

Milestone 1

Project:

**SkillMatch- Resume Matcher
and Skill Recommender**

Project Mentor: Ms. Sangeetha Mahalingam

Presented by

Palash Gupta

4th year, MBA Tech in Computer Engr.

**Narsee Monjee Institute of Management
Studies**

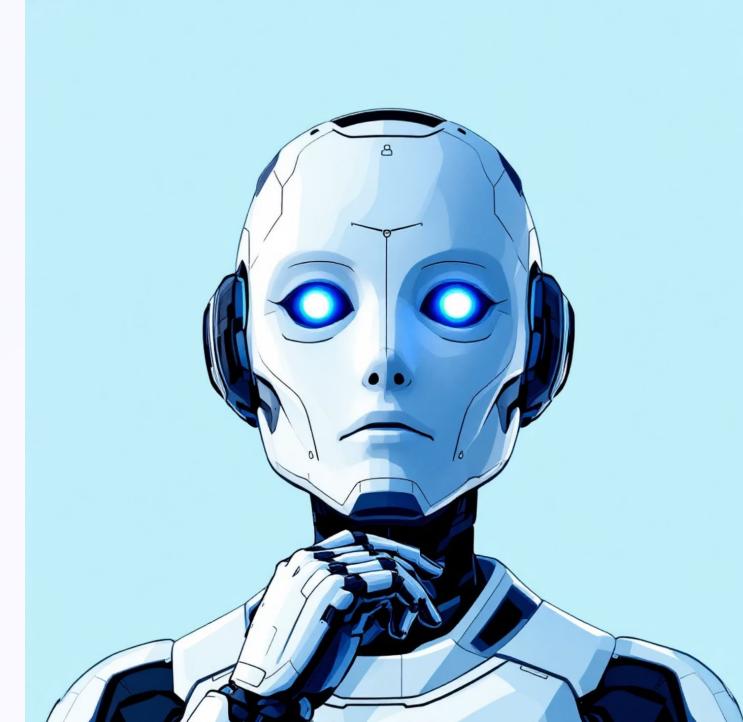


Understanding Artificial Intelligence (AI)

What is AI?

AI systems perform tasks requiring human intelligence:

- Perception
- Reasoning
- Learning
- Decision-making



Weak AI / Narrow AI

Designed for specific tasks only.

E.g., voice assistants like Alexa or Siri,
recommendation engines, spam filters.

John McCarthy (1956)

Coined "[Artificial Intelligence](#)"

Arthur Samuel (1959)

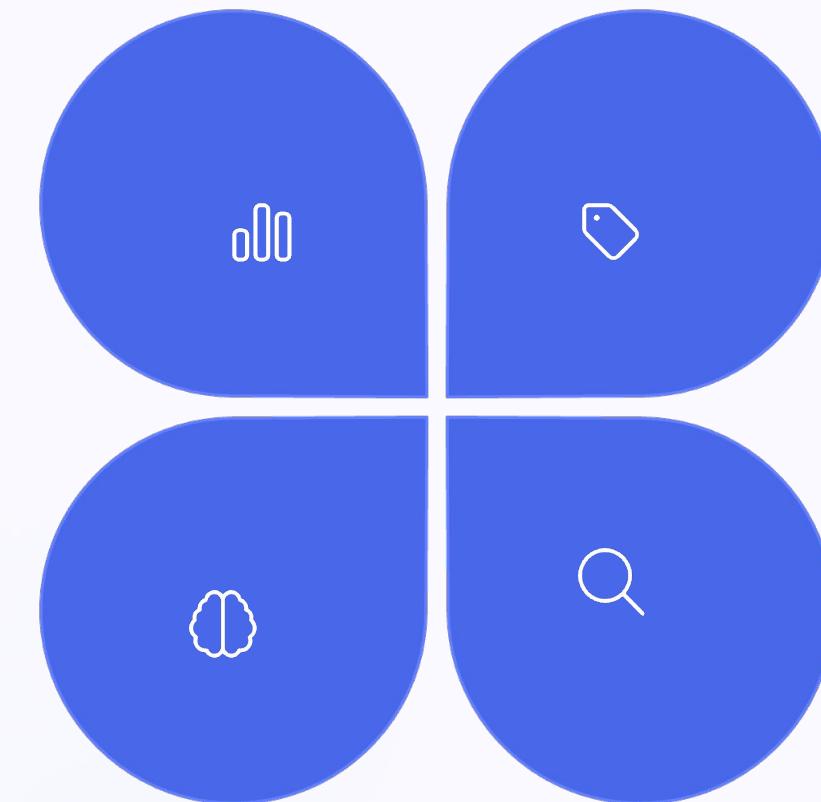
Coined "[Machine Learning](#)"

Diving into Machine Learning (ML)

ML is the subset of AI which provides machines the ability to **learn from data automatically** and **improve from experience without explicit programming**.

Regression
Predicting continuous values.

Deep Learning
Using complex neural networks for advanced pattern recognition.

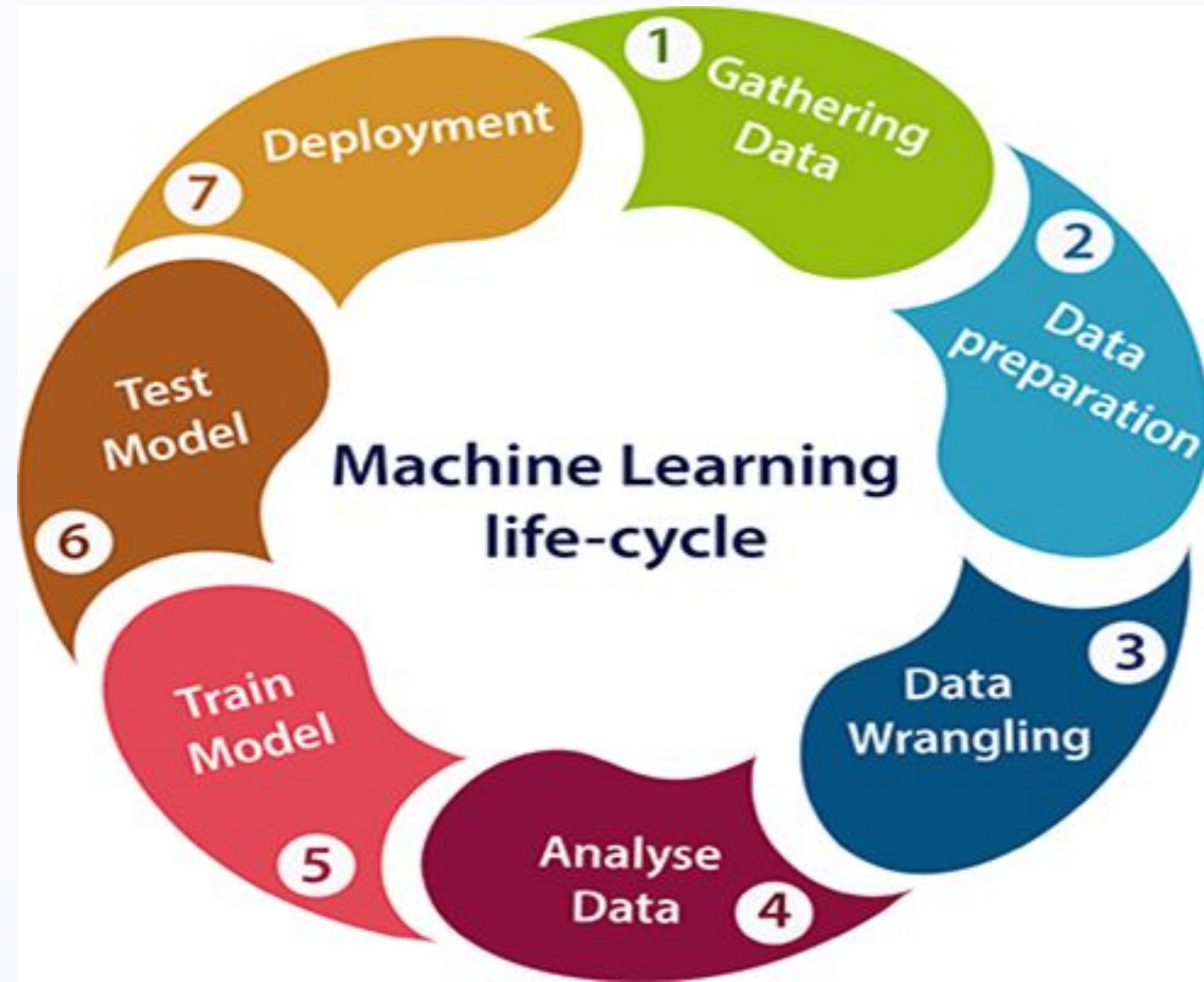


Classification
Categorizing data points (e.g., spam or not-spam).

Clustering
Grouping similar data without predefined labels.

The Machine Learning Project Life Cycle





Core Python Libraries:

Essential tools for data manipulation, computation, and visualization.

12 NumPy

- Numerical Python
- Numerical computing, high-performance array operations, linear algebra.

Matplotlib

- Foundational library for static, interactive, and animated visualizations.



Pandas

- Panel Data
- Efficient data analysis and manipulation using DataFrame and Series.



Seaborn

- Statistical graphics
- builds on Matplotlib to create aesthetically pleasing, informative plots.

NumPy: Numerical Python

The foundation for numerical and matrix operations in Python.

Array Creation Functions

`array()` - Create N-dimensional arrays

`zeros()` - Initialize with zeros

`full()` - Fill array with a specified value

`arange()` - Create range-based arrays

`random` - Generate random numbers

Manipulation & Statistics

`vstack()`, `hstack()` - Stack arrays (Vertical /Horizontal)

`mean()` - Calculate the average

`median()` - Find the middle value

`std()` - Compute standard deviation

Pandas: Data Manipulation Powerhouse

Optimized for fast, efficient data manipulation and analysis using powerful data structures.

1

`Series()`

One-dimensional labeled array (like a column in a spreadsheet).

2

`DataFrame()`

Two-dimensional tabular structure (like a spreadsheet or SQL table).

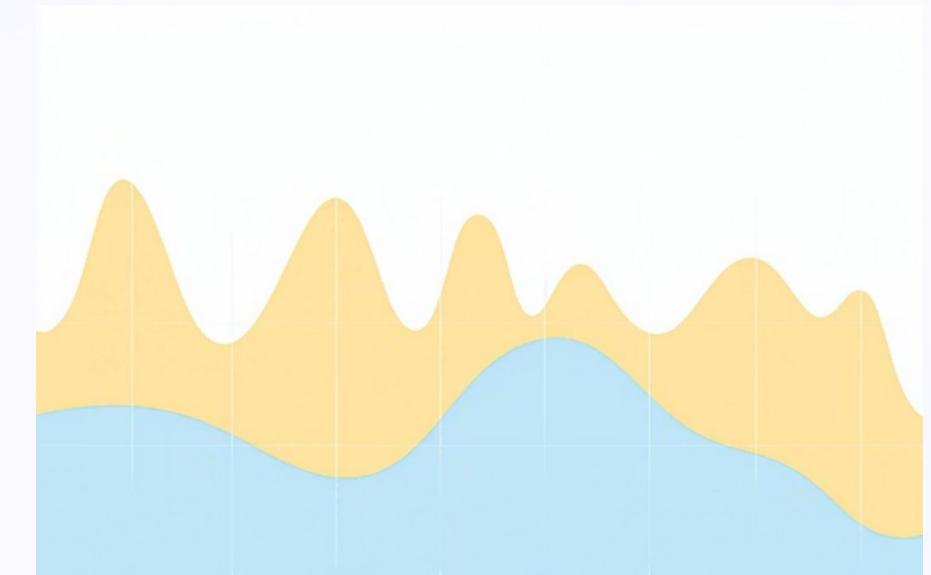
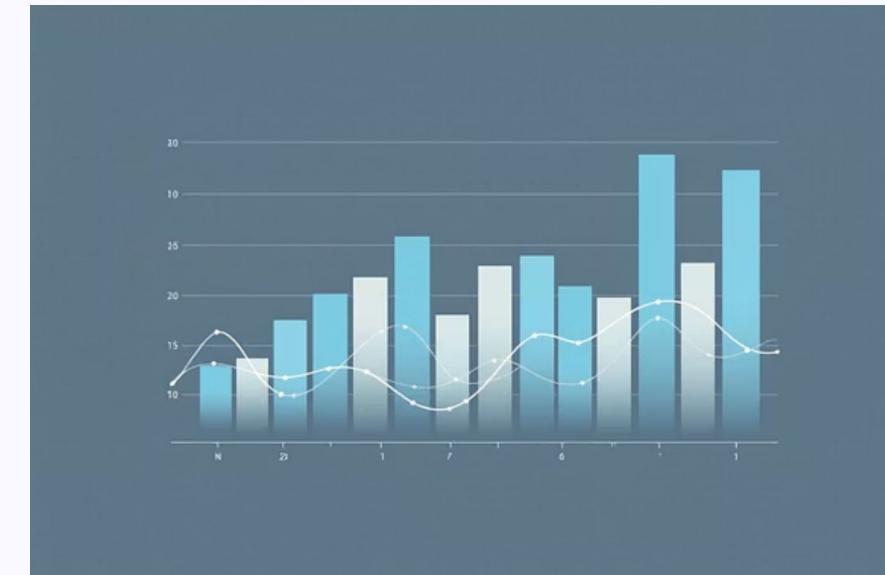
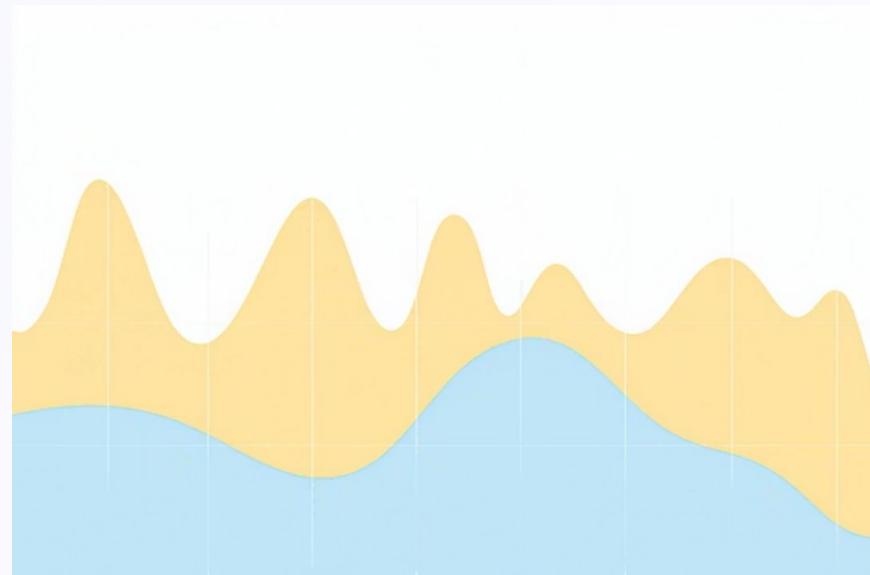
3

`read_csv()`

Function to easily import data from external CSV files.

Matplotlib: Foundational Plotting

for static, animated, and interactive visualizations in Python.



Basic Plots

`plot()`, `scatter()`, `bar()`,
`barh()`

Distributions

`hist()`, `boxplot()`,
`violinplot()`, `pie()`

Customization

`xlabel()`, `ylabel()`, `title()`,
`grid()`, `show()`



Seaborn: Statistical Visualisation

A high-level interface for drawing informative and attractive statistical graphics.

“ Styling & Themes

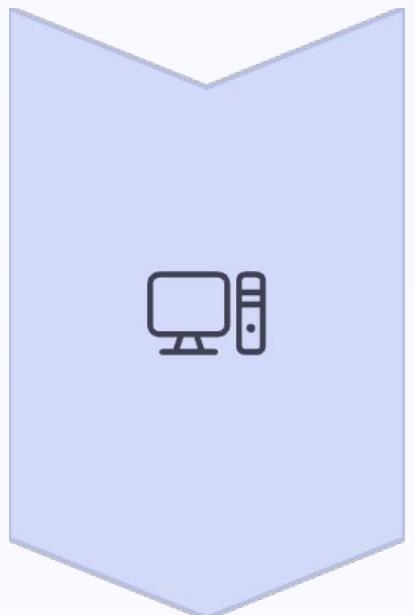
`set()`, `set_theme()` for aesthetic plot adjustments.

“ Load Data

`load_dataset()` provides access to sample datasets for practice.

Key Statistical Plot Types

`lineplot()`, `barplot()`
`boxplot()`, `histplot()`
`replot()`



Now, moving to the Google Colab file
for code and output

Thank you