

### DSM241 Data Mining Using Python

Semester: IV	Course Title: Data Mining Using Python	Credit: 4
Course No.: DSM241		(3 T + 1 P)

Course Outcomes: On successful completion of the course the learner will be able to

CO	COGNITIVE ABILITIES	COURSE OUTCOMES
CO 1	REMEMBERING	Recall the key concepts, techniques, and algorithms used in data mining and their applications.
CO 2	UNDERSTANDING	Understand the various data mining tasks (classification, clustering, regression) and their implementation using Python.
CO 3	APPLYING	Apply Python libraries (such as Scikit-learn, Pandas, and Numpy) for implementing data mining algorithms.
CO 4	ANALYZING	Analyze and preprocess data to extract useful information, handle missing values, and prepare data for mining.
CO 5	EVALUATING	Evaluate the performance of various data mining models using appropriate metrics and validation techniques.
CO 6	CREATING	Design and implement data mining projects, including classification, clustering, and regression models, to solve real-world problems.

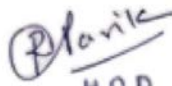
	PSO 1	PSO 2	PSO 3	PSO 4	PSO 5
CO 1	3	0	2	-	-
CO 2	2	1	1	-	-
CO 3	3	1	1	-	-
CO 4	2	1	-	1	-
CO 5	3	1	1	1	-
CO 6	3	1	1	-	1

Unit No.	Detailed Syllabus	Teaching Hours
I	<b>Introduction to Data Mining and Python Libraries</b> Overview of Data Mining: Definition, Tasks, and Applications Introduction to Python for Data Mining: Libraries and Tools (Scikit-learn, Pandas, Numpy, Matplotlib) Understanding Data Types: Structured, Semi-Structured, Unstructured Data Preprocessing: Cleaning, Handling Missing Data, Normalization, and Standardization Exploratory Data Analysis (EDA): Visualization, Correlation Analysis, and Feature Engineering	15
II	<b>Supervised Learning Techniques</b> Classification Algorithms: Decision Trees, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Naive Bayes	15

	Regression Algorithms: Linear Regression, Logistic Regression, and Model Evaluation Model Evaluation: Confusion Matrix, Precision, Recall, F1-Score, Cross-Validation Hyperparameter Tuning: Grid Search and Random Search Case Studies: Practical Implementation of Classification and Regression Problems	
III	<b>Introduction to Clustering and Other Techniques</b> Introduction to Clustering: Concepts, Applications, and Use Cases Dimensionality Reduction: Principal Component Analysis (PCA), t-SNE Association Rule Mining: Apriori Algorithm, Market Basket Analysis Anomaly Detection: Techniques Isolation Forest, Local Outlier Factor (LOF) Neural Networks for Data Mining: Basic Concepts and Applications Text Mining: Text Preprocessing, TF-IDF, Word Embeddings Model Deployment: Saving, Loading, and Evaluating Models with Scikit-learn	15
IV	<b>Practical Applications</b> <ul style="list-style-type: none"> <li>• Data Preprocessing and Exploration: Handling missing data, EDA</li> <li>• Feature Engineering and Data Preprocessing: Encoding, feature extraction</li> <li>• Classification Model Implementation and Evaluation: Decision Trees, KNN</li> <li>• Linear Regression Implementation and Evaluation</li> <li>• Hyperparameter Tuning: Using Grid Search</li> <li>• Dimensionality Reduction: Using PCA</li> <li>• <b>Final Project:</b> End-to-End Data Mining Project using Python (Including Data Preprocessing, Model Building, Evaluation, and Deployment)</li> </ul>	15

### Suggested Reference Books:

1. "Introduction to Data Mining" by Pang-Ning Tan, Michael Steinbach, and Vipin Kumar – Addison-Wesley
2. "Data Mining: Concepts and Techniques" by Jiawei Han, Micheline Kamber, and Jian Pei – Morgan Kaufmann
3. "Python Machine Learning" by Sebastian Raschka – Packt Publishing
4. "Data Science from Scratch" by Joel Grus – O'Reilly Media
5. "Hands-On Data Mining with R" by Manohar Swamynathan – Packt Publishing
6. "Pattern Recognition and Machine Learning" by Christopher M. Bishop – Springer

  
H.O.D  
Dept of DATA SCIENCE & ANALYTICS  
M.G. Science Institute, Ahmedabad-9.

### DSM242 Object Oriented Programming with JAVA

<b>Semester: IV</b>	<b>Course Title:</b> Object Oriented Programming with JAVA	<b>Credit: 4</b>
<b>Course No.:</b> DSM242		<b>(3 T + 1 P)</b>

**Course Outcomes: On successful completion of the course the learner will be able to**

CO	COGNITIVE ABILITIES	COURSE OUTCOMES
CO 1	REMEMBERING	Analyze Java code to identify and correct errors, optimize performance, and understand the flow of control in Java applications.
CO 2	UNDERSTANDING	Evaluate the performance of Java programs by implementing exception handling, threading, and synchronization techniques.
CO 3	APPLYING	Design and implement Java applications involving inheritance, polymorphism, interfaces, packages, and multithreading to solve real-world problems.
CO 4	ANALYZING	Analyze Java code to identify and correct errors, optimize performance, and understand the flow of control in Java applications.
CO 5	EVALUATING	Evaluate the performance of Java programs by implementing exception handling, threading, and synchronization techniques.
CO 6	CREATING	Design and implement Java applications involving inheritance, polymorphism, interfaces, packages, and multithreading to solve real-world problems.

	PSO 1	PSO 2	PSO 3	PSO 4	PSO 5
<b>CO 1</b>	3	0	2	-	-
<b>CO 2</b>	2	1	1	-	-
<b>CO 3</b>	3	1	1	-	-
<b>CO 4</b>	2	1	-	1	-
<b>CO 5</b>	3	1	1	1	-
<b>CO 6</b>	3	1	1	-	1

Unit No.	Detailed Syllabus	Teaching Hours
<b>I</b>	Introduction To OOP And Java Overview of OOP , Object oriented programming paradigms , Features of Object Oriented Programming , Java Buzzwords , Overview of Java , Data Types, Variables and Arrays , Operators , Control Statements , Programming Structures in Java , Defining classes in Java , Constructors Methods ,Access specifiers , Static members, Java Doc comments, Wrapper classes	<b>15</b>

<b>II</b>	Inheritance, Packages And Interfaces Overloading Methods , Objects as Parameters , Returning Objects ,Static, Nested and Inner Classes. Inheritance: Basics, Types of Inheritance ,Super keyword ,Method Overriding , Dynamic Method Dispatch ,Abstract Classes , final with Inheritance. Packages and Interfaces: Packages , Packages and Member Access ,Importing Packages , Interfaces, String and StringBuffer class.	<b>15</b>
<b>III</b>	Exception Handling And Multithreading Exception Handling basics , Multiple catch Clauses , Nested try Statements , Java's Built-in Exceptions , User defined Exception. Multithreaded Programming: Java Thread Model, Creating a Thread and Multiple Threads , Priorities , Synchronization , Inter Thread Communication, Suspending ,Resuming, and Stopping Threads ,Multithreading.	<b>15</b>
<b>IV</b>	Practical Applications <ul style="list-style-type: none"> <li>• Demonstrate object-oriented programming principles using a class with attributes and methods.</li> <li>• Implement method overloading in a class with multiple overloaded methods.</li> <li>• Create a program that defines and uses constructors in Java.</li> <li>• Write a program to showcase inheritance and method overriding, including the use of the super keyword.</li> <li>• Design a program to create and use packages, showcasing access to classes from another package.</li> <li>• Implement interfaces and demonstrate polymorphism with multiple implemented interfaces.</li> <li>• Develop a program to handle exceptions using try, catch, and finally, including user-defined exceptions.</li> <li>• Create a multithreaded program that demonstrates thread creation, synchronization, and communication between threads.</li> <li>• Write a program that demonstrates the use of wrapper classes and their methods.</li> <li>• Build a Java program that manipulates and processes String and StringBuffer objects.</li> </ul>	<b>15</b>

#### **Suggested Reference Books:**

1. **"Programming with Java: A Primer"** by E. Balagurusamy – McGraw Hill
2. **"Java: The Complete Reference"** by Herbert Schildt – McGraw Hill
3. **"Thinking in Java"** by Bruce Eckel – Prentice Hall
4. **"Core Java Volume I – Fundamentals"** by Cay S. Horstmann – Pearson
5. **"Object-Oriented Programming with JAVA"** by M. T. Savaliya – Dreamtech Press

  
H.O.D  
Dept of DATA SCIENCE & ANALYTICS  
M.G. Science Institute, Ahmedabad-9.

### DSM243 Data Warehousing and Data Integration

Semester: IV	Course Title: Data Warehousing and Data Integration	Credit: 4
Course No.: DSM243		(3 T + 1 P)

Course Outcomes: On successful completion of the course the learner will be able to

CO	COGNITIVE ABILITIES	COURSE OUTCOMES
CO 1	REMEMBERING	Recall key concepts, architecture, and methodologies used in data warehousing and data integration.
CO 2	UNDERSTANDING	Understand data warehousing techniques, ETL processes, and their role in business intelligence.
CO 3	APPLYING	Apply tools and frameworks for designing and implementing data warehousing and integration workflows.
CO 4	ANALYZING	Analyze data sources, integration challenges, and the effectiveness of ETL processes.
CO 5	EVALUATING	Evaluate the performance and scalability of data warehouses and integration pipelines.
CO 6	CREATING	Design and implement a data warehouse and integration solutions for real-world problems.

	PSO 1	PSO 2	PSO 3	PSO 4	PSO 5
CO 1	3	0	2	-	-
CO 2	2	1	1	-	-
CO 3	3	1	1	-	-
CO 4	2	1	-	1	-
CO 5	3	1	1	1	-
CO 6	3	1	1	-	1

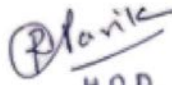
Unit No.	Detailed Syllabus	Teaching Hours
I	<b>Introduction to Data Warehousing</b> <ul style="list-style-type: none"> <li>Definition, Purpose, and Role of Data Warehousing</li> <li>Data Warehouse Architecture: Components and Layers</li> <li>Data Warehouse Design: Star Schema, Snowflake Schema, and Fact Constellations</li> <li>ETL (Extract, Transform, Load) Process Overview</li> <li>Metadata Management and Data Governance</li> </ul>	15

	<ul style="list-style-type: none"> <li>• Online Analytical Processing (OLAP): MOLAP, ROLAP, HOLAP</li> </ul>	
II	<b>Data Integration Concepts</b> <ul style="list-style-type: none"> <li>• Data Integration: Definition, Challenges, and Importance</li> <li>• Tools for Data Integration: Talend, Apache Nifi, Microsoft SSIS</li> <li>• Data Integration Techniques: Batch Integration, Real-time Integration, and Streaming</li> <li>• Data Quality Management and Validation</li> <li>• Data Federation and Virtualization</li> <li>• Integration of Structured, Semi-structured, and Unstructured Data</li> </ul>	15
III	<b>Advanced Topics in Data Warehousing and Integration</b> <ul style="list-style-type: none"> <li>• Data Warehouse Optimization: Techniques and Scalability</li> <li>• Incremental Data Warehousing: Change Data Capture (CDC) Techniques</li> <li>• Big Data and Data Warehousing: Hadoop, Hive, and Spark</li> <li>• Cloud Data Warehousing: AWS Redshift, Google BigQuery, Snowflake</li> <li>• Master Data Management (MDM)</li> <li>• Data Integration in Big Data Ecosystems</li> </ul>	15
IV	<b>Practical Applications</b> <ul style="list-style-type: none"> <li>• Design and Implementation of a Star Schema for a Sample Dataset</li> <li>• Implementation of ETL Workflows using Tools like Talend or Informatica</li> <li>• Real-time Data Integration Using Apache Nifi or SSIS</li> <li>• Data Cleansing and Transformation Exercises</li> <li>• Building an OLAP Cube and Running Queries</li> <li>• Integration of Multiple Data Sources Using Python or Spark</li> <li>• <b>Final Project:</b> <ul style="list-style-type: none"> <li>○ Build an End-to-End Data Warehouse Solution</li> <li>○ Include ETL Workflow, Data Integration, and Query Implementation</li> </ul> </li> </ul>	15

#### Suggested Reference Books:

1. "Data Warehousing: Fundamentals for IT Professionals" by Paulraj Ponniah – Wiley
2. "Building the Data Warehouse" by W. H. Inmon – Wiley

3. **"The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling"** by Ralph Kimball – Wiley
4. **"Data Integration Blueprint and Modeling"** by Anthony David Giordano – IBM Press
5. **"Big Data Integration and Processing"** by Martin Kleppmann – O'Reilly Media

  
H.O.D  
Dept of DATA SCIENCE & ANALYTICS  
M.G. Science Institute, Ahmedabad-9.