Semester: VII						
STREAM PROCESSING AND ANALYTICS **Category: Professional Core Course** **(Common to AI,CD)** **(Theory and Practice)**						
Course Code	**:**	**AI372IA**		**CIE**	**:**	**100+50 Marks**
Credits: L: T: P	**:**	**3:0:1**		**SEE**	**:**	**100+50 Marks**
Total Hours	**:**	**45L+30P**		**SEE Duration**	**:**	**3.00 Hours**

Unit-I	**9Hrs.**
Introducing Streaming Data: What is Real time system – Differences between real time and streaming systems – architectural blue print – security for streaming systems – scaling **Data Ingestion:** Common Interaction patterns – scaling the interaction patterns – Faulty tolerance	
Unit – II	**9Hrs.**
Data Transportation: Message queue – Core concepts – security – application of core concepts to business logic **Analysing Streaming Data:** Inflight data analysis – Distributed stream processing architecture – key features of stream processing frame work	
Unit –III	**9Hrs.**
Algorithms for Data Analysis: Accepting constraints and relaxing – Thinking about time – Summarization Technique **Storing the analysed or collected data:** Long time storage – keeping it in memory	
Unit –IV	**9Hrs.**
Introduction to Kafka: Why Kafka – Kafka Eco System – Kafka Origin - Kafka Producers and Consumers	
Unit –V	**9Hrs**
Building Data Pipe lines – When to use pipe lines – when to use kafka connect vs producer and consumer **Kafka Streams –** Stream Processing design patterns **-** Architecture over view – How to choose Stream processing framework - Kafka streams by example – word count – stock market statistics – click stream enrichment	

Lab Component
Group of two students of same batch are required to build an application using stream processing tools for various real time applications like (i) Real time Sentiment Analysis (ii) Stock Market analysis (iii) Click stream enrichment (iv) In-flight analysis (v) video stream processing etc

Course Outcomes: After completing the course, the students will be able to:-	
CO1	Describe the need and the application of real time and stream processing in real world applications.
CO2	Comprehend and apply the various operations like data ingestion, data communication, data analysis and storage for different streaming data applications.
CO3	Investigate and apply streaming concepts using modern tools to solve problems related to society and industry.
CO4	Demonstrate a prototype application for streaming data using Kafka as a team / individual.
CO5	Demonstrate solutions for societal and environmental concern problems using modern engineering tools through writing effective reports.

References Books	
1.	Streaming Data – Understanding the Real time Pipe Line ,Andrew Psaltis, Manning Publications, 1st Edition: 2017, ISBN: 9781617292286
2.	Kafka: The Definitive Guide: Real-Time Data and Stream Processing at Scale ,Gwen Shapira, Todd Palino, Rajini Sivaram, Krit Petty, , O'Reilly Media, 2nd Edition, November 2021, ISBN: 978-1-492-08736-6
3.	Streaming Systems ,Tyler Akidau, Slava Chernyak, and Reuven Lax, , O'Reilly Media , 1st Edition 2018 , ISBN : 978-1-491-98387-4
4.	Fundamentals of Stream Processing Application Design, Systems, and Analytics ,Henrique C. M. Andrade, Bugra Gedik , Deepak S. Turaga ,Cambridge University Press 2014 , 1st Edition, ISBN 978-1-107-01554-8 Hardback

RUBRIC FOR THE CONTINUOUS INTERNAL EVALUATION		
#	**COMPONENTS**	**MARKS**
1.	**QUIZZES:** Quizzes will be conducted in online/offline mode. TWO QUIZZES will be conducted & Each Quiz will be evaluated for 10 Marks. **THE SUM OF TWO QUIZZES WILL BE THE FINAL QUIZ MARKS.**	**20**
2.	**TESTS:** Students will be evaluated in test, descriptive questions with different complexity levels (Revised Bloom's Taxonomy Levels: Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating). Two tests will be conducted. Each test will be evaluated for 50 Marks, adding upto 100 Marks. **FINAL TEST MARKS WILL BE REDUCED TO 40 MARKS.**	**40**
3.	**EXPERIENTIAL LEARNING:** Students will be evaluated for their creativity and practical implementation of the problem. Case study-based teaching learning (10), Program specific requirements (10), Video based seminar/presentation/demonstration (20) **ADDING UPTO 40 MARKS**.	**40**
4.	**LAB:** Conduction of laboratory exercises, lab report, observation, and analysis (40 Marks) and lab test (10 Marks) adding up to 50 Marks. **THE FINAL MARKS WILL BE 50 MARKS**	**50**
MAXIMUM MARKS FOR THE CIE (THEORY AND PRACTICE)		**150**

RUBRIC FOR SEMESTER END EXAMINATION (THEORY)		
Q.NO.	**CONTENTS**	**MARKS**
	PART A	
1	Objective type of questions covering entire syllabus	20
	PART B (Maximum of THREE Sub-divisions only)	
2	Unit 1 : (Compulsory)	16
3 & 4	Unit 2 : Question 3 or 4	16
5 & 6	Unit 3 : Question 5 or 6	16
7 & 8	Unit 4 : Question 7 or 8	16
9 & 10	Unit 5: Question 9 or 10	16
	TOTAL	**100**

RUBRIC FOR SEMESTER END EXAMINATION (LAB)		
Student is required to perform Computer Simulation/ Develop a prototype or model as the case may be and present the results in the form of a presentation. Further, students have to submit a poster for exhibition and also a report.		
Q.NO.	**CONTENTS**	**MARKS**
1	Design and development of the project	20
2	Presentation of the working model/simulation results/prototype building	20
3	Viva	10
	TOTAL	**50**