



Sardar Patel Institute of Technology

Bhavan's Campus, Munshi Nagar, Andheri(West), Mumbai-400058, India
(Autonomous College Affiliated to University of Mumbai)

ESE

November 2018

Max. Marks: 60

Class: M.Tech(Sem I)

Course Code: CE912

Duration: 180 min.

Time: 10 am-1 pm

Name of the Course: Big Data Analytics and Management

Instructions:

- (1) All questions are compulsory
- (2) Draw neat diagrams
- (3) Assume suitable data if necessary

Q.No.		Max Marks	CO
Q.1 a	Explain different characteristics of Big Data with example? OR Explain the difference between NameNode, Backup Node and Checkpoint NameNode ?	06	CO1
Q.1 b	What is Rack Awareness? Explain with an example? List advantages of Rack Awareness?	06	CO1
Q.2 a	Write a MapReduce program to count occurrence of word "Environment"(case sensitive) in the large text whose sample is as below. Clearly write Map class and Reduce class : World Environment Day is celebrated in many ways. Street rallies, parades, street plays create awareness about world environment. World Environment Day is a day for us to inspect the state of our environment. It calls for us to stop in our tracks one day in the year and examine our surroundings. It asks us to pledge, in a small way at least, to do something for the environment. There is something each of us can do to preserve the environment. So let us all pledge to do something, at least one thing, before the next World Environment Day	06	CO2
Q.2 b	Explain Natural join operation of RDBMS by map reduce with an example? Write its pseudo code. ?	06	CO2
Q.3 a	What are the drawbacks of traditional clustering algorithm? How CURE clustering algorithm overcomes the problems of traditional clustering? OR What are the applications of Stream computing? Explain any two applications in detail with example?	06	CO3 CO3
Q3.b	What is text mining? Explain difference between text mining and text analytics? Explain various applications of text analytics?	06	CO3

Q.4 a	<p>What are different key problems in recommendation? Explain Collaborative filtering based recommendation?</p> <p>Suppose there are 6 users and 5 Movies with 5 star rating is given in a table below, calculate User profile for user 2 and user 5</p> <table><tr><th>Keywords</th><th>Harry Potter 1</th><th>Harry Potter 2</th><th>Harry Potter 3</th><th>Harry Potter 4</th><th>Harry Potter 5</th></tr><tr><td>Users</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>User 1</td><td>5</td><td>4</td><td>3</td><td>2</td><td>-</td></tr><tr><td>User 2</td><td>2</td><td>4</td><td>5</td><td>-</td><td>-</td></tr><tr><td>User 3</td><td>3</td><td>4</td><td>2</td><td>4</td><td>3</td></tr><tr><td>User 4</td><td>-</td><td>1</td><td>-</td><td>3</td><td>2</td></tr><tr><td>User 5</td><td>2</td><td>5</td><td>3</td><td>-</td><td>-</td></tr><tr><td>User 6</td><td>2</td><td>-</td><td>-</td><td>3</td><td>2</td></tr></table>	Keywords	Harry Potter 1	Harry Potter 2	Harry Potter 3	Harry Potter 4	Harry Potter 5	Users						User 1	5	4	3	2	-	User 2	2	4	5	-	-	User 3	3	4	2	4	3	User 4	-	1	-	3	2	User 5	2	5	3	-	-	User 6	2	-	-	3	2	06	CO4
Keywords	Harry Potter 1	Harry Potter 2	Harry Potter 3	Harry Potter 4	Harry Potter 5																																														
Users																																																			
User 1	5	4	3	2	-																																														
User 2	2	4	5	-	-																																														
User 3	3	4	2	4	3																																														
User 4	-	1	-	3	2																																														
User 5	2	5	3	-	-																																														
User 6	2	-	-	3	2																																														
Q.4 b	<p>What is sentiment analysis? Explain sentiments analysis process in detail by considering YouTube user feedback?</p>	06	CO4																																																
Q.5 a	<p>Describe oozieworkflows? Explain some of the workflow elements?</p> <p style="text-align: center;">OR</p> <p>Describe how to start and configure a flume agent in detail?</p>	06	CO5																																																
Q.5 b	<p>Describe in detail how to submit a workflow job and a coordinator job in oozie workflows?</p>	06	CO5																																																