



NIAI

NETSOL Institute of AI

BIG DATA ANALYTICS Training Program

*Hands-on Training & Real-World Skills to Prepare for
the Future*

1. BIG DATA ANALYTICS Curriculum Breakdown

Course Name:	BIG DATA ANALYTICS
Course Duration:	3 Months (12 Weeks)
No. Of Days per week:	4 Days per week (Course only)
Class Hours	3 hours per class
Total In-Class Credit hours	144

Module	Content Breakdown
1. Introduction to Big Data and Big Data Analytics. Types of Data & Data Lakes	Overview: Concepts, applications, techniques, trends, and course roadmap.
	Definitions, Big Data Characteristics, 10 Vs of Big Data
	Types of Big Data, Data Lakes, Landscapes, Categorization of Big Data Analytics
	Categorization of Big Data Analytics, Overview of NoSQL databases
	Quiz and Tasks
2. NoSQL databases Apache Hadoop Ecosystem	Hands on NoSQL Databases, Overview of Apache Hadoop Ecosystem
	Hadoop 2 (Hands-on)
	YARN (Hands-on), HDFS
	Setting up Hadoop clusters
	Quiz and Tasks
3. MapReduce & MapReduce	MapReduce: Theory and Hands-on
	Hands on MapReduce
	Apache Spark with Apache Kafka
	Apache Hive
	Quiz and Tasks
4. Apache Spark with Apache Kafka Apache Hive, Apache HBase and Apache Cassandra	Apache HBase
	Apache Cassandra
	Apache Presto
	Apache Drill
	Hands-on, R&D
	Quiz and Tasks

5. Apache Presto and Apache Drill, Document NoSQL with MongoDB Graph NoSQL with Neo4J	NoSQL
	NoSQL with MongoDB
	Graph NoSQL with Neo4J
	Case studies
	Quiz and Tasks
6. Key Value Stores with Redis	Client Connection, Cluster Initialization, Maintenance, Database Usage
	CURL Command, Data Manipulation
	Redis
	Quiz, Assignments and Tasks
7. Large-Scale Supervised Learning	Introduction to Supervised learning
	Generalized Linear Models and Logistic Regression
	Regularization
	Support Vector Machine (SVM) and the kernel trick
	Outlier Detection
	Spark ML library
	Quiz, Assignments and Tasks
8. Large-Scale Unsupervised Learning	Large-Scale Unsupervised Learning
	K-means / K-medoids
	Gaussian Mixture Models
	Dimensionality Reduction, Spark MLlib for Unsupervised Learning
9. Large Scale Text Mining	Latent Semantic Indexing
	Topic models
	Latent Dirichlet Allocation
	Spark ML library for NLP
	Quiz, Assignments and Tasks
10. Final Project	Final Project

13. Capstone Project End-to-End Business Data Analysis	Project runs parallel focusing on real-world data problems: design, execute, and present a complete data analysis solution. Example domains: sales and revenue analytics, customer behavior, marketing performance, operations/HR analytics etc. Hands-On: Build an end-to-end data analytics pipeline (data collection, cleaning, EDA, analysis/modeling, dashboard). Present insights and business recommendations to stakeholders and evaluate project outcomes.
---	--

Notes:

Activities & Tasks All Tasks are as per NAVTTC Course Annexure-I, however Quiz, assignments and Capstone Project are additional
 Motivational Lecture We have a separate Soft Skill program which will be conducted every week on Soft Skill day.
 An Outlined draft of our Soft Skills program is attached Separately

Item	NIAI	Remarks
Course Duration	3 Months (12 Weeks)	All NAVTTC topics covered + Corporate development Soft Skills Program
Total Class Days	5 Days	4 days/week (Course) + 1day for soft skills.
In-Class Hours	3 hours/day	Content coverage sustained
Total In-Class Hours	144 hours (Excluding notional hours)	In-Class subject: 144 hours In-Class Soft Skills: 24 hours Project work: Approx. 100 hours (guided independent and team-based projects)

2. Soft Skills & Leadership Modules

"Delivered by NIAI's (NETSOL Institute of Artificial Intelligence) Organizational Development experts, these modules enhance corporate readiness and professional competencies:"

- Building Future Ready Tech Leaders
- Communication Improvement Program
- Teamwork & Collaboration
- Emotional Intelligence & Empathy

This stream adds 2 hours/week across 12 weeks (24- credit hours), reinforcing Professionalism and Business Skills competencies vital for workplace integration.