Scheme of Examination - VII Semester

Sr No	Course Code	Course Title	Hours per Week			Credits	Maximum Marks		
			L	T	P		Continual Assessment	End Sem Examination	Total
1	22DS701T	Big Data Analytics & Machine Learning	3	1	-	4	30	70	100
2	22DS701P	Big Data Analytics & Machine Learning Lab	-	-	2	1	25	25	50
3	22DS702T	Data Modeling and Optimization	3	1	-	4	30	70	100
4	22DS702P	Data Modeling and Optimization Lab	-	-	2	1	25	25	50
5	22DS703T	Professional Elective-III	4	-	-	4	30	70	100
6	22DS761O	Open Elective - III	3	-	-	3	30	70	100
7	22DS704P	Project - II	-	-	4	2	100	100	200
8	22DS705P	Summer/Winter Internship*	-	-	-	2	50	--	50
9	22DS706P	Capstone Course – II **	-	-	2	1	50	--	50
Total			**13**	**2**	**10**	**22**	**370**	**430**	**800**

* Summer / Winter Internship (Evaluation of Four weeks Internship Completion till 6th Semester)
** Capstone Course – II (Comprehensive knowledge gained in *CSE(Data Science))*

22DS761O	**Open Elective - III**
22DS761O(i)	OE- III Android App Development

22DS703T	**Elective - III**
22DS703T(i)	PE-III Artificial Intelligence
22DS703T(ii)	PE-III Compiler Construction
22DS703T(iii)	PE-III Mathematical Modelling for Data Science
22DS703T(iv)	PE-III Reinforcement Learning
22DS703T(v)	PE-III Generative AI and Prompt Engineering

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS701T	Big Data Analytics & Machine Learning	3	1	-	4	30	70	100

Course Objectives	Course Outcomes
This course is intended ● To provide understanding of Big data and machine Learning principles ● To enable students to understand stages involved in the development of Big Data Analytics. ● To introduced big data tools information standard formats	**Students will be able to** ● Examine the data analytic life cycle for selected problem statement ● Installation and understanding of Hadoop Architecture and its ecosystems ● Analyze huge data set using Hadoop distributed file systems and MapReduce ● Apply different data processing tools like Pig, Hive and Spark. ● Apply different Machine Learning Algorithm with Big Data Tools

Unit I	**[8Hrs]**
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	**[8Hrs]**
The Hadoop MapReduce, The Hadoop MapReduce fundamentals, writing a Hadoop MapReduce example, learning the different ways to write MapReduce in python. RHIPE architecture and RHadoop. Understanding different Hadoop modes, Understanding Hadoop features, The HDFS and MapReduce architecture.	
Unit III	**[8Hrs]**
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, PIG	
Unit IV	**[8Hrs]**
Learning Data Analytics PIG Introduction, Features, Philosophy, Use Case for Pig, Pig Latin Overview, Pig Primitive Data Types,Running Pig, Execution Modes of Pig, HDFS Comman21DS, Relational Operators, Eval Function, Complex Data Types, Piggy Bank, User-Defined Functions, Parameter Substitution, Diagnostic Operator, Word Count Example using Pig, Pig at Yahoo, Pig Versus Hive.	
Unit V	**[8Hrs]**
Supervised and unsupervised machine learning, Overview of regression analysis, clustering, data dimensionality, clustering method , Introduction to Spark programming model and MLib library, Content based recommendation systems. Generalized Linear Models and Logistic Regression, Regularization, Support Vector Machine (SVM) and the kernel trick, Outlier Detection, Spark ML library	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Big Data and Analytics,	Seema Acharya	2nd Edition 2019	Wiley
2	Practical Big Data Analytics Han21DS-	Nataraj Dasgupta ·	2018	Packt Publishing

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	., "Hadoop The Definitive Guide"	Tom White	Fourth Edition, 2015	O'Reilly Publications

2.	Big Data Fundamentals: Concepts, Drivers & Techniques	: Wajid Khattak, Paul Buhler, Thomas	Inc ISBN: 13: 9780134291079	John Wiley & Sons,
3.	Big Data Analytics with Spark	Mohammed Guller		PacktPublishing,2015

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS701P	Big Data Analytics And Machine Learning Lab			2	1	25	25	50

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

Expt. No.	Title of the experiment
1	Study and implement the installation of Big Data tool HADOOP and JDK1.8 for Java framework in Windows.
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 ·	2018	Packt Publishing
2.	Hadoop The Definitive Guide	Tom White	4th edition 2015	O'Reilly Publication

Reference Books

S. N	Title	Authors	Edition	Publisher
1.	Hadoop for Dummies	Dirk Darooz,, Paul Zikopolus,Roman Melnky		Wiley Publication

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS702T	Data Modeling & Optimization	3	1	-	4	30	70	100

Course Objectives	Course Outcomes
This course is intended ● 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 nee21DS.	**Students will be able to** ● Recall the fundamental principles of data modeling and optimization. ● Analyze a given database schema and identify potential areas for optimization. ● Compare and contrast different data modeling approaches and their implications on database performance. ● Critique the design of a database schema in terms of its adherence to best practices and optimization principles. ● Design a comprehensive data model for a specific business scenario, considering both conceptual and physical aspects.

Unit I	**[8Hrs]**
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	**[8Hrs]**
Data Model Standard Naming Standard of objects, Main object Table, Column, Datatype Database, Schema, Tablespace, Segment, Extent, Privileges, Index, View, Synonym, Normalization process	
Unit III	**[8Hrs]**
Physical Data Model, Database & Scripts: Forward Engineering, Reverse Engineering, Create a data model from a database & script, Database Vs data model. Implementation of Physical data model in a database.	
Unit IV	**[8Hrs]**
Data Warehouse, Data Mart, Design of Data Warehouse & Data Mart. Difference Between OLAP Modeling & OLTP Modeling, Design of Dimension & the Fact Tables. Designing using Inmon's or Kimball's approach. SnowFlake Modeling, Star Schema Modeling,	
Unit V	**[8Hrs]**
Slowly Changing Dimensions Type 1, Type II & Type III, Degenerate Dimension, Causal Dimension, Junk Dimension, Outrigger Dimension, Repository, Meta Data and Maintenance of the Data Model. Introduction to NoSQL Data Modeling, Explanation of JSON,* Explanation of Document Database	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Data Modeling and Database Design	Narayan S. Umanath, Richard W. Scamell		Cengage Learning, 2014
2	Cloud Security and Privacy An Enterprise Perspective onRisks and Compliance	Tim Mather, Subra Kumaraswamy		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 Publications, 2009
2	Beginning Data Science in R: Data Analysis, Visualization, and Modelling for Data Scientists	Thomas Mailund		Apress

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS702P	Data Modeling and Optimization Lab			2	1	25	25	50

Course Objectives	Course Outcomes
Students will be able to ● 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** ● Design and implement data models using Python and MySQL. ● Generate automated ER diagrams using data modeling tools.. ● Apply naming conventions and normalization in database design. ● Create and manage indexes, views, and ETL scripts for databases. ● Design OLTP and OLAP schemas for data warehousing.

*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	Automate the creation of a data model using Python
3	Implement database schema and tables based on naming standards and normalization principles
4	Create and manage indexes and views in a MySQL database.
5	Perform forward and reverse engineering of a data model
6	Implement a physical data model in a database
7	Design and implement a data warehouse schema.
8	Explore the Weka tool.

Text Books

S.N	Title	Authors	Edition	Publisher
1	Data Modeling and Database Design	Narayan S. Umanath, Richard W. Scamell		Cengage Learning, 2014
2	Cloud Security and Privacy An Enterprise Perspective onRisks and Compliance	Tim Mather, Subra Kumaraswamy		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 Publications, 2009
2.	Beginning Data Science in R: Data Analysis, Visualization, and Modelling for Data Scientists	Thomas Mailund		ApresS

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS703T(iv)	PE-III **(iv) Reinforcement Learning**	**4**	**-**	**-**	**4**	**30**	**70**	**100**

Course Objectives	Course Outcomes
This course is intended ● To introduce the basic mathematical foundations of reinforcement learning. ● To highlight some of the recent directions of research. ● To apply reinforcement learning techniques to real-world problems and case studies.	Students will be able to ● Explain the basic concepts of reinforcement learning. ● Analyze dynamic programming algorithms, Monte-Carlo learning, and Temporal-difference learning. ● Describe and apply Incremental Method, Batch Method, and Deep Q-Learning method. ● Develop strategies for balancing exploration and exploitation in reinforcement learning problems. ● Analyze and design multi-agent reinforcement learning systems.

Unit I	**[9Hrs]**
Probability Primer-Brush up of Probability concepts, Introduction to Reinforcement Learning(RL): Recent Advancements and Highlights, The RL Problem and Overall Landscape. Markov Decision Process (MDP): Markov Process, Markov Reward Process, Markov Decision Process and Bellman Equations.	
Unit II	**[9Hrs]**
Dynamic Programming: Policy Evaluation, Value Iteration, Policy Iteration, DP Extensions and Convergence using Contraction Mapping. Monte-Carlo (MC) Learning, Temporal-Difference (TD) Learning, TD(λ) and Eligibility Traces. On-Policy MC Control, On-Policy TD Learning, Off-Policy Learning, Q-Learning.	
Unit III	**[9Hrs]**
Value Function Approximation: Incremental Methods (Linear and Gradient based), Batch Methods (Least Square based).Deep RL: Deep Q-Learning, Deep Q-Networks and Experience Replay.	
Unit IV	**[10Hrs]**
Exploration and Exploitation (Bandits): Exploration Principles (Greedy, Optimistic, Probabilistic, Informative), Multi-arm Bandits, Contextual Bandits and Upper Confidence Bounds (UCB).	

Unit V	**[11Hrs]**
Multi-Agent RL: Cooperative vs. Competitive Settings, MARL Algorithms Applications and Case Studies, Introduction to Game theory.	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Reinforcement Learning: An Introduction	Richard S. Sutton	2nd edition	MIT Press, 2020

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS703T(v)	PE-III **Generative AI and Prompt Engineering**	**4**	**-**	**-**	**4**	**30**	**70**	**100**

Course Objectives	Course Outcomes
This course is intended ● To introduce the foundational concepts of Generative AI & its real world applications ● To explain and compare different general model architectures like VAEs, GANs and transformers ● To introduce the Prompt Engineering concepts and its significance in LLMs ● To explore LLM applications and fine-tuning strategies. ● To explore applications of GenAI in real-world industry domains.	**Students will be able to** ● Explain the core concepts and applications of Generative AI models. ● Analyze different generative model architectures and their use cases. ● Apply prompt engineering techniques for effective LLM outputs. ● Evaluate strategies for designing and customizing LLM applications. ● Exploring practical applications and case study in real world.

Unit I : Foundations of Generative AI	**[10Hrs]**
Python Basics: Basic data types, control structures, functions and OOP concepts. Introduction to Generative AI :Understand the evolution of AI, Overview of Generative AI systems, Types of generative content , concepts in Generative AI :Generative vs Discriminative Models, Creativity in AI.	
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	2nd Edition (2021)	Manning Publications
2	Generative Deep Learning	David Foster	2nd Edition (2022)	O'Reilly Media

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Deep Learning	Ian G, Yoshua B, Aaron Courville	1st Edition (2016)	MIT Press
2	Neural Networks and Deep Learning	Rambhaugh	1st Edition (2015)	Determination Pres

SEVENTH SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
22DS761O(i)	OE-III (i) Android App Development	3	-	-	3	30	70	100

Course Objectives	Course Outcomes
This course is intended ● To develop the ability of students to design android applications. ● To create user friendly android applications. ● Use various features of android like broadcast receivers, services, etc. ● Effectively use a database to store the data.	**Students will be able to** ● Design of basic android applications using the user interface (UI) elements of android OS. ● Implement Android's APIs for menus and intents. ● Use Android's APIs for data storage, retrieval, preferences, databases and content providers to create android apps. ● Implement Android's communication APIs for sending SMS, broadcast receiver.

Unit I	**[8Hrs]**
Android Introduction: Introduction to Android App, set up the development Environment Application Structure: AndroidManifest.xml, Uses-SDK, Layouts and & Drawable, Resources, Activities and Active Life Cycle	
Unit II	**[8Hrs]**
Emulator: Android Virtual Device**,** Layout resource, Running app on Emulator Basic UI Design**:** Form widgets, Text Field , Layouts, [dip, dp, sip, sp] versus px Preferences**:** Shared Preferences, Preferences from XML	
Unit III	**[8Hrs]**
Menu: Option Menu, Context Menu Intents**:** Explicit Intents, Implicit intents	
Unit IV	**[8Hrs]**
Advanced User Interface Design: **Time** and Date, Images Styles & Themes: styles.xml, drawable resources for shapes, gradients Content Providers**:** SQLite Programming	

Unit V	[8Hrs]
Widgets: ListView and ListActivity, Custom listview, GridView Notification: Broadcast Receivers	

Text Books

S.N	Title	Authors	Edition	Publisher
1	Beginning Android Programming with Android Studio	J. F. DiMarzio	4	Wrox publication
2	Professional Android 4 Application Programming	Reto Meier	2	Wiley Publication

Reference Books

S.N	Title	Authors	Edition	Publisher
1	Android Programming for Beginners	by John Horton	2	Packt Publishing Pvt. Ltd.