Department of Artificial Intelligence and Machine Leaning

Date: 23.06.2025	Improvement Test	Max. Marks: 50+10
Semester: VI	UG	Duration: $1\frac{1}{2}$ Hrs + ½ Hr
Course Title: Big Data Technologies		Course Code:AI362IA

Sl No	Questions	M	BT	CO
1.	Define flume	2	2	1
2.	Define Interceptor	2	2	1
3.	State two spark's processing model	2	2	1
4.	List any two programming language APIs supported by Apache Spark	2	2	1
5.	How do you identify whether a given operation is a transformation or an action in spark?	2	2	1
PART- B				
1	Discuss Job submission phase of a Spark Job Run with a diagram?	10	2	1
2(a)	Discuss any two ways of creating RDD's in Spark?	5	2	1
2(b)	Summarize reduceByKey(), foldByKey()transformations for aggregating RDDs with an example for each?	5	2	2
3(a)	Demonstrate the operation of Apache Flume with a supporting diagram and configuration example	6	2	2
3(b)	Compare and contrast two aspects of serialization in apace spark?	4	2	2
4(a)	Discuss how to configure flume for HDFS sink?	6	2	2
4(b)	Define Fanout? Illustrate with a diagram how event delivery occurs from a single source to multiple sinks in Apache Flume	4	2	2
5	Justify the given Statement? "A sink group allows multiple sinks to be treated as one"	10	3	2