Chhattisgarh Swami Vivekanand Technical University

University Teaching Department

Class Test-II (*January-June 2025*)

B. Tech(H)-8th Semester

Branch: Data Science

Subject Name: Data Warehousing **Subject Code:**

Max Marks: 40 **Min Marks:14** **Times: 2 hrs**

Note: Part A is compulsory.

Attempt any two questions from Part B, Part C, and Part D.

ROLL No. 30001282 1042

CO 1: To provide a comprehensive understanding of the concepts and architecture of data warehousing systems.

CO 2: To explore the process of data extraction, transformation, and loading (ETL) and its application in building data warehouses.

CO 3: To understand the design and implementation of OLAP (Online Analytical Processing) systems for business intelligence.

CO 4: To study various data modeling techniques such as star schema, snowflake schema, and fact constellation for effective data warehousing.

CO 5: To familiarize students with tools and technologies used in data warehousing and business intelligence, focusing on real-world applications.

Q.No.		Questions	Marks	BL	CO
		UNIT 4			
1	A	Differentiate between MOLAP and ROLAP.	4	L2	4
	B	Differentiate between drill-down and roll-up with scenarios with suitable examples.	8	L3	4
	C	Evaluate MOLAP, ROLAP, and HOLAP based on performance and scalability.	8	L5	4
	D	Evaluate the use of pivoting for better decision-making.	8	L5	4
		UNIT 5			
2	A	Mention two issues in data warehouse maintenance.	4	L1	5
	B	Analyze the challenges in maintaining a large data warehouse.	8	L4	5
	C	Demonstrate how data governance improves data quality.	8	L3	5
	D	Critically evaluate traditional vs modern warehousing techniques.	8	L5	5

Chhattisgarh Swami Vivekanand Technical University
University Teaching Department
Class Test-II (*Jan-June 2025*)
B. Tech(H)-8th Semester
Branch: Artificial Intelligence, Data Science

Subject Name: Cloud Computing **Subject Code: D127872(022)**
Max Marks: 40 **Min Marks:14** **Times: 2 hrs**

Note: All questions are compulsory

ROLL NO - 300012821042

CO4: To investigate security, privacy, and compliance issues in cloud environments.
CO5: To gain practical experience in deploying and managing cloud-based applications using industry-relevant platforms.

Q.No.	Questions	Marks	BL	CO
	UNIT 4			
1	Illustrate roles of threats and vulnerabilities in Cloud Computing	8	2	CO4
2	Analyze regulatory compliance in cloud computing	8	4	CO4
	UNIT 5			
3	Compare the serverless computing offerings of AWS Lambda, Google Cloud Functions, and Azure Functions	8	5	CO5
4	Explain the process of containerizing a monolithic application using Docker and deploying it on Kubernetes	8	2	CO5
5	Design a cloud architecture for a real-time machine learning model deployed on Microsoft Azure. Incorporate serverless components, data pipelines, and monitoring tools	8	6	CO5

Chhattisgarh Swami Vivekanand Technical University
University Teaching Department
Class Test-2 (*January-June 2025*)
B.Tech(H)-8th Semester
Branch: Artificial Intelligence/Data Science

Subject Name: Artificial Intelligence (AI) for Investments **Subject Code:**
Max Marks: 40 **Min Marks:14** **Times: 2 hrs**
Note: All questions are compulsory

CO1: Understand and recall key concepts in portfolio theory and asset pricing.
CO2: Analyze and apply principles of portfolio optimization and risk management.
CO3: Evaluate portfolio performance using theoretical and numerical techniques.
CO4: Understand and apply technical analysis and algorithmic trading strategies.
CO5: Implement machine learning and time-series models in financial analysis and risk-management.

Q.No.		Questions	Marks	BL	CO
		UNIT 1			
1	a	**The efficient frontier in portfolio theory is best described as:** A. The set of portfolios offering the highest expected return for a given level of risk B. A collection of portfolios with the lowest risk irrespective of return C. The set of portfolios that maximize risk exposure D. A collection of portfolios combining risk-free and risky assets only **Within the Capital Asset Pricing Model (CAPM), the beta of a stock measures:** A. The risk-free rate of return B. The asset's sensitivity to overall market movements C. The diversification benefit of the portfolio D. The stock's potential for pure growth	4	1	CO1
	b	Discuss the following: A. Define the Portfolio Possibility Curve (PPC) and explain how it relates to the Efficient Frontier. B. Describe how the inclusion of a risk-free asset (through lending and borrowing) broadens portfolio choices, leading to the Capital Market Line (CML).	8	4	CO2
	c	Consider two assets with the following attributes: Expected Returns: • Asset A: 8% • Asset B: 12% Standard Deviations: • Asset A: 10% • Asset B: 15% Correlation Coefficient between A and B: 0.3 Tasks: a) Calculate the covariance between Asset A and Asset B. b) Using the formula for a two-asset minimum variance portfolio, determine the optimal weights for Asset A and Asset B (assume no risk-free asset).	8	3	CO2
	d	A portfolio is evaluated using the CAPM framework. Assume that: • The risk-free rate is 3% • The expected market return is 10% • The portfolio's beta is 1.2 **Tasks:** a) Calculate the portfolio's expected return using the CAPM equation. b) For performance trend analysis, a manager computes a 5-day moving average. Given these closing prices: • Day 1: $100, Day 2: $102, Day 3: $101, Day 4: $103, Day 5: $105	8	3	CO3

		Calculate the 5-day moving average and briefly discuss its relevance in evaluating portfolio trends.			
		UNIT 2			
2	a	**Which technical analysis tool is primarily used to smooth out price data to identify trends?** A. Momentum Indicator B. Moving Average C. Bollinger Bands D. Exponential Smoothing **In the context of machine learning for trading, a regression algorithm is best used to:** A. Classify securities into buy or sell categories B. Forecast numerical values such as prices or returns C. Cluster stocks based on historical volatility D. Determine optimal portfolio weights	4	1	CO4
	b	Explain the role of regression and classification algorithms in security analysis and trading execution. In your answer, address: • How these models facilitate forecasting and decision-making in portfolio management. • The benefits and challenges of implementing machine learning techniques in active trading strategies. • Provide a brief case study (either real or hypothetical) that illustrates the impact of these algorithms on market prediction.	8	4	CO4
	c	Assume you are modelling asset returns using an ARIMA(1,0,1) model with the following estimated parameters: • Autoregressive coefficient AR(1): 0.5 • Moving Average coefficient MA(1): -0.3 • Constant term: 0.02 Given: • The previous period's return is 4% (0.04) • The previous error term is –0.01 **Task:** Calculate the expected return for the next period using the ARIMA(1,0,1) model.	8	3	CO5
	d	A risk management team uses a GARCH(1,1) model to estimate volatility for Value-at-Risk (VaR) calculations. The model parameters are: • ω (constant): 0.00001 • α (coefficient for lagged squared residual): 0.05 • β (coefficient for lagged conditional variance): 0.9 For a given day, assume: • Yesterday's squared error term = 0.0004 • Yesterday's conditional variance = 0.0005 **Tasks:** a) Calculate today's conditional variance using the GARCH(1,1) formula. b) Assuming normally distributed returns, briefly outline how you would compute the 95% one-day VaR for a portfolio valued at $1,000,000 using the conditional variance obtained.	8	3	CO5

Chhattisgarh Swami Vivekanand Technical University
University Teaching Department
Class Test-2 (*January-June 2025*)
B.Tech(H)-8th Semester
Branch: Artificial Intelligence/ Data Science

Subject Name: High Performance Scientific Computing **Subject Code:**
Max Marks: 40 **Min Marks:14** **Times: 2 hrs**
Note: All questions are compulsory

CO1: Master the fundamental concepts of parallel programming.
CO2: Analyze and optimize performance aspects in parallel computing environments.
CO3: Implement advanced computational tools and visualization techniques by effectively applying scientific toolkits.

Roll No - 3000128210 42

Q.No.		Questions	Marks	BL	CO
		UNIT 1			
1	a	Which MPI routine is used for point-to-point communication? A. MPI_Send B. MPI_Bcast C. MPI_Reduce D. MPI_Scatter Which MPI function supports non-blocking communication operations? A. MPI_Send B. MPI_Isend C. MPI_Recv D. MPI_Barrier Which of the following is a key feature of GPGPU architecture? A. Serial execution B. Highly parallel processing C. Low memory bandwidth D. Synchronous execution Which tool is commonly used to detect race conditions in parallel programs? A. Valgrind B. GDB C. TotalView D. Perf	4	1	CO1
	b	Discuss the following: **(a)** Describe how MPI facilitates communication in a distributed computing environment and explain one advanced feature that distinguishes it from basic message passing approaches. **(b)** Explain how GPGPU or vector programming can be employed to accelerate computation on a node and what unique challenges arise when integrating GPU operations with MPI.	8	4	CO2
	c	An MPI-based application is executed on a cluster of 16 nodes. The following data was collected: (i) Total serial execution time: 320 seconds. (ii) Total parallel execution time: 40 seconds. (iii) The measured MPI communication overhead is estimated to be 20 seconds for the entire run. Calculate the theoretical speedup achieved by parallelizing the application (ignoring communication overhead). Also, compute the effective speedup when communication overhead is considered.	8	3	CO2

	d	A high-performance application uses a GPGPU to accelerate a vectorized computation. During testing, the following performance data is collected: (i) The theoretical peak performance of the GPU kernel: 500 GFLOPS. (ii) Actual measured performance: 350 GFLOPS. (iii) A parallel debugger log indicates that 8% of the execution time is lost due to synchronization overhead between CPU and GPU. Compute the percentage of peak performance achieved by the application. Also, calculate the adjusted effective performance after accounting for the synchronization overhead.	8	4	CO2
		UNIT 2			
2	a	In parallel computing, what does 'speedup' refer to? A. The ratio of serial execution time to parallel execution time B. The increase in clock speed of processors C. The improvement in memory bandwidth D. The reduction in energy consumption Efficiency in parallel computing is defined as: A. Speedup multiplied by the number of processors B. Speedup divided by the number of processors C. The inverse of speedup D. The difference between serial and parallel run times Which toolkit is best known for providing routines for solving linear algebra problems such as dense matrix operations? A. PETSc B. LAPACK C. FFTW D. OpenMP Which of the following is a primary benefit of advanced scientific visualization in parallel computing? A. Reduced memory usage B. Enhanced interpretation of complex simulation data C. Increased clock speed D. Improved I/O performance	4	1	CO2
	b	Discuss the following: **(a)** Explain the significance of speedup and efficiency as performance metrics in parallel programming. **(b)** Describe the roles of scientific toolkits such as BLAS, LAPACK, and PETSc in optimizing numerical computations.	8	4	CO3
	c	An application is run on a multiprocessor system under the following conditions: (i) Measured serial execution time: 480 seconds. (ii) Measured parallel execution time on 24 processors: 30 seconds. (iii) Additional overhead due to inter-processor communication: 6 seconds. Calculate the theoretical speedup (ignoring overhead). Also, compute the actual speedup when the communication overhead is included.	8	3	CO2
	d	A scientific application uses LAPACK routines to solve a system of linear equations on a high-performance cluster. The following data is obtained from a performance test: (i) For a given problem size, the LAPACK routine takes 50 seconds on 1 processor. (ii) When run on 10 processors using a parallelized environment via PETSc, the runtime is recorded at 8 seconds. (iii) Advanced scientific visualization of performance data shows that the efficiency decreases non-linearly beyond 8 processors. Calculate the theoretical speedup when using 10 processors. Also, compute the actual parallel efficiency.	8	4	CO3