



Ministry of Education  
Government of India

---

---

MHRD SPONSORED OFFLINE GIAN COURSE ON  
**DISTRIBUTED SYSTEMS AND MACHINE LEARNING**

06 - 10 JUNE 2022

**VENUE:**

*Department of Computer Science and Engineering  
Indian Institute of Technology Patna  
Bihta, Patna -801106 (Bihar)*



Prof. Supratik Mukhopadhyay  
Louisiana State University, USA



Prof. Rajiv Misra  
IIT Patna



## Overview

“A distributed system is one in which the failure of a computer you didn't even know existed can render your own computer unusable.” –Leslie Lamport

Modern computing has seen the proliferation of distributed systems and applications in various forms such as geo-replicated data stores, cloud computing systems, networked, fog, and serverless computing systems, etc. that drive entities that range from global enterprises to vehicular systems. The reliable and optimal operation of such systems has become an issue of utmost importance. Correct and optimal operation of distributed systems depends on several factors: appropriate choice of consistency settings based on application contexts, appropriate programming models that enable the development of reliable applications, and predictive execution models that enable optimal usage of resources.

Today, killer applications that utilize distributed computing systems are machine learning engines. Machine learning, in particular deep learning, has seen tremendous success in recent years with applications ranging from web search, product recommendations, healthcare diagnostics, autonomous driving, etc. Training deep learning engines amounts to solving approximately optimization problems that require searching for a “saddle point” in a high dimensional space: a computationally expensive task. Distributed cloud-based systems are routinely used to accelerate the search for saddle points as well as store large volumes of data for training deep learning engines. On the other hand, machine learning engines form the basis of predictive execution models for optimal usage of resources in distributed systems. In addition, machine learning systems find usage in predicting and diagnosing faults, vulnerabilities, and attacks in distributed systems.

This course will focus on the interplay between distributed systems and machine learning. The primary goal is to examine how distributed systems can enable reliable and efficient deployment of machine learning engines and how machine learning can contribute towards the correct and optimal operation of distributed computing systems.

## Objectives

The primary objectives of the course are as follows:

- 1) To provide an understanding of the principles behind the reliable and optimal operation of distributed computing systems,
- 2) To provide an understanding of the interactions between the fields of machine learning and distributed computing,
- 3) To provide an understanding of the use of distributed computing for accelerating machine learning algorithms,
- 4) To provide an understanding of the use of machine learning for reliable and optimal operation of distributed systems,
- 5) To provide hands-on experience in building applications using distributed machine learning frameworks,
- 6) To provide exposure to state-of-the-art distributed machine learning frameworks like Petuum, GraphLab, etc.

**Course Details:**

Date	Lectures/Tutorials	Instructor	Hours
Day 1	Lecture 1: Distributed Programming Abstractions, Safety, Liveness, Process Failures	Prof. Supratik Mukhopadhyay	1
	Lecture 2: Notions of Synchrony and Asynchrony, Causality and Orderings of Events, Happened-Before Relation, Notion of Sequential Consistency	Prof. Supratik Mukhopadhyay	1
	Tutorial 1: Problem solving session with examples: Examples of benign and byzantine faults, Depicting Distributed Systems through Sequence Diagrams, Developing a socket program for communicating between two machines	Prof. Supratik Mukhopadhyay	2
Day 2	Lecture 3 : Notions of Availability, Partition Tolerance, CAP Theorem, Eventual Consistency	Prof. Supratik Mukhopadhyay	1
	Lecture 4: Kadmelia, Chord, Distributed Hash Tables, Geo-replicated stores	Prof. Rajiv Misra	1
	Tutorial 2: Problem solving session with examples: Case study with Dynamo, Cassandra, Spanner	Prof. Rajiv Misra	2
Day 3	Lecture 5: Introduction to Machine Learning: Linear and Logistic Regression, Decision Trees	Prof. Supratik Mukhopadhyay	1
	Lecture 6: Machine Learning for Predicting Client-centric Consistency: OptCon and Consistify	Prof. Supratik Mukhopadhyay	1
	Tutorial 3: Problem solving session with examples: Python Programming, Programming Random Forest: Predict Store Access Pattern, Case study on Google Dapper	Prof. Rajiv Misra	2
Day 4	Lecture 7: Introduction to Cloud Computing, Hypervisors. Distributed Computing Frameworks: Hadoop, Apache Spark	Prof. Rajiv Misra	1
	Lecture 8: Optex: Execution Model for Apache Spark; Infra: Cloud optimization; Scaling machine learning engines using Apache Spark	Prof. Supratik Mukhopadhyay	1
	Tutorial 4: Problem solving session with examples: Case study with AWS	Prof. Rajiv Misra	2
Day 5	Lecture 9: Machine learning on the cloud: AutoML	Prof. Rajiv Misra	1
	Lecture 10: Graph processing on the Cloud	Prof. Rajiv Misra	1
	Tutorial 5: Case study on Petuum and Graphlab frameworks	Prof. Supratik Mukhopadhyay	2

**Date of Examination:** June 11, 2022

### Who can attend

- Executives, engineers and researchers from the software industry, service, and government organizations including R&D laboratories.
- Student students at all levels (BTech/MSc/MTech/PhD) or Faculty from reputed academic institutions and technical institutions.

### How to Register

#### **Step 1: One-time GIAN Registration to get access to all GIAN courses:**

All prospective participants need to do mandatory web registration for this course on the GIAN portal through the link <https://gian.iitkgp.ac.in/GREGN/index> by paying Rs. 500/- (those who have already registered, need not do it again). After that, register for the course titled “**Distributed Systems and Machine Learning**” from the drop-down menu of course registration.

#### **Step 2: The course fee for attending the course is as follows:**

**Participants from abroad: Rs. 10000**

**Industry/ Research Organizations: Rs. 7500**

**Faculty from Academic Institutions: Rs. 5000**

**Research Scholars/Students: Rs. 2500**

The above fee includes all instructional materials (soft copy) and virtual laboratory use for tutorials and assignments.

#### **Payment (State Bank iCollect):**

Payment Link: <https://www.onlinesbi.com/sbicollect/icollecthome.htm?corpID=1968961>

Select Payment Category “GIAN-DSML” and pay for the appropriate category.

*The number of participants for the course will be limited to fifty.*

#### **Step 3: Registration Link: <https://forms.gle/z52HNhi1npMPm4Z76>**

Link for the course will be provided to the participants.

#### **Contact:**

Prof. Rajiv Misra (*Course Coordinator*)

E-mail: [rajivm@iitp.ac.in](mailto:rajivm@iitp.ac.in)

Rohit Kumar Gupta (*Ph.D Scholar*)

E-mail: [rohit10495@gmail.com](mailto:rohit10495@gmail.com), [1821cs16@iitp.ac.in](mailto:1821cs16@iitp.ac.in)

Mobile: +91-7873759080

## Accommodation

There is limited availability of accommodation in IIT Patna hostels at an affordable rate, which will be offered on a first-come-first-served basis. Besides, there are several hotels and guest-houses around IIT Patna where the participants may stay during the GIAN course.

## The Faculty



**Dr Supratik Mukhopadhyay** is a faculty member in Computer Science at Louisiana State University. His research interests lie in the areas of Artificial Intelligence/Machine Learning with applications to Education, Automated Drug Discovery, Satellite Imagery Recognition, Transportation Systems, Sustainable Buildings, Cyber-Physical Human Systems, etc. In these areas, Dr Mukhopadhyay's research has been supported by NSF, NASA, ONR, DARPA, ARO, USDOT, NGA, DOE, NRL, state agencies, and industry. He has more than 110 publications in reputed journals and conferences and has three awarded US Patents. He led the DeepDrug team to the semifinal of the IBM Watson Artificial Intelligence XPRIZE. He is an Associate Editor of IEEE Transactions on Artificial Intelligence and Remote Sensing Letters and a program committee member for AAAI 2021.



**Dr Rajiv Misra** is a Professor at the Department of Computer Science and Engineering, Indian Institute of Technology Patna. Dr Misra did his PhD from IIT Kharagpur, M.Tech from IIT Bombay and B.Tech from NIT Allahabad. His area of interest includes Computer Communications (Networks), BigData, IIOT, Cloud Computing, Distributed Systems and Algorithms, 5G & Beyond Networking, Network Slicing, AI-ML, UAV-assisted MEC etc. His current project is 'AI-based 6G Network Slicing for multi-UAV prototype'.