



Dr. Vishwanath Karad
**MIT WORLD PEACE
UNIVERSITY** | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

School of Computer Engineering & Technology

Report on the hands on Workshop on “*Big Data Analytics using Hadoop: Basic concepts and installation*”

Date: 4th September 2020, Friday

Organized for: T.Y. B.Tech students of Professional Elective 1 (Big Data Analytics) of trimester T9

Attended by: More than 70 students.

Organized by: Prof. Suja Panicker, Prof. Pranali Kosamkar.

Report prepared by: Muskan Ahuja and Vaishnavi Goyal

About the Guest Speakers:

Mr. Yogesh Murumkar is Director at BharatSoft Solutions, Pune and also a Corporate Trainer. He is a Computer Engineer and has completed his Graduation and Post-Graduation from Pune University. He has worked for Capgemini India Pvt Ltd for 4 years and worked for clients like Barclays and S-Pankki bank. He is a Data Science and Big Data Analytics enthusiast with more than 10 years of experience including IT experience and teaching/training experience and is a Corporate Trainer for Machine Learning, Blockchain, Data Science and Big Data Analytics Trainer. He has successfully trained more than 300 participants for Global Certification in Data Science and Big Data Analytics by DELL EMC. He is also a Microsoft Certified Python Developer.

Viresh Chaptre is a computer engineer and has done Master of Technology in Computer Science Engineering, JNTU Hyderabad, 2013. He has several years of teaching experience as Assistant Professor in various Engineering Colleges in

Pune University and Aurangabad. He is skilled in - Big Data (Hadoop), DBMS (SQL), Core java, has completed DELL EMC Certification for EMC Academic Associate in Data Science and Big Data Analytics. His achievements include - Awarded by Indian Embassy in Bangkok,Thailand @Indian Commerce Chamber in Feb 2015, awarded by Asians Institute of technology,Bangkok Thailand as a chief guest in Feb 2015 , awarded by Glows of Thailand in March2016 (Chief guest for Book Launch), awarded by Russian Community in Pattaya for social activities in Russian community Thailand , being an Art Of Living Teacher he has conducted stress management courses for youngsters, lead the world Record Largest human map made by 1800 students, lead world Record Largest human portrait of architects of swarajya made by 4000 students, delivered an expert lecture on Big Data for PG students of Asians Institute of Technology,Bangkok,Thailand 17th Feb 2015 .

Agenda:

- Basic concepts of Big Data, Big Data Analytics
- Basic concepts of Hadoop Distributed File System
- VMware Work Pro 15
- VM Start

About the session:

The session began with student representatives (Anuj Kakade and Muskaan Rajput)welcoming and introducing the guest speaker, followed by Murumkar Sir stating the Agenda of the workshop. He began the session by providing everyone the link for downloading VMWare Work Pro 15 and guiding everyone through the process. He further gave a brief about how to create a virtual session.

Further, he gave time to download the software. Later distributed systems and how Google File Systems differ from Hadoop Distributed File System (HDFS) was explained by Viresh Sir. He also talked about the challenges faced while using Google File Systems and how they can be solved using HDFS. He further gave a real-life example explaining HDFS and telling about the master and slave services included in it. He added on by explaining how structured data can be integrated using HBase, Hive, and Hadoop. Installation of VMware was successfully completed by all students, and run time issues of all students were meticulously solved. The session concluded with Murumkar Sir asking a student to share his work and solving all possible doubts, most of the doubts circulated about the installation of VMware Work Pro 15, and basic concepts of virtual machines.

During the Question-Answer session, several students asked questions and all of these were resolved to their satisfaction by Sir. The webinar concluded with a Vote of Thanks from the student representatives.

The workshop was quite exploratory and informative as Hadoop and Big Data is a very trending topic today and is used for several projects undertaken by students. Students expected the workshop to last longer. Murumkar Sir was very helpful while solving the queries, also he encouraged students to contact him if they need any help in the future. Overall all participants hope for more similar sessions in the future.

The workshop was a beginner's guide towards Hadoop and virtual machines. With the help of this hands on session, many students could actually get started with using virtual machines and thus save in the memory space, they could also start working on Hadoop and enhance or broaden their knowledge in the Big Data field. Overall the students were greatly benefitted with this hands on Workshop.

Ubuntu - VMware Workstation

File View VM Tabs Help Zoom Leave

Home Ubuntu 11:21 PM

saurabh@ubuntu: ~/Downloads

```
saurabh@ubuntu:~$ pwd
/home/saurabh
saurabh@ubuntu:~$ cd Downloads/
saurabh@ubuntu:~/Downloads$ ls
apache-hive-2.3.2-bin          HBASE_HIVE_HADOOP4.mp4
db-derby-10.13.1.1-bin       HBASE_HIVE_HADOOP.docx
eclipse-installer            hdata
eclipse-inst-linux32.tar.gz  input
g                             input~
g.txt                        kbiet~
g.txt~                      kbiet123.txt
hadoop-3.0.0                 New folder
hadoop-3.0.0.tar.gz         New folder.rar
hadoop-file-for-configuration part-00000
hadoop-mapreduce-example-file.txt Untitled Document~
hbase-1.2.0                  whats_app_chat_data.txt
HBASE_HIVE_HADOOP(1).docx    whatsapp.txt
HBASE_HIVE_HADOOP1.mp4      youtubeanalysis
HBASE_HIVE_HADOOP2.mp4     zookeeper-3.4.10
HBASE_HIVE_HADOOP3.mp4
saurabh@ubuntu:~/Downloads$ cd
```

Recording

Unmute Start Video Share Participants More 11:51 04-09-2020

Activate Windows Go to Settings to activate Windows. 1 Downloads

HDFS (Recovered) - Paint

File Home View

Clipboard Image Tools Shapes Colors Edit colors Edit with Paint 3D

demons

NAMENODE
 S.NAME NODE
 JOB TRACKER

master services

DATA NODE
 TASK TRACKER

slave services

namenode 100mb 20mb

64mb	36mb	20mb		
DN	DN	DN		

HADOOP DISTRIBUTED FILE SYSTEM

13kb
 4kb 1kb
 4kb 4kb
 4kb

GOOGLE FILE SYSTEM

- BLOCK SIZE - 64mb /128mb
- REPLICATION
- COMMODITY HARDWARE

Datanode - commodity hardware

540, 177px 1360 x 570px size: 2. MB 100% 12:09 04-09-2020

viresh chapte's screen

12:55 PM

...     4G VoLTE 

Close

Chat

From Kedar Shinde to Everyone

Installed everything successfully

From chinmay ambolikar to Everyone

It worked sir!

From Shivam Agrawal to Everyone

Yes

From Aditi Chavan to Everyone

yes

From Rutuja Bhandigani to Everyone

yes

From Saloni Oswal to Everyone

Send to: **Everyone** 

Your text can be seen by everyone

Send

11:58 AM

...     4G VoLTE 

Close

Participants (65)

S suja panicker (me)   >

 **Yogesh Murumkar** (host)    >

A anuj (co-host)   >

M Muskaan (co-host)   >

PK Pranali Kosamkar (co-host)   >

VC viresh chapte (co-host)    >

AC Aditi Chavan   >

AG Aditi Gupta   >

AW Aditya Wanjari   >

 Aishwarya Bhure   >

Invite