GUVI | Google for Education Partner | IIT Madras

IIT-M Advanced Programming & Master Data Science

5 Months Weekend / 3 Months Weekday
• **Live Classes** with 100% Job Placement Assistance

IIT-M Advanced Programming & Master

About IIT MADRAS

The Indian Institute of Technology Madras or IIT Madras is recognized globally and holds the laureate of being the No.1 engineering university in India. With a faculty of international fame, bolstered with a highly motivated and profound student community, IIT Madras stands true as an Institute of Eminence. It is a public technical and research university established by the Government of India.

GUVI In A Glance

Grab Ur Vernacular Imprint-GUVI (an IIT-Madras Incubated Company) is World's First Vernacular Ed-Tech Learning Platform. Introduced by Ex PayPal Employees, GUVI empowers students to master programming skills with the comfort of their native language. Its mission is to impart technical skills to all through focussed pedagogical tools.

About the Co-Founders

Co - Founder at GUVI
20+ years of Technical Expertise
PayPal Honeywell & more
Tech Women Entrepreneur who was selected For Google Developers' Launchpad Program
Sridevi

CEO & Founder at GUVI
20+ years of Technical Expertise
PayPal Honeywell Symantec & more
Built 7 Products from Scratch Mentored 1000+ students Hosted 200+ sessions & 25+ webinars
Arun Prakash

Co - Founder at GUVI
17+ years of experience with IT industry
PayPal
Technologist with 9+ years of Entrepreneurial experience & Member of the Syllabus Sub-Committee at Anna University
Bala Murugan

Window to Data Science Program

Data Science with Python is a comprehensive program focusing on top-notch, in-demand fields. This decade foresees progressive inflation in the data science and analytics field, projecting a boom from **3.64+Lakh to 27.2+ Lakh** job opportunities. GUVI offers this 100% job supporting Data Science program as a blend of online master classroom sessions and self-paced course learnings, intending to meet these job requirements. Learn from the best of Industry Experts and grace a career as a Data Scientist/Analyst by acquiring esteemed recognition from IIT-M & certification approved by NASSCOM.

Why GUVI's Data Science Class?

- **175%** Highest Hike
- **₹21 Lakhs** Highest Salary
- **40% Hike** in Demand
- **200+ Hiring** Partners
- **360+** Best Industry Experts
- **100% Job** Placement Support
- One day Campus Visit

Data Science Course

Weekdays/Weekend LIVE Classes
↓
Hands-on Industry Projects + Bi-weekly Hackathons
↓
Technical Mentorship by Industry Experts + Periodic developer Meetups
↓
Practise on Coding Practise Platforms CodeKata, WebKata, MicroArc
↓
Mock interviews + Compensation Management
↓
Placement Interview Eligibility Test(by Mentors)
↓
Proceed to Placement Phase

Placement Phase

Tech-Skilled DS Learner
↓
Enters Placement Window
↓
Exclusive Skill-based Job Notifications
↓
Company-specific Expectations Setting (Hits Interviews Arranged)
↓
Tech-Guidance + 100% Placement Support

Top skills you'll learn!

- Best Statistical programming language skills with Python.
- Excellent database querying skills.
- Good understanding of Analytical tools & Statistics.
- Conceptual clarity towards Predictive performance & algorithm optimization.
- Master data visualization & communication skills.

Technologies covered

Data Science, Python, MongoDB, Tableau, Power BI, NLP, Machine Learning, Data Visualization, Computational Thinking, NumPy, Pandas, & more....

Program Curriculum

Module-1

We will go through the basics of python with all essential beginner friendly concepts of python programming like datatypes, loops, data structures and functions, followed by assessments and assignments

Python - Basics

- Why Python
- Python IDE
- Hello World Program
- Variables & Names
- String Basics
- List
- Tuple
- Dictionaries
- Conditional Statements
- For and While Loop
- Functions
- Numbers and Math Functions
- Common Errors in Python

Module-2

Since we have the essential basics of python we will see some advanced concepts like Comprehensions, File handling, Regular Expressions, Object-oriented Programming, Pickling and many more essential concepts.

Python - Advanced

- Functions as Arguments
- List Comprehension
- File Handling
- Debugging in Python
- Class and Objects
- Lambda, Filters and Map
- Regular Expressions
- Python PIP
- Read Excel Data in Python
- Python MySQL
- Iterators
- Pickling
- Python JSON

Module-3

We will explore the need for Algorithmic Thinking and the necessity of efficient coding, we will drive through Data Structures and Algorithms along with Memory Management Techniques

Algorithmic thinking with Python

- Introduction to algorithmic Thinking
- Algorithm Efficiency and time complexity
- Example algorithms - binary search, Euclid's algorithm
- Data structures - stack, heap and binary trees
- Memory Management/Technologies
- Best Practices - Keeping it simple, dry code, naming Conventions, Comments and docs.
- Assessment

Module-4

Since we need to handle huge amounts of data, we will be implementing data handling techniques with Pandas library. And we will explore the different miscellaneous functions of Pandas library in detail.

Data handling in Python - Pandas & MongoDB

- Introduction to Pandas
- Series Data Structure - Querying and Indexing
- DataFrame Data Structure - Querying, Indexing and loading
- Merging data frames
- Group by operation
- Pivot table
- Date/Time functionality
- Example: Manipulating DataFrame

Module-5

Now we are going to explore the popular NoSQL databases like MongoDB and their importance as well as their key components. We will perform several operations like CRUD operations, etc.

MongoDB

- No Schema
- Install MongoDB
- How MongoDB Works?
- Insert First Data
- CRUD Operations
- Insert Many
- Update and Update Many
- Delete and Delete Many

Module-6

We will continue with MongoDB beyond basics like embed documents, dealing with types of data, schema types and Data relationships in MongoDB, etc.

MongoDB - Continued

- Diving Deep into find Difference between update and update many
- Projection
- Intro to Embed Documents
- Embed Documents in Action
- Adding Arrays
- Fetching Data From Structured Data
- Schema Types
- Types of Data in MongoDB
- Relationship between data's
- One to One using Embed Method
- One to One using ReferenceMany
- One to Many Embed
- One to ManyReferenceMethod
- Assessment - MongoDB

Module-7

We will go through Probability and Statistics whereas they are key to understand, process and interpret the vast amount of data, we will deal with the basics of probability and statistics like Probability theory, Bayes theorem, distributions etc and their importance. Besides that, we will do hands-on with Numpy upon those concepts

Probability and Statistics with Numpy

- Why counting and probability theory?
- Basics of sample and event space
- Axioms of probability
- Total Probability theorem and Bayes Theorem
- Random variables, PMF and CDF
- Discrete Distributions - Bernoulli, Binomial and Geometric
- Expectation and its properties
- Variance and its properties
- Continuous Distributions - uniform, exponential and normal
- Sampling from continuous distributions
- Simulation techniques - simulating in NumPy
- Assessment

Module-8

We will continue with statistics and probability and we will deal with descriptive and inferential statistics along with Hypothesis testing and a lot of other relevant statistics methods

Probability and Statistics with Numpy - Continued

- Inferential statistics - sample vs population
- CLT and it's proof
- Chi-squared distribution and its properties
- Point and Interval Estimators
- Estimation technique - MLE
- Interval Estimator of μ with unknown σ
- Examples of estimators
- Hypothesis Testing - I
- Hypothesis Testing - II
- Hypothesis Testing - III
- Assessment

Module-9

Data Visualization is used to understand data in a visual context so that the patterns, trends and correlations in the data can be understood. We will do a lot of visualization with libraries like Seaborn, Matplotlib etc in turn that leads to effective storytelling.

Data Visualisation in Python (Matplotlib, Seaborn)

- Read Complex JSON files
- Styling Tabulation
- Distribution of Data - Histogram
- Box Plot
- Data Visualization - Recap
- Pie Chart
- Donut Chart
- Stacked Bar Plot
- Relative Stacked Bar Plot
- Stacked Area Plot
- Scatter Plots
- Bar Plot
- Continuous vs Continuous Plot
- Line Plot
- Line Plot Covid Data
- Assessment

Module-10

Plotly's Dash is a famous open-source data visualization library which we will be using to build custom data visualization projects which allows for better storytelling.

Data Visualisation with Plotly's dash

- Dash by Plotly setup
- Dash core components
- Style our Dash Application
- Callbacks,
- Adding interactivity to our Dash Apps using Callbacks

Module-11

It is always needed to analyze the data and preprocess it, since the real-world data is not always industry ready, so in this week we will be dealing with a lot of data cleaning and Exploratory data Analysis techniques which is a very crucial stage for any data science project

Data Engineering with Python

- Handling missing data
- Techniques to impute missing values
- Encoding the data
- Outlier detection and correction
- Meaningful data transformation
- Assessment

Module-12

Real-World data will not always be of the numerical form, we need to know how to handle a lot of other forms of data like image and text. In this week we are going to see all essential data analysis techniques on image and text data.

Data Analysis on Image and text data

- How computers process and understand images, Pixel
- Basic Properties of Images
- Greyscale, Processing Pixel Values
- Masking
- Image Processing
- Text data preprocessing
- Cleaning Text Data
- Exploratory Data Analysis on Image and text data
- Assessment

Module-13

We are going to explore the need for machine learning and its types, Algorithms when to use and how to use essential mathematical intuition along with Evaluation metrics. We will see in detail about regression algorithms.

Machine Learning with Sklearn

- Introduction to machine learning
- Expert systems and 6 Jars
- Supervised Learning - Regression and Classification
- Evaluation metrics and measuring accuracy
- Introduction to regression
- Interpreting models
- Feature selection
- Regularisation - Ridge and Lasso
- Assessment

Module-14

We are going to explore the need for machine learning and its types, Algorithms when to use and how to use essential mathematical intuition along with Evaluation metrics. We will see in detail about regression algorithms.

Machine Learning with Sklearn - Continued

- Introduction to classification
- Evaluation metrics - TP, FP and AUC
- Classification using logistic regression
- Classification using KNN
- SVM
- Assessment

Module-15

We are going to explore classification algorithms like tree-based algorithms in detail like how to interpret trees, pruning and ensemble methods like bragging and boosting, etc.

Machine Learning with Sklearn - Continued

- Introduction to decision trees
- Building, pruning and interpreting trees
- Ensemble techniques - Bagging and boosting
- Random forests
- Boosted trees - Gradient boosting
- Assessment

Module-16

After dealing with a lot of Supervised Machine Learning Algorithms we will compare and get to know when to use what, Besides that we will deal with the do's and don'ts while training an ML model.

Machine Learning with Sklearn - Continued

- Comparison of supervised techniques - when to use what?
- Do's and Dont's while training ML models
- Handling imbalanced data
- Undersampling
- Oversampling
- Other methods - ROSE, SMOTE, etc.
- Assessment

Module-17

Now we will explore Unsupervised learning algorithms, why unsupervised ?, when to use them and as well as the essential mathematical intuition

Machine Learning with Sklearn - Continued

- Introduction to unsupervised learning
- Market Basket Analysis
- K means algorithm
- Assessment

Module-18

We are going to explore Natural Language Processing (NLP). At first, we will see Syntactic analysis or parsing to analyze text using basic grammar rules to identify sentence structure, how words are organized, and how words relate to each other. Some of them are Tokenization, Part of Speech Tagging (PoS Tagging), Lemmatization and Stemming, Stop word removal etc

Natural Language Processing

- Syntactic Analysis
- Tokenization
- Part of Speech Tagging (PoS Tagging)
- Lemmatization and Stemming
- Stop word removal

Module-19

Now we will see that Semantic analysis focuses on capturing the meaning of the text. First, it studies the meaning of each individual word (lexical semantics). Then, it looks at the combination of words and what they mean in context. Some of the sub-tasks of semantic analysis which we are going to explore are Word sense disambiguation, Relationship extraction. Besides that, we will explore Sentiment Analysis, Text extraction etc

Natural Language Processing - Continued

- Semantic Analysis
- Word sense disambiguation
- Relationship extraction
- Sentiment Analysis, Text extraction

Module-20

This whole week we are going to work on industry projects which are currently in demand under the guidance of industry experts

Putting it together - Solving DS problems

- Case Study - I : Credit Card Fraud detection
- Case Study - II : Airline Customer segmentation
- Case Study - III : Product recommendation engine

Module-21

Eventually, it's time to attend the mock interviews which will be conducted by the industry experts like Data scientists, IIT professors and renowned HR's in order to mould you in every area possible

Mock Interviews

Final Project and Course wrap-up

20+ projects with Industry experts mentorship

Oil Price Prediction
A time-dependent model capable enough to take the modeling decisions predicting around three months price of the crude oil. Sound understanding of Data engineering and Data visualizing. Hands-on experience on Model fitting and Model Validation.

COVID-19 Data Visualization
A pictorial representation of the current COVID-19 data on a map, presenting a quick update on the situation. Collect and visualize a map version of the coronavirus outbreak. Agile know-how on the crisis, pinned with a glimpse at its history.

Gender and Age Detection
A model that can detect any human being's age & gender through analyses of single face detection via an image. Developed on the UTK Face dataset. It is widely used in applications like access control, human-computer interaction, law enforcement, etc.

Hear it from our learners

Vishally
Tejas Samanthapudi
Gokila
Gokul
Gokak Mohd Ishtiyaque
Sonia kola
Shubham Nehete
Diliban Sibi
Sridharan K
Anbazhagan

Instructors

Learn from India's top Industry Leaders

- Dr Sanatan Sukhija — Doctorate - IIT-Ropar
- Dr Yayati Gupta — Doctorate - IIT-Ropar
- Mr Sudarshan S R, — Associate Professor, IIT-Ropar
- Mr Koushik Krishnan — Data Science Analyst at Credit Suisse
- Mr Suman Auth — International Mentor for NYAS, USA
- Mr Revanth Guthala — Lead Data Analyst, Yulu
- Mr Bala Chandar — Data Scientist, US-based client
- Mr Abhishek — Data Scientist, Bosch

Our Placements

"I got a 57% Hike, Thank you GUVI Team" — Sonia Kola, Data Scientist
Watch Video

"Every topic was covered from scratch" — Rakesh, Python Developer
Watch Video

"GUVI helped me in advancing my Career" — Divyansh Chaudhary, Intern Data Analyst
Watch Video

Program Details

PROGRAM DURATION AND FORMAT
3-Month Weekday/5-Month Weekend

PROGRAM START DATES
Please contact our Data Science coordinator
Deepak: +91-9344419057

PROGRAM FEE

Total Course Fee	₹78,800
Pre-BootCamp Booking Fees	-₹5000
Remaining Fee	₹73,800

Now become a proficient Data Scientist/ Analytics at **Affordable Installments!** Master Data Science at just **₹6,800** /Month

AVAILABLE EMI OPTIONS
Upto 12 Months

ELIGIBILITY

- **No - Eligibility / Restrictions!!!**
- Any interested individual who is aspiring to get a job in the IT industry as a **Data Scientist**
- Working Professionals who wish to **switch** their **career into Data Science.**

"Gain Proficiency in Data Science from GUVI & Become recruiters TOP PICK!"

Begin your Skill Development Journey Today!

For further information:
deepak@guvi.in
+91 9344419057

GUVI
IITM Research park - phase 2 module #9, 3rd floor, D block, Kanagam Rd, Tharamani, Chennai, Tamil Nadu, India. 600113