

**Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY**

2017-2018

**Report
On
Cloudera Academic Partnership (CAP)**

Cloudera – VESIT Collaboration for “**Developing with Spark and Hadoop**” training was conducted from 6th September 2017 to 19th January 2018.

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. Sangeeta Oswal (Department of MCA)
4. 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 final year students of Computers, IT and Second MCA. The list of the students is attached.



A handwritten signature in blue ink, consisting of a stylized 'M' followed by a long horizontal stroke that curves upwards at the end.

**Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY**

2017-2018

List of Students

Sr.No	Computers	IT	MCA
1	Nikil Vatwani	Radhika Raghuvanshi	Rohini Rangnekar
2	Vijay Kataria	Rohini Yedelli	Kapil Trivedi
3	Kajal Khatwani	Vaibhav Goyal	Anjali Nauriyal
4	Sanjana Mulchandani	Rajeevkumar Yadav	Sharan Iyengar
5	Juilee Kulkarni	Raj Ramchandani	Vaidika Narkar
6	Pooja More	Shubham Ghosalkar	Ujwal Suvarna
7	Yohanth Shetty	Pooja V. Singh	Komal Shah
8	Lalit R. Pathak	Anjali Wani	Loveena Bansal
9	Manish Chandwani	Abhinav Valecha	Hiral Khavanekar
10	Ritika Mendjoge	Punit Kulal	Pawandeep Singh
11	Mayuri Kate	Anjali Jaisinghani	Shubham Kamalapur
12	Mugdha Govilkar	Sonia Chhabria	Ajay Nair
13	Tejas Ingole	Ankit Kesharmani	Priyanka Devasthani
14	Anand Vaswani	Rahul Shetty	Vikash Gupta
15	Sameer Udgirkar	-	Mahesh Bhosale
16	Siddham Sharma	-	-
17	Deepika Khithani	-	-
18	Juli Bhagtani	-	-
19	Bhavik Dand	-	-
20	Shruti Zade	-	-
21	Padmja Kolle	-	-
22	Akshya Chidrawar	-	-
23	Chetan Gupta	-	-
24	Piyush K	-	-

**Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY**

2017-2018

**“Developing with Spark and Hadoop”
ODD SEM**

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

Resource Person	Agenda	Task	Dates
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Introduction Motivations for Hadoop Hadoop Ecosystem	Session 1	06/09/2017 (AB) (Wednesday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata khedkar Mrs. Sangeeta Oswal	Hadoop Architecture and HDFS & Yarn	Session 2	07/09/2017 (AB) (Thursday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Sqoop	Session 3	09/09/2017 (SO) (Saturday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Introduction to Impala & Hive	Session 4	13/09/2017 (SO) (Wednesday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Modeling and Managing Data using Hive and Impala	Session 5	16/09/2017 (SO) (Saturday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Data Format	Session 6	19/09//2017 (SKH) (Tuesday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Data File Partitioning	Session 7	20/09/2017 (SKH) (Friday)
Mid Term Break			

**Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY**

2017-2018

EVEN SEM

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

Resource Person	Agenda	Task	Dates
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Revision and Capturing Data with Apache Flume	Session 9	10/01/2018 (SKH) (Wednesday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata khedkar Mrs. Sangeeta Oswal	Distributed Data Processing with Spark Spark Basics	Session 10	11/01/2018 (SO) (Thursday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Working with RDDs in Spark	Session 11	12/01/2018 (AB) (Friday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Aggregating Data with Pair RDDs	Session 12	13/01/2018 (AB) (Saturday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Writing and Deploying Spark Applications	Session 13	15/01/2018 (SO) (Monday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Parallel Data Processing in Spark	Session 14	16/01/2018 (AB) (Tuesday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Spark RDD Persistence	Session 15	17/01/2018 (SKH) (Wednesday)

**Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY**

2017-2018

Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Common Pattern in spark Data Processing	Session 16	18/01/2018 (SO) (Thursday)
Dr. Mrs. M. Vijayalakshmi Mrs. Asha Bharambe Mrs. Sujata Khedkar Mrs. Sangeeta Oswal	Sprak SQL and Data Frame	Session 17	19/01/2018 (SKH) (Friday)

The final exam was schedule on 2nd Feb 2018 and the project Presentation on 14th March 2018 from 8:15 am to 10:15 am @ 505. The project assigned was a kaggle problem statement on **“Toxic Comment Classification Challenge”**



**Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY**

2017-2018



Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY

Cloudera Academic Training
for "Developing with Spark and Hadoop" Academic Year : 2017-18

Department :- Information Technology

List of student

Class:- D20

Attendance Sheet

Total Attend	Sr.No	Name	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date		
			06/09/2017	07/09/2017	08/09/2017	09/09/2017	10/09/2017	11/09/2017	12/09/2017	13/09/2017	14/09/2017	15/09/2017	16/09/2017	17/09/2017	18/09/2017	23/09/2017
11	1	Radhika Raghurwanshi	Radhika	Radhika	Radhika	Radhika	AB	Radhika	Radhika	Radhika	Radhika	Radhika	Radhika	Radhika	Radhika	
10	2	Rohini Yedelli	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	
07	3	Vaibhav Goyal	V	V	V	V	AB	V	V	V	V	A	A	A	A	
13	4	Rajeevkumar Yadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	Ryadav	
13	5	Raj Ramchandani	R	R	R	R	R	R	R	R	R	R	R	R	R	
07	6	Shubham Ghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	Sghosalkar	
06	7	Pooja V. Singh	Pooja	Pooja	Pooja	Pooja	AB	AB	AB	AB	AB	Pooja	Pooja	A	A	
07	8	Anjali Wani	Anjali	Anjali	Anjali	Anjali	AB	AB	AB	AB	Anjali	Anjali	A	A	A	
09	9	Abhinav Valecha	Abhinav	Abhinav	Abhinav	AB	Abhinav	Abhinav	AB	AB	Abhinav	Abhinav	Abhinav	A	Abhinav	
07	10	Punit Kulkarni	Punit	Punit	Punit	Punit	AB	Punit	Punit	AB	Punit	A	A	A	A	
02	11	Anjali Jaisinghani	A	A	A	Anjali	AB	AB	Anjali	AB	A	A	A	A	A	
10	12	SONIA CHHABRA	A	Sonia	Sonia	A	Sonia	Sonia	Sonia	Sonia	Sonia	Sonia	Sonia	Sonia	A	Sonia
08	13	Ankit Kasharneni	A	Ankit	Ankit	Ankit	AB	AB	AB	Ankit	Ankit	Ankit	Ankit	AB	Ankit	
07	14	Rohit Shetty	A	Rohit	Rohit	Rohit	AB	AB	AB	A	Rohit	Rohit	Rohit	Rohit	Rohit	
02	15	Rajeev Yadav	Ryadav	Ryadav	A	AB	AB	AB	AB	AB	A	A	A	A	A	

INSTITUTE OF TECHNOLOGY

Cloudera Academic Training

for "Developing with Spark and Hadoop" Academic Year : 2017-18

Department :- Computer Engineering

List of student

Class:- D17

Attendance Sheet

Total Attendance	Sr.No	Name	Date												
			06/09/2017	07/09/2017	09/09/2017	10/09/2017	11/09/2017	12/09/2017	13/09/2017	14/09/2017	15/09/2017	16/09/2017	17/09/2017	18/09/2017	19/09/2017
03	1	Nikil Vatwani	Nikil	Nikil	A	Nikil	A	A	A	A	A	A	A	A	A
05	2	Vijay Kataria	Vijay	Vijay	Vijay	Vijay	Vijay	A	A	A	A	A	A	A	A
04	3	Kajal Khatwani	Kajal	Kajal	Kajal	Kajal	A	A	A	A	A	A	A	A	A
04	4	Sanjana Mulchandani	Sanjana	Sanjana	Sanjana	Sanjana	A	A	A	A	A	A	A	A	A
12	5	Juilee Kulkarni	Juilee	Juilee	Juilee	Juilee	Juilee	A	Juilee	Juilee	Juilee	Juilee	Juilee	Juilee	Juilee
11	6	Pooja More	Pooja	A	Pooja	Pooja	Pooja	Pooja	Pooja	Pooja	Pooja	Pooja	Pooja	Pooja	Pooja
10	7	Yohanth Shetry	Yohanth	Yohanth	Yohanth	A	A	A	Yohanth	Yohanth	Yohanth	Yohanth	Yohanth	Yohanth	Yohanth
03	8	Lalit R. Pathak	Lalit	Lalit	Lalit	A	A	A	A	A	A	A	A	A	A
03	9	Manish Chandwani	Manish	Manish	Manish	A	A	A	A	A	A	A	A	A	A
00	10	Ritika Mendjoge	A	A	A	A	A	A	A	A	A	A	A	A	A
03	11	Mayuri Kate	Kate	Kate	Kate	A	A	A	A	A	A	A	A	A	A
02	12	Mugdha Govilkar	Mugdha	Mugdha	A	A	A	A	A	A	A	A	A	A	A
06	13	Tejus Ingole	A	Tejus	A	A	A	A	Tejus	Tejus	Tejus	Tejus		A	Tejus
02	14	Anand Vashwanth	A	Anand	Anand	A	A	A	A	A	A	A	A	A	A
01	15	Sameer Ujirkar	A	Sameer	A	A	A	A	A	A	A	A	A	A	A
01	16	Sudharn Sharma	A	Sudharn	A	A	A	A	A	A	A	A	A	A	A

Vivekanand Education Society's
INSTITUTE OF TECHNOLOGY

Cloudera Academic Training
for "Developing with Spark and Hadoop" Academic Year : 2017-18

Sr.No	Name	List of student															23/11
		Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	Date	
11	Rohini Rangnekar	Rohini	Rohini	A		Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	Rohini	A
10	Kapil Trivedi	Kapil	Kapil	A		Kapil	A	A	A	A	A	A	A	A	A	A	A
03	Anjali Nauriyal	Anjali	Anjali	A		Anjali	A	A	A	A	A	A	A	A	A	A	A
12	Sharan Iyengar	Sharan	Sharan	A		Sharan	Sharan	Sharan	Sharan	Sharan	Sharan	Sharan	Sharan	Sharan	Sharan	Sharan	A
05	Vaidika Narkar	Vaidika	Vaidika	A		Vaidika	A	Vaidika	A	A	A	A	A	A	A	A	A
12	Ujwal Suvarna	Ujwal	Ujwal	A		Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	Ujwal	A
12	Komal Shah	Komal	Komal	A		Komal	Komal	Komal	Komal	Komal	Komal	Komal	Komal	Komal	Komal	Komal	A
07	Loveena Bansal	Loveena	Loveena	A		Loveena	A	A	A	A	A	A	A	A	A	A	A
03	Hiral Khavanekar	Hiral	Hiral	A		Hiral	A	A	A	A	A	A	A	A	A	A	A
02	Pawandeep Singh	Pawandeep	Pawandeep	A		Pawandeep	A	A	A	A	A	A	A	A	A	A	A
04	Shubham Kamalpure	Shubham	Shubham	A		Shubham	A	A	A	A	A	A	A	A	A	A	A
12	Ajay Nair	Ajay	Ajay	A		Ajay	Ajay	Ajay	Ajay	Ajay	Ajay	Ajay	Ajay	Ajay	Ajay	Ajay	Ajay
10	Priyanka Devasthali	Priyanka	Priyanka	A		Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	Priyanka	A
12	Vikas Gupta	Vikas	Vikas	A		Vikas	Vikas	Vikas	Vikas	Vikas	Vikas	Vikas	Vikas	Vikas	Vikas	Vikas	A
11	Mahesh Bhojale	Mahesh	Mahesh	A		Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	Mahesh	A