

MASTER OF COMPUTER APPLICATIONS (MCA_NEW)

ASSIGNMENTS OF MCA_NEW (2Yrs) PROGRAMME SEMESTER-II

(July - 2024 & January - 2025)

MCS-218, MCS-219, MCS-220, MCS-221

MCSL-222, MCSL-223



**SCHOOL OF COMPUTER AND INFORMATION SCIENCES
INDIRA GANDHI NATIONAL OPEN UNIVERSITY
MAIDAN GARHI, NEW DELHI – 110 068**

CONTENTS

Course Code	Assignment No.	Submission-Schedule		Page No.
		For January-June Session	For July-December Session	
MCS-218	MCA_NEW(II)/218/Assignment/24-25	31st October, 2024	15th April, 2025	3
MCS-219	MCA_NEW(II)/219/Assignment/24-25	31st October, 2024	15th April, 2025	4
MCS-220	MCA_NEW(II)/220/Assignment/24-25	31st October, 2024	15th April, 2025	5
MCS-221	MCA_NEW(II)/221/Assignment/24-25	31st October, 2024	15th April, 2025	6
MCSL-222	MCA_NEW(II)/L-222/Assignment/24-25	31st October, 2024	15th April, 2025	8
MCSL-223	MCA_NEW(II)/L-223/Assignment/24-25	31st October, 2024	15th April, 2025	9

Important Notes

1. Submit your assignments to the Coordinator of your Study Centre on or before the due date.
2. Assignment submission before due dates is compulsory to become eligible for appearing in corresponding Term End Examinations. For further details, please refer to Programme Guide of MCA (2Yrs).
3. To become eligible for appearing the Term End Practical Examination for the lab courses, it is essential to fulfill the minimum attendance requirements as well as submission of assignments (on or before the due date). For further details, please refer to the Programme Guide of MCA (2yrs).
4. The viva voce is compulsory for the assignments. For any course, if a student submitted the assignment and not attended the viva-voce, then the assignment is treated as not successfully completed and would be marked as ZERO.

Course Code	:	MCS-221
Course Title	:	Data Warehousing and Data Mining
Assignment Number	:	MCA_NEW(II)/221/Assign/2024-25
Maximum Marks	:	100
Weightage	:	30%
Last Date of Submission	:	31st October, 2024 (For July, 2024 Session) 15th April, 2025 (For January, 2025 Session)

This assignment has ten questions. All the questions are compulsory and there is no choice. Rest 20 marks are for viva voce. You may use illustrations and diagrams to enhance the explanations. Please go through the guidelines regarding assignments given in the Programme Guide.

- Q1:** Discuss the role of ETL (Extract, Transform, Load) processes in data warehousing. Provide a detailed explanation of each phase and its importance. Illustrate your answer with examples of common tools used in ETL and the challenges that may arise during these processes. **(8 Marks)**
- Q2:** (a) Explain the concept of Data Warehousing architecture. Compare and contrast the different types of architectures such as Single-tier, Two-tier, and Three-tier. Provide examples of scenarios where each architecture might be most beneficial. **(4 Marks)**
- (b) Analyze the concept of OLAP (Online Analytical Processing) and its significance in data warehousing. Describe the differences between MOLAP, ROLAP, and HOLAP. Discuss the advantages and disadvantages of each type with respect to data analysis and querying performance. **(4 Marks)**
- Q3:** Design a data warehouse schema for a retail company. Include fact tables, dimension tables, and consider the star schema and snowflake schema designs. Justify your design choices and discuss how your schema supports efficient query processing and business intelligence needs. **(8 Marks)**
- Q4:** Explain the use of metadata in data warehousing. Discuss the different types of metadata and their roles. Provide examples of how metadata can enhance the usability, maintenance, and performance of a data warehouse. **(8 Marks)**
- Q5:** Evaluate the role of data warehousing in supporting business intelligence and analytics. Discuss the process of transforming raw data into actionable insights. Provide examples of business intelligence tools and techniques that leverage data warehousing to enhance decision-making processes. **(8 Marks)**
- Q6:** Analyze various data pre-processing techniques such as data cleaning, data integration, data transformation, and data reduction. Explain the significance of each technique in improving the quality of data for mining and provide examples of scenarios where each technique would be applied. **(8 Marks)**
- Q7:** Compare and contrast the various classification algorithms used in data mining, such as Decision Trees, Naive Bayes, Support Vector Machines, and Neural Networks. Discuss the strengths and weaknesses of each algorithm and provide examples of appropriate use cases for each. **(8 Marks)**
- Q8:** Evaluate the different clustering techniques, including K-means, hierarchical clustering and DBSCAN. Explain the underlying principles of each technique, and discuss their advantages, limitations, and practical applications. **(8 Marks)**

- Q9:** Examine the role of association rule mining in data mining. Describe the Apriori algorithm and its variations. Discuss the challenges associated with association rule mining, such as the generation of large numbers of rules and the need for efficient computation. **(8 Marks)**
- Q10:** Analyze the role of feature selection and dimensionality reduction in data mining. Discuss techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and feature selection algorithms. Explain how these techniques help in improving model performance and reducing computational complexity. **(8 Marks)**