

A Report On

WEBINAR

On

BIG DATA

Organized by

The Department of Computer Science & Engineering

June 24th, 2021

At

ICFAI University, Tripura

A Webinar on "Big Data" was organized by the CSE department and was delivered by Mr. Hrishav Bakul Barua, Robotics and AI Researcher, TCS Research, Kolkata on Thursday, June 24, 2021. The session was held online through Google meet. The event was focussed on one of the most important contemporary topics in the domain of Data Analytics. The participants were the Faculties from various departments and students from various programmes.

Prof. (Dr.) Biplab Halder, Pro-VC , ICFAI University highlighted the need of learning latest technologies like Big Data and Analytics. This will help the students to increase their efficacy along with real time processing skills. Making engineers aware of big data and analytics tools helps speed up problem identification and troubleshooting. Big data analytics is the use of advanced analytic techniques against very large, diverse data sets that include structured, semi-structured and unstructured data, from different sources, and in different sizes from terabytes to zettabytes. Analysis of big data allows analysts, researchers and business users to make better and faster decisions using data that was previously inaccessible or unusable. Businesses can use advanced analytics techniques such as text analytics, machine learning, predictive analytics, data mining, statistics and natural language processing to gain new insights from previously untapped data sources independently or together with existing enterprise data.

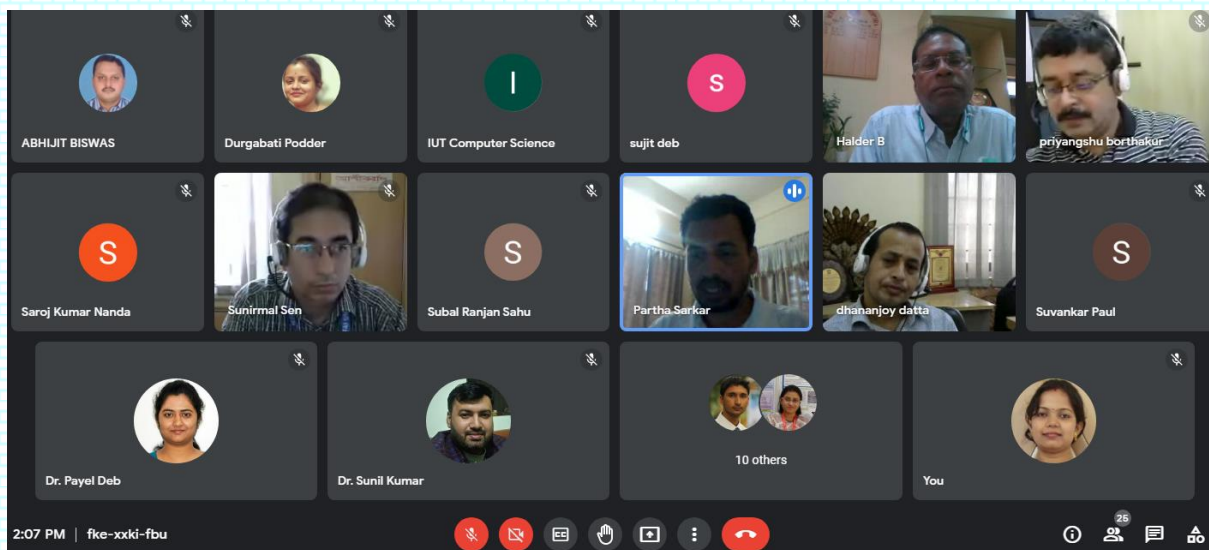
The speaker Mr. Hrishav Bakul Barua nicely explained the concept of Big Data and why it has been a burning subject in today's world. As we are encountering big volumes of data of different variations on a day to day basis, our method of study is gradually being based on the understanding of the type and trend of data which is the sole objective of Data Science.

The session continued from 14.00 hrs to 16.00 hrs. This has been a highly informative session where Mr. Barua explained the main concepts of Big Data , its life cycle, about the architectures followed, the storage mechanism and the taxonomy and may more.

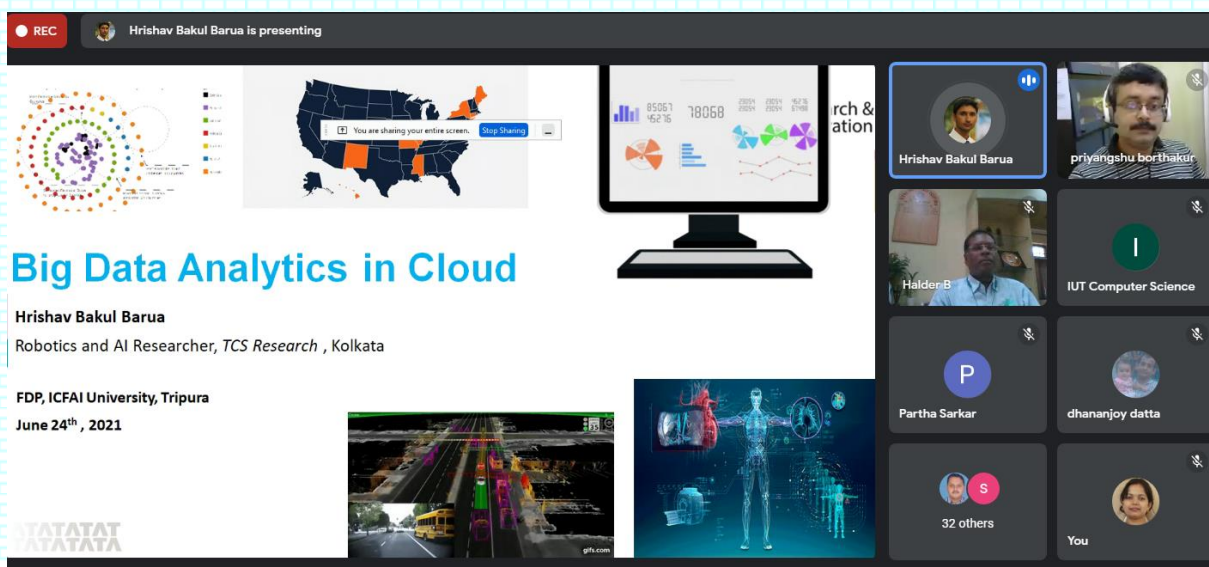
Further he explained the concept of green analytics when the main objective is to reduce the code execution time and thereby reducing the utilization of electrical energy.

At the end there was some questions being asked on the topic and Mr. Barua answered with clarity.

Dr. Priyangshu Rana Borthakur, Dean, Faculty of Science and Technology concluded the session by giving vote of thanks to all dignitaries present, the Authorities, the organizing department and students. He also thanked Mr. Hrishav Bakul Barua for such a wonderful session and Mr. Partha Sarkar, TCS for enabling such a session.



Picture 1: Prof. (Dr.) Biplab Halder giving his valuable speech.



Picture 2: Mr. Hrishav Bakul Barua presenting his session.

DATA DNA

It is very important to understand the so called “Data DNA” of various Data-driven organizations, sectors, industries, institutes, and enterprises. By Data DNA of Big data, we need to understand that we are exposed to an extremely massive amount of data and our responsibility is to get the information, insight and potential trends out of it. This is what Data Science is all about and we hope to take it to the next stage. The more experienced a person gets the more knowledgeable he becomes. Similarly, more the historical data we are exposed to, more is the scope of gaining insight and knowledge out of it!

5

TATA CONSULTANCY SERVICES

Picture 3: Some moment of presentation by Mr. Hrishav Bakul Barua

Basic components

Machine Learning, Mathematics, Statistics, Algorithms and Programming, Data Science, Data Structure and Database

14

TATA CONSULTANCY SERVICES

Picture 4: Some moment of presentation by Mr. Hrishav Bakul Barua

How big?

Experiments/scenarios	Data generated
Computational Fluid Dynamics turbulence simulations	100 TB
BaBar Particle Physics Experiments	1 TB/day
CERN Large Hedron Collider	1 GB/sec ~ 10 PB/yr
VLBA Radio Telescope	1 GB/sec
NCBI/EMBL bio-informatics databases	0.5 TB, doubles every year
Brain imaging	4 TB/brain (full colour, 10mm resolution)
Telecom (one provider)	~50-80M CDR/day => 20-30GB/day

6

TATA CONSULTANCY SERVICES

Picture 5: Some moment of presentation by Mr. Hrishav Bakul Barua