



Data Science Program

Course Curriculum



About Edureka Learning Center

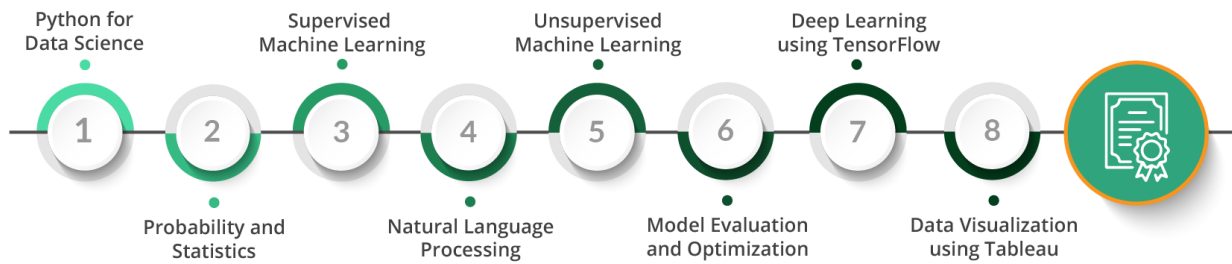
'Your go-to partner for education upgrading and upskilling with current, industry-relevant courses that unleash your employment potential and make you production-ready to fit the tech market demand!'

Edureka Learning Center is democratizing access to high-quality education at an affordable price. We are playing our part in helping the country meet the growing demand for talent across segments. We follow an outcome-oriented approach while not compromising on either quality or affordability. Our emphasis is to become a bridge and fill the employment gap in our country by increasing the employment quotient of students and enabling them to access premium jobs.

About the Program

Edureka Learning Center Data Science Program prepares you to get job-ready and enhance your skills to match the industry standards. This program will enable you to build end-to-end Machine Learning models using the Python Programming Language. This program starts with the basics of Python, guides you to perform data analysis using Pandas, NumPy, Matplotlib, and Seaborn libraries in Python, helps you implement Supervised and Unsupervised learning algorithms, and enables you to evaluate and optimize your models. It also covers Natural Language Processing and Deep Learning concepts such as RNN and CNN using TensorFlow and helps you create interactive visualizations of your data using Tableau while increasing your employability manifold.

Data Science Program Roadmap



Index

1. Python for Data Science
2. Probability and Statistics
3. Supervised Machine Learning
4. Natural Language Processing
5. Unsupervised Machine Learning
6. Model Evaluation and Optimization
7. Deep Learning using TensorFlow
8. Data Visualization using Tableau

**Depending on industry requirements, Edureka may make changes to the course curriculum*



Python for Data Science



About the Module

This module is designed to build the foundation of Data Science with the help of Python as a programming language. You will learn the basic concepts of Python and use various libraries in Python for data processing and analysis.

Module Outline

Learning Units:

- Day 1
 - LU1 - Applications of Data Science
 - LU2 - Introduction to Python
- Day 2
 - LU1 - Operators and Variables in Python
 - LU2 - Data Types in Python
- Day 3
 - LU1 - Control Flow in Python I
 - LU2 - Control Flow in Python II
- Day 4
 - LU1 - Functions in Python I
 - LU2 - Functions in Python II
- Day 5
 - LU1 - Packages and Modules in Python
 - LU2 - File Handling in Python

- Day 6
 - LU1 - Introduction to NumPy Arrays
 - LU2 - Basic NumPy Operations
- Day 7
 - LU1 - NumPy Functions
 - LU2 - Indexing and Slicing of NumPy Arrays
- Day 8
 - LU1 - Array Manipulation in Python I
 - LU2 - Array Manipulation in Python II
- Day 9
 - LU1 - File Handling using NumPy
 - LU2 - NumPy Case Study
- Day 10
 - LU1 - Introduction to Pandas Library in Python
 - LU2 - Pandas Data Structures
- Day 11
 - LU1 - Importing and Exporting Data Using Pandas
 - LU2 - Functionality of Pandas Series
- Day 12
 - LU1 - Functionality of Pandas DataFrames I
 - LU2 - Functionality of Pandas DataFrames II
- Day 13
 - LU1 - Combining Data using Pandas I
 - LU2 - Combining Data using Pandas II
- Day 14
 - LU1 - Data Cleaning using Pandas I
 - LU2 - Data Cleaning using Pandas II
- Day 15
 - LU1 - Grouping Data using Pandas I
 - LU2 - Grouping Data using Pandas II
- Day 16
 - LU1 - Data Visualization Library – Matplotlib
 - LU2 - Data Visualization Library – Seaborn

- Day 17
 - LU1 - Visualizing Matplotlib Plots and Charts
 - LU2 - Customizing Visualizations and Saving Plots

- Day 18
 - LU1 - Introduction to Web Scraping
 - LU2 - Web Scraping using BeautifulSoup

Probability and Statistics

About the Module

This module is designed to brush up your concepts of Probability and Statistics. You will learn how to perform Statistical Analysis that allows you to summarize your data and draw useful insights from it.

Module Outline

Learning Units:

- Day 19
 - LU1 - Statistical Analysis in Data Science
 - LU2 - Measures of Central Tendency
- Day 20
 - LU1 - Measures of Dispersion
 - LU2 - Measures of Position
- Day 21
 - LU1 - Univariate Non-Graphical EDA
 - LU2 - Univariate Graphical EDA
- Day 22
 - LU1 - Multivariate Non-Graphical EDA
 - LU2 - Multivariate Graphical EDA
- Day 23
 - LU1 - Introduction to Probability Theory
 - LU2 - Probability Events

- Day 24
 - LU1 - Types of Probabilities
 - LU2 - Bayes' Theorem
- Day 25
 - LU1 - Probability Distributions
 - LU2 - Skewness and Kurtosis
- Day 26
 - LU1 - Types of Probability Distributions
 - LU2 - Sampling Distributions
- Day 27
 - LU1 - Inferential Statistics
 - LU2 - Confidence Interval
- Day 28
 - LU1 - Statistical Hypothesis Testing
 - LU2 - P-Value and Critical Value
- Day 29
 - LU1 - Hypothesis Tests
 - LU2 - T-Tests
- Day 30
 - LU1 - Chi-Squared Tests
 - LU2 - Probability and Statistics Case Study

Supervised Machine Learning

About the Module

From this module onwards you will start your Machine Learning journey by taking the first steps of preparing data for Machine Learning. You will learn about Supervised Machine Learning techniques and work with Classification and Regression algorithms.

Module Outline

Learning Units:

- Day 31
 - LU1 - Introduction to Machine Learning
 - LU2 - Types of Machine Learning
- Day 32
 - LU1 - Data Pre-processing Techniques I
 - LU2 - Data Pre-processing Techniques II
- Day 33
 - LU1 - Testing and Training Data
 - LU2 - Supervised Learning: Regression
- Day 34
 - LU1 - Linear Regression
 - LU2 - Calculation of R Square
- Day 35
 - LU1 - Gradient Descent
 - LU2 - Regularization Techniques

- Day 36
 - LU1 - Regression Case Study
 - LU2 - Classification Algorithms
- Day 37
 - LU1 - Logistic Regression
 - LU2 - Decision Tree
- Day 38
 - LU1 - Decision Trees with CART Algorithm - I
 - LU2 - Decision Trees with CART Algorithm – II
- Day 39
 - LU1 - Random Forest
 - LU2 - Performance Measurements
- Day 40
 - LU1 - Naïve Bayes Classification
 - LU2 - How Naïve Bayes Works?
- Day 41
 - LU1 - K Nearest Neighbor
 - LU2 - K in KNN Algorithm
- Day 42
 - LU1 - Support Vector Machine
 - LU2 - Non-Linear SVMs

Natural Language Processing

About the Module

This module is designed to build the foundation of Natural Language Processing. You will learn how to perform text pre-processing and feature extraction here.

Module Outline

Learning Units:

- Day 43
 - LU1 - Introduction to Natural Language Processing (NLP)
 - LU2 - Natural Language Tool-Kit (NLTK)
- Day 44
 - LU1 - Text Pre-processing - I
 - LU2 - Text Pre-processing – II
- Day 45
 - LU1 - Text Pre-processing - III
 - LU2 - Text Pre-processing – IV
- Day 46
 - LU1 - Feature Extraction I
 - LU2 - Feature Extraction II
- Day 47
 - LU1 - Sentiment Analysis
 - LU2 - Case Study - Sentiment Analysis

Unsupervised Machine Learning

About the Module

This module is designed to teach you about Unsupervised Machine Learning techniques. You will work with Dimensionality Reduction and Clustering algorithms here. You will also learn Association Rule Mining and start building a Recommender System.

Module Outline

Learning Units:

- Day 48
 - LU1 - Dimensionality Reduction
 - LU2 - Principal Component Analysis (PCA)
- Day 49
 - LU1 - Linear Discriminant Analysis (LDA)
 - LU2 - Other Techniques of Dimensionality Reduction
- Day 50
 - LU1 - Unsupervised Learning Using Clustering
 - LU2 - Hierarchical Clustering
- Day 51
 - LU1 - K Means Clustering - I
 - LU2 - K Means Clustering – II
- Day 52
 - LU1 - Fuzzy C Means Clustering
 - LU2 - DBSCAN Clustering

- Day 53
 - LU1 - Association Rule Mining
 - LU2 - Generating Association Rules
- Day 54
 - LU1 - Apriori Algorithm
 - LU2 - Market Basket Analysis
- Day 55
 - LU1 - Recommendation Engine
 - LU2 - Types of Recommender System
- Day 56
 - LU1 - Recommending Similar Movie to the User
 - LU2 - Introduction to Time Series
- Day 57
 - LU1 - Types of Data
 - LU2 - Checks for Stationarity Of Data
- Day 58
 - LU1 - Convert Non-Stationary Data to Stationary Data
 - LU2 - Time Series Models
- Day 59
 - LU1 - Time Series Models Using Python
 - LU2 - Case Study - Association Rule Mining and Time Series

Model Evaluation and Optimization

About the Module

This module is designed to teach you about techniques that validate the performance of Machine Learning models. You will also learn to evaluate the performance of Classification and Regression Models.

Module Outline

Learning Units:

- Day 60
 - LU1 - Model Selection
 - LU2 - K-Fold Cross Validation
- Day 61
 - LU1 - Model Evaluation
 - LU2 - Model Evaluation Metrics for Regression
- Day 62
 - LU1 - Model Evaluation Metrics for Classification
 - LU2 - Calculating A Confusion Matrix
- Day 63
 - LU1 - ROC and AUC
 - LU2 - Precision, Recall and F1 Score
- Day 64
 - LU1 - Hyperparameter Tuning
 - LU2 - Hyperparameter Optimization

- Day 65
 - LU1 - Perform Grid Search
 - LU2 - Ensemble Learning
- Day 66
 - LU1 - Bagging
 - LU2 – Boosting
- Day 67
 - LU1 - AdaBoost I
 - LU2 - AdaBoost II
- Day 68
 - LU1 - Gradient Boosting
 - LU2 – XGBoost
- Day 69
 - LU1 - Model Optimization
 - LU2 - Linear Programming
- Day 70
 - LU1 - Formulating Optimization Problem
 - LU2 - Predicting Promotion Using Boosting Techniques

Deep Learning Using TensorFlow

About the Module

This module is designed to teach you about Deep Learning and its networks using TensorFlow 2.x. You will learn about Deep Learning for Sequences using RNN and Deep Learning for Image Recognition using CNN.

Module Outline

Learning Units:

- Day 71
 - LU1 - Introduction to Deep Learning
 - LU2 - Introduction to Neural Networks
- Day 72
 - LU1 - Single Layer Perceptron
 - LU2 - Multilayer Perceptron (MLP)
- Day 73
 - LU1 - How Does a Neural Network Learn?
 - LU2 – Backpropagation
- Day 74
 - LU1 - Introduction to TensorFlow
 - LU2 - MNIST Digit Classification Using TensorFlow 2.x
- Day 75
 - LU1 - Understanding CNN
 - LU2 - Image Recognition

- Day 76
 - LU1 - Introduction to RNN
 - LU2 - Architecture of RNN

- Day 77
 - LU1 - Understanding RNN
 - LU2 - Drawback of Backpropagation

- Day 78
 - LU1 - LSTM – Long Short-Term Memory Networks
 - LU2 - Understanding LSTM Structure

- Day 79
 - LU1 - Introduction to Reinforcement Learning (RL)
 - LU2 - Understanding Reinforcement Learning

- Day 80
 - LU1 - RL Agent Taxonomy
 - LU2 - OpenAI Gym

Data Visualization Using Tableau

About the Module

This module is designed to give you a brief idea on Data Visualization and explore the various features and techniques used for Visualization in Tableau. You will learn various advanced visualizations and techniques for analyzing data using Tableau.

Module Outline

Learning Units:

- Day 81
 - LU1 - Introduction to Tableau
 - LU2 - Data Connections and Charts
- Day 82
 - LU1 - Data Granularity and Sorting
 - LU2 - Data Grouping and Filtering
- Day 83
 - LU1 - Data Blending
 - LU2 - Joins and Unions
- Day 84
 - LU1 - Calculations in Tableau
 - LU2 - Functions in Tableau
- Day 85
 - LU1 - Table Calculations and Parameters
 - LU2 - LOD Calculations

- Day 86
 - LU1 - Trend Lines and Reference Lines
 - LU2 - Forecasting and Clustering
- Day 87
 - LU1 - Introduction to Mapping
 - LU2 - Web Mapping Service (WMS)
- Day 88
 - LU1 - Using Charts Effectively
 - LU2 - Introduction to Dashboards in Tableau
- Day 89
 - LU1 - Dashboard Layouts and Formatting
 - LU2 - Interactive Dashboards
- Day 90
 - LU1 - Story Points in Tableau
 - LU2 - Visual Best Practices