SEMESTER V

Sr No	Course Category	Course Code	Course Title	Hours per Week			Credits	Maximum Marks		
				L	T	P		Continual Assessment	End Sem Examination	Total
1	PCC	23DS501T	Theory of Computation	3	-	-	3	30	70	100
2	PCC	23DS502T	Machine Learning for Data Science	3	-	-	3	30	70	100
3	PCC	23DS502P	Machine Learning for Data Science Lab	-	-	2	1	25	25	50
4	PCC	23DS503T	Software Engineering and Project Management	2	-	-	2	15	35	50
5	PCC	23DS503P	Software Engineering and Project Management Lab	-	-	2	1	25	25	50
6	PCC	23DS504P	Data Analytics Lab	-	-	2	1	25	25	50
7	PEC	23DS505T	Professional Elective -I	2	-	-	2	15	35	50
8	OE	23DS561O	Open Elective - II	3	-	-	3	30	70	100
9	MDM	23DS531M	Multidisciplinary Minor III	3	-	-	3	30	70	100
10	VSC	23DS506T	Technical Skill Development - II	2	-	-	2	50	-	50
11	ELC	23DS507P	Mini Project	-	-	2	1	25	25	50
12	SEC	23DS541P	Career Development - V	-	-	2	1	50	-	50
			Total	18		10	23	350	450	800

Open Elective - II	
23DS561O(i)	OE-II Business Intelligence

Multidisciplinary Minor III	
23DS551M	MDM-III Data Engineering

Professional Elective -I	
23DS505T(i)	PE-I Distributed Databases
23DS505T(ii)	PE-I Network Security
23DS505T(iii)	PE-I Image Analytics

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS501T	Theory of Computation	3	-	-	3	30	70	100

Course Objectives	Course Outcomes
This course is intended	**Students will be able to**
• Understand fundamental concepts of formal languages, automata theory, and computational models including finite automata, pushdown automata, and Turing machines. • Analyze and apply theoretical models and principles to classify languages and solve problems related to computability and language recognition.	• Explain the concepts of formal languages, grammars, and the Chomsky hierarchy • Construct and minimize deterministic and nondeterministic finite automata, and demonstrate their equivalence • Apply the pumping lemma to prove properties of regular languages and design regular expressions and grammars equivalent to finite automata • Analyze context-free grammars, pushdown automata, and their equivalence, and transform grammars into normal form • Construct and evaluate Turing machine models

Unit I	**[9Hrs]**
Basics of Theory of Computation: Basics of Sets and Relation, Countability and Diagonalisation, Principle of mathematical induction, Pigeon- hole principle. Fundamentals of formal languages and grammars, Chomsky hierarchy of languages.	
Unit II	**[9Hrs]**
Finite Automata: Deterministic finite automata (DFA), Nondeterministic finite automata (NFA) and equivalence with DFA, Minimization of finite automata, NFA with Epsilon Transitions, Finite Automata with output.	
Unit III	**[9Hrs]**
Regular Grammar: Regular expressions and Regular languages, Regular grammars and equivalence with finite automata, properties of regular languages, pumping lemma for regular languages	
Unit IV	**[9Hrs]**
Context-Free Grammar: Context-free grammars (CFG) and language(CFL), parse trees, ambiguity in CFG, Reduction of CFGs, Chomsky and Greibach normal forms. **Push Down Automata:** Deterministic pushdown automata and Non-deterministic push down automata, Acceptance by two methods: Empty stack and Final State, Equivalence of PDA with CFG, closure properties of CFLs.	
Unit V	**[9Hrs]**
Turning Machine: The basic model for Turing machines (TM), Turing recognizable recursively enumerable) and Turing-decidable (recursive) languages, variants of Turing machines, unrestricted grammars and equivalence with Turing machines, TMs as enumerators.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Introduction to Automata Theory, Languages and Computation	J. E. Hopcraft,R. Motwani, J. D Ullman	2nd Edition	Pearson Education, Aisa
2	Theory of Computer Science, Automata, Languages and Computation	K. L. P. Mishra and N. Chandrasekaran	3rd Edition	PHI Learning.

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Introduction to Theory of Computation	Sipser	2nd Edition	Cengage publications
2	An Introduction to Formal Languages and Automata	Peter Linz	7 Edition	Jones & Bartlett Learning
3	Introduction to Langauges and the theory of Automata	John Martin	5 Edition	McGraw-Hill Education
4	Elements of Theory of Computation	Lewis H.P and Papadimition C.H	2nd Edition	Pearson

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS502T	Machine Learning for Data Science	3	-	-	3	30	70	100

Course Objectives	Course Outcomes
this course is intended • To introduce the basic concepts and techniques of machine learning. • To understand major machine learning algorithms. • To identify machine learning techniques suitable for different problems. a given problem	**Students will be able to** • Describe the fundamental components of machine learning • Apply supervised learning techniques and explain key issues like overfitting, generalization, and regularization. • Implement and evaluate probabilistic learning methods , along with clustering techniques. • Apply ensemble learning techniques such as bagging, boosting, and stacking. • Explore advanced learning paradigms and critically review current research trends in machine learning.

Unit I	**[9Hrs]**
History of Machine Learning, Machine Learning definition, Components of a learning, Different Types of Learning, Concept Learning Task, Concept learning as search, Inductive Learning Bias, FIND-S and Candidate-Elimination algorithm. Decision Trees, Hypotheses Space, Overfitting, Generalization and Occam's Razor.	
Unit II	**[9Hrs]**
Linear Regression and error terms, Artificial Neural Networks, Perceptron, Multilayer networks and Backpropagation algorithm. Methods to avoid overfitting regularization etc, Introduction to Deep Neural networks, Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs)	
Unit III	**[9Hrs]**
Bayesian learning and Bayesian networks, EM algorithm, Naive Bayes classifier, Baysian Belief Networks, Hidden Markov Model. Clustering, Generative Models, Mixture Models, Hierarchical Clustering, K-means algorithm and complexity.	
Unit IV	**[9Hrs]**
Ensemble of Machine Learning models, Boosting, Bagging, Stacking. Dimensionality Reduction and Principal Component Analysis (PCA).	
Unit V	**[9Hrs]**
Other Learning Paradigms: Reinforcement Learning, Deep Learning, Transfer Learning, Semi-supervised Learning, Active Learning. ML in Summary and Path Ahead - Research Survey Articles.	

Text Books

S.N	Title	Authors	Edition	Publisher
1.	Machine Learning- an Artificial Intelligence Approach	Tom Mitchell	2nd Edition	Morgan Kaufmann
2.	Algorithms for Clustering Data	A. K. Jain and R. C. Dubes	5th Edition	Prentice Hall PTR
3.	Introduction to Machine Learning	Ethem Alpaydin	2nd Edition	Prentice Hall PTR

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS502P	Machine Learning for Data Science Lab	-	-	2	1	25	25	50

Course Objectives	Course Outcomes
this course is intended • To implement basic machine learning algorithm for solving problem. • To understand the usage of datasets in implementing machine learning problems. • To learn various modern tools, packages and techniques for machine learning.	**Students will be able to** • implement procedures for machine learning algorithms. • Design python programs for various learning algorithms. • Apply appropriate machine learning algorithms to various data sets. • Apply machine learning algorithms to solve real world problems.

Expt . No.	Title of the experiment
1	To implement the concept of data pre-processing (Working with various data types, fixing values, handling missing values and outliers, etc.) using public dataset.
2	Perform regression on a given dataset.
3	Build a decision tree classifier for the dataset.
4	Apply a Naïve Bayes classifier to a dataset.
5	Apply a SVM on a given dataset.
6	Use the Apriori algorithm to discover frequent itemsets from a dataset.
7	Perform K-means clustering on a given dataset.
8	A micro-project based on real world problems.

Text Books

S.N	Title	Authors	Edition	Publisher
1.	Understanding Machine Learning: From Theory to Algorithms	Shalev-Shwartz,S., Ben-David,S	2nd Edition	Cambridge University Press
2.	Pattern Recognition and machine learning	Christopher Bishop	5nd Edition	Springer Verlag

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Machine Learning- an Artificial Intelligence Approach	Tom Mitchell	2nd Edition	Morgan Kaufmann
2.	Algorithms for Clustering Data	A. K. Jain and R. C. Dubes	5th Edition	Prentice Hall PTR
3.	Introduction to Machine Learning	Ethem Alpaydin	2nd Edition	PHI

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS503T	Software Engineering and Project Management	2	-	-	2	15	35	50

Course Objectives	Course Outcomes
This course is intended ● Understand the key concepts of software engineering, including its purpose, phases, and the role of process models in software development. ● To teach students how to effectively gather, model, and analyze software requirements and how to design solutions that fulfill these requirements. ● To provide students with the knowledge and skills needed to ensure the quality of software through rigorous testing, planning, and quality assurance practices.	**Students will be able to** ● Select and justify the use of process models for different software projects. ● Gather and analyze software requirements, create SRS, and design modular and scalable software. ● Apply software testing methods, plan software projects, and ensure software quality and reliability.

Unit I	**[10Hrs]**
Software Engineering & Process Models-What is Software? Characteristics and Types, The Role of a Software Engineer in the SDLC, Phases of Software Development, Software Process Models, Evolutionary Models, Incremental Model, Spiral Model, Extreme Programming (XP).	
Unit II	**[10Hrs]**
Requirements Engineering & Software Design- Initiating the Requirements Process, Eliciting Requirements from Stakeholders, Requirements Analysis Techniques, Software Requirements Specification (SRS), Software Design: What is Design, Design Principles, User Interface Design: User Interface Principles, User-Centered Design.	
Unit III	**[10Hrs]**
Fundamentals of Software Testing-Types of Testing: White Box Testing, Black Box Testing, Unit Testing, Integration Debugging Techniques, Software Project Planning, Concepts and Importance of Software Quality Assurance, Software Quality Factors ,Software Reviews, Configuration Management (SCM).	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Software Engineering, A practitioner's approach	Roger Pressman	7th Edition	Tata Mcgraw Hill
2	Object Oriented Software Engineering Using UML Patterns and Java	Bernd Bruegge & Allen H. Dutoit.	2nd Edition,	

Reference Books

S.N	Title	Authors	Edition	Publisher
1	OOA and Design	Grady Booch	3nd Edition,	Ad. Wesly
2	OO Modeling and design	Rambhaugh	2nd Edition,	PHI

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS503P	Software Engineering and Project Management Lab	-	-	2	1	25	25	50

Course Objectives	Course Outcomes
This course is intended • Identify Project Requirements and Develop a Formal Specification • Apply Design Principles in System Development: • Implement Effective Testing Methods to Ensure Software Quality.	**Students will be able to** • Gather and Analyze Project Requirements, Create Formal Specifications, and Plan Development Tasks. • Implement Design Principles Throughout the System Development Process. • Ensure quality of software by implementing testing methods.

Expt . No.	Title of the experiment
1A	**To Study Software Development Life Cycle**
1B	**Identifying the Requirements from Problem Statements** Requirements | Categorization of Requirements | Functional Requirements | Non-Functional Requirements | Other Requirements
2	**Estimation of Project Metrics** Project Estimation Technique| COCOMO Model | Advantages of COCOMO | Drawbacks of COCOMO | Basic COCOMO model to estimate project parameters | Project cost estimation
3	**Scheduling Project** Identifying Tasks | Identifying Resources | Schedule Project Use GANTT chart
4	**Modeling UML Use Case Diagrams and Capturing Use Case Scenarios** Identifying Actors | Identifying Use cases | Draw Use Case diagrams
5	**Modeling Data Flow Diagram & Control Flow Diagram** Draw Data Flow Diagram | Control Flow Diagram
6	**Modeling UML Class Diagrams** Structural and Behavioral aspects | Class diagram | Elements in class diagram | Class | Relationships | Draw Class Diagram
7	**Modeling Sequence Diagrams** Sequence diagram | Elements in sequence diagram | Object | Life-line bar | Messages | Draw Sequence Diagram
8	**Designing Test Suites** Software Testing | Need for Software Testing | Testing Frameworks | | Test Cases | Types of Software Testing | Unit Testing | Integration Testing| Write Test Cases (format specified CASE study tab on V-Lab)

Text Books

S.N	Title	Authors	Edition	Publisher
1	Software Engineering, A practitioner's approach	Roger Pressman	7th Edition	Tata Mcgraw Hill
2	The Art of Software Testing	Glenford J. Myers	4th Edition	**John Wiley & Sons**.

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS504P	Data Analytics Lab	-	-	2	1	25	25	50

Course Objectives	Course Outcomes
This course is intended • Understand key techniques and tools for data cleaning, transformation, and analysis in modern data workflows. • Apply machine learning and statistical methods for building and evaluating predictive and analytical models. • Analyze performance and efficiency of data engineering and machine learning pipelines using scalable tools and frameworks.	**Students will be able to** • Explain data pre-processing concepts and perform data profiling and cleansing using modern Python • Apply EDA techniques, statistical analysis, and machine learning libraries to extract insights and build models • Develop and deploy data-driven solutions such as predictive models pipelines using Flask • Analyse time series trends and performance bottlenecks in large-scale datasets using high-performance frameworks

Syllabus: Unit 1: Data Wrangling and Cleaning, Unit 2: Exploratory and Statistical Data Analysis, Unit 3: Predictive Modeling , Unit 4: Time Series and Scalable Data Analysis

Expt . No.	Title of the experiment
1	Perform data wrangling and cleaning using Pandas and Pyjanitor
2	Build a data cleansing pipeline **using** Apache NiFi or Talend
3	Conduct interactive exploratory data analysis using Sweetviz or Pandas Profiling
4	Analyze correlation and multicollinearity using Seaborn heatmaps and VIF analysis
5	Deploy a predictive machine learning model using Flask API
6	Build a sentiment analysis pipeline using SpaCy and Scikit-learn
7	Perform time series modeling using ARIMA (Statsmodels)
8	Execute descriptive data analysis on large datasets using Dask **or** Polars

Text Books

S.N	Title	Authors	Edition	Publisher
1.	Python Data Science Handbook: Essential Tools for Working with Data	Wes McKinney	Second Edition (2023)	O'Reilly Media
2.	Data Engineering with Python	Paul Crickard	First Edition (2020)	Packt Publishing

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Data Science and Data Analytics Using Python	Munesh Chandra Trivedi and Anil Kumar Dubey	First Edition (2025)	Khanna Publishing House

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS505T(i)	Professional Elective –I (**Distributed Databases)**	**2**	**-**	**-**	**2**	**15**	**35**	**50**

Course Objectives	Course Outcomes
This course is intended ● Understand distributed database principles, design, and architectures. ● Explore query processing, transaction management, and concurrency control. ● Analyze security, fault tolerance, and emerging trends in distributed databases.	**Students will be able to** ● Discuss the fundamental concepts, architectures, and advantages of distributed databases. ● Analyze distributed query processing, optimization techniques, and cost estimation strategies. ● Apply transaction management and concurrency control mechanisms in a distributed environment. ● Evaluate security challenges, fault tolerance, and reliability aspects in distributed databases. ● Explore emerging trends such as NoSQL, NewSQL, cloud-based databases, and blockchain technology.

Unit I	**[10Hrs]**
Fundamentals of Distributed Databases :Concepts & Characteristics, Centralized vs. Distributed Databases, Data Fragmentation, Replication, and Allocation, Distributed Database Architectures Distributed Query Processing (08 Hours),Distributed Database Design, Query Decomposition & Optimization, Query Execution Strategies, Cost Estimation in Distributed Queries	
Unit II	**[10Hrs]**
Transaction Management & Concurrency Control (08 Hours),Distributed Transactions & ACID Properties, Concurrency Control: Two-Phase Locking, Timestamp Ordering, Deadlock Detection & Prevention Two-Phase & Three-Phase Commit Protocols	
Unit III	**[10Hrs]**
Security, Fault Tolerance & Reliability ,Security Challenges & Access Control, Data Integrity & Confidentiality, Fault Tolerance & Recovery Mechanisms, Case Studies: Google Spanner, Amazon Dynamo DB, Apache Cassandra Emerging Trends & Future Scope ,Cloud-Based & NoSQL Databases,NewSQL & Distributed Ledger Technology (Blockchain),Edge Computing & Big Data Storage, Future of Distributed Databases	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Distributed Databases: Principles and Systems	Stefano Ceri, Giuseppe Pelagatti	2nd	Mcgraw Hill Education
2	Database Systems: Models, Languages, Design And Application Programming	Ramez Elmasri, Shamkant B. Navathe,	3rd	Pearson

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS505T(iii)	Professional Elective –I (**Image Analytics)**	**2**	**-**	**-**	**2**	**15**	**35**	**50**

Course Objectives	Course Outcomes
This course is intended • To provide a foundational understanding of digital image formation and representation. • To introduce methods for feature extraction, image segmentation, and pattern recognition. • To familiarize students with advanced topics such as image classification, object detection, and deep learning applications in image analytics.	**Students will be able to -** • Learn the fundamentals of image formation and representation. • Apply various image processing techniques for enhancement, restoration, and compression. • Implement algorithms for feature extraction, image segmentation, and pattern recognition. • Develop and evaluate models for image classification and object detection using machine learning and deep learning approaches.

Unit I	**[10 Hrs]**
Digital Image Fundamentals: Introduction to digital images, Color image processing, Basic geometric transformations, Image Enhancement techniques. Image Transforms and Frequency Domain Processing: Fourier Transform and its properties, Frequency domain filtering, Image restoration.	
Unit II	**[10 Hrs]**
Feature Extraction and Image Segmentation: Edge detection, Line and shape detection, Morphological operations, Segmentation techniques.	
Unit III	**[10 Hrs]**
Image Compression and Representation: Lossy and lossless compression techniques, Image compression standards, Video compression standards and formats, Feature representation. Advanced Topics in Image Analytics: Pattern recognition techniques, Machine learning in image analytics, Deep learning approaches. Applications: Object detection, face recognition, image captioning, and scene understanding.	

Text Books

S.N	Title	Authors	Edition	Publisher
1.	Digital Image Processing	Gonzalez, R.C., & Woods, R.E. (2011)	2nd	Pearson Education.
2.	Fundamentals of Digital Image Processing	Jain, A.K.	3rd	Prentice-Hall

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Computer Vision and Image Processing: Fundamentals and Applications	Bhuyan, M.K.	2nd	CRC press

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS505T(ii)	Professional Elective –I (**Network Security)**	**2**	**-**	**-**	**2**	**15**	**35**	**50**

Course Objectives	Course Outcomes
This course is intended • To learn the fundamentals of computer network . • To learn the key techniques cryptography • To explore the network security techniques.	Students will be able to - • Remember the network components • Explore structures of network topologies • Examine cryptography and hash function • Apply cryptography algorithm • Implement network security

Unit I	**[10 Hrs]**
Introduces the concept of computer networks, their uses, and different types (LAN, WAN, etc.). Discusses the components of a network, including network interface cards (NICs), cables, hubs, switches, routers, and the software that manages them. various network topologies like bus, star, ring, and mesh. Open System Interconnection (OSI) model and the Transmission Control Protocol/Internet Protocol (TCP/IP) model, outlining their structure and functions.	
Unit II	**[10 Hrs]**
Introduction to Cryptography: conventional, public key cryptography , hash functions	
Unit III	**[10 Hrs]**
Cryptography Algorithm: DES, AES symmetric-key algorithms, SHA, Diffie-Hellman asymmetric-key algorithms, SHA Hash function. Network Security Protocols: Network access control, IP security	

Text Books

S.N	Title	Authors	Edition	Publisher
1.	Network Security Essentials	William Stalling	5	Pearson Education

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Network Security: Private Communications in a Public World	Mike Speciner, Radia Perlman, Charlie Kaufman	2	Pearson Education

ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING & TECHNOLOGY,NAGPUR

(An autonomous institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

B. Tech. Scheme of Examination & Syllabus 2023-24

COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS561O	Open Elective – II (**Business Intelligence)**	**3**	**-**	**-**	**3**	**30**	**70**	**100**

Course Objectives	Course Outcomes
This course is intended • To introduce the foundational concepts of Business Intelligence (BI) and Business Analytics (BA), along with the types of analytics used in business decision-making. • To equip students with practical knowledge of data management, data warehousing, data mining, and visualization tools used in BI systems. • To enable students to analyze business problems through case studies and apply BI tools and techniques for effective decision support,	**Students will be able to** • Explain the fundamentals of business intelligence • Design Business Intelligence system architecture and apply data management techniques to perform data mining and analyze customer lifetime value to support strategic business decisions. • Implement statistical learning and data mining techniques to solve business problems using decision trees, ensemble methods, • To design and implement Decision Support Systems (DSS) using BI tools • Evaluate emerging trends in BI, including the role of AI, machine learning, big data, and real-time analytics in strategic decision-making

Unit I	**[9Hrs]**
Introduction to Business Intelligence and Analytics- Overview of BI and BA, Types of Analytics: Descriptive, Diagnostic, Predictive, Prescriptive, Data Management Fundamentals, Data Warehousing and Architecture, ETL Process, OLTP vs. OLAP.	
Unit II	**[9Hrs]**
Analytical Techniques and Data Management- BI System Architecture, Database Design for BI Data Mining Techniques, Data Visualization Tools, Customer Lifetime Value.	
Unit III	**[9Hrs]**
Case study and project- Case Studies on Data Mining in Introduction to Statistical Learning in Business Contexts, Data Pre-processing for BI Projects, Ensuring Data Quality in BI Systems, , Evaluating Model Performance with ROC and PR Curves, Case-based Decision Trees and Tree Induction, Pruning and Ensemble Methods in BI Solutions.	
Unit IV	**[9Hrs]**
Decision Support Systems (DSS) and BI Applications- Overview of DSS, Types of DSS, Information Gathering for BI, Customer Analytics and Segmentation , System Implementation- Introduction to BI Tools: Tableau.	
Unit V	**[9Hrs]**
Advanced BI Topics and Emerging Trends- Advanced Analytics Models, Knowledge Management in BI, Collaborative BI Systems, Strategic BI for Competitive Advantage, Social Media and Web Analytics, Big Data in BI, Emerging Trends: AI, Machine Learning, Real-Time Analytics.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Business Intelligence: The Savvy Manager's Guide	David Loshin	2nd	Newnes
2	Business Intelligence: Practices, Technologies, and Management	1. Rajiv Sabherwal, Irma Becerra-	3rd	Wiley & Sons.

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Business Intelligence: Making Decisions Through Data Analytics, ,	Jerzy Surma,	2nd	Business Expert Press

		JULY 2025	1	Applicable for 2025-26
Chairman - BoS	Dean – Academics	Date of Release	Version	

ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING & TECHNOLOGY,NAGPUR

(An autonomous institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

B. Tech. Scheme of Examination & Syllabus 2023-24

COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS531M	Multi-Disciplinary Minor-III (Data Engineering)	3	-	-	3	30	70	100

Course Objectives	Course Outcomes
This course is intended • Build a strong foundation in data structures, algorithms, and database management for data science applications. • Develop expertise in data preprocessing, pattern recognition, visualization, and statistical analysis for effective decision-making.	**Students will be able to:** • Apply fundamental data structures and algorithms to solve real-world problems in data science. • Design and implement relational databases using SQL, including advanced queries and procedures. • Perform data cleaning and preprocessing, including feature engineering and dimensionality reduction techniques. • Analyse patterns in data using clustering methods such as K-means, hierarchical clustering, and density-based clustering. • Visualize data effectively using charts and plots, and apply statistical hypothesis testing to extract insights from case studies.

Unit I **[9Hrs]**

Data structures: array, matrix, stack, queue, dictionary, graph; Basics of algorithm, pseudo-code, testing-debugging; Searching, sorting; Graph traversal: depth-first, breadth-first; Trees, Binary search; Problems related to data science.

Unit II **[9Hrs]**

Purpose of Database System -– Views of data – data models, database management system, three-schema architecture of DBMS, components of DBMS. E/R Model , entities, entity types, attributes relationships, relationship types, E/R diagram notation, examples. RDBMS basics, SQL: basics.

Unit III **[9Hrs]**

Data Cleaning, Transformation and Feature Engineering- Importance of Data Cleaning and Preprocessing, Handling Missing Data, Outliers, Noise ,Data Encoding, Normalization, and Standardization, Feature Engineering: Text and Graph Data, Dimensionality Reduction: PCA, t-SNE, Autoencoders, Introduction to Data Pipelines and Data Wrangling (using Python tools like Pandas).

Unit IV **[9Hrs]**

Data Mining and Pattern Discovery: Introduction to Data Mining and Knowledge Discovery, Classification, Clustering, and Association Rules, Clustering Techniques: K-means, Hierarchical, DBSCAN ,Association Rule Mining: Apriori, FP-Growth,Basics of Recommendation Systems ,Case Studies in Market Basket Analysis and Customer Segmentation

Unit V **[9Hrs]**

Data Analytics and Visualization:- Descriptive Statistics: Mean, Median, Mode, Variance, Std. Deviation,Probability Distributions (Normal, Binomial, Poisson),Inferential Statistics: Hypothesis Testing, p-value, Confidence Intervals,Statistical Tests: t-test, Chi-Square Test, ANOVA,Data Visualization Tools and Techniques: Histograms, Boxplots, Scatterplots, Heatmaps,Introduction to Experiment Design and A/B Testing, Visualization Tools: PowerBi (basics)

Text Books

S.N	Title	Authors	Edition	Publisher
1	Fundamentals of Database Systems	Ramez Elmasri	2nd	Pearson Education.
2	Data Visualization: A Practical Introduction	Kieran Healy	2nd	Prentice-Hall
3	Introduction to the Design and Analysis of Algorithms	Anany Levitin	2nd	Pearson Education.

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Statistics for Data Science	James D. Miller	2nd	CRC press
2.	"Fundamentals of Data Visualization"	By Claus O. Wilke	3rd	Pearson Education.

ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING & TECHNOLOGY,NAGPUR

(An autonomous institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)

B. Tech. Scheme of Examination & Syllabus 2023-24

COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)

FIFTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS506P	Technical Skill Development-II	2	-	-	1	50		50

Course Objectives	Course Outcomes
This course is intended • Equip students with fundamentals in Java and Object-Oriented Programming. • Develop skills for solving real-time coding problems relevant to technical interviews. • Support logical thinking through structured programming and data structures. • Prepare students for campus recruitment drives, coding rounds, and technical interviews.	**Students will be able to** • Learn Java programming language with object-oriented concepts. • Apply control structures and data structures to solve computational problems. • Develop modular, reusable code using OOP principles. • Analyze time and space complexity of algorithms. • Implement Java collections and basic algorithms used in coding interviews.

Unit I : Java Fundamentals and Programming Foundations	**[8Hrs]**
Java Runtime Environment, Java Basics, Advanced Data Types: Wrapper Classes, Autoboxing/Unboxing, Java Input/Output Coding best practices, Quick coding tasks	
Unit II : Advanced Arrays, Strings & Problem Solving	**[8Hrs]**
Arrays:2D Arrays: Matrix operations, Strings in Java, Problem Solving, Competitive-style problems using arrays/strings	
Unit III : Object-Oriented Programming & Design Thinking	**[5Hrs]**
Classes, Objects, Access Specifiers, this and super keywords, static members, Inheritance, Method Overloading & Overriding, Polymorphism, Abstraction, Hands-on: Design a Library or Employee Management system.	
Unit IV: Exception Handling, File I/O & Collections	**[5Hrs]**
Exception Handling, File I/O, Java Collections Framework, Use cases like student data, attendance logs, analytics, Mini-Projects.	
Unit V: Multithreading, Synchronization & Debugging	**[6Hrs]**
Thread Creation, Thread lifecycle and priorities, Synchronization, Exception Handling, Threading, and Testing Different Case study.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Programming with Java	E. Balagurusamy	2nd	McGraw Hill.
2	Head First Java	Kathy Sierra, Bert Bates	2nd	O'Rei;;y Media

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Core Java Volume I – Fundamentals	Cay S. Horstmann	11th	Pearson
2	Data Structures and Algorithms in Java	Robert Lafore	2nd	Pearson

		JULY 2025	1	Applicable for 2025-26
Chairman - BoS	Dean – Academics	Date of Release	Version	