

# **INFOSOFT IT SOLUTIONS**

**Training | Projects | Placements**

Revathi Apartments, Ameerpet, 1<sup>st</sup> Floor, Opposite Annapurna Block, Infosoft It solutions,  
Software Training & Development Institute, +91 - 9059683947 | +91 - 9182540872

## **DATA SCIENCE WITH PYTHON**

### **COURSE CONTENT**

#### **Introduction to Python for Data Science**

- Introduction to Python programming language
- Installation and setup of Python environment (Anaconda)
- Basic data types, variables, and operators
- Introduction to NumPy and Pandas for data manipulation

#### **Data Manipulation with Pandas**

- Data structures in Pandas: Series and DataFrame
- Indexing, slicing, and filtering data
- Data cleaning and preprocessing techniques
- Handling missing data and outliers

#### **Data Visualization with Matplotlib and Seaborn**

- Introduction to data visualization

- Basic plotting techniques with Matplotlib
- Advanced visualization with Seaborn
- Customizing plots and adding annotations

## **Exploratory Data Analysis (EDA)**

- Understanding the structure and distribution of data
- Statistical summaries and descriptive statistics
- Univariate and bivariate analysis
- EDA case studies and practical exercises

## **Introduction to Machine Learning**

- Overview of machine learning concepts and algorithms
- Supervised vs. unsupervised learning
- Model evaluation and validation techniques
- Introduction to scikit-learn library for machine learning in Python

## **Supervised Learning: Regression**

- Introduction to regression analysis
- Simple linear regression
- Multiple linear regression
- Polynomial regression and regularization techniques

## **Supervised Learning: Classification**

- Introduction to classification algorithms
- Logistic regression
- Decision trees and random forests
- Support Vector Machines (SVM)

## **Unsupervised Learning**

- Introduction to unsupervised learning algorithms
- Clustering techniques: K-means clustering, hierarchical clustering
- Dimensionality reduction techniques: Principal Component Analysis (PCA), t-Distributed Stochastic Neighbor Embedding (t-SNE)

## **Model Evaluation and Hyperparameter Tuning**

- Cross-validation techniques for model evaluation
- Hyperparameter tuning using grid search and random search
- Model selection and performance metrics
- Overfitting, underfitting, and bias-variance trade-off

## **Feature Engineering and Selection**

- Introduction to feature engineering

- Handling categorical variables: one-hot encoding, label encoding
- Feature scaling and normalization
- Feature selection techniques: filter methods, wrapper methods, embedded methods

## **Introduction to Deep Learning with TensorFlow and Keras**

- Overview of deep learning concepts
- Introduction to TensorFlow and Keras libraries
- Building and training neural networks
- Convolutional Neural Networks (CNNs) for image classification

## **Capstone Project**

- Final project where students apply their knowledge and skills to solve a real-world data science problem
- Project planning, data exploration, modeling, and evaluation
- Presentation of projects and peer feedback