Question Paper

Exam Date & Time: 29-Jun-2022 (02:00 PM - 05:00 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

Manipal School of Information Sciences (MSIS), Manipal
Second Semester Master of Engineering - ME (Cloud Computing) Degree Examination - June 2022

Data Streaming and Visualization [CDC 5203]

Marks: 100 **Duration: 180 mins.**

Wednesday, June 29, 2022

Answer all the questions.

1)	With example, show how machine, people and organization contribute to Big Data. List the challenges associated with Big Data. (TLO 1.1) (6+4 marks)	(10)
2)	Differentiate the requirements, responsibility and purpose of Batch, Serving and Speed layer. (TLO 2.1) (4+2+4 marks)	(10)
3)	With sketch list the steps involved in read and write operations in HDFS. (TLO 2.1) (4+6 marks)	(10)
4)	Discuss different ways of job failures in MapReduce and how failures are handled. (TLO 2.1) (10 marks)	(10)
5)	Write a python script to scrape book data from http://books.toscrape.com/catalogue/category/books/fantasy_19/index.html. Data may be present is multiple pages. "Next" button is provided to move to next page. Tag name of "Next" is **li** with attribute name 'next'. Data need to be scraped are "Book Title", "Author" and "Price". Tags for Title, Author and Price are **h2**, **h3** and **h4**. (TLO 3.1) (10 marks)	(10)
6)	Write python code to generate stacked bar chart and donut chart. (TLO 3.2) (5 + 5 marks)	(10)
7)	Explain any five use cases of stream processing. Write Short notes on Data Sources and Data Sinks of Structured Streaming. (TLO 1.2) (5 + 5 marks)	(10)
8)	Discuss the difference between informative, persuasive and visual art. Where they are used? (TLO 3.2) (7 + 3 marks)	(10)
9)	Differentiate between Spark and Hadoop MapReduce. List the areas where Spark and Hadoop MapReduce are good. (TLO 2.1) (6 + 4 marks)	(10)
10)	Text file contain numbers. Numbers are separated by one tab. There is no order to store the numbers. One line may contain one or more numbers. Find the maximum, minimum, sum and mean of numbers using Spark RDD. (TLO 2.2) (10 marks)	(10)

-----End-----