23SSP524/22SSP531

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc (SS) DEGREE EXAMINATION DECEMBER 2025
(Fifth Semester)

Branch – **SOFTWARE SYSTEMS (Five Years Integrated)**

DATA MINING

Time: Three Hours Maximum: 75 Marks

SECTION-A (10 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** marks (10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	A data warehouse in data mining context is: a) Transactional database b) Integrated repository for analysis c) Real-time processing system d) Single-user storage	K1	CO1
	2	Major issues in data mining affect knowledge discovery through: a) Simplifying processes b) Including scalability and privacy concerns c) Eliminating preprocessing d) Focusing only on visualization	K2	CO1
2	3	The purpose of data warehouse implementation is to: a) Delete data b) Build efficient storage for OLAP c) Limit access d) Ignore metadata	K1	CO2
	4	Data warehouse architecture differs from traditional databases by: a) Using the same structure b) Supporting historical and summarized data c) Being for analysis only in traditional databases d) Having no differences	K2	CO2
3	5	One basic concept in mining frequent patterns is: a) Support threshold b) Data deletion c) Single instance d) Noise addition	K1	CO3
	6	The role of boosting in classification techniques is: a) Weakening strong learners b) Sequentially improving weak learners c) Randomizing data d) Clustering data	K2	CO3
4	7	A source of Multivariate Time Series (MVTs) data is: a) Static images b) Sensor networks c) Text files only d) Single variable logs	K1	CO4
	8	The importance of constrained sequences in mining lies in: a) Ignoring patterns b) Incorporating domain knowledge for better results c) Increasing noise d) Limiting to batch data	K2	CO4
5	9	Graph mining involves: a) Extracting patterns from graph structures b) Text extraction only c) Spatial ignoring d) Web exclusion	K1	CO5
	10	A trend in data mining applications includes: a) Decreasing use of machine learning b) Increasing focus on big data integration c) Limiting to small datasets d) Avoiding predictions	K2	CO5

Cont...

SECTION - B (35 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks (5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Outline the types of data in data preprocessing.	K2	CO1
		(OR)		
	11.b.	Explain the data smoothing techniques in preprocessing.		
2	12.a.	Apply the multidimensional data model in OLAP technology.	K3	CO2
		(OR)		
	12.b.	Identify & explain the components of data warehouse architecture.		
3	13.a.	Identify the correlation measures in association mining with examples.	K3	CO3
		(OR)		
	13.b.	Apply density-based clustering for a dataset.		
4	14.a.	Examine the sources of MVTs data in sequence mining.	K4	CO4
		(OR)		
	14.b.	Analyze approaches for mining MVTs data.		
5	15.a.	Compare spatial data mining with traditional mining.	K4	CO5
		(OR)		
	15.b.	Examine the categories of web mining with examples.		

SECTION -C (30 Marks)

Answer **ANY THREE** questions

ALL questions carry **EQUAL** Marks (3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	Examine the role of data cleaning and its impact on mining accuracy.	K4	CO1
2	17	Compare and contrast OLAP operations in data warehouse models.	K4	CO2
3	18	Inspect the use of boosting versus stacking in prediction.	K4	CO3
4	19	Analyze the constraints and their effects in sequence mining.	K4	CO4
5	20	Analyze the applications of text mining in trend analysis.	K4	CO5

Z-Z-Z END