

**Duration:** 52 Classroom hours and 58 Lab hours

**Objective:** The objective of this module is to make students aware with the latest AI tools and Techniques used by IT Industry

**Prerequisites:** Knowledge of fundamentals of programming and basic mathematics. ML & statistical knowledge

**Evaluation method:** Theory exam– 40% weightage

Lab exam – 40% weightage

Internal exam – 20% weightage

**List of Books / Other training material**

**Courseware:** No specific courseware for the module, faculty may share some coursematerials

**Reference:**

- Deep Learning, Ian Goodfellow, Yoshua Bengio, Aaron Courville
- Neural Networks and Learning Machines, Simon Haykin
- Pattern Recognition and Machine Learning, Christopher M. Bishop
- Cloud Computing: Master The Concepts, Architecture and Applications With Real-World Examples And Case Studies by Mehul Mahrishi Kamal Kant Hiran, Ruchi Doshi, Dr. Fagbola Temitayo
- Kubernetes: Up and Running - Dive into the Future of Infrastructure, Second Edition [Paperback] Brendan Burns; Kelsey Hightower and Joe Beda by Brendan Burns (Author), Kelsey Hightower (Author), Joe Beda (Author)
- Hands-on Pipeline as Code with Jenkins: CI/CD Implementation for Mobile, Web, and Hybrid Applications Using Declarative Pipeline in Jenkins (English Edition) by Ankita Patil Mitesh Soni
- Introducing GitHub 2e by Brent Beer

**Note:**

- **40 hours for PySpark including both theory and lab sessions.**
- **70 hours for cloud computing including both theory and lab sessions.**
- **Applications/case studies are advised to be covered under Lab Hours.**
- **Each session is having 2 Hours for theory only. Lab assignments are indicatives.**
- **Faculty is advised to give appropriate assignment for sufficient hands-on.**
- **Faculty is advised to give at-least one end-to-end problem statement where students can execute their learning in a proper manner.**

**PySpark (20 Hrs of Theory and 20 Hrs of Lab)**

**40 Hours**

**Session 1 & 2:**

- Apache Spark APIs for large-scale data processing
- Overview, Linking with Spark,
- Initializing Spark
- Spark session & spark context

- Resilient Distributed Datasets (RDDs), External Datasets, RDD Operations,
- Passing Functions to Spark, Working with Key-Value Pairs, Shuffle operations,
- RDD Persistence, Removing Data, Shared Variables, Deploying to a Cluster
- Spark Data Frames

**Lab Assignments:**

- Run the provided Hadoop Streaming program using python
- Install Spark

**Session 3 & 4:**

- EDA using PySpark
- ETL jobs using sparks

**Lab Assignments:**

- Explore Data using PySpark

**Session 5:**

- Introduction to Kafka
- Working with Kafka using Spark
- Spark streaming

**Lab Assignments:**

- Install Kafka
- Execute the spark streaming with Kafka

**Session 6:**

- Integration of setting Spark and Kafka
- Setting up Kafka Producer and Consumer
- Kafka Connect API

**Lab Assignments:**

- Connect the Kafka with API

**Session 7:**

- Machine Learning using Spark's MLlib

**Lab Assignments:**

- Explore MLlib

**Session 8:**

- Deep Learning using Spark

**Lab Assignments:**

- Explore Deep Learning

**Session 9:**

- Spark SQL

**Lab Assignments:**

- Run the Spark SQL programs using step –runs for each and every code line
- Analyze the election data using spark and provide analysis

**Session 10:**

- Connecting DB's with Spark
- Working with DB's

**Lab Assignments:**

- Connecting DB's with Spark
- Accessing and manipulating the DB's
- Create a complex workflow using bash operator, a simple workflow using python

**DevOps for AI/ML (32 Hrs of Theory and 38 Hrs of Lab)**

**70 Hours**

**Session 11 & 12: Git / GitHub**

- Introduction to Version control systems
- Creating GitHub repository
- Using Git – Introduction to Git commands

**Lab Assignment:**

- Create a GitHub repository from the command line

**Session 13 & 14: Introduction to containers (4T + 6L)**

- Introduction to DevOps
- Introduction to Containers
- Advantages of using container-based applications
- Installing Docker and using basic Docker commands
- Build your own container-based application image
- Networking in Docker
- Managing containers – Logs / Resources

**Lab Assignments:**

- Create an httpd container running on port 8000. Name the container as web5. Do not map any directory. Create an “index.html” file & copy this file inside the container and check if the website displays your webpage. Again, modify the “index.html” file and copy it inside the container, and check if the website is updated or not.
- Test any program on-base machine. Then create an image using Dockerfile to execute this exe program. Give image name as “Name: v1”.

**Session 15 & 16: Introduction to Kubernetes (4T + 6L)**

- Need for Kubernetes
- Introduction to Kubernetes cluster – Basic terms - Management node, Worker Nodes, Pods, Deployment, and Service Types etc.
- Working with Kubernetes Cluster – Creating deployment, Exposing Deployment as a service, managing your applications. Rolling application updates etc.

**Lab assignment:**

- Create a Kubernetes cluster, deploy and explore any app.

**Session 17 & 18: CI/CD with Jenkins (4T + 6L)**

- Introduction to CI/CD
- Using Jenkins to build a CI/CD pipeline.

**Lab Assignments:**

- Create a CI/CD pipeline using Jenkins, Docker, and GitHub to automate the web by using the NGINX image.
- Create a CI/CD pipeline using Jenkins, Docker, and GitHub to automate the web by using the httpd image.

**Cloud Computing:**

**Session 19 & 20:**

- Cloud Computing Basics, Understanding Cloud Vendors (AWS/Azure/GCP), Definition, Characteristics, Components

**Lab Assignments:**

- Study about cloud and other similar configuration
- Explore available solutions
- Cloud Architecture

**Session 21 & 22:**

- Cloud provider, SAAS, PAAS, IAAS and other Organizational scenarios of clouds, Administering & Monitoring cloud services, benefits, and limitations

**Lab Assignments:**

- Provide a solution on cloud as SAAS using available systems.

**Session 23 & 24:**

- Deploy application over cloud. Comparison among SAAS, PAAS, IAAS, Cloud Products and Solutions, Cloud Pricing, Compute Products and Services, Elastic Cloud Compute, Dashboard.

**Lab Assignments:**

- Deploy any application over the cloud platform.

**Session 25 & 26:**

- Exploring cloud services for AI/ML

**Lab Assignments:**

- Explore the different cloud computing platforms for AI/ML.

**Self-Study:** AI applications in Financial Services including Insurance banking, stock markets & other financial markets like Forex—and Artificial Economics, AI applications in Health Sciences & other Scientific Applications, AI in Cloud Environment , Deployment of Models on a distributed platform.