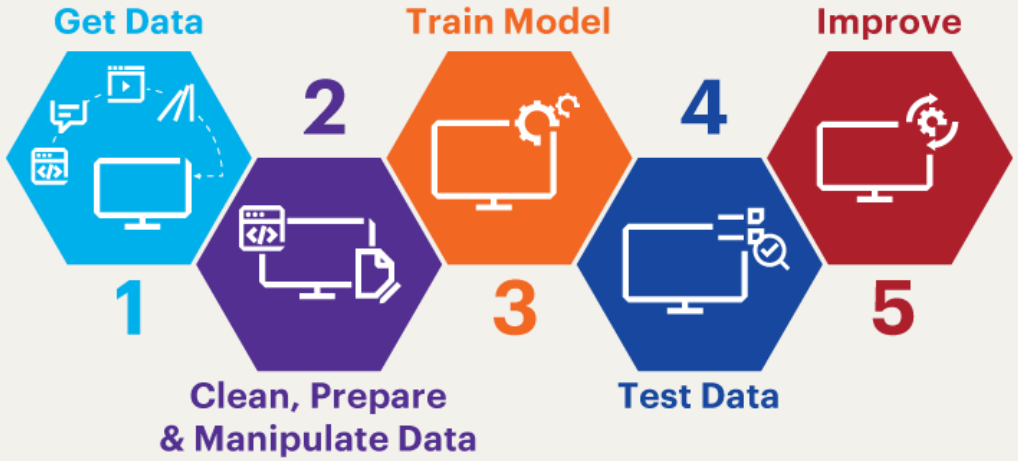
Machine Learning

Free Machine Learning Training – Session 3

Steps to Building a Machine Learning Model

- Data Gathering
- Data Preparation
- Data Wrangling
- Data Analysis
- Model Training
- Model Testing
- Deployment



Stage 1: Data Gathering

The first step in any machine learning project is collecting data. Without data, it is impossible to train a machine learning model.

Data can be obtained from various sources. quantity and quality of the collected data significantly impact the model's performance—the more high-quality data we have, the more accurate our predictions will be.
The collected data is called a **Dataset**, which we will use in the following steps.

Stage 2: Data Preparation

Once we have collected data, the next step is preparing it for training. Data is placed in a structured format, such as CSV or Excel files.

Features (columns) should be properly organized and formatted for analysis. The dataset is randomly shuffled to ensure the model learns effectively.

Stage 3: Data Wrangling (Cleaning)

Data wrangling involves cleaning and preprocessing raw data to make it suitable for machine learning models.

Common issues that need to be addressed:

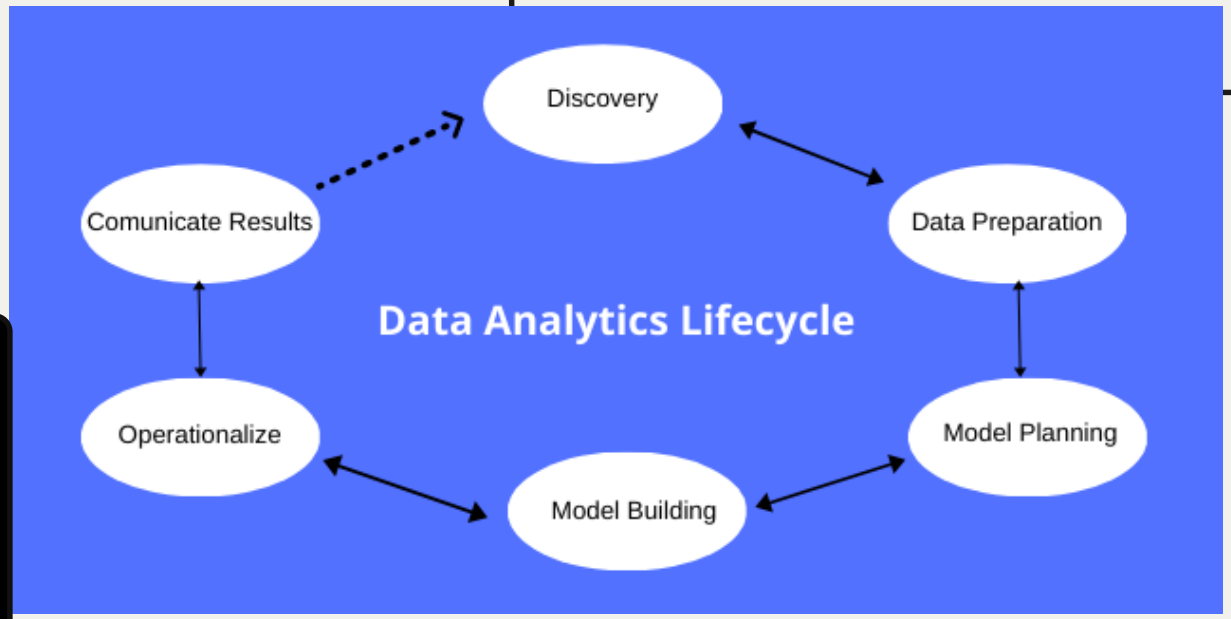
1. **Missing Values** – Some data points might be incomplete (e.g., missing age in a dataset of individuals).
2. **Duplicate Data** – Removing repeated entries to avoid redundant information.
3. **Invalid Data** – Incorrect values (e.g., alphabetical characters in a numerical column).
4. **Noise** – Irrelevant or erroneous data that can affect model performance.

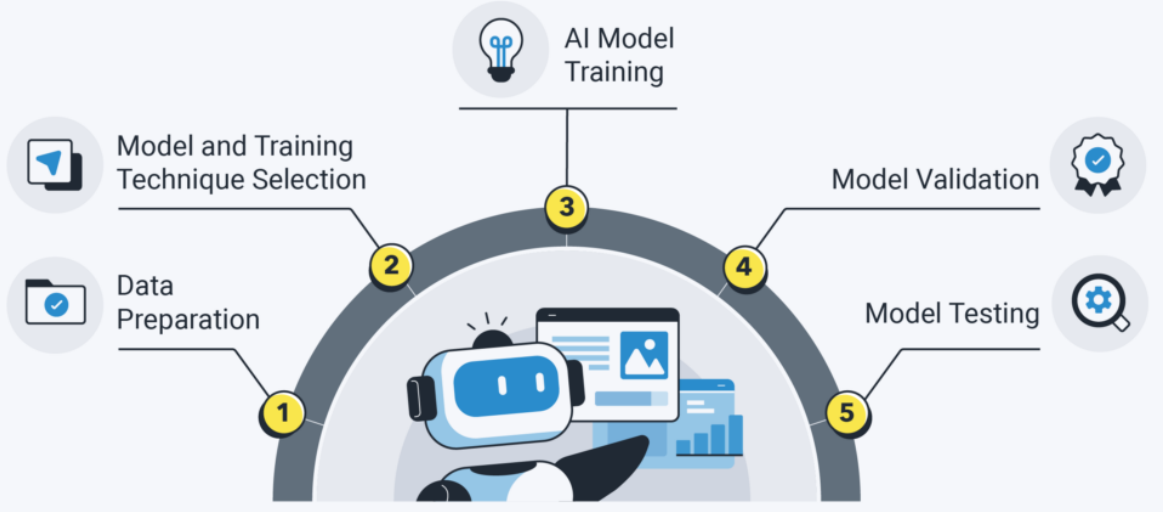
Stage 4: Data Analysis

After data is cleaned and prepared, we need to analyze it to determine the best machine learning technique to apply.

Identify whether the problem is **classification, regression, clustering**, or another type of machine learning task.

The goal is to **select the appropriate algorithm** based on the data characteristics and project objectives.





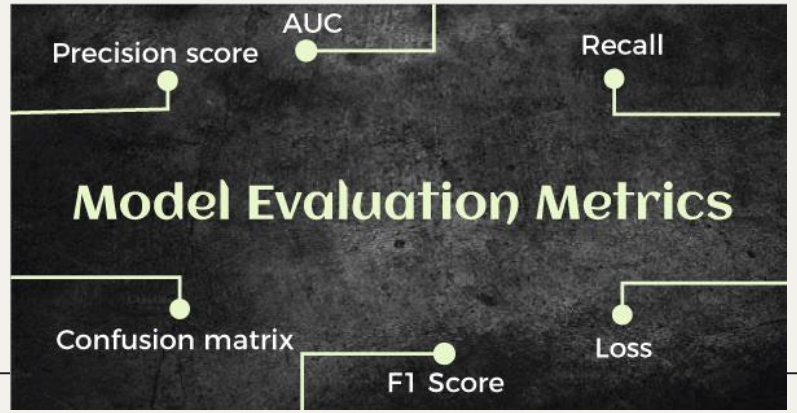
Stage 5: Model Training

Now that the data is ready, we train our machine learning model.

The dataset is fed into machine learning algorithms to learn patterns and relationships within the data.

The goal is for the model to understand the hidden rules and structures within the dataset, allowing it to make accurate predictions.

Stage 6: Model Testing



Once the model is trained, we must test it to evaluate its performance.

- **A separate portion of the dataset (test data) is used to measure the model's accuracy.**

If the model performs poorly, we may need to:

- **Use different algorithms**
- **Fine-tune parameters**
- **Gather more data**
- **Improve data cleaning and preprocessing**

Stage 7: Deployment

The final step is deploying the trained model into a real-world system.

- If the model provides accurate results efficiently, it is implemented in a production environment.
- Deployment is similar to generating a final report for a project.

“

💡 Mastering Machine Learning requires dedication, practice, and continuous learning.

Stay with ComeToMachine ♥

Thank you!

Do you have any questions?

Alizadeh.c2m@gmail.com
ComeToMachine.ir

ComeToMachine