UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, Dec 2021

Course: Stream Processing **Semester: VII**
Program: B.Tech CSE-SPZ-BD **Time: 03 hrs.**
Course Code: CSBD4001 **Max. Marks: 100**

Instructions:

SECTION A
1. Each Question will carry 4 Marks
2. Instruction: Explain in short. (60-70 words)

S. No.		Marks	CO
Q1.	Explain the difference in between batch and real time data processing.	4	CO1
Q2.	Describe the concept of transformations in Spark with a suitable example.	4	CO2
Q3.	Briefly discuss the difference in between stream-stream and stream-static join.	4	CO3
Q4.	Explain real time streaming ETL with the help of an example.	4	CO3
Q5.	Differentiate in between error and trace log with the help of an example.	4	CO4

SECTION B
1. Each question will carry 10 marks
2. Instruction: Write brief notes. (100-150 words)

Q6.	Define and discuss the operations RDD supports and how are they different from D-Stream.	10	CO1
Q7.	Explicate the concept of fault tolerance and recovery in case of stream based processing.	10	CO1
Q8.	Distinguish in between stateful and stateless streaming in detail. **OR** Describe the concept of window transformations in Spark with a suitable example.	10	CO2

Q9.	Illustrate arbitrary stateful computations in spark.	10	CO2

SECTION-C

1. Each Question carries 20 Marks.
2. Instruction: Write long answer. (Up to 350 words while explaining)
3. Attempt any part of question no. 10 as there is an option "a" OR "b".
4. There is no choice for question no.11.

Q10	a) Demonstrate Apache Flume architecture and its data flow components in detail. **OR** b) Exemplify the concept of executor memory in spark with a suitable example and calculate the executor memory in a cluster considering the following information: Number of nodes = 10 Number of cores in each node = 15 cores RAM of each node = 61GB Marks distribution:10+10=20	20	CO4
Q11	Clarify late data handling and watermarking in streaming and its role with the help of a suitable example. (10) Discuss the case study of Kappa architecture implementation at Ericsson. (10)	20	C03