Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15

Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
1	I	1	What is KDD?
2	I	2	What is Databases?
3	I	3	What are steps in DataMimng Process?
4	I	4	What is difference between data and information?
5	I	5	What is difference between information and Knowledge?
6	I	6	Give some Applications of DataMining.
7	I	7	Explain classification of data mining.
8	I	8	What are Data Mining Techniques?
9	I	9	How can we identify Outlier detection ?
10	I	10	What is Clustering?
11	I	11	Define Regression.
12	I	12	Expain with example role of Prediction in Datamining.
13	I	13	What are benefits of DataMining?
14	I	14	What are impacts of Datamining in Business.
15	I	15	What is Datamining?
16	I	16	Explain Data Mining Functionalities.
17	I	17	What do you maen by Characterization of Data?
18	I	18	What is Data Discrimination?
19	I	19	What is Clustering Analysis?
20	I	20	What is Outlier Analysis?
21	I	21	What is Steps In Data Mining Process?
22	I	22	What is Data Cleaning?
23	I	23	What is Data integration?
24	I	24	Define Data selection.
25	I	25	What is Data transformation?
26	I	26	What is Pattern evaluation ?
27	I	27	What are Knowledge presentation & give name of any two techinques used in Knowledge Representation.
28	I	28	What are sourses of Data?
29	I	29	What is Data Mining Engine?
30	I	30	What are Major Issues in Datamining?
31	I	31	In classification class labels are ?
32	I	32	In clustering labels are?
33	I	33	To reduce the data set size which tehnique is used?
34	I	34	Which of the following does not affect quality of data ?
35	I	35	Who founded the term KDD?
36	I	36	Aprior algorithm is invented by?
37	I	37	What are the types of data in data mining?
38	I	38	What are the limitations of data mining?
39	I	39	How data mining is different from DBMS?

Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15

Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
40	I	40	What is the other name for data mining?
41	I	41	What is Data Mining Tool?
42	I	42	What are the five major types of data mining tools?
43	I	43	Why do we need a Data Warehouse?
44	I	44	What are the differences between structured and unstructured data?
45	I	45	Give any example where unstructred data is used?
46	I	46	Give any example where structred data is used?
47	I	47	What do you understand by Data Mining?
48	I	48	State some of the advantages of data mining.
49	I	49	What according to you are the disadvantages of data mining?
50	I	50	Explain the techniques of Data Mining?
51	II	1	Why we need to Preprocess the Data?
52	II	2	What are Missing Values in data?
53	II	3	What are the common methods to handle missing values?
54	II	4	What is Noisy Data?
55	II	5	What are the Data smoothing techniques?
56	II	6	Explain Data binning with example.
57	II	7	Explain data Smoothing by bin means with example.
58	II	8	Explain data Smoothing by bin boundaries with example.
59	II	9	What is Data Integration Issues with data integration?
60	II	10	What do you mean by Data Reduction?
61	II	11	What are Data Reduction Strategies?
62	II	12	What is Data Transformation?
63	II	13	What are Data Transformation Strategies?
64	II	14	Define Discretization.
65	II	15	What is Generalization?
66	II	16	What is Aggregation?
67	II	17	What is Market Basket Analysis?
68	II	18	What are purpose of Association Rules?
69	II	19	Define Support and Confidence with example.
70	II	20	How to calculate Support?
71	II	21	What is Frequent Set?
72	II	22	Expalin in short Apriori Algorithm?
73	II	23	Where Apriori is used?
74	II	24	How does classification work?
75	II	25	What is candidate item set?
76	II	26	What is minimum support threshold?
77	II	27	What is formula to generate association rule?

Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15

Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
78	II	28	Give any example of market basket analysis.
79	II	29	What is the difference between agglomerative and divisive hierarchical clustering?
80	II	30	What is the purpose of cluster analysis
81	II	31	What is the difference between data cleaning and data transformation?
82	II	32	Name Data mining techniques?
83	II	33	What is the Decision Tree Algorithm?
84	II	34	List down the different types of nodes in Decision Trees.
85	II	35	List down the advantages of the Decision Trees.
86	II	36	What is data Overfitting?
87	II	37	What is data Underfitting?
88	II	38	What are decision trees commonly used for?
89	II	39	Are parent node and root node the both same in the decision tree?
90	II	40	Explain The Bayes Theorem
91	II	41	What is posterior probability?
92	II	42	What is prior probability?
93	II	43	What is formula for calculating posterior probability?
94	II	44	What is decision class?
95	II	45	Define Entropy.
96	II	46	What is formula for information gain?
97	II	47	What are major Issues in Cluster Analysis?
98	II	48	What are features of Cluster Analysis?
99	II	49	Explain K-Means algorithm.
100	II	50	What is cluster seed?
101	III	1	What is decision tree?
102	III	2	How are decision trees used for classification?
103	III	3	What is Information Gain?
104	III	4	What is Gini Index?
105	III	5	What is Bayes' Theorem?
106	III	6	What is Requirements for Cluster Analysis?
107	III	7	What are Basic Clustering Methods?
108	III	8	What is partitioning method?
109	III	9	What is Hierarchical methods?
110	III	10	What is Density-based methods?
111	III	11	What is Grid-based methods?
112	III	12	What are Applications of clustering?
113	III	13	What is K means clustering?
114	III	14	what is difference between database and datamart?

Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15

Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
115	III	15	what is difference between datamart and datawarehouse?
116	III	16	How classification relates with supervised learning?
117	III	17	How clustering relates with unsupervised learning?
118	III	18	What is the Naive Bayes Algorithm in Data Mining?
119	III	19	Name methods of clustering?
120	III	20	What Is Data Purging?
121	III	21	What Are Cubes in datamining?
122	III	22	What is Prediction?
123	III	23	What do you understand by predictive analytics?
124	III	24	How does predictive analysis work?
125	III	25	What is the purpose of predictive analytics?
126	III	26	What are the benefits of predictive analytics?
127	III	27	How do the predictive analytics being used by the companies/organizations?
128	III	28	Why is predictive analytics important?
129	III	29	What is 'Training set' and 'Test set'?
130	III	30	What is 'Training set' ?
131	III	31	What is 'Test set' data?
132	III	32	What is cluster?
133	III	33	What is outliers in data?
134	III	34	What is missing values in data?
135	III	35	Define Dimensionality Reduction.
136	III	36	What is Numerosity reduction.
137	III	37	List any four features of Cluster Analysis.
138	III	38	What is posterior probability?
139	III	39	Removing duplicate records is a process called
140	III	40	Removing duplicate records is a process called?
141	III	41	Which algorithm is most sensitive to outliers?
142	III	42	Which method is used to solve the problem of overfitting?
143	III	43	What is ETL? Name some of the best ETL tools.
144	III	44	What does ETL stand for?
145	III	45	Differences Between Star And Snowflake Schemas?
146	III	46	What Is The Use Of Regression?
147	III	47	Define Wave Cluster?
148	III	48	Define Density Based Method?
149	III	49	What Is A Dbscan?
150	III	50	How many tiers are there in data warehouse architecture?
151	IV	1	What is DataWarehouse?

Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15
Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
152	**IV**	2	What are features of Datawarehouse?
153	**IV**	3	What is OLAP?
154	**IV**	4	What is OLTP?
155	**IV**	5	Difference between OLTP and OLAP?
156	**IV**	6	What is Datacube?
157	**IV**	7	Which are tools used for OLAP & OLTP.
158	**IV**	8	Where OLAP is used?
159	**IV**	9	Where OLTP is used?
160	**IV**	10	What is fact table?
161	**IV**	11	What is star schema?
162	**IV**	12	Explain star schema with example.
163	**IV**	13	What is snowflakes schema?
164	**IV**	14	Give any example of snowflakes schema.
165	**IV**	15	What is Fact constellation?
166	**IV**	16	What is Galaxy schema?
167	**IV**	17	Fact constellation schema is also called?
168	**IV**	18	What are steps in data warehouses?
169	**IV**	19	What is ETL?
170	**IV**	20	What is Multidimensional Data Model?
171	**IV**	21	Dfine Relational OLAP (ROLAP) servers?
172	**IV**	22	What is Multidimensional OLAP (MOLAP) ?
173	**IV**	23	Where ROLAP sever is used?
174	**IV**	24	Where MOLAP sever is used?
175	**IV**	25	What is Hybrid OLAP (HOLAP) servers?
176	**IV**	26	What are the tools available for ETL?
177	**IV**	27	What is the relation between data warehousing and data mining?
178	**IV**	28	What are the 4 key components of a data warehouse?
179	**IV**	29	What Is a Subject-Oriented Data Warehouse?
180	**IV**	30	What Is a Real-Time Data Warehouse?
181	**IV**	31	In Data Warehousing, What Is the Concept of a Cube?
182	**IV**	32	Is it Possible to Include a Numeric Value in a Dimension Table?
183	**IV**	33	What are the three major areas in the data warehouse?
184	**IV**	34	What is the definition of Cube in Datawarehousing?
185	**IV**	35	What is the difference between Datawarehouse and OLAP?
186	**IV**	36	Out of star schema and snowflake schema, whose dimension table is normalized?
187	**IV**	37	How many fact tables are there in a star schema?

Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15

Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
188	**IV**	38	Which one is faster, Multidimensional OLAP or Relational OLAP?
189	**IV**	39	List the types of OLAP server
190	**IV**	40	Explain data mart.
191	**IV**	41	Define dimension?
192	**IV**	42	How does a Data Cube help?
193	**IV**	43	What does Metadata Respiratory contain?
194	**IV**	44	Define metadata?
195	**IV**	45	What do you mean by Data Extraction?
196	**IV**	46	List the process that are involved in Data Warehousing.
197	**IV**	47	What do OLAP and OLTP stand for?
198	**IV**	48	What is the very basic difference between data warehouse and operational databases?
199	**IV**	49	List the Schema that a data warehouse system can implements.
200	**IV**	50	List any five applications of data warehouse.
201	**V**	1	Explain the scope of data mining?
202	**V**	2	List out the types of data mining?
203	**V**	3	Can you please tell, which problems, in general, the data mining can solve?
204	**V**	4	Explain the major elements of Data Mining?
205	**V**	5	How is data warehousing related to business intelligence?
206	**V**	6	Give a brief introduction to data mining process?
207	**V**	7	What is text mining?
208	**V**	8	Define Spital Mining.
209	**V**	9	What is WWW mining?
210	**V**	10	What do you mean by DB Miner?
211	**V**	11	What is Semi-structured data?
212	**V**	12	What is difference between strucutred and unstructured data?
213	**V**	13	What is difference between unstrucutred and semi-structured data?
214	**V**	14	What are Text mining techniques?
215	**V**	15	Define Information retrieval.
216	**V**	16	What is NLP?
217	**V**	17	What is Information extraction?
218	**V**	18	Define Tokenization.
219	**V**	19	Explain Summarization Technique.
220	**V**	20	What is Sentiment analysis?
221	**V**	21	What is Feature selection?
222	**V**	22	What is Feature extraction?

Medicaps University
Oral Examination December - 2021

Paper Code:CA5CO15

Paper Name:Data Warehousing and Mining

Sl No	Unit	Question No	Question
223	V	23	What is Spatial data?
224	V	24	What is Spatial continuity?
225	V	25	Give any example of Spatial Data Mining.
226	V	26	What are the steps to detect spatial patterns?
227	V	27	What are Spatial objects?
228	V	28	What are Social Impacts of Data Mining?
229	V	29	List any two applications of text mining.
230	V	30	List any two applications of web mining.
231	V	31	List any two applications of spital mining.
232	V	32	What are some real-life NLP applications?
233	V	33	What is Discrete and Continuous data in Data Mining?
234	V	34	Which are the most popular areas of applications of Data Mining?
235	V	35	What is time series data?
236	V	36	What do you understand by data aggregation and data generalization?
237	V	37	What are the most significant advantages of Data Mining?
238	V	38	What are the most significant disadvantages of Data Mining?
239	V	39	Which are the main prominent fields and areas where Data Mining is used?
240	V	40	Give a brief introduction to data mining process?
241	V	41	Name the steps used in data mining?
242	V	42	Explain steps involved in data mining knowledge process?
243	V	43	What are issues in data mining?
244	V	44	What is Data visualization?
245	V	45	Why Data visualization is important?
246	V	46	What Are The Different Problems That "data Mining" Can Solve?
247	V	47	How Does The Data Mining And Data Warehousing Work Together?
248	V	48	What Is Smoothing?
249	V	49	What Are The Advantages Data Mining Over Traditional Approaches?
250	V	50	What Are The Foundations Of Data Mining?