Question Paper

Exam Date & Time: 19-May-2022 (10:00 AM - 01:00 PM)



VI Semester End Semester Examination Big Data Analytics (ICT 4034)

BIG DATA ANALYTICS [ICT 4034]

Marks: 50

Duration: 180 mins.

Descriptive Questions

Ans	wer all the	Section Duration: 180 mins		
ANS	WER ALL (QUESTIONS		
1)		Explain the components of HDFS		(5)
	A)			
	B)	Explain the primary architectural components of spark		(3)
	C)	What is business intelligence?		(2)
2)		Explain with a pseudo code the concept of Map Reduce for matrix multiplication		(5)
	A)			
	B)	Explain the characteristics that makes a data, a big data?		(3)
	C)	Differentiate between data analytics and data analysis		(2)
3)		With an example explain the working of Spark stream.		(5)
	A)			
	B)	A data file named people.json is available with details of employees. Write the spar perform the following tasks: a) start a spark session b) read the contents of the given the contents of the file.	k script to n file c) display	(3)
	C)	With a suitable example differentiate Transformer from Estimator		(2)
4)	A)	A database of movie has to be maintain with name of movie, year, director, boxoffice rating and genre. Write suitable mongoDb queries for completing the following task. database and collection, ii)Add 5 documents into this database. (iii)	e collection, (i) create a	(5)
	,	Update rating of all movies directed by 'Spielberg' to 5 (iv)Delete all movies of genre "crime" (v) Find the total boxoffice collection of movie directed by 'Spielberg'.	"horror" or	
	B)	What are the different categories available based on the way the NOSQL devices st	ore data?	(3)
	C)	Taking an example for Amazon online store, explain various types of data analysis		(2)
5)		Explain in detail MLlib Pipeline for simple text document processing		(5)
	A)			

- B) How does Spark become a solution to the issues seen in Hadoop?
- C) MLlib supports sparse matrices, whose non-zero entry values are stored in the Compressed Sparse (2) Column (CSC) format in column-major order. Write the scala/ python code using MLlib to get the following matrices.

			[0]	0	44]	
			22	0	0	
[11	0	ן0	0	0	33	
i) 🕻 👩	12	ol ii)	Lo	55	0	

-----End-----