



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai
NAAC Accredited Institution

"Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.

ANNA UNIVERSITY COUNSELLING CODE : 1101



DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

Dated on: 10/12/2019

CIRCULAR

FACULTY MEMBERS AND STUDENTS

We have planned to conduct Value Added course in our department on the topic "BIG DATA ENGINEERING USING HADOOP & SPARK" from 11/12/2019 to 14/12/2019 therefore, I request all the Faculty members and students to attend the course and make utilization of the course for your future.

Resource Person:


BALAMURGAN S

STANFORD INDIA BIODESIGN

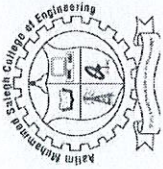
FELLOW SCHOOL OF INTERNATIONAL BIODESIGN

AIIMS NEW DELHI




10/12/2019
Prof. Dr. A.S.SALMA BANU
HoD/ECE


PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING
 Approved by AICTE and Affiliated to Anna University Chennai
NAAC ACCREDITED INSTITUTION.

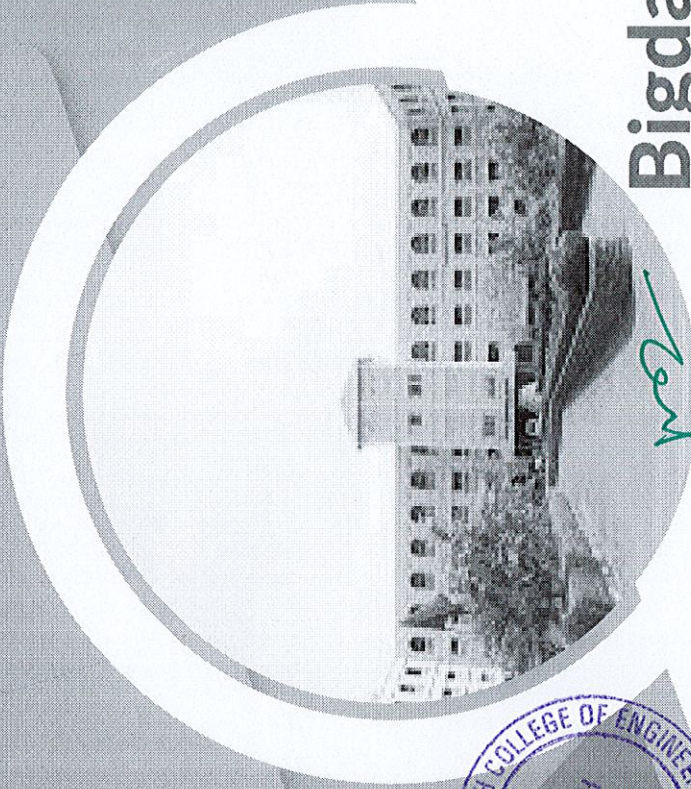
Hadoop is a high latency computing framework, which does not have an interactive mode. Spark is a low latency computing and can process data interactively



11th Dec to 14th Dec 2019

**VALUE
 ADDED
 COURSE**

Dr. A.S. Salma Banu
 Head of the Department



Bigdata Engineering using Hadoop & Spark

PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING



RESOURCE PERSON

- Er. BALAMURUGAN. S
 Stanford- India
 Bio-design fellow
 School of International Bio-design, AIIMS
 New Delhi
-

Dr. S. Sathish
 Principal

Ally S. Segu Jamaludeen
 Secretary & Correspondent



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai
NAAC Accredited Institution

"Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.

ANNA UNIVERSITY COUNSELLING CODE : 1101



DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING VALUE-ADDED COURSE-SUMMARY REPORT

COURSE NAME: BIGDATA ENGINEERING USING HADOOP & SPARK

Date: 11/11/2019 to 14/11/2019

Time: Morning Session: 9.00 am to 1.00 pm

Afternoon Session: 1.30 pm & 5:00 pm

No. of Days: 4

No. of Modules: 4

Course Hours: 30

Mode of Conduction: Offline

Speaker: S. BALAMURGAN

Attended Students: 64

Objective of the Course:

- This course aims to provide a comprehensive understanding of foundational Big Data technologies, with a focus on Apache Spark. Students will learn the fundamental concepts of Big Data, the architecture and management of HDFS and Zookeeper, and how to develop, deploy, and optimize Spark applications.
- By the end of the course, students will be able to apply these technologies to solve real-world data processing challenges.

Summary of the course:

- This course introduces students to the core components of the Big Data ecosystem, focusing on Apache Spark and its related technologies.
- The course begins with an overview of Big Data concepts and the environment, including HDFS (Hadoop Distributed File System) and Zookeeper.
- Students will then explore the basics of Apache Spark, including its architecture, core functionalities, and performance considerations. The course will cover writing and deploying Spark applications and common patterns used in Spark data processing.
- Practical labs and projects will enable students to apply their knowledge to real-world scenarios, enhancing their ability to handle and analyze large-scale datasets effectively.

Course Outcomes:

By the end of the course, students will be able to:

1. Understand Big Data Fundamentals:

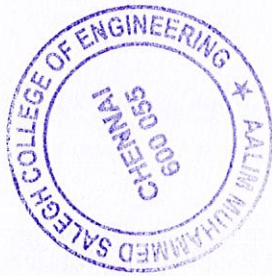
- Define the key characteristics and challenges of Big Data.


M02

PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING



- Describe the components and architecture of the Big Data ecosystem, including Hadoop and Spark.
- 2. **Manage Big Data Environments:**
 - Configure and manage HDFS for scalable storage and data processing.
 - Explain the role of Zookeeper in distributed systems and set up a Zookeeper ensemble.
- 3. **Work with Apache Spark:**
 - Describe Spark's architecture and components.
 - Implement basic operations using Spark RDDs and DataFrames.
 - Utilize Spark SQL for querying structured data.
 - Apply performance tuning and optimization techniques for Spark applications.
- 4. **Develop and Deploy Spark Applications:**
 - Write and test Spark applications using Scala, Python, or Java.
 - Deploy Spark applications in various environments (local, standalone, YARN, Kubernetes).
 - Manage and configure Spark clusters for optimal performance.




PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING
4VADI - IAF, MUTHAPUDUPE
CHENNAI 600 055

Place: Chennai
Date: 02 /12/2019

From,

Head of the Department
Department of ECE
Aalim Muhammed Salegh College of Engineering
Avadi IAF, Muthapudupet
Chennai - 600055

To,

The Principal
Aalim Muhammed Salegh College of Engineering
Avadi IAF, Muthapudupet
Chennai - 600055

Respected Sir,

Sub: Requisition for the approval to conduct Value Added course in our ECE department – Reg.

We wish to conduct Value Added course in our ECE department on the topic
"BIG DATA ENGINEERING USING HADOOP & SPARK" a Guest Lecture Series From 11/12/2019 to
14/12/2019.

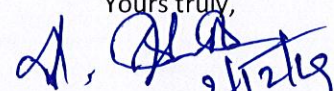
Resource Person:

S.BALAMURUGAN
STANFORD INDIA BIODESIGN
FELLOW SCHOOL OF INTERNATIONAL BIODESIGN
AIIMS NEW DELHI

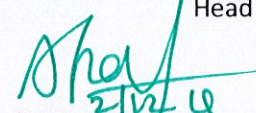
The objective of conducting this Value Added course will give the industry focus in developing skills required by using latest technology advancement in "BIG DATA ENGINEERING USING HADOOP & SPARK". This effort will benefit our Student's community and bring core strength in engineering field.

Thank you,

Yours truly,


Professor Dr. A. S. SALMABANU
Head of the Department/ECE



Permitted

PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai
NAAC Accredited Institution

"Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.

ANNA UNIVERSITY COUNSELLING CODE : 1101



DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING CERTIFICATE COURSE-BIGDATA ENGINEERING USING HADOOP & SPARK

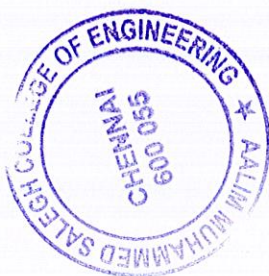
COURSE CONTENT:


Module No.	CONTENT	DURATION(Hrs)
Module 1	<p>Introduction to Big Data</p> <ul style="list-style-type: none">• Introduction to Big Data:<ul style="list-style-type: none">• Definition and characteristics of Big Data (Volume, Velocity, Variety, Veracity, Value)• The importance and impact of Big Data in various industries	7 hours 30 minutes
Module 2	<p>Foundation & Environment</p> <p>Big Data Ecosystem:</p> <ul style="list-style-type: none">• Overview of key components (Hadoop, Spark, HBase, etc.)• Understanding the role of distributed computing <p>Zookeeper, HDFS</p> <ul style="list-style-type: none">• Introduction to Zookeeper and its role in distributed systems• Basic concepts (nodes, znodes, sessions)• Zookeeper architecture and use cases• Setting up and managing a Zookeeper ensemble <p>Hadoop Distributed File System (HDFS):</p> <ul style="list-style-type: none">• Architecture of HDFS (NameNode, DataNodes, Secondary NameNode)• HDFS file operations (read, write, replication)• HDFS configuration and management	7 hours 30 minutes
Module 3	<p>Scale Basics, Spark Basics</p> <p>Introduction to Apache Spark:</p> <ul style="list-style-type: none">• Overview of Spark and its advantages over traditional Big Data processing tools• Key components of Spark (Spark Core, Spark SQL, Spark Streaming, MLlib, GraphX) <ul style="list-style-type: none">• Spark Architecture:<ul style="list-style-type: none">• Spark's architecture (Driver, Executors, Cluster Manager)	7 hours 30 minutes

PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING

	<ul style="list-style-type: none"> Understanding Resilient Distributed Datasets (RDDs) and DataFrames Basic Operations in Spark: <ul style="list-style-type: none"> Creating and manipulating RDDs (transformations and actions) Introduction to DataFrames and Spark SQL Using Spark's built-in functions and libraries Performance and Scaling: <ul style="list-style-type: none"> Spark's performance considerations (in-memory computing, lazy evaluation) Tuning and optimizing Spark applications Understanding partitioning and parallelism 	
Module 4	<p>Writing and Deploying Spark Applications, Common Patterns in Spark Data Processing</p> <ul style="list-style-type: none"> Writing Spark Applications: <ul style="list-style-type: none"> Setting up the Spark development environment (using IDEs, configuring build tools like Maven or SBT) Writing basic Spark applications in Scala, Python, or Java Utilizing Spark's libraries (Spark SQL, MLlib, GraphX) in applications Deploying Spark Applications: <ul style="list-style-type: none"> Deployment options (local mode, standalone cluster mode, YARN, Kubernetes) Configuring and managing Spark clusters Deploying applications on cloud platforms (AWS, Azure, GCP) Common Patterns in Spark Data Processing: <ul style="list-style-type: none"> Data ingestion patterns (batch processing, streaming) Data transformation patterns (filtering, mapping, reducing) Common use cases (ETL processes, data aggregation, machine learning workflows) Best practices for error handling, logging, and debugging in Spark applications 	7 hours 30 minutes

TOTAL: 30 HRS




PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING
4VADI - IAF, MUTHAPUDUPE
CHENNAI 600 055



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai
NAAC Accredited Institution

"Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.

ANNA UNIVERSITY COUNSELLING CODE : 1101



Value Added Course Tentative Time Table And Schedule

Value Added Course Name: BIG DATA ENGINEERING USING HADOOP & SPARK

Date of Commencement: From 11/12/2019

Timing: Morning Session: 9.00 am to 1.00 pm

Afternoon Session: 1.30 pm & 5:00 pm

Tentative Date of Exam- 14/12/2019

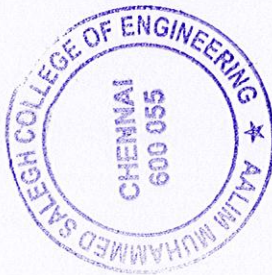
Tentative Course Schedule - 11/12/2019 to 14/12/2019 (As per following)


No. Of Days: 4

Course Hours: 30

Speaker: Mr.S.BALAMURUGAN

1. Module -1:Introduction to Big Data – 11-12-2019 – 7 hours 30 minutes
2. Module -2:Foundation & Environment– 12-12-2019 – 7 hours 30 minutes
3. Module -3: Introduction to Electric Circuits – 13-12-2019 – 7 hours 30 minutes
4. Module -4: Sensors, Actuators and Manipulators – 14-12-2019 – 7 hours 30 minutes




PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING
AVADI - IAF, MUTHAPUDUPE
CHENNAI 600 055



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai
NAAC Accredited Institution

"Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.

ANNA UNIVERSITY COUNSELLING CODE : 1101



Department of Electronics and Communication Engineering Value Added Courses: Big Data Engineering Using Hadoop and Spark

Year : IV

Sem : VII

Sec:A

Sl.No.	Roll.No	Name	11/12/19	12/12/19	13/12/19	14/12/19
1	110115106001	AAKIFAH JABEEN	Jabeen	Jabeen	Jabeen	Jabeen
2	110115106002	ABDUL BARI N	Bari	Bari	Bari	Bari
3	110115106003	ABDUL KADAR K M	Xade	Xade	Xade	Xade
4	110115106004	ABDUL KALAM AZAD J H	AZAD	AZAD	AZAD	AZAD
5	110115106005	ABDUL RAHIM R	Rahim	Rahim	Rahim	Rahim
6	110115106006	ABIVANDHANA B S	ABIV	ABIV	ABIV	ABIV
7	110115106007	AFRIDI V S	Afridi	Afridi	Afridi	Afridi
8	110115106008	AHAMED RASIM H	Rasim	Rasim	Rasim	Rasim
9	110115106009	AHAMED RASMI Z	Rasmi	Rasmi	Rasmi	Rasmi
10	110115106010	AJITH KUMAR M	Ajith	Ajith	Ajith	Ajith
11	110115106011	ALFIA FATHIMA D	Alfia	Alfia	Alfia	Alfia
12	110115106012	ALIYA SHEERIN K M	Aliya	Aliya	Aliya	Aliya
13	110115106013	ANBARASU E	Anbarasu	Anbarasu	Anbarasu	Anbarasu
14	110115106014	ASLAM FARHAN M	Aslam	Aslam	Aslam	Aslam
15	110115106015	ATHIF AHAMED J	Athif	Athif	Athif	Athif
16	110115106016	BALAJI P	Balaji	Balaji	Balaji	Balaji
17	110115106017	BALAMURUGAN G	Bala	Bala	Bala	Bala
18	110115106018	BALUROBINSON	Balu	Balu	Balu	Balu
19	110115106020	FAHIM MALIM F	Fahim	Fahim	Fahim	Fahim
20	110115106021	FAIZAL HASAN S	Fazil	Fazil	Fazil	Fazil
21	110115106022	FAIZAN AHMED O	Faizan	Faizan	Faizan	Faizan
22	110115106023	FARHANA SHAHIN E	Shahin	Shahin	Shahin	Shahin
23	110115106024	FATHIMA BEEVI M	Fathima	Fathima	Fathima	Fathima
24	110115106026	FAYAZ KHAN F M	Fayaz	Fayaz	Fayaz	Fayaz
25	110115106027	GEETHA LAKSHMI V	Geetha	Geetha	Geetha	Geetha
26	110115106029	GULNAAZ FATHIMA J	Gulnaaz	Gulnaaz	Gulnaaz	Gulnaaz
27	110115106030	HAFIZUL RAHMAN R	Hafiz	Hafiz	Hafiz	Hafiz
28	110115106031	HAJIRA BEGUM M S	Hajira	Hajira	Hajira	Hajira
29	110115106032	HAJIRA FATHIMA A J	Hajira	Hajira	Hajira	Hajira
30	110115106033	HAMEEDA BANU I	Banu	Banu	Banu	Banu



PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING



AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

Approved by All India Council for Technical Education - New Delhi, Affiliated to Anna University, Chennai
NAAC Accredited Institution
"Nizara Educational Campus", Muthapudupet, Avadi - IAF, Chennai - 600 055.
ANNA UNIVERSITY COUNSELLING CODE : 1101



Department of Electronics and Communication Engineering

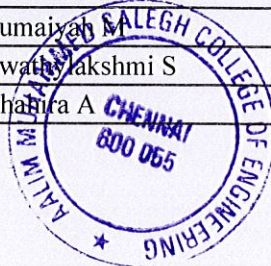
Value Added Courses: Big Data Engineering Using Hadoop and Spark

Year : IV

Sem : VII

Sec: B

Sl.No.	Roll.No	Name	11/12/19	12/12/19	13/12/19	14/12/19
1	110115106048	Kavya pradeep	Kat	Kat	Kat	Kat
2	110115106049	Kuraisha shabana. P.A	She	She	She	She
3	110115106050	Madhubala S	She	She	She	She
4	110115106051	Manikandan R	Mand	AB	Manikandan	Manikandan
5	110115106052	Md affan C	affan	affan	affan	affan
6	110115106053	Md arshath khan K.A	arshath	arshath	arshath	arshath
7	110115106054	Mohamed ejaz M	ejaz	ejaz	ejaz	AB
8	110115106055	Mohamed eliyas S	elijah	elijah	elijah	elijah
9	110115106056	Mohamed jaffer S.M	jaffer	jaffer	jaffer	jaffer
10	110115106057	Mohamed mujammil M	Mujammil	Mujammil	Mujammil	Mujammil
11	110115106058	Mohamed nainar zubair M A	Zubair	Zubair	AB	Zubair
12	110115106059	Mohamed rasif M	Rasif	Rasif	Rasif	Rasif
13	110115106060	Mohammed anas S	Anas	Anas	Anas	Anas
14	110115106061	Mohammed fazul rahman B	Rahman	Rahman	Rahman	Rahman
15	110115106062	Mohammed ibrahim A H	Ibrahim	Ibrahim	Mush	Mush
16	110115106063	Mohammed junaid	Junaid	Junaid	Junaid	Junaid
17	110115106064	Mohammed razeen S M	Razeen	Razeen	Razeen	Razeen
18	110115106065	Mohammed usama nadeem I	Nadeem	Nadeem	Nadeem	Nadeem
19	110115106066	Mohammed yasheer M	Yasheer	Yasheer	Yasheer	Yasheer
20	110115106067	Mookambigai S	Mookambigai	Mookambigai	Mookambigai	Mookambigai
21	110115106068	Nagaraj R	Nagaraj	Nagaraj	AB	Nagaraj
22	110115106069	Naseema parveen K	Naseema	Naseema	AB	Naseema
23	110115106070	Naveeth mohamed M B	Naveeth	AB	Naveeth	Naveeth
24	110115106071	Pavithra M	Pavithra	Pavithra	Pavithra	Pavithra
25	110115106072	Preethiga E	Preethiga	Preethiga	AB	Preethiga
26	110115106073	Raghavakumar G	Raghavakumar	Raghavakumar	Raghavakumar	Raghavakumar
27	110115106074	Rajesh R	Rajesh	Rajesh	Rajesh	Rajesh
28	110115106075	Raj kumar M	Rajkumar	AB	Rajkumar	Rajkumar
29	110115106076	Roja priyadarshini K	Roja	Roja	Roja	Roja
30	110115106078	Roshini R	Roshini	Roshini	Roshini	Roshini
31	110115106079	Ruksar M	Ruksar	Ruksar	Ruksar	Ruksar
32	110115106080	Sadikha fathima S	Sadikha	Sadikha	Sadikha	Sadikha
33	110115106081	Safiq noor mohamed V N M	AB	Safiq	Noor	Noor
34	110115106082	Salma begum N	Salma	Salma	Salma	Salma
35	110115106084	Santhanavignesh K	Santha	Santha	Santha	Santha
36	110115106085	Santhosh mohana prakash T	AB	Santhosh	Santhosh	Santhosh
37	110115106086	Shahin fathima M	Shahin	Shahin	Shahin	Shahin
38	110115106087	Sithi fathima S	Sithi	Sithi	Sithi	Sithi
39	110115106088	Siva saravanan S	Siva	Siva	Siva	Siva
40	110115106089	Sumaiyah	Sumaiyah	Sumaiyah	Sumaiyah	Sumaiyah
41	110115106090	Swathy lakshmi S	AB	Swathy	Swathy	Swathy
42	110115106091	Tharuna A	Tharuna	Tharuna	AB	Tharuna



PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING

C.ID.-9e39ce1



CERTIFICATE

Of Competition

This Certificate is Presented to: _____

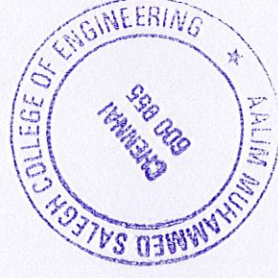
This Certificate is Presented to **ABDUL BARI N** has Successfully Completed 4 days Value added Course in "BIG DATA ENGINEERING HADOOP AND SPARK" from 11.12.2019 to 14.12.2019. We Appreciate your hard work and Sincerity towards your

Course and wish you good luck for your future endeavor.

Balamurgan

BALAMURGAN.S
TRAINER
Soft logic group

Eshilarasan



PRINCIPAL

EZHILARASANMICHAEL
AALIM MUHAMMAD SALEH COLLEGE OF ENGINEERING

CEO at Soft logic group

C.ID.-9e39ce2



CERTIFICATE

Of Competition

This Certificate is Presented to: _____

This Certificate is Presented to **ABDUL KADAR K M** has Successfully Completed 4 days Value added Course in “BIG DATA ENGINEERING HADOOOP AND SPARK” from 11.12.2019 to 14.12.2019. We Appreciate your hard work and Sincerity towards your

Course and wish you good luck for your future endeavor.

Balamurgan

BALAMURGAN.S
TRAINER

Softlogic group



Ezhilarasan
PRINCIPAL
AALIM MUHAMMED SALEGH
COLLEGE OF ENGINEERING

EZHILARASAN MICHAEL
CEO at Soft logic group

C.ID.-9e39ce3



CERTIFICATE

Of Competition

This Certificate is Presented to: _____

This Certificate is Presented to **ABDUL KALAM AZAD J H** has Successfully Completed 4 days Value added Course in "BIG DATA ENGINEERING HADOOP AND SPARK" from

11.12.2019 to 14.12.2019. We Appreciate your hard work and Sincerity towards your

Course and wish you good luck for your future endeavors.

BALAMURGAN.S
TRAINER
Softlogicgroup



PRINCIPAL
EZHILARASANMICHAEL
CEO at Soft logic group

C.ID.-9e39ce4

softlogic
BEST IN THE BUSINESS

CERTIFICATE

Of Competition

This Certificate is Presented to: _____

This Certificate is Presented to **ABDUL RAHIM R** has Successfully Completed 4 days Value added Course in "BIG DATA ENGINEERING HADOOP AND SPARK" from 11.12.2019 to 14.12.2019. We Appreciate your hard work and Sincerity towards your Course and wish you good luck for your future endeavor.

Balamurgan

BALAMURGAN.S
TRAINER
Softlogic group

Ezhilarasan



PRINCE EZHILARASAN MICHAEL
AALIM MUHAMMAD SALEGH
COLLEGE OF ENGINEERING
CEO at Soft logic group