


Name: Enrolment No:			
UNIVERSITY OF PETROLEUM AND ENERGY STUDIES End Semester Examination, May 2023			
Course: Stream Processing Program: B.Tech-CSE-BD Course Code: CSBD 3010		Semester: VI Time: 03 hrs. Max. Marks: 100	
Instructions: Explain in short. (60-70 words)			
SECTION A (5Qx4M=20Marks)			
S. No.		Marks	CO
Q 1	Explain the difference in between micro batch based and continuous data processing.	4	CO1
Q 2	Differentiate in between stream and event processing with the help of a suitable example of each.	4	CO2
Q 3	Briefly discuss the difference in between structured and unstructured data in context with data streaming.	4	CO3
Q 4	Explain the concept of real time ETL with the help of an example.	4	CO3
Q 5	Discuss the concept of check pointing with the help of an example.	4	CO4
SECTION B (4Qx10M= 40 Marks)			
Instruction: Write brief notes. (100-150 words)			
Q 6	Define and deliberate the operations RDD supports and how are they different from D-Stream with the help of an example.	10	CO1
Q 7	Clarify the concept of stream processor despite of having computing platforms like Spark or MapReduce help of a suitable example.	10	CO1
Q 8	Discourse about varied analytics performed streaming data. OR Discuss any two output operations which can be performed on DStreams.	10	CO2
Q 9	Discuss the concept of RDD lineage in association with Spark streaming with the help of an example.	10	CO2
SECTION-C (2Qx20M=40 Marks)			
Instruction: Write long answer. (Up to 350 words while explaining) Attempt any part of question no. 10 as there is an option “a” OR “b”. There is no choice for question no.11.			

Q 10	Differentiate in between stateless and stateful streaming in detail with the help of an example. OR Discuss varied challenges in distributed systems in terms of continuous streaming of data.	20	CO4
Q 11	Deliberate any two use cases of stream processing in weather forecasting and stock analytics domain.	20	CO3