



# NEW HORIZON COLLEGE OF ENGINEERING

## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

### REPORT ON FDP ON “BIG DATA ANALYTICS WITH HADOOP- APACHE SPARK”

---

**Speaker:** Mr. Sachin Mudholkar, Chief Technology Officer, Talentpod

**Date:** 8<sup>th</sup> March 2019

**Venue:** HP Centre of Excellence Lab, Subash Chandra Bose Block

**Audience:** GIC Industry Members and NHCE Faculty

#### **Gist of the Event:**

HP Center of Excellence, Dept of CSE in association with Global IT Commune and Computer Society of India organized a FDP on “**Big Data Analytics using Hadoop-Apache Spark**” from 10:00 am -5:00 pm in HP COE Lab on 8th March 2019.

The speaker Mr. Sachin Mudholkar is currently working as Chief Technology Officer in Talentpod

Mr. Sachin gave a deep insight on how machine learning is playing an important role in big data.

Some of topics discussed in the session were:-

- Four stages in data analysis pipeline
- Architecture of Hadoop –hdfs ,mapreduce and spark
- How spark has taken over mapreduce
- How spark can be installed on windows/linux and run spark program using python/java
- Batch processing v/s stream processing
- Running of spark examples using pyspark
- Various real time applications of big data in industry and what are the challenges in their analysis were discussed by industry members from different domain like healthcare, telecom, e-commerce etc.

**Glimpses of the event:**



