|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Register No.** |  |  |  |  |  |  |  |  |  |  |

**SET A**

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY FACULTY OF ENGINEERING & TECHNOLOGY SCHOOL OF COMPUTING

Cycle Test – I Academic Year: 2023-2024

Program : M.Tech Int - DS

Course Code and Title : 21CSE373T – Streaming Analytics

Year / Sem: III/VI Max. Marks: 50

Duration : 100 mins Date of Exam: 22-02-2024

|  |  |  |
| --- | --- | --- |
| **Course Learning Outcomes (CO):** | | *At the end of this course, learners will be able to:* |
| CO-1: | *illustrate the concepts and terminologies in stream processing* | |
| CO-2: | *Interpret stream processing applications using Apache Spark Streaming* | |
| CO-3: | *Summarize real-time streaming data pipelines and applications that adapt to the data streams using Kafka* | |
| CO-4: | *Interpret stream processing applications using Apache Storm Streaming* | |
| CO-5: | *inquire real time data using NoSQL databases & MongoDB* | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Question No.** | **Reference to Outcome** | **Marks Allotted (Total 25)** | **Marks Scored** | **Outcomes Met**  **Yes / No** |
| 1 | 1 | 1 |  |  |
| 2 | 1 | 1 |  |  |
| 3 | 1 | 1 |  |  |
| 4 | 1 | 1 |  |  |
| 5 | 2 | 1 |  |  |
| 6 | 2 | 1 |  |  |
| 7 | 2 | 1 |  |  |
| 8 | 1 | 1 |  |  |
| 9 | 2 | 1 |  |  |
| 10 | 2 | 1 |  |  |
| 11 | 1 | 5 |  |  |
| 12 | 1 | 5 |  |  |
| 13 | 2 | 5 |  |  |
| 14 | 2 | 5 |  |  |
| 15 | 2 | 5 |  |  |
| 16 | 2 | 10 |  |  |
| 17 | 2 | 10 |  |  |

Faculty Name: Signature:

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

**Set A**

FACULTY OF ENGINEERING & TECHNOLOGY SCHOOL OF COMPUTING

Cycle Test – I Academic Year: 2023-2024

Program : M.Tech Int - DS

Course Code and Title : 21CSE373T – Streaming Analytics

Year / Sem: III/VI Max. Marks: 50

Duration : 100 mins Date of Exam: 22-02-2024

**PART-A**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sl. No** | **Question** | **CLO** | **Bloom’s Taxonomy** | **Marks** |
| 1 | Which of the following is NOT a transformation?  a) map  b) filter  c) average  d) flatmap | 1 | L1 | 1 |
| 2 | Which of the below Spark entities is built on Spark SQL?   1. Structured Streaming 2. Spark MLib 3. Spark Streaming 4. Spark Core | 1 | L1 | 1 |
| 3 | Spark follows which of the below stream processing programming model?   1. Element-at-a-time 2. Micro Batching 3. Cloud Computing 4. Grid Analysis | 1 | L1 | 1 |
| 4 | Out of the below components, which represents the portion where output is written to, i.e an abstraction used to write a data stream outside of Spark’s control?  a) Source  b) Task  c) Operators  d) Sink | 1 | L1 | 1 |
| 5 | Select the feature NOT supported by HDFS   1. Lots of small files   b) Few large files  c) Streaming data access  d) Write once read many times | 2 | L1 | 1 |
| 6 | Aggregations in Spark has --------------- dependencies  a) narrow  b) wide  c) infinite  d) zero | 2 | L1 | 1 |
| 7 | The time periods in --------------- window are non-overlapping  a) tumbling  b) sliding  c) stateless  d) stateful | 2 | L1 | 1 |
| 8 | This approach, in which we are simultaneously trying to understand  new data in the context of data already seen, often leads us ------------  a) stateful processing  b) stateless processing  c) big data processing  d) machine learning | 1 | L1 | 1 |
| 9 | The architecture that supports only batch processing and no stream processing is called ------------ architecture  a) Kappa  b) Lambda  c) Single tier  d) Multi tier | 2 | L1 | 1 |
| 10 | Which one of the following are the Spark’s internal cluster managers?  a) local, standalone  b) local, distributed  c) multi, local  d) remote, multi | 2 | L1 | 1 |

PART – B

Answer **any** **Two** questions **3 X 4 = 12 Marks**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sl. No** | **Question** | **CLO** | **Bloom’s Taxonomy** | **Marks** |
| 11 | Compare the different types of windows in stream processing with necessary illustrations and examples. | 1 | L2 | 5 |
| 12 | Analyze the three data delivery semantics with example scenarios | 1 | L2 | 5 |
| 13 | Compare and contrast micro batching and one-at a time stream processing. | 2 | L2 | 5 |
| 14 | Discuss on the factor of uncertainty in stream processing | 2 | L2 | 5 |
| 15 | List any three external popular cluster managers that could be used with Spark and how they are used. | 2 | L2 | 5 |

PART C

Answer any **One** question **1\*13=13**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sl.No** | **Question** | **CLO** | **Bloom’s Taxonomy** | **Mar ks** |
| 16. | Detail on the below with respect to Stream processing  a) Streaming Architectures  b) Stateless and Stateful Stream Processing | 2 | L3 | 10 |
| (OR) | | | | |
| 17. | With sample code snippets and flow diagram explain how you will be applying aggregation on streaming data like network word count or any other streaming application example. | 2 | L3 | 5  5 |