

Data Science & Machine Learning Internship

(15 Days | Practical-Oriented | Industry Ready)

INTERNSHIP SYLLABUS

Internship Objectives

By the end of this internship, students will be able to:

- Understand complete **Data Science lifecycle**
- Perform **data analysis, visualization & preprocessing**
- Build **Machine Learning models from scratch**
- Evaluate models and improve accuracy
- Develop a **mini real-world ML project**
- Gain confidence to do **final year / real-time projects**

DAY 1 – Introduction to Data Science

- What is Data Science & ML?
- Industry applications & career paths
- DS lifecycle & project workflow
- Tools overview: Python, Jupyter, Anaconda

Practical:

Install Python & Jupyter

Run first Python notebook

DAY 2 – Python for Data Science

- Python basics (variables, loops, functions)
- Lists, tuples, dictionaries
- NumPy basics

Practical:

Python programs for data operations

NumPy array operations

DAY 3 – Data Handling with Pandas

- DataFrames & Series
- Reading CSV, Excel datasets
- Data inspection (head(), info(), describe())

Practical:

Load real dataset

Basic data exploration

DAY 4 – Data Cleaning & Preprocessing

- Handling missing values
- Data type conversion
- Duplicate removal
- Encoding categorical data

Practical:

Clean messy dataset

Prepare ML-ready data

DAY 5 – Exploratory Data Analysis (EDA)

- EDA concepts
- Univariate & multivariate analysis
- Statistical insights

Practical:

Analyze dataset trends
Identify patterns & outliers

DAY 6 – Data Visualization

- Matplotlib basics
- Seaborn charts
- Bar, line, histogram, box plot

Practical:

Visualize dataset
Storytelling using graphs

DAY 7 – Introduction to Machine Learning

- ML types: Supervised, Unsupervised
- Training vs testing data
- Model evaluation concepts

Practical:

Split dataset
First ML workflow

DAY 8 – Linear Regression

- Regression concepts
- Simple & multiple linear regression
- Evaluation metrics (MAE, MSE, R^2)

Practical:

Build linear regression model
Predict outputs

DAY 9 – Classification Algorithms

- Logistic Regression
- K-Nearest Neighbors (KNN)

- Confusion matrix & accuracy

Practical:

Classification model on real dataset

Accuracy analysis

DAY 10 – Decision Tree & Random Forest

- Decision Tree working
- Overfitting & pruning
- Random Forest basics

Practical:

Train tree-based models

Compare results

DAY 11 – Unsupervised Learning

- Clustering concepts
- K-Means clustering
- Elbow method

Practical:

Customer / data segmentation

Cluster visualization

DAY 12 – Model Optimization

- Feature scaling
- Hyperparameter tuning
- Cross-validation

Practical:

Improve model accuracy

Compare tuned vs normal models

DAY 13 – Mini Project Development

- Problem statement selection

- Dataset finalization
- Model selection

Practical:

Start mini ML project

Trainer guidance

DAY 14 – Project Completion & Deployment Intro

- Model testing
- Result interpretation
- Intro to Flask / Streamlit

Practical:

Finalize project

Simple ML app demo

DAY 15 – Presentation & Evaluation

- Project presentation
- Viva & technical review
- Career guidance

Assessment:

Project demo

MCQ + Practical test Feedback & certification

Mini Project Examples

- Sales Prediction System
- Student Performance Analysis
- Disease Prediction System
- Customer Segmentation
- House Price Prediction