NavAdr Technologies Solutions

A-04, Street of Europe, Hinjewadi-Phase1
Pune-411057 ,Ph: +91-8857019010
Email : training@navadr.com

This curriculum is designed to help participants reach an intermediate level of expertise in data science, equivalent to 2–3 years of experience. It includes foundational topics, core data science concepts, advanced techniques, and hands-on practice, delivered in a structured 36-week format.

Week	Topic	Key Concepts and Activities
1	**Introduction to Data Science**	Overview of Data Science lifecycle, Python basics (NumPy, Pandas), and role understanding.
2	**Data Manipulation and Visualization**	Data wrangling with Pandas, EDA using Matplotlib and Seaborn, handling missing data, and outliers.
3	**Statistics for Data Science**	Descriptive statistics, probability distributions, hypothesis testing (t-tests, chi-square).
4	**Linear Algebra and Calculus**	Matrix operations, gradient descent, eigenvalues, eigenvectors, and dimensionality reduction basics.
5	**Supervised Learning: Regression**	Linear regression, polynomial regression, feature engineering, and evaluation metrics (RMSE, R^2).
6	**Supervised Learning: Classification**	Logistic regression, k-Nearest Neighbours, decision trees, and metrics (Precision, Recall, F1-Score).
7	**Unsupervised Learning**	Clustering (k-Means, hierarchical), PCA, t-SNE, and applications like anomaly detection.
8	**Model Evaluation and Validation**	Train-test split, cross-validation, hyperparameter tuning, avoiding overfitting and underfitting.
9	**Ensemble Learning**	Random Forest, Gradient Boosting (XGBoost, LightGBM), and stacking/blending techniques.
10	**Feature Engineering**	Handling categorical data (One-hot encoding, target encoding), feature selection (RFE), interaction terms.
11	**Time Series Analysis**	Time series components (trend, seasonality), ARIMA, SARIMA, and forecasting techniques.
12	**Natural Language Processing Basics**	Text preprocessing (tokenization, stemming), vectorization (Bag of Words, TF-IDF), and text classification.
13	**Introduction to Deep Learning**	Neural networks (perceptron, backpropagation), feedforward networks, and basic TensorFlow/PyTorch usage.
14	**Convolutional Neural Networks (CNNs)**	Understanding convolutions, pooling layers, and building CNNs for image classification.
15	**Recurrent Neural Networks (RNNs)**	RNNs, LSTMs for sequential data, applications in time series and text, TensorFlow/PyTorch implementation.
16	**Advanced NLP**	Word embeddings (Word2Vec, GloVe), transformers (BERT introduction), and sequence-to-sequence models.

NavAdr Technologies Solutions

A-04, Street of Europe, Hinjewadi-Phase1
Pune-411057 ,Ph: +91-8857019010
Email : training@navadr.com

Week	Topic	Key Concepts and Activities
17	**Big Data for Data Science**	Introduction to Hadoop, Spark, and PySpark for large-scale data processing.
18	**MLOps Concepts: Fundamentals**	Overview of MLOps, CI/CD for machine learning, pipeline orchestration, and introduction to MLFlow.
19	**MLOps Concepts: Deployment**	Model serving using Flask/FastAPI, Dockerizing ML models, introduction to Kubernetes for scaling.
20	**MLOps Concepts: Monitoring and Retraining**	Monitoring models for drift, automated retraining workflows, and model versioning in MLFlow.
21	**Capstone Preparation**	Dataset selection, problem formulation, designing an end-to-end pipeline, data cleaning, and feature engineering.
22	**Capstone Project Development**	Building models, evaluating performance, optimizing hyperparameters, and documenting results.
23	**Capstone Deployment**	Deploying the model (Flask/Streamlit), creating dashboards for real-time insights, presenting results.
24	**Certification and Career Preparation**	Preparing for certifications (e.g., TensorFlow, AWS ML)
25-36	**LIVE Project Exposure**	3 Months – Live Project Experience