

BIG DATA CURRICULUM

6 MONTH CERTIFICATE PROGRAMME



FOUNDATION



SPECIALIZATION



SUPER SPECIALIZATION



BIG DATA INTEGRATION &
PROCESSING



MACHINE LEARNING
WITH SPARK

Duration : 6 MONTHS



FOUNDATION

Week 1 Module 1: Introduction to Big Data

Module 2: Python Essentials - I

Week 2 Module 3: Python Essentials - II

Module 4: Introduction to Hadoop

Week 3 Module 5: Hadoop Architecture & Ecosystem

Module 6: Hadoop Distributed File System(HDFS)

Week 4 Module 7: Introduction to Map Reduce

Module 8: MapReduce Programming with Python

Week 5 Module 9: Apache Pig

Module 10: Apache Pig Hands - on

Week 6 Module 11: Apache Hive

Module 12: Apache Hive Hands - on

Duration : 6 Weeks



SPECIALIZATION

Week 7 Module 13: Programming with Scala-I

Module 14: Programming with Scala-II

Week 8 Module 15: Introduction to Apache Spark architecture

Module 16: RDD, transformations and actions

Week 9 Module 17: Spark Programming

Module 18: Spark SQL

Week 10 Module 19: Spark Streaming - I

Module 20: Spark Streaming - II

Week 11 Module 21: Machine Learning with Spark MLlib

Module 22: Live Practice session

Duration : 11 Weeks



BIG DATA CURRICULUM

SUPER SPECIALIZATION

- Week 12** Module 23: Introduction to NoSQL
 Module 24: Introduction to MongoDB
-
- Week 13** Module 25: Deep Dive into MongoDB
 Module 26: HBase
-
- Week 14** Module 27: SQOOP
 Module 28: FLUME

Duration : 14 Weeks



BIG DATA INTEGRATION & PROCESSING

- Week 15** Module 29: Data lakes
 Module 30: Processing big data
-
- Week 16** Module 31: Integration of Hadoop with MongoDB
 Module 32: Integration of Spark with MongoDB
-
- Week 17** Module 33: Apache Kafka
 Module 34: Integration of Kafka with Spark
-
- Week 18** Module 35: Splunk and Big data management
 Module 36: Big data processing using Splunk
-
- Week 19** Module 37: Big data processing pipelines
 Module 38: Live Practice session

Duration : 19 Weeks



MACHINE LEARNING WITH SPARK

- Week 20** Module 39: Introduction to Machine learning
 Module 40: Statistics
-
- Week 21** Module 41: Overview of Machine Learning algorithms
 Module 42: Spark MLlib datatypes
-
- Week 22** Module 43: Classification
 Module 44: Linear regression, Logistic regression
-
- Week 23** Module 45: Spark MLlib Decision tree
 Module 46: Spark MLlib KNN classification
-
- Week 24** Module 47: Spark Clustering
 Module 48: Recommendation system
-
- Week 25** Module 49: Recommendation system Hands - on
 Module 50: Live Practice Session

Duration : 25 Weeks

