



Data Science & Machine Learning

Career Path

Blueprint



Aksadur Rahman

Adjunct Faculty
United International University (UIU)
Data Science Consultant,
Bdjobs.com Ltd
Lead Trainer of EDGE Projects,
ICT Division, BCC

Kazi Hasib

Founder, CEO
Inception

Sabbir Hossain Rossi

Senior BI Analyst at a
German Retail Company

Araf Mustavi

Analyst, IDT Operations
BAT

Blueprint

রেগুলার লাইফে মেশিন লার্নিং আর ডেটা সাইন্সের ডিমাল্ড কিন্তু দিন দিন বেড়েই চলেছে। বিশ্বের টপ টেকনোলজিগুলোর মধ্যে **ডেটা সায়েন্স** এবং **মেশিন লার্নিং (ML)** আজ সবচেয়ে বেশি ডিমাল্ডেড। এই ডিমাল্ড ফুলফিল করার জন্য ইন্টারেক্টিভ কেয়ার্স launch করতে যাচ্ছে **ডেটা সায়েন্স** এবং **মেশিন লার্নিং (ML)** ক্যারিয়ার পাথের প্রথম ব্যাচ!

ওয়ার্ল্ড ওয়াইড **ডেটা সায়েন্স** এবং **মেশিন লার্নিং**- এর ডিমাল্ড বিশাল। 2030 সালের মধ্যে গ্লোবাল AI মার্কেটের বার্ষিক গ্রোথ রেট (CAGR) 37.3% হবে বলে ধারণা করা হচ্ছে। 2027 সালের মধ্যে এই মার্কেটের সাইজ প্রায় 267 বিলিয়ন ডলারে পৌঁছাবে। এছাড়া, 2030 সালের মধ্যে A-এর মোট ইকোনমিক ইনফ্লুয়েন্স হবে 15.7 ট্রিলিয়ন ডলার। তাই এই হিউজ ফিল্ডে ক্যারিয়ার বিল্ড আপ করা হবে আপনার বেস্ট চয়েস।

ইন্টারেক্টিভ কেয়ার্স নিয়ে এসেছে **ডেটা সায়েন্স** এন্ড **মেশিন লার্নিং** ক্যারিয়ার পাথ ব্যাচ ওয়ান। এই ক্যারিয়ার পাথে আপনারা পেয়ে যাবেন একদম বেসিক থেকে শুরু করে SQL, Python Statistics, Data Visualization, Data Processing, Supervised & Unsupervised learning, Time Series Analysis সহ আরও অনেক কিছু। আর এই পুরো লার্নিং প্রসেসটাই থাকবে থিওরি আর প্রজেক্ট বেসড।

একদম জিরো থেকে শুরু করে কীভাবে এই ফিল্ডে শাইন করে জব রেডি হতে পারবেন তার কমপ্লিট গাইডলাইন পাচ্ছেন এই ক্যারিয়ার পাথে। এই ক্যারিয়ার পাথে আপনি পাচ্ছেন Database Fundamentals & SQL (BigQuery) for Data Scientists, Python Fundamentals, Statistics with Python for Data Scientists, Machine Learning with Python, MLOps, Deep Learning with Python এই সব কিছু তাও আবার দেশসেরা মেন্টর প্যানেলের কাছে।

ক্যারিয়ার পাথে যা যা শিখবেন

- ✓ Database Fundamentals & SQL (BigQuery) for Data Scientists
- ✓ Python Fundamentals
- ✓ Statistics with Python for Data Scientists
- ✓ MLOps, Deep Learning with Python

ক্যারিয়ার পাথে কী কী পাচ্ছেন



১৫০+
প্রিরেকর্ডেড ভিডিও



৬০+
লাইভক্লাস



ডেইলি ২টি
সাপোর্ট সেশন



১০০+ কোম্পানিতে
চাকরির সুযোগ



২৪+
কনসেপচুয়াল ক্লাস



মক
ইন্টারভিউ



লাইফটাইম
অ্যাকসেস



প্র্যাক্টিস
ম্যাটেরিয়ালস

ক্যারিয়ার প্যাথটি কাদের জন্যে?

- ☑ Industry-Specific Specialists
- ☑ Decision-Makers and Strategists
- ☑ Technology Enthusiasts
- ☑ Programming Aficionados
- ☑ Math and Logic Enthusiasts

ডেটা সাইন্স এন্ড AI ফিল্ডে ক্যারিয়ার ডেভেলপ করতে চাইলে এই Data Science & Machine Learning Career Path হতে পারে আপনার জন্য বেস্ট চয়েস। একদম বেসিক লেভেল থেকে শুরু করে অ্যাডভান্সড পর্যন্ত সবকিছুই শিখবেন আমাদের সাথে। নিজেকে প্রিয়ার করতে পারবেন ফ্যাবিউলাস এক ক্যারিয়ারের জন্য।

☑ ক্যারিয়ার পাথে এনরোলমেন্টের জন্য আগে থেকে কি কিছু জানা লাগবে?

১. প্রোগ্রামিং ল্যাঙ্গুয়েজ

ডেটা সাইন্স এবং মেশিন লার্নিংয়ের মূল বেইজ হলো প্রোগ্রামিং।

- **Python এবং R:** সবচেয়ে জনপ্রিয় ভাষা। সহজ সিনট্যাক্স এবং ডেটা প্রসেসিং লাইব্রেরি (NumPy, Pandas, Scikit-learn) থাকায় এগুলো সিলেক্ট করা হয়।
- **SQL:** ডেটাবেস থেকে ডেটা কালেক্ট এন্ড অ্যানালাইসিসের জন্য অপরিহার্য।

২. ম্যাথ এন্ড স্ট্যাটিস্টিকস

ডেটা সাইন্স এবং মেশিন লার্নিংয়ের অনেক কাজ গাণিতিক মডেল এবং স্ট্যাটিস্টিকাল অ্যানালাইসিসের উপর ভিত্তি করে।

- **লিনিয়ার অ্যালজেব্রা এবং ক্যালকুলাস:** মেশিন লার্নিং অ্যালগরিদমের প্রাইমারি কম্পোন্ট বোঝার জন্য প্রয়োজন।
- **স্ট্যাটিস্টিকস এবং প্রোবাবিলিটি:** ডেটার প্যাটার্ন এবং টেন্ডেন্সি অ্যানালাইসিসের জন্য মেম্বারি।

৩. ডেটা প্রসেসিং এন্ড ভিজুয়ালাইজেশন

- ডেটা ক্লিয়ার এন্ড রেডি করার (Data Wrangling) প্রসিডিউর জানতে হবে।
- **Visualization Tools:** Matplotlib, Seaborn, Power BI, বা Tableau-তে স্কিল গুরুত্বপূর্ণ, কারণ ডেটার রিপ্রেসেন্টেশন বোঝাতে ভিজুয়ালাইজেশন প্রয়োজন।

৪. মেশিন লার্নিং অ্যালগরিদমের বেসিক কন্সেপ্ট

- **Supervised ও Unsupervised Learning:** বিভিন্ন মডেল যেমন লিনিয়ার রিগ্রেশন, ক্লাস্টারিং, ডিসিশন ট্রি ইত্যাদির ধারণা থাকা প্রয়োজন।
- **ডেটা মডেলিং:** মেশিন লার্নিং মডেলের কাজ এবং লিমিটেশন বুঝতে হবে।

৫. ক্লাউড এন্ড বিগ ডেটা প্ল্যাটফর্মের সাথে পরিচিত

- **Tools:** Hadoop, Spark, বা AWS-এ কাজ করার জ্ঞান ডেটা হ্যান্ডলিংয়ের জন্য দরকার হতে পারে।

৬. সফট স্কিলস

- **Problem-Solving Skills:** ডেটা বিশ্লেষণের মাধ্যমে সলিউশন বের করতে পারা।
- **কমিউনিকেশন স্কিল:** টেকনিক্যাল ডেটা সহজ ভাষায় প্রেসেন্ট করতে পারা।

FAQ (Frequently Asked Question) ক্যারিয়ার পাথ

? ক্যারিয়ার পাথের ক্লাস কবে থেকে শুরু?

⇒ ক্লাস শুরু হবে.....

? কোর্সটি কি বিগিনার ফ্রেন্ডলি?

⇒ জি, কোর্সটি বিগিনার ফ্রেন্ডলি।

কোর্সের প্রথম মাসে, আপনি ডাটা সাইন্স এন্ড মেশিন লার্নিং এর ইম্পর্ট্যান্ট সেক্টর সম্পর্কে ধারণা পাবেন। যেমন পাইথন ফাউন্ডামেন্টালস, এক্সিকিউটেল ফাউন্ডামেন্টালস কিভাবে কাজ করে তা শেখান হবে। এছাড়া স্টাটিস্টিক্স উইথ পাইথন, মেশিন লার্নিং উইথ পাইথন, ভেরিএবেলস, কন্ডিশনাল এসেটমেন্ট, এমেলোপিএস, ডিপ লার্নিং উইথ পাইথন সহ সব ফাউন্ডামেন্টাল বিষয়গুলি জানান হবে।

? ক্যারিয়ার পাথের কোনো ফ্রি ডেমো ভিডিও আছে?

⇒ হ্যাঁ, আমাদের ইন্ট্রাকটরদের একটি ফ্রি ডেমো ভিডিও দেখতে- **এখানে ক্লিক করুন।**

? আমাদের এই ক্যারিয়ার পাথে কী কী বার লাইভ ক্লাস হবে?

⇒ "পরবর্তীতে জানিয়ে দেওয়া হবে"

? ক্যারিয়ার পাথের সবকিছুই কি বাংলায় থাকবে?

⇒ জি। আমাদের ক্যারিয়ার পাথে ডিজিটাল মার্কেটিং এর সমস্ত কিছু শিখাবেন বাংলায়।

? কোন ধরনের ক্যারিয়ার সেশন/ইন্টারভিউ প্রিপারেশন থাকবে?

⇒ প্রতিটি ইম্পরট্যান্ট টপিক শেষে আপনার নলেজ ব্যালিয়ে নিতে থাকছে মক ইন্টারভিউ, এর সাথে জব মডিউলের সময় পেয়ে যাচ্ছেন একটি ইন্টারভিউ হ্যান্ডবুক সম্পূর্ণ ফ্রি। এছাড়াও ক্যারিয়ার সেশন থাকছেই। তাই ক্যারিয়ার পাথে ভর্তি হওয়ার পর আপনাকে জব রেডি করে তোলার জন্য যা যা প্রয়োজন সবকিছু নিয়ে আমরা রেডি।

? প্রিরেকর্ডেড ভিডিও কিংবা লাইভ ক্লাস দেখে কোন কনসেপ্ট যদি না বুঝি তাহলে কী করবো?

⇒ ডেইলি ৪টি করে সাপোর্ট সেশন থাকবে। সাপোর্ট ক্লাসে জয়েন করে যে প্রব্লেমগুলো ফেইস করছেন সেগুলো সল্ড করতে পারবেন।

? ক্যারিয়ার পাথ শেষে জব প্লেসমেন্টের ব্যবস্থা আছে?

⇒ হ্যাঁ, আমাদের এই ৬ মাসের ক্যারিয়ার পাথের শেষেই থাকছে জব প্রিপারেশন মডিউল। এই মডিউলে আপনাকে CV আর পোর্টফোলিও বানানো থেকে শুরু করে বিভিন্ন ইন্টারভিউ হ্যান্ডেল করা পর্যন্ত সবকিছু আপনাকে শেখানো হবে এক্সটেনসিভলি। পাশাপাশি আমাদের অ্যাফিলিয়েটেড ১০০+ কোম্পানিতে পাচ্ছেন জব আর ইন্টার্নশিপ প্লেসমেন্ট অপরচুনিটি।

? কীভাবে জব প্লেসমেন্টে হেল্প করা হবে?

⇒ দেশ ও দেশের বাইরের প্রায় ১০০টিরও বেশি কোম্পানির সাথে আমরা সরাসরি ইনভল্ভড আছি, প্লেসমেন্ট নিয়ে। অন্যদের মতো শুধু মাত্র রেকমেন্ডেশন কিংবা রেফারেই আমরা থেমে নেই। বরং আমাদের অ্যাফিলিয়েটেড কোম্পানিগুলোর হয়ে আমরাই তাদের সিভি শর্টলিস্ট করে দেই, অ্যাসেসমেন্ট রাউন্ডও আমরা কন্ডাক্ট করে থাকি। সিভি শর্টলিস্ট থেকে শুরু করে প্লেসমেন্ট পর্যন্ত পুরো প্রসেসেই আমরা ডিরেক্টলি ইনভল্ভড থাকি।

❓ **ক্যারিয়ার পাথে এনরোল করতে কত টাকা লাগবে?**

☞ ক্যারিয়ার পাথের প্রাইস ৳৫০০ টাকা।

❓ **ক্যারিয়ার পাথটি কীভাবে কিনবো?**

☞ ওয়েবসাইটে অ্যাড টু কার্ট থেকে কিনতে পারবেন, তাছাড়াও রেজিস্ট্রেশন ফর্মের মাধ্যমে ম্যানুয়ালিও কিনতে পারবেন।

ওয়েবসাইট লিংক: **এখানে ক্লিক করুন**

গুগল রেজিস্ট্রেশন ফর্ম: **এখানে ক্লিক করুন**

❓ **কীভাবে পেমেন্ট করবো?**

☞ বিকাশে 01727659043 / 01763881476 -এই ২টি নম্বরের সেন্ড মানি অপশনের মাধ্যমে পেমেন্ট করতে পারবে। আর নগদে 01727659043-এই নম্বরের সেন্ড মানি অপশনের মাধ্যমে পেমেন্ট করতে পারবে। ওয়েবসাইট থেকে কিনলে ডিসা কার্ড, মাস্টার কার্ড, বিকাশ, নগদ যে কোন মাধ্যমেই পেমেন্ট করা যাবে। এছাড়াও আপনি দেশের বাইরে থেকে পেমেন্ট করতে চাইলে Stripe -এ পেমেন্ট করতে পারবেন।

❓ **ক্যারিয়ার পাথের ভর্তি কি বাতিল করা সম্ভব?**

☞ না, ক্যারিয়ার পাথে এনরোলমেন্টের পরে তা আর বাতিল করার অপশন নেই।

কোনো টেকনিকাল সমস্যা কীভাবে রিপোর্ট করবো?

☐ ওয়েবসাইটের হেল্প অপশনে টেক্সট দিতে পারেন কিংবা আমাদের ফেইসবুক পেইজে আমাদের রিচ করতে পারেন। এছাড়াও যেকোনো প্রয়োজনে 01958622155 এই নাম্বারে কল করুন।

কেনার পর ক্যারিয়ার পাথটি কীভাবে শুরু করবো?

☐ ক্যারিয়ার পাথটি কেনার পরে আপনাকে এনরোলমেন্ট লিংক ও প্রয়োজনীয় ইনিশিয়াল ও ইন্সট্রাকশন পাঠিয়ে দেওয়া হবে; সেগুলো ফলো করে মডিউল অনুযায়ী পেয়ে যাবেন ক্যারিয়ার পাথের প্রিরেকর্ডেড ভিডিও।

ক্যারিয়ার পাথ কেনার পর রিফান্ড পাবো কীভাবে?

☐ না, ক্যারিয়ার পাথ কিনে ফেলার পরে আমরা কোনো রিফান্ড অপশন অফার করছি না।

ক্যারিয়ার পাথের মেয়াদ কতদিন?

☐ ক্যারিয়ার পাথের মেয়াদ ৬ মাস, তবে রেকর্ডেড ভিডিওগুলোতে থাকবে লাইফটাইম অ্যাকসেস।

ক্যারিয়ার পাথে যেকোনো সমস্যায় পড়লে কল করুন আমাদের কাস্টমার কেয়ার নাম্বারে: 01958622155



Data Science & Machine Learning

Career Path

Detailed Career Path Outline

Python Fundamentals

Week 1

- **Class 1: Basics of Python Programming Language**
 - Introduction to Python Programming Language.
 - Introduction to Google Colab & Jupyter Notebook for Writing and Executing Python Codes.
 - Getting used to the User Interface (UI) of Jupyter Notebook/Google Colab.
 - Python Data Types: Numeric, Strings, Boolean, List, Dictionary, Tuple, Set, None.
 - Iterables & Mutability; List vs. Tuples.
- **Class 2: Variables, Type Conversion, Operators & Methods**
 - Assigning & Overwriting Variables; Naming Convention for Variables.
 - Type Conversion.
 - Operators: Arithmetic & Logical/Comparison.
 - Numeric & String (Text) Functions.
 - String Indexing & Slicing.
 - String Methods.
 - F String.
- **Class 3: Conditional Statement, User Defined Functions & Iteration**
 - IF & ELIF Statements.
 - How to Define/Create a Function in Python?
 - How to Create a Function with Parameter/Argument(s)? Function within a Function.
 - Use Conditional Statements and Functions Together.
 - Using FOR & WHILE Loops.
 - Combining Conditional Statements & Loops.
 - Combining Conditional Statements, Functions & Loops.
 - How to Iterate over a Dictionaries.

- **Learning**

*** By the end of the Python Fundamentals module, students will have built a solid foundation in core Python programming concepts and practices. This is critical for anyone pursuing a career in data science or as a machine learning (ML) engineer, as Python is the primary language used in these fields.

Outcomes:

- Mastery of Data Types.
- Understanding Python Basics.
- Variables and Type Operations.
- String Handling and Methods.
- Control Flow and Logic.
- Functions and Modular Programming.
- Iterative Processing.

Database Fundamentals & SQL for Data Scientists

Week 2

- **Class 4: Fundamentals of DBMS-I**

- What is a Relational Database Management System?
- ACID Property.
- How is data stored in a relational database?
- Concept of Normalization.
- OLTP vs. OLAP.
- Database vs. Data Warehouse vs. Data Lake/Lakehouse.
- What is a NoSQL Database and BASE Property?
- Difference between Relational and NoSQL Database.
- Which one should you choose in which case?
- For BI Solution, what should you choose to store your data?
- What is Snowflake and what advantage does it provide? (Micro-Partitioning)

- **Class 5: Fundamentals of DBMS-II**

- Installing MySQL Locally.
- MySQL Workbench UI Tour.
- Concepts of Database, Schema, Table and Fields.
- Types of SQL Commands: DDL, DML, DCL, TCL, DQL.
- What are the Commands under each of these categories and what does each of these commands do?
- As a Data Analyst & Data Scientist which Commands do you only need to Master?
- **Constraints:** Primary Key, Foreign Key, Not Null, Unique, Check, Default.
- **Data Types:** DATE, DATETIME, VARCHAR, INT, NUMBER, FLOAT, Auto Increment.

- **Class 6: Designing a Relational Database in MySQL & Bulk Insertion**

- Designing and Creating Database, Schema, Tables.
- Masterclass on DDL, DML, TCL & DQL Commands/Statements.
- Bulk Insert of Data in MySQL.
- Error Handling during Bulk Insertion.

Week 3

- **Class 7: Exploratory Data Analysis using SQL**

- Basic Queries: SELECT, FROM, WHERE, LIKE, ILIKE, IN, DISTINCT, BETWEEN, GROUP BY, ORDER BY, LIMIT, OFFSET, ALIAS.
- Aggregate Functions: COUNT, SUM, AVG, MIN, MAX.
- Difference between WHERE and HAVING Clause.
- Some built-in Functions: EXTRACT, DATE_PART, TO_DATE, TO_CHAR.
- CASTING, SUBSTRING, POSITION, COALESCE, NULLIF.
- Removing Duplicates.

- **Class 8: JOINING, UNION & CASE WHEN**

- SQL Join: Left, Right, Inner & Full Join.
- Be aware of **Cross Join!**
- UNION.
- **SQL Code Order of Execution.**
- Wide use of **CASE WHEN** Statement during data Cleaning, Analysis & Feature Engineering.

- **Class 9: Subqueries, CTEs & RFM Segmentation**

- Use of Subqueries & CTEs in long SQL Codes.
- Segmenting Customers based on their Recency, Frequency & Monetary value using SQL.
- Between Subqueries & CTEs one is more efficient?

Week 4

- **Class 10: Window Functions**

- Window Functions, the most widely used SQL Commands used by Data Analysts & Data Scientists.
- Window Functions: RANK, DENSE_RANK, ROW_NUMBER, LEAD, LAG, FIRST VALUE, AGGREGATE WINDOW FUNCTION, FRAME SPECIFICATION, WINDOW CHAINING.

- **Class 11: Cohort Analysis & Creating VIEWS**

- Cohort Analysis using Transactional Customer Data.
- Customer Lifetime Value, Retention, and Churn Rate Calculation.
- Use Case of Views; How to Create Views?

- **Learning**

*** By the end of this module, Students would be able to confidently perform any sort of Data Analysis and Reporting for different Departments in any Organisations using SQL. They would be well-equipped to understand any complex queries; validating data and help any Business by generating important KPIs.

They would also be able to help Businesses by Performing Complex Analyses like Cohort and RFM Segmentation. They would understand the importance of VIEWS, Stored Procedures, Triggers which would help them when they would be working on Python/Power BI/Tableau, etc. with Database/Data Warehouse as Data Sources.

Statistics with Python for Data Science

Week 4

- **Class 12: Descriptive Statistics**

- Types of Variables.
- Level of Measurement.
- Frequency Distribution.
- Getting Familiar to Different Statistical Charts.
- Which Charts/Tables/Frequency Distribution to choose for which Type of Variables?

Week 5

- **Class 13: Measures of Central Tendency & Dispersion**

- Measures of Central Tendency [Mean, Median, Mode], and Location [Quartile, Decile, Percentile].
- Which Measure is the Best one in which situation?
- Measures of Dispersion [Range, Variance, Standard Deviation, Coefficient of Variation].

- **Class 14: Shape Characteristics & Causality**

- Measures of Shape Characteristics: Skewness, Kurtosis.
- Correlation Analysis.
- Concepts of Regression Analysis.
- What is causality and what is the difference between Correlation and Regression Analysis?

- **Class 15: Inferential Statistics**

- Normal Distribution & Standard Normal Distribution.
- Central Limit Theorem and where this is applied?
- Estimators and Estimates.
- Confidence Interval and Margin of Error.
- Hypothesis Test Concepts: Null & Alternative Hypothesis, Type 1 & Type 2 Error, Test Statistics, Level of Significance & p-value.

Week 6

- **Class 16: Parametric Hypothesis Test**

- Assumptions for Choosing a Parametric Hypothesis Test.
- T-tests, F-test, Chi-squared Test.
- Executive the Hypothesis Testing in Python and Interpreting the Results.
- When to choose which one?

- **Class 17: A/B Testing & Non-parametric Hypothesis Tests**

- What is A/B Testing and why is it so popular among the Data Scientists?
- Performing A/B Testing in a Dynamic way in Python.
- When to pick Non-parametric Hypothesis Tests?
- Mann-Whitney U Test (Wilcoxon Rank-Sum Test), Wilcoxon Signed-Rank Test, Kruskal-Wallis Test, Kolmogorov-Smirnov Test (for Normality Checking).

- **Learning**

*** By completing the Statistics with Python for Data Science module, students will gain a robust understanding of essential statistical concepts and how to apply them to real-world data science problems using Python. This module aims to bridge statistical theory and practical implementation, which is critical for data analysis and predictive modelling.

Overall Achievement:

By the end of this module, students will have the ability to:

- Analyse and interpret data using descriptive and inferential statistical methods.
- Apply Python to execute statistical analyses, enhancing their data science skills.
- Select and implement the appropriate hypothesis tests for various data types and scenarios.
- Understand the role of statistical analysis in data-driven decision-making processes, preparing them for roles in data science and machine learning.

Data Analysis using NumPy, Pandas

Week 6

- **Class 18: Data Analysis using NumPy**

- A brief introduction
- Jupyter notebook Installation & exploring.
- Google Colab
- NumPy arrays, arrange, linspace
- Array methods and attributes.
- Indexing, slicing
- Broadcasting
- Boolean masking
- Arithmetic Operations
- Universal Functions

Week 7

- **Class 19: Data Analysis using Pandas -1**

- Pandas Introduction.
- Pandas Data Structures - Series
- Pandas Data Structures – DataFrame
- Creating DataFrame
- Grab data (column wise)
- Grab data (row wise)
- Grabbing an element or a sub-set of the dataframe
- Adding new column
- Deleting the column
- Boolean mask
- reset_index(), set_index(), head(), tail(), info(), describe()

- **Class 20: Data Analysis using Pandas -2**
 - Handling Missing Data
 - Data Wrangling
 - Combining, Merging, Joining, Group by
 - Useful Methods and Operations
 - `unique()`, `nunique()`, `value_counts()`, `count()`, `sort_value()`
 - Different Ways of Creating DataFrame
- **Class 21: Exploratory Data Analysis (EDA) using Pandas**

Week 8

- **Class 22: Project on Data Analysis using Pandas**
- **Learning**

*** NumPy and Pandas are two fundamental Python libraries that form the cornerstone of data analysis and manipulation. Learning these libraries equips you with the tools to efficiently work with large datasets, extract meaningful insights, and make data-driven decisions.

Key Purposes: Data Cleaning and Preparation:

- Handling Missing Data: Imputing missing values, removing outliers, and handling inconsistencies.
- Data Formatting: Converting data types, standardising formats, and cleaning text data.
- Data Integration: Merging and joining datasets from various sources.
- Efficient Analysis: Structured and organised data facilitates efficient analysis and visualisation.

Data Visualization using the most widely used Python library

- **Class 23: Data Visualization using Matplotlib**

- Basic Plotting
- Creating multiple plots on the same canvas.
- Matplotlib "Object Oriented" approach
- Creating a figure and a set of subplots
- Saving figures
- Decorating the figures

- **Class 24: Data Visualisation Using Pandas**

- Histograms
- Style Sheets
- Area plot
- Bar plot
- Line plot
- Hexbin
- Scatter plot
- Box Plot
- Pie Plot

Week 9

- **Class 25: Data Visualization using Seaborn**

- Distribution Plots: Visualising the distribution of a dataset
- Joint plot, Pair plot, KDE plot
- Plotting with categorical data
- Swarm plot and Strip plot
- Box plot and Violin plot
- Bar plot and Point plot
- Matrix Plots, Heatmap

- **Class 26: Data Visualization using Plotly**

- Histogram
- Bar Chart
- Stacked Bar Chart
- Bubble Chart
- Donut Chart
- Scatter Chart
- Line Chart
- Spider Chart
- Sunburst Chart
- Waterfall Chart

- **Class 27: Exploratory Data Analysis (EDA) with data Visualisations**

- **Learning**

***Data visualisation is a powerful technique that helps us understand and communicate complex information through visual representations. Matplotlib, a versatile Python library, is a key tool in this process.

Key Purposes of Data Visualization:

Exploratory Data Analysis (EDA):

- Identifying Patterns: Spotting trends, seasonality, and outliers.
- Understanding Distributions: Visualising the shape and spread of data.
- Detecting Anomalies: Identifying unusual data points.

Communicating Insights:

- Storytelling: Creating compelling narratives from data.
- Simplifying Complex Ideas: Making complex data accessible to a wider audience.
- Highlighting Key Findings: Emphasising important conclusions.

Machine Learning Data Processing, Transforming and Feature Engineering

Week 10

- **Class 28: Introduction to Machine Learning**
 - Introduction to ML - What, Why
 - Machine Learning Applications
 - Supervised Learning
 - Unsupervised Learning
 - What is a Machine Learning Model?
 - Training data and Test data
 - Splitting Data, Train set & Test set
 - Underfitting and Overfitting
 - KFold Cross Validation
- **Class 29: Data processing, Transforming, Extractions**
 - Feature Scaling Theory
 - Feature Scaling - Hands-on
 - Feature Extractions
 - Image to Pixel using CV2
 - Resize, bitwise not and Pixel to Photo
 - How to prepare a dataset using photos
 - Linear Discriminant Analysis (LDA)
 - Bag of words, vectorization
 - Bag of N-grams

- **Class 30: Feature Selection, Outlier Detection and Removal**

- Feature Importance, Feature Selection
- Label Encoding
- Ordinal Encoding
- One Hot Encoding
- Outlier Detection and Removal Using IQR
- Outlier Detection and Removal using Standard Deviation
- Outlier Detection and Removal using Z-Score

- **Learning**

***By the end of this module, students will have the ability to

- Features Scaling: Scale features as per model requirements (e.g., Normalizations Standardization and Absolute).
- Selecting Relevant Features: Identifying the most important features for a specific task.
- Transforming Features: Applying transformations to improve model performance (e.g., log transformation, square root transformation).
- Encoding Categorical Variables: Converting categorical variables into numerical representations (e.g., one-hot encoding, label encoding).
- Outlier Detection & Removal: Finding and removing outliers

Why is Data Processing Important?

- Improved Model Performance: Well-prepared data leads to more accurate and reliable models.
- Enhanced Insights: Clean and transformed data allows for deeper understanding of the underlying patterns.
- Efficient Analysis: Structured and organised data facilitates efficient analysis and visualisation.
- Better Decision Making: Data-driven decisions based on accurate and reliable insights.

Supervised Learning (Regression)

Week 11

- **Class 31: Regression**

- Simple vs. Multiple Linear Regression Model.
- Model Assumptions
- Multiple Linear Regression Model in Python using Statsmodels (Statistical Approach for Estimating the Causality/Dependency).
- Interpreting the Model Outputs.
- R-squared and Adjusted R-squared.
- Hypothesis Tests of Model Coefficients & Goodness of Fit of the Model.
- Residual Plots.

- **Class 32: Regression**

- Multiple Linear Regression Model in Python using Scikit-learn (Machine Learning Approach for Prediction Purpose)
- Splitting Data into Train-Test-Validation Set
- Prediction
- Model Evaluation Metrics: R-squared, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), etc.
- Feature Importance Plot

- **Class 33: Regression**

- Introduction to Regularized Regression Methods
- Ridge Regression Concept
- LASSO Regression Concept
- Elastic Net Regression Concept
- Which one to choose?
- Fitting a Elastic Net Regression with ElasticNetCV

Supervised Learning (Classifications)

Week 12

- **Class 34: Class Imbalance, Confusion Matrix & Bias Variance Trade off**
 - Class Imbalance Problem
 - Class Imbalance Problem Handling
 - Confusion Matrix (Precision, Recall, f1 score)
 - Bias and Variance, Bias Variance Trade off
- **Class 35: Binary Logistic Regression Model**
 - Binary Logistic Regression Theory
 - Confusion Matrix (Precision, Recall, f1 score)
 - Binary Logistic Regression Algorithm
 - Binary Logistic Regression Pen & Paper Exercise
 - Hands-on: Model development using Binary Logistic Regression
- **Class 36: K Nearest Neighbours**
 - K Nearest Neighbors Theory
 - K Nearest Neighbors Algorithm
 - K Nearest Neighbors Pen & Paper Exercise
 - Model development using K Nearest Neighbors - Hands-on
 - Hands-on: Handwritten digits recognition using KNN

Week 13

- **Class 37: Decision Tree**
 - Decision Tree Theory
 - Decision Tree Algorithm
 - Decision Tree Pen & Paper Exercise
 - Hands-on: Model development using Decision Tree
- **Class 38: Support Vector Machines (SVMs)**
 - Support Vector Machines Theory
 - Support Vector Machines Algorithm
 - Support Vector Machines (SVMs)
 - Hyper-parameter tuning
 - GridsearchCV
 - Hands-on: Model development using Support Vector Machines
- **Class 39: Naïve Bayse Classifications**
 - Naïve Bayse Classifications Theory
 - Naïve Bayse Classifications Algorithm
 - Naïve Bayse Classifications Pen & Paper Exercise
 - Naïve Bayse Classifications
 - Hands-on: Ham Spam Detection using Naïve Bayse Classification

Week 14

- **Class 40: Project on Supervised Learning (Classification)**

Unsupervised Learning (Clustering)

- **Class 41: K Means Clustering**

- K Means Clustering Theory
- K Means Clustering Algorithm
- K Means Clustering, Elbow method
- K Means Clustering - Hands-on
- Hands-on: RFM Based Customer Segmentation using K Means Clustering

- **Class 42: Principal Component Analysis (PCA)**

- What is Principal Component Analysis?
- PCA vs LDA Analysis
- What are Principal Components?
- Principal Component Analysis (PCA) Examples
- Hands-on: Implement PCA in Python
- Predictive Modelling with PCA Components

Week 15

- **Class 43: Project on Unsupervised Learning (Clustering)**

Ensemble Learning

- **Class-44 Ensemble Learning Models - Random Forest**

- Ensemble Learning
- What is Bagging
- What is Bagged Tree
- Random Forest
- Random Forest Theory
- Random Forest Algorithm
- Hands-on: Random Forest

- **Class-45 Ensemble Learning Models – XGBoost**

- Boosting
- XGBoost Theory
- XGBoost Algorithm
- Hands-on: XGBoost

Time Series Analysis

Week 16

- **Class 46: Introduction to the Basics of Time Series Analysis**

- Introduction to Time Series Analysis
- Moving Averages
- Exponential Smoothing
- Decompositio

Class 47: Time Series Analysis with Statistical Modeling

- Different Approach to TRAIN-TEST Split for Time Series Forecasting.
- ARIMA
- SARIMAX
- AUTO ARIMA

Class 48: Time Series Analysis with Prophet library

- Forecasting Future Data using Facebook Prophet Library.
- Comparing the Results between the Statistical Approach and ML Approach.

Deep Learning (7 Classes + 1 Project)

Week 17

• Class 49: Deep Learning Fundamentals and Model Training

- Introduction to deep learning, neural network basics, and model architecture.
- Fundamentals of training neural networks, including loss functions and optimization.

Hands-on: Training a basic neural network on a simple dataset.

- **Class 50: Convolutional Neural Networks (CNNs) and Convolution**

- Introduction to CNNs, the concept of convolution, and layers used in CNNs.
- Applications of CNNs in image classification and object detection.

Practical: Building and training a CNN model on an image classification task.

- **Class 51: Hyper-parameter Tuning and Optimization**

- Techniques for tuning deep learning models and managing overfitting.

Hands-on: Optimizing a CNN or OCR model with hyperparameter adjustments.

Week 18

- **Class 52: Model Evaluation and Validation Technique**

- Evaluation metrics and validation techniques for image classification and OCR models.
- Practical: Evaluating the performance of the OCR pipeline and CNN model.

- **Class 53: Introduction to Image Processing**

- Overview of Computer Vision and its role in ML and deep learning.

Practical: Preprocessing images including resizing, filtering, and edge detection.

- **Class 54: Object Detection and Recognition**

- Introduction to Object detection and principles

Week 19

- **Class 55: Real-World Applications**

- Advanced image processing for region detection.

Practical: Detecting objects in complex images and extracting information.

- **Class 56: Project Presentation and Discussion**

- Students work on a project combining learnings
- Project presentations, peer feedback, and instructor critique.

MLOps (4 Classes)

- **Class 57: Introduction to MLOps Concepts**

- Overview of MLOps and the ML model lifecycle.
- Importance of MLOps in production environments and introduction to CI/CD in ML.

Week 20

- **Class 58: Building and Managing ML Pipelines**

- Creating an end-to-end ML pipeline, including preprocessing, training, and inference.

Practical: Using python to build a preprocessing pipeline within a full ML pipeline.

- **Class 59: Cloud Deployment of ML Models**

- Deploying a deep learning model to the cloud and creating an API endpoint.

Hands-on: Deploying a model for real-time access.

- **Class 60: Monitoring and Maintenance of Deployed Models**

- Techniques for monitoring model performance and handling model drift.

Practical: Setting up monitoring for the deployed model

Job Preparation

Extra Week taken by the Support Instructor:

- **Class 1:**

- CV Making & How to Write a Cover Letter?
- Portfolio Building.

- **Class 2:**

- Guidance on Freelancing Career
- Interview Skill Development & How to Effectively Apply for a Position.
- Roadmap for Future Ahead.

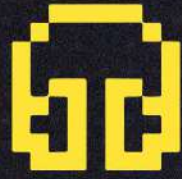
Data Science & Machine Learning

Career Path

যেকোনো সমস্যায় পড়লে

কল করুন

আমাদের কাস্টমার কেয়ার নাম্বারে:



+880 1958622155

Scan to visit our website

