

Report
On
Cloudera Academic Partnership (CAP)

Cloudera – VESIT Collaboration for “**Developing with Spark and Hadoop**” training was conducted from 17th September 2019 to 1st February 2020

The schedule of the course is attached. The course span was 36 hrs including theory and lab session which was conducted by following faculty members

1. Dr. M. Vijayalakshmi (Department of Information Technology)
2. Mrs. Asha Bharambe (Department of Information Technology)
3. Mrs. Jayshree Hajgude (Department of Information Technology)
4. Mrs. Sangeeta Oswal (Department of MCA)
5. Mrs. Sujata Khedkar (Computer Engineering Department)

The course included topics on basics of Hadoop, its ecosystem, and components and MapReduce. The lab was conducted on the Virtual machine provided by Cloudera.

The course was attended by Third year students of Computers, IT and Second year MCA. The list of the students is attached.



A handwritten signature in blue ink, consisting of a stylized 'M' followed by a long horizontal line.

Sr.No.	Computers	IT	MCA
1	Yogesh Tekwani	Aakash Barhate	Sneha Hirnaik
2	Sahil T	Jay Bendre	Samiksha Dhavale
3	Jayesh Samtani	Shubham Darekar	Pranav More
4	Avinash Bhawnani	Hrishikesh Deshpande	Jaspreet Kaur
5	Ekta Kithani	Ayesha Gulrajani	Swati Mishra
6	Divya Kalwani	Mahesh Keswani	Abhishek Shukla
7	Rishabh Sah	Vignesh Pillai	Vishal Bhandari
8	Mahadevan K	Atharva Potdar	Raj Shetye
9	Jaya Tanwani	Mitali Ranawat	Shivam Kumari
10	Aishwarya Sahoo	Tejas Shinde	Suraj Singh
11	Nikita Achhra	Harshita Singh	Ritesh Yadav
12	Adithya Shrivastava		Aakash Gupta
13	Sujitkumar Singh		Divya Pokhriyal
14	Tina Chandwani	-	Medhank Pathare
15	Pratiksha Wadibhasme	-	Nikhil Bhoite
16	Anjali Amin	-	Bhavesch Mandavkar
17	Gayatri Patil	-	
18	Shreyas Talole	-	
19	Yash Goyal	-	-
20	Hrushabhsingh Chouhan	-	-
21	Aditya Raute	-	-
22	Abhay Tiwari	-	-
23	Raj Chavan	-	-
24	Bhavesch Chatnani	-	-
25	Sahil Talreja	-	-
26	Kanchan Shownkeen	-	-

“Developing with Spark and Hadoop”

ODD SEM

2019-2020

Schedule: Timings: 8:10 to 10:20 (each day)

Resource Persons:

1. Dr. Mrs. M. Vijayalakshmi
2. Mrs. Sujata Khedkar
3. Mrs. Sangeeta Oswal
4. Mrs. Asha Bharambe
5. Mrs Jayshree Hajgude

Sr No	Agenda	Task	Dates
1	Introduction Motivations for Hadoop Hadoop Ecosystem	Session 1	17/09/2019 (AB)
2	Hadoop Architecture and HDFS	Session 2	18/09/2019 (AB)
3	Hadoop Architecture Continued (Yarn)	Session 3	19/09/2019 (AB)
4	Introduction to Impala & Hive	Session 4	20/09/2019 (SKH)
5	Sqoop	Session 5	21/09/2019 (JH)
6	Modeling and Managing Data using Hive and Impala	Session 6	23/09/2019 (SKH)
7	Data Format	Session 7	24/09//2019 (SO)
8	Data File Partitioning	Session 8	25/09/2019 (SO)
9	Capturing Data with Apache Flume	Session 9	26/09/2019

		(JH)	
--	--	------	--

Sr No	Agenda	Task	Dates
1	Distributed Data Processing with Spark Spark Basics	Session 10	20/01/2020 (SO) Mon
2	Working with RDDs in Spark	Session 11	21/01/2020 (AB) Tue
3	Aggregating Data with Pair RDDs	Session 12	22/01/2020 (AB) Wed
4	Writing and Deploying Spark Applications	Session 13	23/01/2020 (SO) Thur
5	Parallel Data Processing in Spark	Session 14	24/01/2020 (JH) Fri
6	Spark RDD Persistence	Session 15	25/01/2020 (SKH) Sat
7	Common Pattern in spark Data Processing	Session 16	27/01/2020 (JH) Mon
8	Spark SQL and Data Frame	Session 17	01/02/2020 (SKH) Sat

Students demonstrated the application of techniques learned through a project on following topics.

1. Credit Card Risk Analysis
2. Facebook Social Circle
3. Chicago Restaurant Recommendations

Project evaluation is scheduled on 7th March 2020.

The final exam was conducted on 8th Feb 2020.



**Vivekanand Education Society's
Institute Of Technology
Cloudera Academic Training**

for "Developing with Spark and Hadoop" Academic Year - 2019-20

Class: D12B

Department : CMPN

List of students

Attendance sheet

Sr.No.	Name	Date							
		20/01/2020	21/01/2020	22/01/2020	23/01/2020	24/01/2020	25/01/2020	27/01/2020	1/2/2020
1	Adithya Shrivastava	AB	AB	AB	AB	AB	AB	AB	
2	Pratiksha Wadibhasme	Pratiksha	Pratiksha	Pratiksha	Pratiksha	AB	AB	AB	
3	Gayatri Patil	AB	AB	AB	AB	AB	AB	AB	
4	Anjali amin	Anjali	AB	AB	Anjali	AB	AB	AB	
5	Shreyas Talole	Shreyas	Shreyas	Shreyas	Shreyas	Shreyas	Shreyas	Shreyas	
6	Sujitkumar singh	AB	AB	AB	AB	AB	AB	AB	
7	Tina Chandwani	AB	AB	AB	AB	AB	AB	AB	

Department : CMPN

List of students

Attendance sheet

Class: D12C

Sr.No.	Name	Date							
		20/01/2020	21/01/2020	22/01/2020	23/01/2020	24/01/2020	25/01/2020	27/01/2020	1/2/2020
1	aditya Raute	Raute	Raute	Raute	Raute	Raute	Raute	Raute	
2	Abhay Tiwari	Abhay	Abhay	Abhay	Abhay	Abhay	Abhay	Abhay	
3	Sahil Talreja	AB	Sahil	Sahil	AB	AB	AB	AB	
4	Yash Goyal	Yash	Yash	Yash	Yash	Yash	Yash	Yash	
5	Hrushabhsingh Chouhan	Hrushabhsingh	Hrushabhsingh	Hrushabhsingh	Hrushabhsingh	Hrushabhsingh	Hrushabhsingh	Hrushabhsingh	
6	Bhavesh Chatnani	Bhavesh	Bhavesh	Bhavesh	Bhavesh	AB	AB	AB	
7	Kanchan Shownkeen	Kanchan	Kanchan	Kanchan	Kanchan	AB	AB	AB	
8	Raj Chavan	Raj	Raj	Raj	Raj	AB	AB	AB	

**Vivekanand Education Society's
Institute Of Technology
Cloudera Academic Training**

for "Developing with Spark and Hadoop" Academic Year - 2019-20
Attendance sheet

Class: MCA2A

Department : MCA

List of students

Date

Sr.No.	Name	Date							
		20/01/2020	21/01/2020	22/01/2020	23/01/2020	24/01/2020	25/01/2020	27/01/2020	1/2/2020
1	Jaspreet Kaur	AB	AB	AB	AB	AB	AB	AB	
2	Abhishek shukla	AB	ASNAKAR	ASNAKAR	ASNAKAR	ASNAKAR	ASNAKAR	ASNAKAR	
3	Swati Mishra	AB	Swati	Swati	Swati	Swati	Swati	Swati	
4	Vishal Bhandari	Vishal	Vishal	Vishal	Vishal	Vishal	Vishal	Vishal	
5	Pranav More	PP.MORE	PP.MORE	PP.MORE	PP.MORE	PP.MORE	PP.MORE	PP.MORE	
6	Ritesh Yadav	AB	AB	AB	AB	AB	AB	AB	
7	Samiksha Shavale	AB	AB	AB	AB	AB	AB	AB	
8	Sneha Hirnaik	AB	AB	AB	AB	AB	AB	AB	

Department : MCA

List of students

Attendance sheet

Class: MCA2B

Sr.No.	Name	Date							
		20/01/2020	21/01/2020	22/01/2020	23/01/2020	24/01/2020	25/01/2020	27/01/2020	1/2/2020
1	Aakash Gupta	AB	AB	AB	AB	AB	AB	AB	
2	Divya Pokhriyal	AB	AB	AB	AB	AB	AB	AB	
3	Medhank Pathare	AB	AB	AB	AB	AB	AB	AB	
4	Bhavesh Mandavkar	AB	AB	AB	AB	AB	AB	AB	
5	Raj Shetye	AB	RShetye	RShetye	RShetye	RShetye	RShetye	RShetye	
6	Suraj Singh	AB	AB	AB	AB	AB	AB	AB	
7	Nikhil Bhoite	AB	AB	AB	AB	AB	AB	AB	
8	shivam Kumari	Shivam	Shivam	Shivam	Shivam	Shivam	Shivam	Shivam	

Vivekanand Education Society's
Institute Of Technology
Cloudera Academic Training

for "Developing with Spark and Hadoop" Academic Year - 2019-20
 Attendance sheet

Class: D12A

Department : CMPN

List of students

Sr.No.	Name	Date							1/2/2020
		20/01/2020	21/01/2020	22/01/2020	23/01/2020	24/01/2020	25/01/2020	27/01/2020	
1	Sahil Talreja	AB	AB	AB	AB	AB	AB	AB	
2	Jayesh Samtani	AB	Samtani	AB	AB	AB	AB	AB	
3	Avinash Bhawnani	AB	Bhawnani	AB	AB	AB	AB	AB	
4	Ekta Kithani	AB	AB	AB	AB	AB	AB	AB	
5	Yogesh Tekwani	AB	AB	AB	AB	AB	AB	AB	
6	Divya Kalwani	AB	AB	AB	AB	AB	AB	AB	
7	Rishbha Sah	AB	AB	AB	AB	AB	AB	AB	
8	M. Krish Venkateshwaran	AB	AB	AB	AB	AB	AB	AB	
9	Jaya Tanwani	AB	AB	AB	AB	AB	AB	AB	
10	Aishwarya Sahoo	AB	AB	AB	AB	AB	AB	AB	
11	Nikita Achhra	AB	AB	AB	AB	AB	AB	AB	

**Vivekanand Education Society's
Institute Of Technology
Cloudera Academic Training**

for "Developing with Spark and Hadoop" Academic Year - 2019-20

Department : INFT

List of students

Attendance sheet

Class: D15

Sr.No.	Name	Date							
		20/01/2020	21/01/2020	22/01/2020	23/01/2020	24/01/2020	25/01/2020	27/01/2020	1/2/2020
1	Aakash Barhate	AB	AB	AB	AB	AB	AB	AB	
2	Jay Bendre	AB	AB	AB	AB	AB	AB	AB	
3	Shubham Darekar	AB	AB	AB	AB	AB	AB	AB	
4	Hrishikesh Deshpande	AB	AB	AB	AB	AB	AB	AB	
5	Ayesh Gulrajani	AB	AB	AB	AB	AB	AB	AB	
6	Mahesh Keswani	AB	AB	AB	AB	AB	AB	AB	
7	Vignesh Pillai	AB	AB	AB	AB	AB	AB	AB	
8	Atharva Potdar	AB	AB	AB	AB	AB	AB	AB	
9	Mitali Ranawat	Mitali	Mitali	AB	Mitali	Mitali	AB	Mitali	
10	Tejas shinde	AB	AB	AB	AB	AB	AB	AB	
11	Harshita Singh	AB	AB	AB	AB	AB	AB	AB	