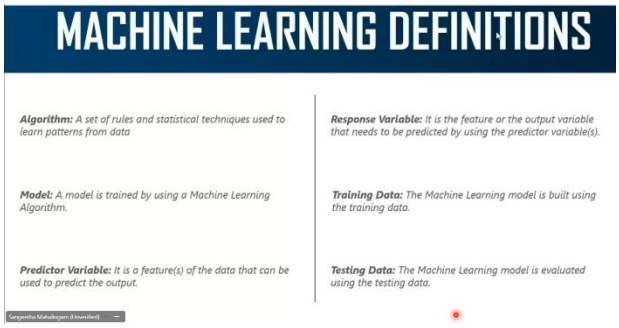
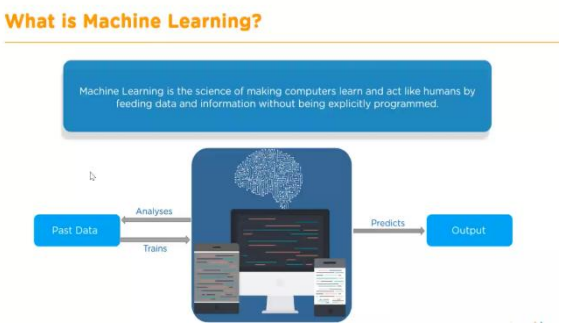
Report on SkillMatch Resume Matcher and Skill Recommender

1.What is Machine Learning (ML)?

Machine Learning is a crucial component of Artificial Intelligence.

- **Core Definition:** Machine learning is a **subset of Artificial Intelligence (AI)** which provides machines the ability to **learn automatically & improve from experience** without being explicitly programmed.
- **Alternative Definition:** Machine Learning is the science of making computers **learn and act like humans** by feeding data and information without being explicitly programmed.



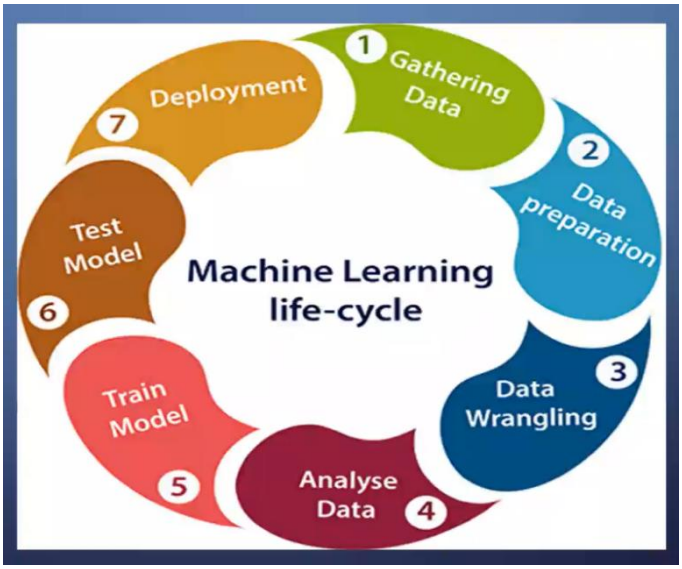


- **The Learning Rule:** "A computer program is said to learn from **experience E** with respect to some class of **tasks T** and **performance measure P** if its performance at tasks in T, as measured by P, **improves with experience E**."
- **History: Arthur Samuel** first coined the term Machine Learning in the year **1959**.
- **Basic Flow:** Data $\rightarrow$ Training the Machine $\rightarrow$ Building a Model $\rightarrow$ Predicting Outcome.
    - The machine takes **Past Data**, **Analyses** it, and **Trains** on it to **Predict** an **Output**.

2. The Machine Learning Process and Life-Cycle

The Machine Learning process involves building a **Predictive model** that can be used to find a **solution for a Problem Statement**.

The 7-Step Machine Learning Life-Cycle



The process is cyclical and involves:

1. **Gathering Data**
2. **Data preparation**
3. **Data Wrangling**
4. **Analyse Data**
5. **Train Model**
6. **Test Model**
7. **Deployment**

Alternative ML Process Steps

Another view of the process emphasizes model development:

- Define Objective
- Data Gathering
- Preparing Data
- Data Exploration
- Building a Model
- Model Evaluation
- Predictions

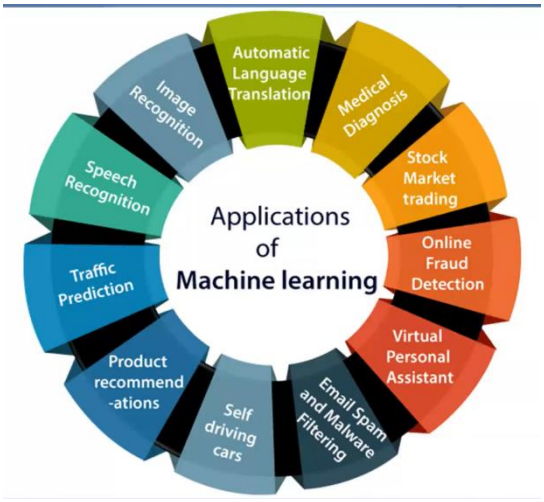
---

3. Real-World Applications of Machine Learning

ML is widely applied across various domains, including:

- Automatic Language Translation
- Medical Diagnosis
- Stock Market trading
- Online Fraud Detection
- Virtual Personal Assistant
- Email Spam and Malware Filtering
- Self driving cars

- Product recommendations
- Traffic Prediction
- Speech Recognition
- Image Recognition
- Face Recognition
- Siri and Cortana
- Healthcare Industry
- Weather Forecasting
- Produce a Web Series (e.g., Netflix recommendations)



Certainly! I can expand section 4 to include **Matplotlib** and **Seaborn**, which are essential libraries for data visualization in the ML process, often used in the **Analyze Data** and **Test Model** phases.

Here is the updated and detailed report.

🤖 Detailed Report: Machine Learning and Data Science Fundamentals

This report is compiled from the provided instructional slides and expanded with essential information about core Python data visualization libraries.

---

1. What is Machine Learning (ML)?

Machine Learning is a crucial component of Artificial Intelligence.

- **Core Definition:** Machine learning is a **subset of Artificial Intelligence (AI)** which provides machines the ability to **learn automatically & improve from experience** without being explicitly programmed.
- **Alternative Definition:** Machine Learning is the science of making computers **learn and act like humans** by feeding data and information without being explicitly programmed.
- **The Learning Rule:** "A computer program is said to learn from **experience E** with respect to some class of **tasks T** and **performance measure P** if its performance at tasks in T, as measured by P, **improves with experience E**."

- **History: Arthur Samuel** first coined the term Machine Learning in the year **1959**.
- **Basic Flow:** Data $\rightarrow$ Training the Machine $\rightarrow$ Building a Model $\rightarrow$ Predicting Outcome.
    - The machine takes **Past Data**, **Analyses** it, and **Trains** on it to **Predict** an **Output**.

---

2. The Machine Learning Process and Life-Cycle

The Machine Learning process involves building a **Predictive model** that can be used to find a **solution for a Problem Statement**.

The 7-Step Machine Learning Life-Cycle

The process is cyclical and involves:

1. **Gathering Data**
2. **Data preparation**
3. **Data Wrangling**
4. **Analyse Data**
5. **Train Model**
6. **Test Model**
7. **Deployment**

Alternative ML Process Steps

Another view of the process emphasizes model development:

- Define Objective
- Data Gathering
- Preparing Data
- Data Exploration
- Building a Model
- Model Evaluation
- Predictions

---

3. Real-World Applications of Machine Learning

ML is widely applied across various domains, including:

- Automatic Language Translation
- Medical Diagnosis
- Stock Market trading
- Online Fraud Detection
- Virtual Personal Assistant
- Email Spam and Malware Filtering
- Self driving cars
- Product recommendations
- Traffic Prediction
- Speech Recognition
- Image Recognition
- Face Recognition
- Siri and Cortana
- Healthcare Industry
- Weather Forecasting
- Produce a Web Series (e.g., Netflix recommendations)

---

4. Key Python Libraries for Data Science

The Python ecosystem relies on several libraries for data handling and analysis in ML. **NumPy** and **Pandas** handle the data itself, while **Matplotlib** and **Seaborn** handle its visual representation.

Library	Full Meaning	Core Functionality	Data Structures / Focus
NumPy	**Numerical Python**	The core library for **numeric and scientific computing**.	**Multi-dimensional array objects** and routines for processing them.
Pandas	**Panel Data**	The core library for **data manipulation and data analysis**.	**Single and multi-dimensional data-structures** (like DataFrames).

Library	Full Meaning	Core Functionality	Data Structures / Focus
Matplotlib	Matplotlib	A **low-level** library for creating **static, animated, and interactive visualizations**.	Used for **basic plotting** (line charts, scatter plots, bar charts) and offers **high customizability**.
Seaborn	Seaborn	A **high-level** library **built on top of Matplotlib** for drawing **attractive and informative statistical graphics**.	Simplifies creating complex plots (like heatmaps and violin plots) with **beautiful default themes** and statistical functions.