Sardar Patel Institute of Technology

Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai-400058, India

(Autonomous College Affiliated to University of Mumbai)

End Semester Examination

November 2019

Max. Marks: 60 Duration: 3 Hr

Class: M.Tech Semester:I

Course Code: CE912 Branch: Computer

Name of the Course: Big Data Analytics and Management

Instructions:

(1) All questions are compulsory

(2) Assume suitable data if necessary

(3) Draw neat diagram wherever required.

Q No.		Max. Marks	CO-BL-PI
Q.1a	Write a Map reduce program using Java - To count the words excluding following words :- the, this, there, that, these. Note : No need to write the driver class	6	2-3-2.2.3
Q.1b	Write a pseudocode for computing Difference operation using Map reduce.	6	1.2-3-2.1.2
Q.2a	Justify - Combiner can be viewed as mini-reducers in the map phase. Draw a scenario to explain combiner in Map and Reduce phase OR What is partitioning phase – Draw and explain need of partitioner's in map reduce you can take any relevant scenario and explain this	6	2-3-2.3.2
Q.2b	Write a pseudocode for computing Intersection operation using Map reduce	6	2-3-2.1.2
Q.3a	What is Reducer side join ? Write a pseudocode for the data set having following schema : SchCustomerSale(TranID, Date, CustID, Amount, Equipment, city, state, mode)	6	2-3-2.3.2
Q.3b	What is the use of Oozie coordinator.	6	5-3-1.1.2
Q.4a	Write CURE Algorithm. Explain initialization phase and completion phase. OR How amazon reviews can be used for sentiment analysis? Write the basic steps for it.	6	3-3-3.1.2

<table>
<tr><td>Q.4b</td><td>Discuss how item based collaborative filtering can be applied for the example given.

<table><tr><th>User</th><th>Item</th><th>Rating</th></tr><tr><td>u1</td><td>m1</td><td>2</td></tr><tr><td>u1</td><td>m3</td><td>3</td></tr><tr><td>u2</td><td>m1</td><td>5</td></tr><tr><td>u2</td><td>m2</td><td>2</td></tr><tr><td>u3</td><td>m1</td><td>3</td></tr><tr><td>u3</td><td>m2</td><td>3</td></tr><tr><td>u3</td><td>m3</td><td>1</td></tr><tr><td>u4</td><td>m2</td><td>2</td></tr><tr><td>u4</td><td>m3</td><td>2</td></tr></table></td><td>8</td><td>3-3-3.1.2</td></tr>
<tr><td>Q.5a</td><td>Write any three varieties in Social Network graph.

OR

Write any three applications of Data Streams.</td><td>3</td><td>4-3-3.1.4</td></tr>
<tr><td>Q.5b</td><td>Why are dead-ends and spider traps a problem and why do teleports solve the problem?</td><td>7</td><td>4-3-3.14</td></tr>
</table>