Department of Artificial Intelligence and Machine Learning

Date: 12-01-26	Test – 3	Max. Marks: 10 + 50
Semester: VII	UG	Duration: 2 Hrs.
Course Title: Stream Processing and Analytics		Course Code: AI372TA

Common to AIML and CSE (Data Science)

PART A

S. No	Questions	M	BT	CO
1	_____is tuple-at-a-time stream-processing framework designed for real-time processing of data streams, while _____ is a stage-wise stream-processing framework	2	2	1
2	Identify three key differences between traditional query model and streaming query model.	2	2	3
3	What is Concept drift in streaming systems ? Give an example.	2	2	2
4	What is a tumbling window? List the two types of tumbling windows	2	2	2
5	Enumerate the techniques used for stream summarization.	2	2	1

PART B

S. No	Questions	M	BT	CO
1a.	Analyze the need for distributed stream processing and explain the key components of a distributed stream-processing architecture with a neat diagram.	10	2	2
2 a	Explain the importance of state management in stream-processing systems using an example.	4	3	3
2b	Briefly discuss the three ways to write data to long-term storage.	6	2	2
3 a	A video-analytics platform continuously ingests CCTV camera feeds and needs to store the processed results for long-term historical analysis, regulatory compliance, and future model training. Explain the three possible ways this system can write streaming data to long-term storage	6	2	2
3 b	Compare the State-Machine and Rollback Recovery approaches for fault tolerance in stream-processing systems.	4	2	2
4	Identify widely used stream-processing frameworks and illustrate the working of any two by applying their core concepts. Use a labelled diagram to explain how these frameworks enable distributed stream-processing.	10	4	3
5.	Analyze any two stream summarization techniques in real-time data processing. Compare their characteristics and justify with suitable examples	10	4	2