

THE KAVERY ENGINEERING COLLEGE
(An Autonomous Institution)

M.E/M.C.A DEGREE END SEMESTER THEORY EXAMINATIONS, NOV /DEC 2024

Third Semester

Computer Science and Engineering

DS4015 Big Data Analytics

(Common to Master of Computer Application)

Regulations - 2021

Duration: 3 Hrs

Maximum: 100 Marks

PART – A(ANSWER ALL QUESTIONS) (Marks 10*2=20)

<i>Q.No</i>	<i>Questions</i>	<i>BLT</i>	<i>CO</i>	<i>Marks</i>
1.	List the main characteristics of Big Data	RE	1	2
2.	Justify the statement “Accuracy in big data is beneficial”	EV	1	2
3.	Examine the various dimensions of growth of search strategies.	AN	2	2
4.	Explain the role of visualization tools in big data.	UN	2	2
5.	Measure out the few challenges of data stream mining algorithms	UN	3	2
6.	Show the features of what is Real Time Analytics Platform	RE	3	2
7.	Label the components of Hadoop framework	RE	4	2
8.	Show how does Map Reduce computation execute.	RE	4	2
9.	Infer the difference with R data structures	UN	5	2
10.	Compare recursive and iterative functions in R	UN	5	2

PART – B (ANSWER ALL QUESTIONS) (Marks 5*13 = 65)

<i>Q.No</i>	<i>Questions</i>	<i>BLT</i>	<i>CO</i>	<i>Marks</i>
11.	a i How re-sampling methods applied to improve model accuracy and reliability. Give example.	RE	1	7
	ii Illustrate the statistical inference contribution to decision-making in data-driven industries.	UN	1	6
	Or			
	b i Analyze the impact of different types of prediction errors on model performance	AN	1	7
	ii Explain why conventional data processing systems struggle with big data. Identify specific challenges with example.	UN	1	6
12.	a i Describe the different interaction techniques that supports the analysis of large and complex datasets.	UN	2	7
	ii Compare and contrast genetic algorithms and simulated annealing for optimization.	UN	2	6

Or

	b	i	In what way visualization techniques enables effective data analysis without overwhelming users.	RE	2	7
		ii	Outline the importance of specific visual data analysis techniques with examples.	RE	2	6
13.	a	i	Describe Data Stream Model with a neat architecture diagram.	UN	3	7
		ii	Evaluate briefly how to count the distinct elements in a stream.	EV	3	6
			Or			
	b	i	Discuss the mining concept is used in real time sentiment analysis with an example.	UN	3	7
		ii	Generalize how data analysis is used in stock market predictions.	RE	3	6
14.	a	i	With suitable block diagram explain architecture of HDFS. Discuss role of Data node and Name node with examples.	RE	4	7
		ii	Give commands with appropriate arguments to perform data transfer between local file system and HDFS.	RE	4	6
			Or			
	b	i	Write Map Reduce steps for counting occurrences of specific numbers in the input text file(s).	UN	4	7
		ii	Illustrate the process of creating and managing the database and tables using Hive.	UN	4	6
15.	a	i	Describe data frames and their importance in data analysis. Specify its data operations.	UN	5	7
		ii	Provide examples of how string manipulation can be useful in data cleaning and analysis tasks.	AP	5	6
			Or			
	b	i	Discuss considerations around memory management and efficient data manipulation techniques in R.	UN	5	7
		ii	Discuss various methods for input and output in R. How to handle data input errors.	UN	5	6

*PART – C (Marks 1*15 = 15)*

<i>Q.No</i>		<i>Questions</i>	<i>BLT</i>	<i>CO</i>	<i>Marks</i>	
16.	a	i	Developing a Simple Calculator Using Control Statements and Operators	AP	5	7
		ii	Demonstrate how to extract all employee names and calculate the average salary across all employees.	AP	5	8
		Or				
	b	i	How big data analytics is transforming the medical device industry. Explain	RE	4	7
		ii	Discuss the key data sources and types of analyses that support device innovation and quality control.	UN	4	8