



## **IIIT-Delhi and IBM Research join hands to introduce Data Lifecycle Management Course for CSE students**

**New Delhi, 20<sup>th</sup> July 2020** - IIIT-Delhi, in association with IBM Research, is introducing a ground-breaking course in Data Science 'Data Lifecycle Management' for the CSE students of the institute. The course will be offered to the pre-final year and final year UG, PG and Doctorate students. Although, IIIT-Delhi already has many data science courses in its program structure, the proposed course is unique and focuses more on data management issues for AI pipeline.

The course will cover the different components and challenges for the Data Lifecycle. It will help the students to understand the evolution of Data Research from Business Intelligence to Artificial intelligence to Hybrid Cloud. The course will take student's knowledge on Data Science and AI a step further, by explaining the different challenges involved in the management and preparation of data for ML applications. The course will also cover state-of-the-art algorithms and best practices to handle data to construct better ML pipelines. The course will be taught by Dr Sameep Mehta and Hima Patel from IBM Research.

Introducing the course, Dr. Vikram Goyal, HOD, Department of CSE, IIIT-Delhi said, *"We are very excited to offer this first-of-a-kind course at IIIT-Delhi. The course will expose students to the data side of the AI pipeline with a healthy mix of theoretical concepts and hands-on labs. I believe this course will prepare our students well for tackling real-world AI problems."*

Dr. Sameep Mehta, IBM Research AI, added, *"It is important for our next generation of AI researchers to understand the data lifecycle. In a typical AI project, around 80% of the effort is spent on data acquisition, cleaning and preparation, whereas model learning accounts for 20%. This course will focus on teaching these concepts in a principled fashion to the students."*

The course 2 credit course is divided into 6 sections which will cover *AI Background Refresher, Framework For Operationalizing Data For AI Tasks, Data Exploration, Data Quality Analysis For ML, Getting Data Ready For AI, How It All Comes Together In A Practical Setting, and Re-imagining Data in Hybrid Cloud Environments.*

The course will also feature a couple of guest lectures from industry experts to showcase how these principles are applied to build Large Scale Data Lakes.

For more details, kindly contact:

Simran Jain +91-8368504309

Sumit Kumar: \_91- 9999560239