Data Science Master Course

Data Science & Machine Learning using Python

NASSCOM CERTIFIED

4,500+
Batches

55+
Countries

59,000+
Participants

10+
Years

python™ MACHINE LEARNING
In-depth Coverage
Curriculum Certified by **NASSCOM**®

Digital Vidya

Data Science | Master Course

Why Learn Data Science Using Python?
Stay Ahead of the Change

- Average salary base for a Data Scientist is $128,750

- Demand For Data Scientists Will Soar 28% By 2020

- Data Science Job Openings are expected to increase to 2,720,000

- In India, the initial salary ranges between 4,63,000 – 10,22,000 INR

What Will You Get?
Learning That's Not Confined

27	**100+**	**54+**
Number of Modules	Training & Exercise Hours	Assignment Hours
6+	**15+**	**10+**
Projects	Assignments	Trainers
	24*7	
	Forum Support	

Salient Features

Prepares you for 2,800 Companies

Q&A Sessions bi-weekly

3 Weeks of Project Work

Class Labs/Home Assignment (05 hrs/Week Learning Time)

Placement Support

Individual Attention to Each Learner

Lifetime Access to Updated Content and Videos

Top Rated Advisors

Industry and Academia Faculty

Multiple Tools Covered

Internal Competitions with Prizes

Money Back Guarantee

Active Q/A Forum

Doubt Clearing Sessions

Hands-on Approach

Who Should do this Course?
Get Ready to Transform

- Freshers aspiring an exemplary career in Data Science

- Non-IT Professionals desirous of making a career shift to the Data Science Industry

Get Trained & Certified for Data Science Jobs in 2800 Companies

NASSCOM has certified Digital Vidya's Data Science Master Course as it covers 100% of the Data Science skills across Industries.

NASSCOM Data Science curriculum is based on skills required by 2800+ NASSCOM member companies.

Every course participant will get NASSCOM Certificate on successful course completion.

20+ leading companies were involved in defining NASSCOM Course Curriculum.

SME Statistics

40+
Industry Expert

20+
Organizations

SME Leaders

HR

L&D

Talent Acquisition

Engineering

Members of the Industry Working Group (Indicative list)

IBM   Microsoft   Fractal   wipro

Mindtree   accenture   genpact   BANCA SELLA

Fidelity INVESTMENTS   TATA   GE Digital

Other Members part of the review

intuit   ADP   ideas A sas COMPANY   SAMSUNG

Course Curriculum

Python Programming

Python - Fundamentals

- Installation
- Python – Syntax
- Python – Variables and Datatypes
- Python – Numbers
- Strings
- Sequences
- List
- Tuples
- Ranges
- Dictionary
- Sets
- Operators
- If..Else.. Statements
- For Loop
- While Loop
- Break
- Continue
- Pass
- Date & Time
- Functions
- Packages and modules
- Reading a File
- Writing into File
- Class & Objects
- Python – Exceptions
- Regular Exp
- Mathematics

Python MySQL

- Environment Setup

- Database Connection

- Creating a New Database

- Creating Tables

- Insert Operation

- Read Operation

- Update Operation

- Join Operation

- Performing Transactions

Numpy

- ndarray

- Array Creation

- Data Type Objects

- Data type Object (dtype) in NumPy

- Indexing

- Basic Slicing and Advanced Indexing

- Iterating Over Array

- Binary Operations

- Mathematical Function

- String Operations

- Linear Algebra

- Sorting, Searching and Counting

- Set 1 (Introduction)

- Set 2 (Advanced)

- Multiplication of two Matrices in Single line using Numpy in Python

Pandas DataFrame

- Creating a Pandas DataFrame

- Dealing with Rows and Columns in Pandas DataFrame

- Indexing and Selecting Data with Pandas

- Boolean Indexing in Pandas

- Conversion Functions in Pandas DataFrame

- Iterating over rows and columns in Pandas DataFrame

- Working with Missing Data in Pandas

- Working With Text Data

- Working with Dates and Times

- Merging, Joining and Concatenating

Data Analysis

- Data visualization using Bokeh

- Exploratory Data Analysis in Python

- Data visualization with different Charts in Python

- Data Analysis and Visualization with Python

- Math operations for Data analysis

Object-Oriented Concepts

- Class, Object and Members
- Data Hiding and Object Printing
- Inheritance, examples of an object, subclass and super
- Polymorphism in Python
- Class and static variable in Python
- Class method and static method in Python
- Changing class members
- Constructors in Python
- Destructors in Python

- First-class function
- str() vs repr()
- str() vs vpr()
- Metaprogramming with metaclasses
- Class and instance attribute
- Reflection
- Barrier objects
- Timer objects
- Garbage collection

Functions

- Functions in Python
- class method vs static method in Python
- Write an empty function in Python – pass statement
- Yield instead of Return
- Return Multiple Values
- Partial Functions in Python
- First Class functions in Python
- Precision Handling

- *args and **kwargs
- Python closures
- Function Decorators
- Decorators in Python
- Decorators with parameters in Python
- Memoization using decorators in Python
- Python bit functions on int (bit_length, to_bytes and from_bytes)

Machine learning

Statistics Fundamentals

- Graphically Displaying Single Variable

- Measures of Location

- Measures of Spread

- Displaying relationship – Bivariate Data

- Scatterplot

- Measures of association of two or more variables

- Covariance and Correlation

- Probability

- Joint Probability and independent events

- Conditional probability

- Bayes' Theorem

- Prior, Likelihood and Posterior

- Discrete Random Variable

- Probability Distribution of Discrete Random Variable

- Binomial Distribution

- Continuous Random Variables

- Probability Distribution Function

- Uniform Distribution

- Normal Distribution

- Point Estimation

- Interval Estimation

- Hypothesis Testing

- Testing a one-sided Hypothesis

- Testing a two-sided Hypothesis

Machine Learning with Python

- Applications of Machine Learning

- Supervised vs Unsupervised Learning

- Python libraries suitable for Machine Learning

Regression - Intro and Data

- Regression – Features and Labels

- Regression – Training and Testing

- Regression – Forecasting and Predicting

- Regression – Theory and how it works

- Regression – How to program the Best Fit Slope

- Regression – How to program the Best Fit Line

- Regression – R Squared and Coefficient of Determination Theory

- Model evaluation methods

Classification

- Classification Intro

- Applying K Nearest Neighbors to Data

- Euclidean Distance theory

- Decision Trees

- Regression Trees

- Random Forests

- Boosting Algorithm

- Principal Component Analysis

- Linear Discriminant Analysis

Support Vector Machine Inroduction

- Vector Basics

- Support Vector Machine Fundamentals

- Constraint Optimization with Support Vector Machine

- Beginning SVM from Scratch in Python

- Support Vector Machine Optimization in Python

- Visualization and Predicting with our Custom SVM

- Kernels Introduction

- Soft Margin Support Vector Machine

Machine Learning - Clustering Introduction

- Handling Non-Numerical Data for Machine Learning

- K-Means with Titanic Dataset

- K-Means from Scratch in Python

- Finishing K-Means from Scratch in Python

- Hierarchical Clustering with Mean Shift Introduction

- Introduction Naive Bayes Classifier

- Naive Bayes Classifier with Scikit

- Introduction into Text Classification using Naive Bayes

- Python Implementation of Text Classification

Recommender Systems

- Content-based recommender systems

- Collaborative Filtering

Introduction to NLP

- Text Preprocessing

- Noise Removal

- Lexicon Normalization

- Lemmatization

- Stemming

- Object Standardization

- Text to Features (Feature Engineering on text data)

- Syntactical Parsing

- Dependency Grammer

- Part of Speech Tagging

- Entity Parsing

- Phrase Detection

- Named Entity Recognition

- Topic Modelling

- N-Grams

- Statistical features

- TF – IDF

- Frequency / Density Features

- Readability Features

- Word Embeddings

- Important tasks of NLP

- Text Classification

- Text Matching

- Levenshtein Distance

- Phonetic Matching

- Flexible String Matching

- Important NLP libraries

Assignments

Digital Vidya's Python Programming Course assignments are created with an aim to provide the best possible practical exposure to the learner. These assignments are designed in a way that you can apply all your Python programming skills gained from your instructor-led online lessons.

- Introduction to Python

- Dive Deep into Python

- Introduction to NumPy Library

- Data Manipulation using Pandas Library

- Analyzing & Manipulating Data

- Data Visualization

- Merge Multiple Datasets into One

Machine Learning

Each assignment of Digital Vidya's Machine Learning Course is designed with a focus to provide the best practical experience.

- Statistics: Probability, Hypothesis testing

- Multiple Linear Regression & Quadratic regression analysis

- Introduction to Trees, Decision Trees, Ensemble Learning (Random Forest)

- Classification introduction, Logistics Regression & Text analysis using classification algorithms

- Unsupervised Learning, Unsupervised Learning Techniques-K Means Clustering, Hierarchical Clustering

- Bias-Variance tradeoff, Model evaluation techniques

- Logistic Regression Model Tuning

Certifications

**Digital Vidya's
Data Science Certification**
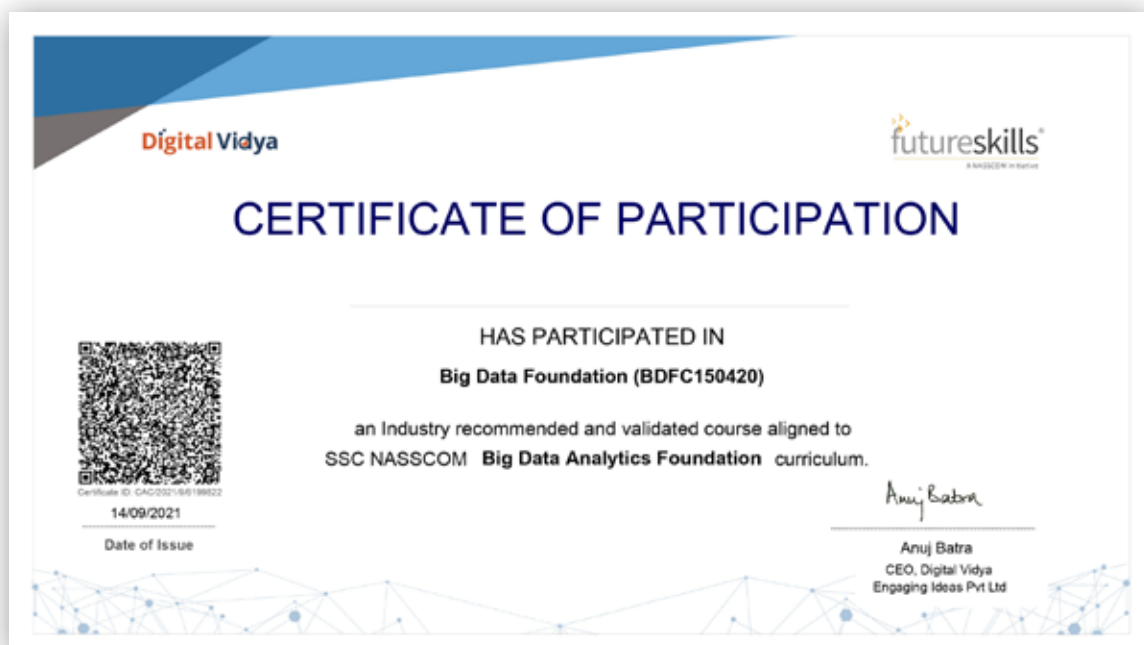
Nasscom Certification

Machine Learning

Natural Language Processing

Project Description:

This is one of the most applied areas for AI, Data Science, and Machine Learning across domains and industries. The real world is filled with mostly messy text data, and handling text is an important step towards making smarter algorithms. Using IMDB dataset from the movie domain, the learner will apply the most common concepts of NLP.

Key Takeaway:

This project will empower the learners to build intermediate skills in the natural language processing domain. A few of the fundamentals of working with textual data covered in this project are:

- Remove stop words

- Apply Stemming and Lemmatization

- Create a cluster of words

- Build a sentiment analysis model and a clustering mode

Bank Marketing

Project Description:

The banking industry is working in a very competitive environment and needs to strategize to grow its business. This project is related to the marketing campaigns related to term deposits, making an interesting multi-disciplinary work that mixes both the finance and the marketing domain.

Key Takeaway:

The approach to this project is to think, define, design, code, test and tune your solution, in such a way that you apply all aspects of the data science process. The data is a real-world data with unclean and null values.

- Build the model to predict if a customer will subscribe

- Identify influential factors to form marketing strategies

- Improve long-term relationship with the clients

Healthcare Analysis

Project Description:

Electroencephalography (EEG) is an electrophysiological monitoring method to record the electrical activity of the brain. For this project, we will use the large EEG database at UCI Machine learning repository. This data arises from a large study to examine EEG correlates of genetic predisposition to alcoholism. One fascinating question is whether the patterns are different for an alcoholic and regular subject?

Key Takeaway:

This capstone project focuses on EEG data analysis, giving an opportunity for students to learn through complexities in dealing with such complex real-world data. The project contains the following exercises:

- Parse and store in an easily understandable and readable form

- Exploratory data analysis to better understand the data

- Using Statistical concepts like Hypothetical testing

- Identify features to predict whether a subject is alcoholic or not

- Use machine learning algorithms to develop a suitable classifier

Deep Learning Based Project

Project Description:

E-Commerce has experienced considerable growth since the dawn of the internet as a commercial enterprise. Deep Learning excels at identifying patterns in unstructured data and can predict the class of an uploaded image applied on eCommerce context. This project is an attempt to replicate virtual store assistance through image recognition over an eCommerce Fashion MNIST dataset.

Key Takeaway:

This project focuses on the implementation of Neural Networks to solve complex unstructured data problems.
The objective is to:

- Build the model to classify the various categories (analytic vertical) of clothing/fashion related images.

- Understanding the implementation of deep learning concepts through Tensorflow and Keras.

- Model optimization by tuning hyper-parameters and implementing dropout layers.
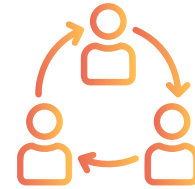
Tools You Will Learn

Placement Services

Resume Creation

On successful completion of the course, which includes submission of assignments & attaining necessary certifications, we work with the candidates to create an effective resume.

Job Application

The updated resume is shared with relevant organisations and agencies including our partners. On shortlisting, we help the candidate to pass the initial round of discussion.

Interview Readiness

Based on the organizations needs & candidates ability, we train them to maneuver themselves to crack the interview. This stage helps the candidate to be 100% ready.

Selection & Joining

After a successful interview, we guide the candidate from accepting the offer to joining the organization for a successful career. We help him to stand out at his workplace.

Course Advisors & Instructors

DILNOOR SINGH

Dilnoor Singh is SAS Base and SAS Advanced Certified Professional having a total of 6 years of IT industry experience. The first 3 years were with pharmaceutical and Insurance Consulting firms and in the later 3, he has built 2 startups in web and mobile domains.

MANDAR MULAY

Mandar Mulay is a Corporate Software Trainer and Consultant in Databas-es and Business Intelligence area. He has 19 years of training experience and is passionate about executing trainings in Excel and VBA. Within Databases, he executes trainings in Oracle, DB2, SQL Server & Tera-data

RUSHABH SHAH

Rushabh Shah is the Founder & CEO of DLTC.co and Blogger at ExcelRush.com. His strengths are data & business analysis. To share his knowledge and experience of working with various businesses areas, he has trained 1300+ working professionals and students in Excel Advanced & others.

SHAHEER AHMAD KHAN

Shaheer is a Data Analytics professional. His deep learning for perception makes him a perfect Data Analytics professional. He holds expertise in Data Mining, Machine Learning, Information Retrieval, Concurrent and Distributed Systems, Web Programming and Advanced Programming.

VAISHALI GARG

Vaishali Garg is a self-taught data analyst with a health-care background. She uses Python with Pandas, Numpy, Matplotlib and Scikit. She has a keen interest in data analysis using Pandas and is actively answer Pandas related questions on StackOverflow (Vaishaligarg, alias: A-Za-z). Some of her analysis is available on Kaggle.

PRITESH SHRIVASTAVA

Pritesh is a Data Science enthusiast with an ability to turn data into actionable insights and meaningful stories. He possesses solid knowledge and hands-on experience of both quantitative and qualitative analysis and data mining. Apart from his profession, he also procures passion and talent in Dramatics, Travel, Story-telling, Martial Artist.

Course Advisors

SHWETA GUPTA
Vice President, Tech.

Shweta Gupta has 19+ years of Technology Leadership experience. She holds a patent and number of publications in ACM, IEEE and IBM journals like Redbook and developer Works.

Digital Vidya

MANAS GARG
Architect

Manas Garg heads the Analytics for Marketing at Paypal. He takes Data Driven Decisions for Marketing Success.

PayPal

VISHAL MISHRA
CEO & Co-Founder

Vishal is a Technology Influencer and CEO of Right Relevance. (A platform used by millions for content & influencer discovery)

Right Relevance

Management Team

ANUJ BATRA
CEO

Anuj Batra has spent more than 15 years in senior leadership roles at organizations such as Genpact, CPA Global, Electronic Data Systems (EDS) and COLT Telecommunications. He strongly feels that industry interface in professional education makes learning more accessible and students more equipped to take on new roles and jobs.

An astute commercial leader, Anuj has been responsible for the design and implementation of commercial and business strategy as well as key financial roles including leading Corporate Training Programs across several companies. He envisions professional education as a key enabler for enhancing employability.

PRADEEP CHOPRA
Co-Founder

Among the pioneers of Digital Marketing in India, Pradeep has been part of the Internet Industry since 1999.

Pradeep is an international speaker & an author on Digital Marketing & Entrepreneurship. He is one of the most sought after speakers in Digital Marketing and his speaking assignments include ad-tech, Search Engine Strategies (USA, India), World Social Media Summit at Kuala Lumpur, Social Media Workshops in Singapore, Click Asia Summit, TiEcon, Global Youth Marketing Forum and NASSCOM India Leadership Forum in India. He writes on Entrepreneurship and Digital Marketing for Wall Street Journal, Inc & Entrepreneur Magazines.

A graduate from IIT Delhi, Pradeep is a core member of the Global Committee at SEMPO (Search Engine Marketing Professional Organization), a non-profit that nurtures the growth of Digital Marketing globally. He also sits on the advisory board of Social Media Chambers Malaysia. His obsession for the digital medium has led him to believe that Internet connectivity is more important than the water supply.

KAPIL NAKRA
Co-Founder

Kapil Nakra, a graduate from IIT Delhi, is a first-generation Serial Entrepreneur. As a pioneer of Digital Marketing in India, Kapil has grown along with the Internet Industry as a User, a Service Provider and now as an Educator.

He started his entrepreneurial journey in 2000 when he started his 1st venture, Whizlabs Software. At Whizlabs, he pioneered the efforts in building an online market for Whizlabs'

assessment solutions for enterprises and helped Whizlabs acquire over 100 Enterprise Customers including blue-chip accounts such as Cisco purely using Digital Marketing. Whizlabs won the 'Most Innovative Indian IT Company' award from NASSCOM in 2004. Kapil is known for his simplistic and inquisitive leading style, which motivates people to think out-of-the-box. It goes without that saying that Kapil has a deep passion for Entrepreneurship and Digital Marketing. He is a regular speaker at various Digital Marketing conferences and his previous speaking engagements include speaking at SearchCamp and OMCAR.

Industry Experts Speak

"

Good to see Digital Vidya becoming increasingly more involved in covering data science vertical, look forward to collaborate with DV to help shape this industry.

Naresh Mehta
AVP – Data Science & Analytics

CARS24.com

"

Yes, I like the huge investment Digital Vidya is doing to create the next generation of talent. Initial feedback suggests Digital Vidya produces high-quality Data Analysts.

Ajay Ohri
Data Scientist

KOGENTiX

"

I can see a good course structure and well-designed syllabus for those who are passionate enough to enter into the analytics world. The platform helps people grow professionally and in very less time.

Madhu Vadlamani
Lead Analytics

kony
Stay Ahead

Participant Reviews

"

This course helped me in understanding Python from the basics & got me ready to implement machine learning models successfully.

Shubham Parab
Software Developer

"

The course Data science using python at Digital Vidya provided me with a turning track for my career which was fruitful in curving my role as a Data Scientist.

Priytosh Sahoo
Data analyst

Interested? Contact Us!

📞 +91-84680-02880

✉ www.digitalvidya.com

🌐 info@digitalvidya.com

**Lifetime
Content Access**

Fee
Rs. 25,000+GST

Weekend Batch
Sat - Sun

Digital Vidya