

Question Bank

Specialization : Business Analytics

Course Code : 206

Course Name – Data Mining

Sr.no	Question Type	Question	Marks
1	REMEMBERING	Define data types Big data management and data mining	2
2	REMEMBERING	Define business intelligence	2
3	REMEMBERING	Define Hadoop clusters	2
4	REMEMBERING	What is Open source	2
5	REMEMBERING	Define Big Data.	2
6	REMEMBERING	Discuss the needs of Big data.	2
7	REMEMBERING	Define Data Mining.	2
8	REMEMBERING	Describe Business Intelligence.	2
9	REMEMBERING	Define Data Mining Process.	2
10	REMEMBERING	Elaborate relation to Business Intelligence techniques.	2
11	REMEMBERING	List any two Data Mining Tasks	2
12	REMEMBERING	Define Classification in data mining	2
13	REMEMBERING	Describe Clustering in data mining.	2
14	REMEMBERING	Define Association Analysis	2
15	REMEMBERING	List any two Data Mining Tasks	2
1	UNDERSTANDING	Discuss the needs of Big data.	5
2	UNDERSTANDING	Explain the implementation of Data Mining.	5

3	UNDERSTANDING	What is Business Intelligence? how it can be useful in sells and distribution .	5
4	UNDERSTANDING	Explain Data Mining Process in details .	5
5	UNDERSTANDING	Elaborate relation to Business Intelligence techniques.	5
6	UNDERSTANDING	Explain Data Mining Tasks in detail.	5
7	UNDERSTANDING	Discuss the needs of Big data in Patient history management .	5
8	UNDERSTANDING	Explain the implementation of Data Mining in HR department .	5
9	UNDERSTANDING	Explain Clustering in data mining.	5
10	UNDERSTANDING	Define Association Analysis with suitable Example.	5
11	UNDERSTANDING	Explain Anomaly Detection.	5
12	UNDERSTANDING	What is model? Explain different type of Model	5
13	UNDERSTANDING	Explain descriptive model with suitable example.	5
1	APPLY	A shoes company wants to extract the leads of their customer who online visit the product , brows the information , and check out the process but not purchase the product , a company wants to find out such cutomer who vist but not converted illustrate the use of Business Intelligence.	10
2	APPLY	A Red Chief company wants to computerised all their activities , Manoj Gyanchandani started a leather shoe export business in his early 20s when he was also involved in his family's business. In 1995, he set up Leayan Global Private Limited to export leather shoes to Europe. However, two years later, Manoj realised that the Indian footwear market isn't organised in the leather shoe sector. Wrapping up the export business, he studied the market and launched Red Chief brand under the parent company Leayan Global Private Limited in 1997. Initially, Manoj decided to set a firm foot in Kanpur. He started by showcasing products in multi-branded outlets across the city. This continued till 2010 when Red Chief expanded to multi-branded outlets in other states. In 2011, the entrepreneur launched the first exclusive Red Chief outlet in Kanpur. Today, Red Chief has 175 stores located across 16 states	10

		including Uttar Pradesh, Madhya Pradesh, Gujarat, Rajasthan, Maharashtra, and Chhattisgarh. It is also present in over 3,000 multi-brand outlets. According to the Registrar of Companies (RoC) filings, the company makes an annual turnover of over Rs 324 crore suggest data mining tool.	
3	APPLY	<p>Indian Railways (IR) is India's national railway system operated by the Ministry of Railways. It is one of the public facilities given by the government and manages the fourth largest railway network in the world by size, with a route length of 67,368-kilometre (41,861 mi) as of March 2017. About 57.91% of the routes are electrified with 25 kV 50 Hz AC electric traction while 33% of them are double or multi-tracked.</p> <p>In the fiscal year ending March 2018, IR carried 8.26 billion passengers and transported 1.16 billion tonnes of freight. In the fiscal year 2017–18, IR is projected to have revenue of ₹1.874 trillion (US\$26 billion), consisting of ₹1.175 trillion (US\$16 billion) in freight revenue and ₹501.25 billion (US\$7.0 billion) in passenger revenue, with an operating ratio of 96.0 percent. By using data mining prepare various report of trains, passengers</p>	10
4	APPLY	Explain Anomaly detection, how anomaly detection is implemented in bank to identify the fraud in bank	10
5	APPLY	Discuss the needs of Big data in health care management system.	10
6	APPLY	Define Data Mining . how it is used in education sector ?	10
7	APPLY	Describe Business Intelligence and it's implementation phases .	10
8	APPLY	Define Data Mining Process in detail.	10
9	APPLY	Elaborate relation to Business Intelligence techniques.	10
1	ANALYSE	Explain market basket analysis technique which is used in supermarket	10
2	ANALYSE	Explain Classification in data mining , an e-mail program might attempt to classify an e-mail as "legitimate" or as "spam". Justify the statement	10
1	EVALUATE	Defend Clustering – is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data.	10

1	CREATE	A supermarket gather data on customer purchasing habits. Using association rule learning, the supermarket wanted to determine which products are frequently bought together.	10
---	--------	--	----

Unit -2

Sr.no	Question Type	Question	Marks
1	REMEMBERING	Define Continuous attribute	2
2	REMEMBERING	Define Ordinal attribute	
3	REMEMBERING	Define Numeric attribute	2
4	REMEMBERING	Define Nominal attribute	2
5	REMEMBERING	Define smooth by bin boundaries	2
6	REMEMBERING	Define model and it's features.	2
7	REMEMBERING	Describe different type of Model	2
8	REMEMBERING	Define descriptive models	2
9	REMEMBERING	Explain descriptive model with suitable example	2
10	REMEMBERING	Define predictive mode	2
11	REMEMBERING	List any two Real-world data mining applications	2
12	REMEMBERING	Define Big Data Analytics in Mobile Environments.	2
13	REMEMBERING	Explain Fraud Detection and Prevention	2

14	REMEMBERING	Describe Data Mining Techniques.	2
15	REMEMBERING	Big Data Analytics in Business Environments.	2
1	UNDERSTANDING	Illustrate any two Real-world data mining applications	5
2	UNDERSTANDING	explain Big Data Analytics in Mobile Environments.	5
3	UNDERSTANDING	Explain Fraud Detection and Prevention with Data Mining Techniques.	5
4	UNDERSTANDING	Discuss Big Data Analytics in Business Environments.	5
5	UNDERSTANDING	Explain different types of data	5
6	UNDERSTANDING	What is Data Normalization.	5
7	UNDERSTANDING	Explain various stages of Normalization.	5
1	APPLY	Explain Association Analysis how it is implemented in real world scenario.	10
2	APPLY	Explain Anomaly Detection	10
3	APPLY	Describe different type of Model with suitable example.	10
4	APPLY	Define descriptive model with suitable example .	10
5	APPLY	Explain descriptive model with its implacability in commodity Prices in APMC Market.	10
6	APPLY	Define predictive mode it's advantages and disadvantages.	10
7	APPLY	Effective use of Data types is very important during questionnaire designing data collection,. Comment	10
1	ANALYSE	Explain various data types and there applicability	10
2	ANALYSE	Explain the feature of data quality, why it is important to focus on data quality comment	10
1	CREATE	What are the sampling technique ? explain it in detail,	10

2	CREATE	Big data may contain a huge number (from hundreds to millions) of features and often most of the features could be unimportant, irrelevant or redundant, which can cause poor efficiency and/or over-fitting in data analysis and machine learning. Therefore, it is necessary to employ some feature selection methods (FS) to remove irrelevant and redundant features to reduce the complexity of analysis and the generated models and also to improve the efficiency of the whole modelling process justify the statement	10
3	CREATE	The ALL approach uses all the instances in a given dataset in its feature selection step, while the PART approach only uses training instances partitioned from the dataset in feature selection. Elaborate both the approach with example.	10

Unit -3

Sr.no	Question Type	Question	Marks
1	REMEMBERING	Define SMTOP	2
2	REMEMBERING	Define OLTP	
3	REMEMBERING	Define FTP	2
4	REMEMBERING	Define OLAP	2
5	REMEMBERING	Define snow flake schema	2
6	REMEMBERING	Define Data and its attribute .	2
7	REMEMBERING	Explain different types of attributes.	2
8	REMEMBERING	Define properties of attribute values.	2

9	REMEMBERING	List different types of data	2
10	REMEMBERING	Define data quality.	2
11	REMEMBERING	Define Sampling.	2
12	REMEMBERING	What is Data Normalization?	2
13	REMEMBERING	Explain various stages of Normalization.	2
14	REMEMBERING	Explain the characteristics of Normalization.	2
15	REMEMBERING	Define Data Cleaning	2
1	UNDERSTANDING	Explain the characteristics of Normalization	5
2	UNDERSTANDING	What do you mean by Similarity Measures.	5
3	UNDERSTANDING	What is Feature Selection/Instance Selection?	5
4	UNDERSTANDING	What is Decision-Tree Based Approach?	5
5	UNDERSTANDING	What are the different type of approaches?	5
6	UNDERSTANDING	What is Support Vector Machines.	5
7	UNDERSTANDING	Discuss any two implementation of SVM	5
8	UNDERSTANDING	Explain the characteristics of Normalization	5
1	APPLY	Rule-based Approach, Support Vector Machines, Ensemble Learning, Classification Model Selection and Evaluation, Applications: B2B customer buying stage prediction, Recommender Systems.	10
2	APPLY	Use Decision Tree to predict whether a person is fit or unfit, given their information like age, eating habits, physical activity, etc. The decision nodes are the questions like 'What's the age?', 'Does he exercise?', 'Does he eat a lot of pizzas'? And the leaves represent outcomes like either 'fit', or 'unfit'.	10
3	APPLY	Use regression tree to predict whether a company insured and not their driver marital status, gender and male or female ,age	10

		Assume suitable condition	
1	ANALYSE	Predictive prospecting in B2B space and how can we assess user intent using 1st party, 2nd party and 3rd party data use in dell inc.	10

Unit -4

Sr.no	Question Type	Question	Marks
1	REMEMBERING	Define clustering	2
2	REMEMBERING	Define K-Means clustering	
3	3. REMEMBERING	Define agglomerative clustering	2
4	REMEMBERING	Define conceptual clustering	2
5	REMEMBERING	Define Adaptive system management	2
6	REMEMBERING	What do you mean by Similarity Measures.	2
7	REMEMBERING	What is Feature Selection/Instance Selection.	2
8	REMEMBERING	Explain the importance of feature selection/instance selection in various big data scenarios.	2
9	REMEMBERING	Define Decision-Tree Based Approach.	2
10	REMEMBERING	What is Decision-Tree Based Approach?	2
11	REMEMBERING	Define Rule-based Approach	2
12	REMEMBERING	Define different type of approaches?	2
13	REMEMBERING	Define Instance-based classifiers	2

14	REMEMBERING	What is Support Vector Machines.	2
15	REMEMBERING	Discuss any two implementation of SVM	2
1	UNDERSTANDING	Classification Model Selection and Evaluation.	5
2	UNDERSTANDING	Explain B2B customer buying stage prediction	5
3	UNDERSTANDING	What is Recommender systems?	5
4	UNDERSTANDING	What is Clustering?	5
5	UNDERSTANDING	Explain various types of clustering	5
6	UNDERSTANDING	Differentiate between Partitional and Hierarchical Clustering Methods.	5
7	UNDERSTANDING	Explain Partitional Clustering Methods.	5
8	UNDERSTANDING	Explain Hierarchical Clustering Methods.	5
9	UNDERSTANDING	Explain Graph-based Methods.	5
10	UNDERSTANDING	Clustering: Partitional and Hierarchical Clustering Methods, Graph-based Methods, Density-based Methods, Cluster Validation, Applications: Customer Profiling, Market Segmentation.	5
1	APPLY	Illustrate any two Applications of Customer Profiling.	10
2	APPLY	What is Apriori Algorithm and its Extensions.	10
3	APPLY	B2B Customer Buying Path Analysis.	10
4	APPLY	Explain Medical Informatics.	10
5	APPLY	Any two uses of Telecommunication alarm diagnosis.	10
6	APPLY	What is Anomaly Detection? Consider an example of Super Store Sales data set that can be used to find patterns in Sales and Profit separately that do not conform to expected behavior.	10

		That is, spotting outliers for one variable at a time.	
--	--	--	--

Unit-5

Sr.no	Question Type	Question	Marks
1	REMEMBERING	Define machine-learning techniques	2
2	REMEMBERING	Define Algorithm	2
3	REMEMBERING	Define Biotope	2
4	REMEMBERING	Define Ensemble Learning.	2
5	REMEMBERING	Classification Model Selection and Evaluation.	2
6	REMEMBERING	Applications: B2B customer buying stage prediction	2
7	REMEMBERING	What is Recommender Systems.	2
8	REMEMBERING	What is Clustering:	2
9	REMEMBERING	Explain various types of clustering	2
10	REMEMBERING	Differentiate between Partitional and Hierarchical Clustering Methods.	2
11	REMEMBERING	Explain Partitional Clustering Methods.	2
12	REMEMBERING	Explain Hierarchical Clustering Methods.	2
13	REMEMBERING	Explain Graph-based Methods.	2

14	REMEMBERING	Define Cluster Validation.	2
1	UNDERSTANDING	Illustrate any two Applications of Customer Profiling.	5
2	UNDERSTANDING	Illustrate any two Applications of Customer Profiling.	5
3	UNDERSTANDING	What is Apriori Algorithm and its Extensions.	5
4	UNDERSTANDING	B2B Customer Buying Path Analysis.	5
5	UNDERSTANDING	Explain Medical Informatics.	5
6	UNDERSTANDING	Any two uses of Telecommunication alarm diagnosis.	5
7	UNDERSTANDING	What is Anomaly Detection?	5
8	UNDERSTANDING	Distinguish Statistical-based and Density-based Methods .	5
9	UNDERSTANDING	Explain Ethics of data mining	5
1	APPLY	Illustrate any two Applications of Customer Profiling.	10
2	APPLY	What is Apriori Algorithm and its Extension	10
3	APPLY	What is Medical Informatics. How it can be Implemented in today scenario Covid-19	10
4	APPLY	Differentiate between Statistical-based and Density-based Methods .	10
5	APPLY	What is Ethics of data mining , Advantages and disadvantages of ethics of data mining	10