techniques.
- Sharpen their analytical skills by getting exposure to computer-based marketing models and tools for decision making.
CA6B3 – FINANCIAL RISK ANALYTICS AND MANAGEMENT

Objectives:
- To identify the different risks involved in Finance arena.
- To understand and solve the different risks pertaining to the stock market and its instruments.
- To analyze the legal issues affecting the business.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:
- Identify and categorize the various risks faced by an organization.
- Explore the tools and practices needed to assess and evaluate financial risks.
- Explore risk management practices in an industry.
- Identify and solve legal issues that impact financial and other risk affecting business.
CA6C1 - INTELLIGENT SYSTEMS AND PROCESS AUTOMATION

Objectives:
- To know about AI and problem solving techniques
- To understand various search techniques and knowledge representation schemes
- To learn the concepts and working principles of RPA
- To gain knowledge in RPA related tools, technologies and services


Introduction to RPA Tool - The User Interface - Variables - Control Flow - The Assign Activity - The Delay Activity - The Do While Activity - Data Manipulation - Gathering and Assembling Data


REFERENCES:

**COURSE OUTCOMES:**
The students will be able to:

- Understand the Significance of AI and various techniques used in intelligent applications.
- Apply basic principles of Expert systems and their fundamentals, and learning.
- Understand the RPA and various platforms available.
- To show the importance of RPA in solving real world problems.
CA6C2 - COGNITIVE SCIENCE

Objectives:

- To understand the structure, features and functions of brain
- To know about brain’s ability to understand symbols, visuals, language and communication
- To understand the algorithms used in information processing at various levels involving thinking process.
- To model artificial intelligent systems involving cognitive aspects.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:
• Understand the principles and techniques behind cognitive science.
• Design applications involving cognitive science.
CA6C3 - BLOCKCHAIN TECHNOLOGY

Objectives:
- To understand blockchain systems (Bitcoin and Ethereum)
- To design, build, and deploy smart contracts and distributed applications
- To integrate ideas from blockchain technology into projects.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:
- Explain design principles of Bitcoin and Ethereum.
- Explain the Simplified Payment Verification protocol.
- Interact with a blockchain system by sending and reading transactions.
- Design, build, and deploy a distributed application.
• Evaluate security, privacy, and efficiency of a given blockchain system.

CA6D1 - NATURAL LANGUAGE PROCESSING

Objectives:
• To introduce language processing techniques to enable Text data processing
• To impart knowledge on text data processing using Statistical and Machine learning models
• To describe the various Embeddings and Deep learning models for NLP
• To introduce real world applications of Language processing


REFERENCES:
2. Delip Rao, Brian McMahan, Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning, O'Reilly, 2019
COURSE OUTCOMES:
The students will be able to:

- Identify the patterns in text and pre-process the large text corpus
- Describe and work with basic NLP tasks
- Use statistical and machine learning models with feature selection for NLP
- Adopt Deep learning models for NLP along with embeddings
- Apply the concepts for solving NLP Applications.
CA6D2 - SOCIAL NETWORK ANALYTICS

Objectives:

- To enable students to understand and visualize the social network structure.
- To impart knowledge on mining the user link structure in the social network.
- To enable students to analyse the social network content

Introduction – Properties of Social Networks – Network Analysis – Network Measures – Discussion Networks – Blogs and Online Communities – Semantic Web


Mining Communities – Aggregating and Reasoning with Social Network Data – Detecting Communities in Social Networks – Evaluating Communities – Community Welfare – Collaboration networks – Co-Citation Networks - Link Prediction – Rank Aggregation and Voting Theory – Tools: Ghephi – Neo4J – Mathematica


REFERENCES:

COURSE OUTCOMES:
The students will be able to:

- Interpret the social network structure
- Model and visualize the social network.
- Mine the user relations in the social network.
- Analyse social network content
CA6D3 - SUPPLY CHAIN ANALYTICS

Objectives:

- To provide the knowledge and necessary skills for carrying out job roles in the domain of Supply chain analytics

Demand Planning: Demand Planning- Review of Forecasting and planning concepts - Defining KPIs- Forecasting Model Building - Discrete and continuous manufacturing

Supply planning: Supply planning - Procurement and Strategic Sourcing - Inventory Modelling – aggregate planning and resource allocation decisions - Procurement Analytics - Production modelling

Demand Fulfilment: Demand Fulfilment - DC location and network design - optimizing inventory levels in distribution network - Logistics & Network Modelling - Transportation modelling - delayed differentiation - mass customization

Integrated supply chain: Advanced and business supply chain related topics like CPFR - DDSN - Make/Buy Case Study - Total Supply Chain Cost - computation of transfer prices - revenue management-yield management - product changes/economies of scale

Project Development: Undertaking projects with industry inputs- validation of models – frameworks – Review of data analytics techniques - choice of tools and designing solution approach to specific applications

REFERENCES:

COURSE OUTCOMES:
The students will be able to:

- Understand the concepts of supply chain analytics.
- Apply various supply chain management concepts.
- Improve an existing supply chain and design an efficient supply chain in alignment with the strategic goals of the company.
CA6E1 - WEB ANALYTICS

Objectives:

- To know the methodologies for collecting, reporting and analysis of website data
- To understand various analysis associated with website and web usage data
- To understand the features of Web Analytics to improve the business in the web, mobile and cloud domains.


Google Tag Managers – Building blocks – Enhancing websites with Tag Manager – Using Google Analytics with BigQuery – Tag Manager and Analytics APIs.

REFERENCES:
2. Beatriz Plaza “Advanced Web Metrics with Google Analytics”, Koros Press, 2017

COURSE OUTCOMES:
The students will be able to:

- Understand the factors influencing the business through the web site data
- Understand various types of analysis involved in web site data (users, business data, etc.)
- Practice applying Google Analytics APIs to improve the business process.
CA6E2 - FEDERATED LEARNING

Objectives:
- To know the basics of federated learning.
- To learn the applications of federated learning.
- To know the privacy preserving deep learning.


Horizontal Federated Learning – Vertical Federated Learning and Federated Transfer Learning.

Federated Optimization for Heterogeneous Networks – Deep Networks from Decentralized Data – Federated Multi-task Learning – Personalized Federated Learning.


REFERENCES:

COURSE OUTCOMES:
The students will be able to:
- Knowledge of the basic concepts, architecture and applications of Federated Learning.
- Understanding of new research and application trends in Federated Learning.
- Ability to deploy real-world Federated Learning projects.
- Hands-on experience in applying Federated Learning tools to solve privacy-preserving AI challenges.
CA6E3 - HUMAN RESOURCE ANALYTICS

Objectives:

- To provide the knowledge on various frameworks and hands-on analytical approaches for identifying the business contributions of the HR function in the organizations.

Introduction to HR Analytics: Overview of HR Process - HR Analytics – An overview - Role of analytics in HR – HRM in changing context - Transition from transaction orientation to analytics orientation - HR Analytics Framework – People Capability Maturity Model – LAMP framework – HCM 21 framework – Talent analytics maturity model.


HR Metrics and Audits: Formulation of key performance indicators and key result areas; HR Metrics – Recruitment metrics – Training and development metrics - Talent retention metrics – HR cost benefit metrics – Career Progression Metrics - Performance metrics – Diversity and Inclusion Metrics - Human capital ROI - Designing and Implementing HR Scorecard - Conducting HR Practice Audits


Predictive and Prescriptive Analytics: Predicting future performance - Techniques to capture the fallouts of HR Practices – Data driven decision making - Organization change and improvement.

REFERENCES:


COURSE OUTCOMES:

The students will be able to:

- Identify necessary skills to carry out the personnel roles in the domain of HR.
- Identify and develop metrics to improve employer-employee relationship and improve employee retention.
- Identify skilled personnel and job tasks to achieve mission-critical goals.
- Align organization's mission and goals with key metrics and benchmarks.
- Apply HR analytics to improve organizational performance by providing better insights on human resources data.
CA6F1 - DATA VISUALIZATION

Objectives:

- To introduce students with fundamental concepts, approaches, and applications of data visualization
- To familiarize students with different visualization and analytic tools and techniques
- To develop skills and competencies required by the professionals


Data-Ink and Graphical Redesign – Data-Ink – Maximizing the Share of Data-Ink – Two Erasing Principles – Application of the Principles in Editing and Redesign - Data-Ink Maximization and Graphical Design – Redesign of Box Plot – Histograms and Scatterplot – Data Density


Mapping – Connections and correlations – Multivariate Analysis – Scatterplot Maps – Trees – Hierarchies and Recursion – Networks and Graphs


REFERENCES:

2. Claus O. Wilke, Fundamentals of Data Visualization, O’Reilly, 2019
5. Alex Campbell, Data Visualization: Ultimate Guide to Data Mining and Visualization, 2020
6. A. Loth, Visual Analytics with Tableau, Willey, 2019
7. Kieran Healy, Data Visualization: A Practical Introduction, Princeton University Press, 2018
COURSE OUTCOMES:
The students will be able to:

- Understand and apply the fundamental design principles of data visualization
- Present data by applying different data visualization techniques for better communication
- Analyse data and communicate information through visual representation
CA6F2 - PATTERN RECOGNITION

Objectives:

- To develop the mathematical tools required for the pattern recognition.
- To learn about supervised and unsupervised pattern classifiers.
- To familiarize about different feature extraction techniques


Sequential Pattern Recognition – Hidden Markov Models (HMMs) – Discrete HMMs - Continuous HMMs – Nonparametric techniques for density estimation – Parzen-window method – K-Nearest Neighbour method


REFERENCES:
COURSE OUTCOMES:
The students will be able to:

- Summarize the various techniques involved in pattern recognition
- Categorize the various pattern recognition techniques into supervised and unsupervised.
- Illustrate the artificial neural network-based pattern recognition
- Discuss the applications of pattern recognition in various applications
CA6F3 - IMAGE AND VIDEO ANALYTICS

Objectives:

- To teach the fundamentals of digital image processing, image and video analysis.
- To understand the real time, use of image and video analytics.
- To demonstrate real time image and video analytics applications and others.


REFERENCES:

COURSE OUTCOMES:

The students will be able to:

- Describe the fundamental principles of image and video analysis and have an idea of their application.
- Apply image and video analysis in real world problems.