

## Big Data, Hadoop, MapReduce, R, Hive, Pig and Machine Learning

### Big Data Impact and DSRI

**Big Data** is storming across all the Industry Sectors. Almost every industry is planning to monetize the benefits from this technology and also provide their clients value additions and enhance own business.

Globally, job creation for Big Data Analytics roles will reach 960K in 2021 with an average CAGR of 23%. NASSCOM predicts that Global Analytics is expected to touch \$16 billion in 2025 from the current \$2 billion.

During the last decade, the momentum coming from both academia and industry has lifted the **R Programming language** tool for computational statistics, visualization and decision science. Hive and Pig are used for data analysis of large amount of structured and unstructured data. Facebook uses Hive for data analysis and yahoo uses pig for its search engine. Sqoop is used for data injection from databases like MySQL etc. Flume is used for data extraction from social media like Facebook, Twitter etc. Mahout is used for machine learning techniques like clustering, classification and recommendations.

**Master Speaker: Dr. S. Basavaraj Patil, Ph.D** Chief Data Scientist of **Predictive Research** ([www.predictive-research.com](http://www.predictive-research.com)). Worked as Financial Data Scientist, Technical Architect and Independent Consultant having 20 years of experience in Big Data Analytics.

**Master Speaker: Dr. Hemanth Kumar P, Ph.D** Data Scientist of **Predictive Research**. He has worked in Data Mining, Simulation, Predictive Modeling and Product Development fields for more than 8 years in industry. He has conducted several workshops at Bangalore, Cochin and Coimbatore in collaboration with MSME.

**Associate Speaker: Mr. Harish L, M.Tech, (Ph.D)** Junior Data Scientist of **Predictive Research**. He has experience in Big Data Technologies.

Come to take a wonderful Ride on this technology get the feel of it..!



Day	Time	Workshop Content
Day 1	09:00-10:30	Introduction to Big Data. Basics of Data Science
	10.30-11.00	Coffee Break
	11.00-1:00	What is Hadoop? MapReduce & Hadoop Eco system
	1:00 – 2:00	Lunch Break
	2:00 – 3:30	Installation of Hadoop
	3:30 – 3:45	High Tea
	3:45 – 5:00	Single Cluster Multi Cluster MapReduce Examples
Day 2	09:00-10:30	Statistics & Data Mining. Why R?
	10.30-11.00	Coffee Break
	11.00-1:00	Data Analytics-R
	1:00 – 2:00	Lunch Break
	2:00 – 3:30	Data Analytics –R (contd...)
	3:30 – 3:45	High Tea
Day 3	09:00-10:30	Hive programming introduction
	10.30-11.00	Coffee Break
	11.00-1:00	Hands on experience on hive
	1:00 – 2:00	Lunch Break
	2:00 – 3:30	Introduction to Pig programming
	3:30 – 3:45	High Tea
Day 4	3:45 – 5:00	Hands on experience on Pig programming
	09:00-10:30	Sqoop Introduction
	10.30-11.00	Coffee Break
	11.00-1:00	Sqoop Hands on Experience
	1:00 – 2:00	Lunch Break
	2:00 – 3:30	Flume Introduction
Day 5	3:30 – 3:45	High Tea
	3:45 – 5:00	Flume configuration and hands on experience
	09:00-10:30	Introduction to Hbase
	10.30-11.00	Coffee Break
	11.00-1:00	Hands on experience
	1:00 – 2:00	Lunch Break
Day 5	2:00 – 3:30	Introduction to Mahout
	3:30 – 3:45	High Tea
	3:45 – 5:00	Hands on experience

**Limited seats  
Contact Immediately for Registration**

Come &  
Benefit

Workshop Expenses per Person:  
\$1550/- (Inclusive of all Taxes)  
Enquire For Group Discount

Contact Person: Ashvini Patil Ph. No: +1 415-233-2408

Mail to: [workshop@dsresearchinstitute.com](mailto:workshop@dsresearchinstitute.com) Visit us at: [www.dsresearchinstitute.com](http://www.dsresearchinstitute.com)

Don't Wait  
Register