

Reg.No.:



VIVEKANANDHA COLLEGE OF ENGINEERING FOR WOMEN
[AUTONOMOUS INSTITUTION AFFILIATED TO ANNA UNIVERSITY, CHENNAI]
Elayampalayam – 637 205, Tiruchengode, Namakkal Dt., Tamil Nadu.

Question Paper Code: 60002

B.E. / B.Tech. DEGREE END-SEMESTER EXAMINATIONS – NOV. / DEC. 2024

Seventh Semester

Information Technology

U19IT724 – BIG DATA ANALYTICS

(Regulation 2019)

Time: Three Hours

Maximum: 100 Marks

Answer ALL the questions

Knowledge Levels	K1 – Remembering	K3 – Applying	K5 - Evaluating
(KL)	K2 – Understanding	K4 – Analyzing	K6 - Creating

PART – A

(10 x 2 = 20 Marks)

Q.No.	Questions	Marks	KL	CO
1.	Define Big data Analytics.	2	K1	CO1
2.	What are the key factors than an organization must consider before investing on Big data applications?	2	K1	CO1
3.	What is Hadoop?	2	K1	CO2
4.	Differentiate between SQL and NoSQL databases.	2	K2	CO2
5.	List out the data types in MongoDB.	2	K1	CO3
6.	What are the features of Cassandra ?	2	K1	CO3
7.	State the limitations of Map Reduce model.	2	K2	CO4
8.	What is the need for Apache spark and Apache kafka?	2	K2	CO4
9.	What is Pig in Hadoop?	2	K1	CO5
10.	What are the various execution modes of Pig?	2	K1	CO5

PART – B

(5 x 13 = 65 Marks)

Q.No.	Questions	Marks	KL	CO
11.	a) i. Explain the main characteristic features and structure of Big data in detail.	4	K1	CO1
	ii. Explain big data architecture in detail with a neat schematic diagram.	9	K1	CO1
	(OR)			
	b) Explain in detail about various analytics tools used in Bigdata.	13	K2	CO1
12.	a) With suitable examples, differentiate between the applications, structure, working and usage of different NoSQL databases.	13	K2	CO2
	(OR)			
	b) Explain in detail about managing resources and applications with Hadoop YARN	13	K2	CO2
13.	a) Explain in detail about MongoDB database and its CRUD operations with an example.	13	K2	CO3
	(OR)			
	b) How Cassandra is integrated with Hadoop and also explain in detail about its features and operations with an example.	13	K3	CO3
14.	a) i. What is the role of combiner and partitioner in a map reduce application?	8	K1	CO4
	ii. List the algorithms used in Map Reduce.	5	K1	CO4
	(OR)			
	b) i. Explain in detail about Hive data manipulation, queries, and data types.	7	K1	CO4
	ii. Explain the difference between hive and map reduce.	6	K1	CO4
15.	a) Explain Pig data model in detail with neat diagram.	13	K1	CO5
	(OR)			
	b) Evaluate a procedure to find the number of occurrences of a word in a document using pig.	13	K1	CO5

PART – C

(1 x 15 = 15 Marks)

Q.No.	Questions	Marks	KL	CO
16. a)	<p>Consider a collection of literature survey made by a researcher in the form of a text document with respect to cloud and big data analytics. Analyze Using Hadoop and Map Reduce. Write a program to count the occurrences of pre-dominant key words</p> <p>(OR)</p>	15	K4	CO3
b)	<p>Perform analysis on web server report.</p> <p>Sample Data: teleman.pr.mcs.net, -, , [01/Jul/2005:00:03:57,0400], "GET,/images/KSC- logosmall.gif, HTTP/1.0", 304, 0teleman.pr.mcs.net,-,- , [01/Jul/2005:00:03:57,0400], "GET, /images/KSClogosmall.gif, HTTP/1.0", 304, 0.</p> <p>The data is comma separated. It consists of the user IP address, time at which the request is received, time zone, request type, requested link, request details, response code and bytes transferred. Usually the scale of these datasets is quite huge and running queries in a conventional method is not possible. Hence use Pig programming on this dataset to retrieve the necessary statistics which helps us to understand the load and usage of the server, user visit frequency, webpage popularity and the total bytes transferred</p>	15	K4	CO5