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)

SEMESTER VII

Sr No	Course Category	Course Code	Course Title	Hours per Week			Credits	Maximum Marks				Minimum Marks	No of Hours for ESE
				L	T	P		Mid Semester Examination	Continual Assessment	End Sem Examination	Total		
1	PCC	**23DS701T**	Big Data Analytics	**3**	**-**	**-**	**3**	**15**	**15**	**70**	**100**	**45**	**3**
2	PCC	**23DS701P**	Big Data Analytics Lab	**-**	**-**	**2**	**1**	**-**	**25**	**25**	**50**	**25**	**-**
3	PCC	**23DS702T**	Deep Learning	**3**	**-**	**-**	**3**	**15**	**15**	**70**	**100**	**45**	**3**
4	PCC	**23DS702P**	Deep Learning Lab	**-**	**-**	**2**	**1**	**-**	**25**	**25**	**50**	**25**	**-**
5	PCC	**23DS703T**	Data Modelling and Optimization	**3**	**-**	**-**	**3**	**15**	**15**	**70**	**100**	**45**	**3**
6	PCC	**23DS703P**	Data Modelling and Optimization Lab	**-**	**-**	**2**	**1**	**-**	**25**	**25**	**50**	**25**	**-**
7	PEC	**23DS704T**	Program Elective – III	**3**	**-**	**-**	**3**	**15**	**15**	**70**	**100**	**45**	**3**
8	MDM	**23DS731M**	MDM – V (Refer MDM Basket)	**3**	**-**	**-**	**3**	**15**	**15**	**70**	**100**	**45**	**3**
9	ELC	**23DS705P**	Project – II	**-**	**-**	**4**	**2**	**-**	**50**	**50**	**100**	**45**	**-**
10	ELC	**23DS706P**	Summer / Winter Internship*	-	-	-	**2**	**-**	**50**	**-**	**50**	**25**	**-**
			Total	**15**	**0**	**10**	**22**	**75**	**250**	**475**	**800**	-	-

*** Summer / Winter Internship (Evaluation of Four weeks Internship Completion till 6th Semester)**

Multidisciplinary Minor - V	
23DS731M	Tools for Data Science

Program Elective -III	
23DS704T (i)	Data Warehousing and Mining
23DS704T (ii)	Social and Information Network Analysis
23DS704T (iii)	Generative AI and Prompt Engineering

Chairman - BoS	**Dean – Academics**	**July 2026**	**NEP 1.0**	**Applicable for 2026-27**
		Date of Release	**Version**	

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits		Evaluation		
						MSE	CA	ESE	Total
23DS701T	Big Data Analytics	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. To provide understanding of Big data and machine Learning principles 2. To enable students to understand stages involved in the development of Big Data Analytics. 3. To introduced big data tools information standard formats	**Students will be able to** 1. Examine the data analytic life cycle for selected problem statement 2. Interpret Hadoop Architecture and it's ecosystem components. 3. Analyze huge data set using Hadoop distributed file systems and MapReduce 4. Apply different data processing tools like Pig, Hive and Spark. 5. Execute machine learning algorithms using big data tools and platforms.

Unit I	**[9Hrs]**
Basics of Big Data: Introduction to Big Data, Big Data Characteristics, Types of Big Data, Traditional Versus Big Data Approach, Technologies Available for Big Data, Infrastructure for Big Data, Use of Data Analytics, Big Data Challenges, architectures Data analysis process, Data analytics lifecycle.	
Unit II	**[9Hrs]**
The Hadoop MapReduce: The Hadoop MapReduce fundamentals, writing a Hadoop MapReduce example, RHIPE architecture and RHadoop, Understanding different Hadoop modes, Understanding Hadoop features, The HDFS and MapReduce architecture.	
Unit III	**[9Hrs]**
Hive Architecture and Installation: Comparison with Traditional Database, HiveQL Querying Data, Sorting and Aggregating, Map Reduce Scripts, Joins & Sub queries, HBase concepts, Advanced Usage, Schema Design, Advance Indexing.	
Unit IV	**[9Hrs]**
Learning Data Analytics: PIG Introduction, Use Case for Pig, Pig Latin Overview, Pig Primitive Data Types, Execution Modes of Pig, HDFS Commands, Complex Data Types, User-Defined Functions, Parameter Substitution, Diagnostic Operator, Word Count Example using Pig.	
Unit V	**[9Hrs]**
Supervised and unsupervised machine learning: Introduction to Spark programming model and MLib library, Content based recommendation systems. Generalized Linear Models.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Big Data and Analytics	Seema Acharya	2	Wiley
2	Practical Big Data Analytics Hands	Nataraj Dasgupta	3	Packt Publishing

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Hadoop The Definitive Guide	Tom White	4	O'Reilly
2.	Big Data Fundamentals: Concepts, Drivers & Techniques	Wajid Khattak, Paul Buhler, Thomas	3	John Wiley & Sons
3.	Big Data Analytics with Spark	Mohammed Guller	2	Packt

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS701P	Big Data Analytics Lab	-	-	2	1	25	25	50

Course Objectives	Course Outcomes
This course is intended 1. To provide understanding of Big data and machine Learning principles 2. To enable students to understand stages involved in the development of Big Data Analytics. 3. To introduced big data tools information standard formats	**Students will be able to** 1. Installation and understanding of Hadoop Architecture and its ecosystems 2. Analyze huge data set using Hadoop distribute file systems and MapReduce 3. Apply different data processing tools like Pig, Hive and Spark. 4. Execute machine learning algorithms using big data tools and platforms.

Expt. No.	Title of the experiment
1	Implement the installation of Big Data tool HADOOP.
2	Implementing the basic Hadoop HDFS Commands like File/Directory creation, deletion, update, and many more operations.
3	To Develop a MapReduce program to calculate the frequency of a given word in a given file
4	Installation of Hive along with practice examples.
5	To implement Hive for Create, Alter, and Drop databases query in tables using HQL
6	Installation of PIG with Write Pig Latin scripts sort, group, join, project, and filter your data.
7	Implement any one ML- algorithm using Apache Spark.

Text Books

S.N	Title	Authors	Edition	Publisher
1.	Practical Big Data Analytics Hands-on Techniques to Implement Enterprise Analytics and Machine Learning Using Hadoop, Spark, NoSQL and R	Nataraj Dasgupta	3	Packt
2.	Hadoop The Definitive Guide	Tom White	4	O'Reilly

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Hadoop for Dummies	Dirk Darooz,, Paul Zikopolus,Roman Melnky	2	Wiley
2.	Big Data Analytics with Spark	Mohammed Guller	2	Packt

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation			
						MSE	CA	ESE	Total
23DS702T	Deep Learning	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. To introduce basic deep learning algorithms. 2. To understand real-world problems which can be solved by deep learning methods. 3. To identify deep learning techniques suitable for a real-world problem	**Students will be able to** 1. Illustrate perceptrons and multilayer perceptrons to basic learning problems. 2. Evaluate feedforward neural networks and apply gradient-based optimization techniques for training deep learning models. 3. Apply activation functions, regularization techniques, and hyperparameter tuning methods to improve model generalization. 4. Employ convolutional neural network architectures for complex problems. 5. Create recurrent neural network–based models for various applications.

Unit I	**[9Hrs]**
Basics of Deep Learning History of Deep Learning, McCulloch Pitts Neuron, Thresholding Logic, Perceptrons,, Perceptron Learning Algorithm and Convergence, Multilayer Perceptrons (MLPs), Representation Power of MLPs, Sigmoid Neurons.	
Unit II	**[9Hrs]**
Feedforward Neural Networks, Representation Power of Feedforward Neural Networks, Training of Feedforward Neural Networks, Gradient Descent (GD), Momentum Based GD, Nesterov Accelerated GD, Stochastic GD, AdaGrad, RMSProp, Adam	
Unit III	**[9Hrs]**
Optimization Algorithm Activation Function and Initialization Methods: Sigmoid, Tanh, ReLU. Regularization: Bias and Variance, Overfitting, Hyperparameters Tuning, L1 and L2 Regularization, Data Augmentation and Early Stopping, Parameter Sharing and Tying.	
Unit IV	**[9Hrs]**
Convolutional Neural Networks, 1D and 2D Convolution, Visualizing Convolutional Neural Networks, Guided Backpropagation.	
Unit V	**[9Hrs]**
Recurrent Neural Networks, Vanishing and Exploding Gradients, Long Short-Term Memory, Gated Recurrent Units (GRUs). Variants of CNN and RNN Encoder-Decoder Models.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Introduction to Deep Learning	Sandro Skansi	2	Springer
2	Neural Networks and Deep Learning	Charu C. Aggarwal	3	Springer

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Pattern Recognition and Machine Learning	Bishop, C. M	3	Springer
2	Artificial Neural Networks	Yegnanarayana, B	3	PHI Learning

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS702P	Deep Learning Lab	-	-	2	1	25	25	50

Course Objectives	Course Outcomes
This course is intended 1. Introduce fundamental deep learning architectures and training mechanisms. 2. Apply deep learning techniques to real-world datasets. 3. Select and implement suitable deep learning models for given problem domains.	**Students will be able to** 1. Implement fundamental deep learning models including multilayer perceptrons and feedforward neural networks for supervised learning tasks. 2. Apply gradient-based optimization and regularization techniques to train neural networks with improved convergence and generalization. 3. Construct convolutional neural network architectures to analyze and process one-dimensional and two-dimensional data. 4. Develop sequence learning models using recurrent neural networks for temporal and sequential data applications.

Expt. No.	Title of the experiment
1	To implement Perceptron for Linearly Separable Problems
2	To implement a Multilayer Perceptron for classification on a given dataset
3	To develop a Feedforward Neural Network and analyze its representation capability
4	To apply Gradient Descent and its variants for training neural networks
5	To implement regularization techniques to reduce overfitting in neural networks
6	To design and implement a 1D Convolutional Neural Network for sequential data
7	To design and implement a 2D Convolutional Neural Network for image classification
8	To implement a Recurrent Neural Network for sequence modeling

Text Books

S.N	Title	Authors	Edition	Publisher
1	Deep Learning	Ian Goodfellow	2	MIT Press
2	Neural Networks and Deep Learning	Charu C. Aggarwal	3	Springer

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow	Aurélien Géron	3	O'Reilly Media
2.	Deep Learning with Python	François Chollet	2	Manning

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation			
						MSE	CA	ESE	Total
23DS703T	Data Modeling and Optimization	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. To equip students with the knowledge and skills to design efficient database schemas, apply normalization and optimization techniques, evaluate performance metrics, and address ethical considerations, preparing them to effectively model and optimize data systems for diverse business needs.	**Students will be able to** 1. Demonstrate basic knowledge of fundamental principles, concepts, standards, and practices of data modeling. 2. Analyze a given database schema and identify potential areas for optimization. 3. Compare and differentiate various data modeling approaches and assess their impact on database performance and scalability. 4. Critique the design of a database schema in terms of its adherence to best practices and optimization principles. 5. Design a comprehensive data model for a specific business scenario, considering both conceptual and physical aspects.

Unit I	**[9Hrs]**
Data Modeling Foundations: Introduction to Data Modeling, Overview of data modeling concepts, Importance of data modeling in database design, Data Modeling development life cycle Types of data models (Conceptual, Logical, Physical, Enterprise) Data Modeling tools.	
Unit II	**[9Hrs]**
Data Model Standards & Enterprise Database Objects : Data Model Standards (Industry Practices) , Database Objects (Enterprise View), Normalization as a Design Optimization Tool.	
Unit III	**[9Hrs]**
Physical Data Modeling & Automation. Implementation of Physical data model in a database, Forward & Reverse Engineering, Script-based Modeling.	
Unit IV	**[9Hrs]**
Advanced Data Warehouse Modeling :Data Warehouse & Data Mart Strategy for Analytics, Design of Data Warehouse & Data Mart. OLAP Modeling & OLTP Modeling, Design of Dimension & the Fact Tables. Designing using Inmon's or Kimball's approach, schema Modeling.	
Unit V	**[9Hrs]**
Advanced Dimensional Modeling & NoSQL:Slowly Changing Dimensions Type 1, Type II & Type III, Advanced Dimensions Meta Data and Maintenance of the Data Model. Introduction to NoSQL Data Modeling.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Data Modeling and Database Design	Narayan S. Umanath	2	Cengage Learning
2	Cloud Security and Privacy An Enterprise Perspective onRisks and Compliance	Tim Mather, Subra Kumaraswamy	3	O'Reilly

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Data Modeling Made Simple: A Practical Guide for Business & IT Professionals	Steve Hoberman	2	Technics
2	Beginning Data Science in R: Data Analysis, Visualization, and Modelling for Data Scientists	Thomas Mailund	3	Apress

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
23DS703P	Data Modeling and Optimization Lab	-	-	2	1	25	25	50

Course Objectives	Course Outcomes
Students will be able to 1. To equip students with the knowledge and skills to design efficient database schemas, apply normalization and optimization techniques, evaluate performance metrics, and address ethical considerations, preparing them to effectively model and optimize data systems for diverse business needs.	**Students will be able to** 1. Design professional data models for real-world business applications. 2. Apply data model standards and naming conventions to database schemas. 3. Perform forward and reverse engineering for database implementation. 4. Evaluate and Optimize database performance using schema design and indexing. 5. Design OLTP, OLAP, and hybrid SQL–NoSQL data models.

*Practical data modeling and optimization-on Techniques to Implement different data models Using Python (Libraries) and MySQL.

Expt. No.	Title of the experiment
1	Design a comprehensive data model for a business scenario.
2	Apply data model standards and naming conventions to an existing database schema and convert it into an industry-compliant professional model.
3	Perform forward engineering & reverse engineering by generating SQL scripts from a data model and deploying the schema in a database.
4	Optimize database performance by identifying slow queries and improving them using schema design techniques.
5	Design both OLTP and OLAP data models.
6	Design and implement a Star Schema model and Snowflake Schema for a Business Intelligence system.
7	Implement Slowly Changing Dimensions (Type 1, Type 2, and Type 3) for managing historical data changes.
8	Design a hybrid data model using both relational (SQL) and document-based (NoSQL) approaches for modern applications.

Text Books

S.N	Title	Authors	Edition	Publisher
1	Data Modeling and Database Design	Narayan S. Umanath, Richard W. Scamell	2	Cengage Learning,
2	Cloud Security and Privacy An Enterprise Perspective onRisks and Compliance	Tim Mather, Subra Kumaraswamy	3	O'Reilly

Reference Books

S. N	Title	Authors	Edition	Publisher
1	Data Modeling Made Simple: A Practical Guide for Business & IT Professionals	Steve Hoberman	2	Technics
2	Beginning Data Science in R: Data Analysis, Visualization, and Modelling for Data Scientists	Thomas Mailund	2	ApresS

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation			
						MSE	CA	ESE	Total
23DS704T(i)	PE – III Data Warehousing and Mining	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. Familiar with the concepts of data warehouse and data mining. 2. Acquainted with the tools and techniques used for knowledge discovery in databases.	**Student will be able to** 1. Illustrate the stages in building a Data Warehouse. 2. Apply various data preprocessing & data mining techniques to large data sets. 3. Demonstrate the various classification algorithm for prediction tasks. 4. Perform various clustering methods to analyze complex dataset. 5. Use appropriate data mining techniques to extract patterns from various dataset.

Unit I	**[9Hrs]**
Data Warehousing and Business Analysis: - Data warehousing Components, Building a Data warehouse, Architecture, Transformation Tools, Data Cube Computation methods, Online Analytical Processing, Multidimensional Data Analysis, Data Warehouse implementation.	
Unit II	**[9Hrs]**
Data Mining: - Fundamentals of Data Mining, Data Mining Functionalities, Data Preprocessing, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation, Association Rule Mining, Frequent Item set Mining Methods, Correlation Analysis, Constraint-Based Association Mining.	
Unit III	**[9Hrs]**
Classification and Prediction: - Classification by Decision Tree Induction, Bayesian Classification, Rule Based Classification, Lazy Learners, Classification by Back propagation, Support Vector Machine, Other Classification Methods.	
Unit IV	**[9Hrs]**
Cluster Analysis: - Types of Data in Cluster Analysis, Measuring Data Similarity and Dissimilarity, Partitioning Methods, Hierarchical methods, Model-Based Clustering Methods, Clustering High-Dimensional Data, Constraint-Based Cluster Analysis, Outlier Analysis.	
Unit V	**[9Hrs]**
Mining Object, Spatial, Multimedia, Text and Web Data: Spatial Data Mining, Multimedia Data Mining, Text Mining, Mining the World Wide Web, Graph Mining, Mining biological data, social networks, Mining time series and sequence data.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Data Mining Concepts and Techniques	Jiawei Han, Micheline Kamber and Jian Pei	3	Elsevier
2	Data Mining Techniques	Arun K. Pujari	3	Universities Press

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Introduction to Data Mining	Pang-Ning Tan, Michael Steinbach	2	PE
2	Data Mining and Data Warehousing	Alex Berson, Stephen Smith	2	McGraw-Hill

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation			
						MSE	CA	ESE	Total
22DS704T(ii)	PE – III Social and Information Network Analysis	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. Introduce foundational concepts of graphs, networks, and knowledge representation 2. Develop skills in metrics and algorithms to discover pattern 3. To examine community detection and its evaluations 4. Use extract and organize information into knowledge graphs 5. Apply real-world applications in social media, recommendations, and information retrieval	**Students will be able to** 1. Identify network structures, properties, and how knowledge is organized 2. Apply metrics and algorithms to discover patterns in social networks 3. Assess community detection issues and evaluation 4. Illustrate knowledge graph, its properties and examples 5. Applications of Social network and knowledge graph solutions for practical issues in both.

Unit I	**[9Hrs]**
Introduction to Network: Real-world examples: social media, web pages, biological systems, Graphs Basics, Types of Networks, Network Representation, Basic Network Properties, Clustering coefficient: Tendency of friends to be friends with each other	
Unit II	**[9Hrs]**
Importance of Centrality Matter, PageRank Algorithm, Finding and Understanding Network Communities, Core-Periphery Structure, Structural Balance: Understanding tension and cooperation in social networks, Visualizing Influence: Creating meaningful network diagrams	
Unit III	**[9Hrs]**
Community Detection Problem Explained, Modularity in community partition, Practical Algorithms (Louvain Method, Girvan-Newman Algorithm, Label Propagation), Evaluating Communities	
Unit IV	**[9Hrs]**
Introduction to Knowledge Graph, Representing Knowledge: Entities, Relationships, Properties, Knowledge Graph Models: Building Knowledge Graphs, Real-World Knowledge Graphs and examples	
Unit V	**[9Hrs]**
Social Network Applications: Influencer Identification, Recommendation Systems, Fraud Detection, Viral Marketing, Misinformation Detection **Knowledge Graph Applications**: Semantic Search, Question Answering Systems, Recommendation Systems, Information Retrieval, Chatbots and Virtual Assistants, **Integration and Challenges.**	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Networks, Crowds, and Markets: Reasoning About a Highly Connected World	David Easley and Jon Kleinberg	1	Cambridge University Press
2	Knowledge Graphs	Amit Sheth, Prateek Jain	1	Morgan Kaufmann

Reference Books

S. N	Title	Authors	Edition	Publisher
1	Network Science: Theory and Practice	Ted G. Lewis	1	Wiley-Blackwell
2	Graph Databases: New Opportunities for Connected Data	Ian Robinson, Jim Webber, and Emil Eifrem	2	O'Reilly Media

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation			
						MSE	CA	ESE	Total
23DS704T(iii)	PE – III Generative AI and prompt Engineering	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. To introduce the foundational concepts of Generative AI and its real-world applications. 2. To explain and compare different model architectures such as VAEs, GANs, and Transformers. 3. To introduce the concepts of Prompt Engineering and their significance in Large Language Models (LLMs). 4. To explore LLM applications and fine-tuning strategies. 5. To explore the applications of Generative AI in real-world industry domains.	**Students will be able to** 1. Explain the core concepts and applications of Generative AI models. 2. Analyze different generative model architectures and their use cases. 3. Apply prompt engineering techniques for generating effective LLM outputs. 4. Evaluate strategies for designing and customizing LLM applications. 5. Explore practical applications and case studies in the real world.

Unit I : Foundations of Generative AI	**[10Hrs]**
Python Basics: Basic data types, control structures, functions, and object-oriented programming (OOP) concepts. Introduction to Generative AI: Understanding the evolution of AI, overview of Generative AI systems, and types of generative content. Core Concepts in Generative AI: Generative vs. discriminative models, creativity in AI, and foundational principles of generative systems.	
Unit II : Generative Models and Architectures	**[10Hrs]**
Introduction to Generative Models, Variational Autoencoders (VAEs):working, loss function, latent space, Generative Adversarial Networks (GANs): Architectures, training dynamic, Transformer Models: self attention,encoder-decoder structure, BERT vs GPT: Understanding Key Architectures, difference and application, Comparative Analysis of generative Models.	
Unit III : Prompt Engineering and LLM Basics	**[10Hrs]**
Foundations of Prompt Engineering, Prompting Techniques and Strategies: Zero-shot prompting, Few-shot prompting, Chain-of-Thought (CoT) prompting, Role-based prompting, ChatGPT Functionality and API Overview, Limitations and Bias in LLM Outputs, Prompt Tuning vs Fine-Tuning, Case Studies (Text-based Examples)	
Unit IV : Building Applications using LLMs	**[9Hrs]**
Introduction to LangChain and Agent Design, Retrieval-Augmented Generation (RAG): Concept and Use, LLM Fine-Tuning: Parameter-efficient Fine-Tuning (PEFT), LoRA, Designing LLM Applications for Chatbots, Search, Code Assistants Prompt Debugging and Evaluation Techniques	
Unit V : Industry Applications and Case Studies	**[9Hrs]**
Applications in healthcare : Diagnosis assistance, Summarizing EMRs, Applications in finance: Fraud detection, report generation, Applications in education: tutoring, automated grading, GenAI in business intelligence and marketing, Future of GenAI : AGI, regulation, job implications, Prototype for GenAI solution for a chosen domain.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Deep Learning with Python	Francois Chollet	1	Manning Publications
2	Generative Deep Learning	David Foster	2	O'Reilly Media

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Deep Learning	Ian G, Yoshua B, Aaron	1	MIT Press
2	Neural Networks and Deep Learning	Rambhaugh	2	Determination Pres

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation			
						MSE	CA	ESE	Total
23DS731M	MDM – V Tools for Data Science	3	-	-	3	15	15	70	100

Course Objectives	Course Outcomes
This course is intended 1. To understand the role and applications of modern data science tools across the data lifecycle. 2. To develop the ability to use appropriate tools for data processing, visualization, and basic analytics. 3. To gain awareness of advanced tools and platforms used in industry for machine learning, big data, and deployment.	**Students will be able to** 1. Identify key tools used in the data science lifecycle. 2. Apply Python-based tools for data handling and analysis 3. Create meaningful visualizations and dashboards using power BI and Tableau 4. Use big data and cloud tools for handling large dataset

Unit I	**[9Hrs]**
Introduction to Data Science Tools: Overview of Data Science Tool chain**,** Categories of tools: Data Collection, Processing, Visualization, Deployment **Introduction** to Python ecosystem: Python for Data Science, Libraries: NumPy, Pandas Development Environments: Jupyter Notebook, Google Colab.	
Unit II	**[9Hrs]**
Data Preprocessing and Analysis Tools: Data cleaning and transformation, Handling missing data, outliers ,Data manipulation using Pandas ,Exploratory Data Analysis (EDA) ,Data integration from multiple sources Tools & Libraries: Pandas NumPy.	
Unit III	**[9Hrs]**
Data Visualization Tool – Tableau: Introduction to Tableau, Tableau workspace and components, Data Handling in Tableau, Creating Visualizations, Dashboards and Storytelling, Sharing and Publishing **.**	
Unit IV	**[8Hrs]**
Data Visualization Tool – Power BI: Importance of visualization in decision making ,Types of charts , Introduction to Power BI, Power BI Interface, Data Transformation, Data Modeling, Visualization and Reports, Sharing and Collaboration.	
Unit V	**[10Hrs]**
Big Data and Cloud Tools: Introduction to Big Data concepts, Distributed computing basics Tools: Apache Hadoop ,Apache Spark ,Cloud platforms: Amazon Web Services ,Google Cloud Platform , Data storage and processing in cloud. Use cases of Big Data in various industry sectors such as e-commerce, banking, and social media platforms.	

Text Books

S.N	Title	Authors	Edition	Publisher
	Python Data Analytics	Fabio Nelli	1	Apress
	Data Science for Dummies	Lillian Pierson	3	John Wiley & Sons

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Analyzing Data with Power BI and Power Pivot for Excel	Alberto Ferrari ,Marco Russo	1	Microsoft Press
2.	Big Data For Dummies	Judith S. Hurwitz ,Alan Nugent , Fern Halper , Marcia Kaufman	1	John Wiley & Sons