

24ML501- MACHINE LEARNING

UNIT 1-FOUNDATIONS OF MACHINE LEARNING AND DATA PREPARATION

Introduction to Machine Learning – AI vs ML vs Deep Learning – Types of Machine Learning: Supervised, Unsupervised, Semi-supervised and Reinforcement Learning – Real-world ML pipeline – Data collection and preprocessing – Feature engineering and feature selection – Curse of dimensionality – Bias-variance tradeoff – Overfitting and underfitting – Cross validation – Regularization (L1/L2) – Learning curves – Dimensionality reduction – Introduction to Responsible AI, Explainable AI (XAI), fairness and bias in AI systems.

INTRODUCTION TO MACHINE LEARNING

What is Machine Learning?

“Learning is any process by which a system improves performance from experience.”

- Herbert Simon

Definition by Tom Mitchell (1998):

- Machine Learning is the study of algorithms that
 - • improve their performance P
 - • at some task T
 - • with experience E.
- A well-defined learning task is given by $\langle P, T, E \rangle$

Traditional Programming



Machine Learning





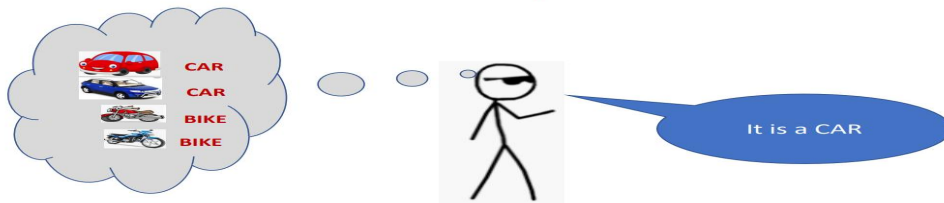
Human can learn from past experience
and make decision of its own



What is this object?



What is this object?

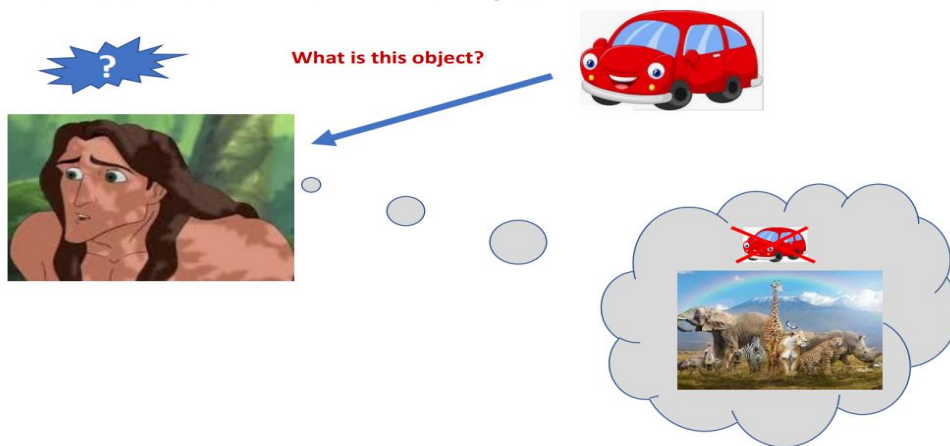


Let us ask the same question to him

What is this object?

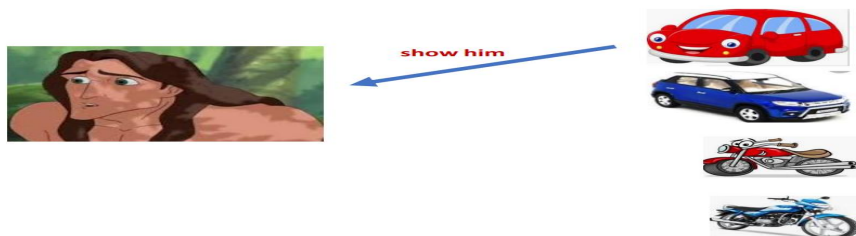


Let us ask the same question to him



[But, he is a human being. He can observe and learn]

Let us make him learn



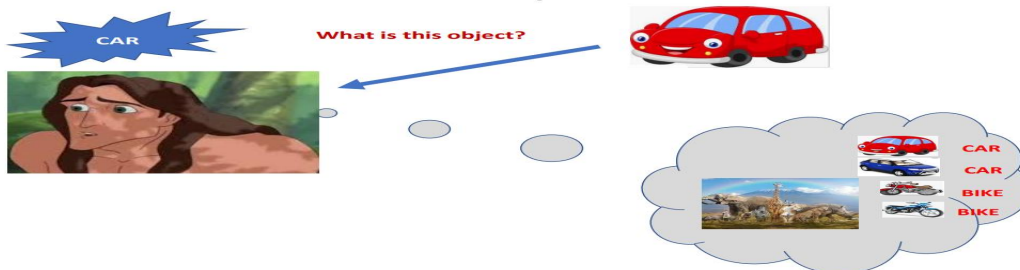
Let us make him learn



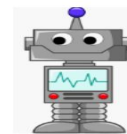
Let us ask the same question now



Let us ask the same question now



What about a Machine ?



Machines follow instructions

[It can not take decision of its own]

What about a Machine ?

We can ask a machine

- To perform an arithmetic operations such as
 - Addition
 - Multiplication
 - Division



Machines follow instructions

What about a Machine ?

- Comparison
- Print
- Plotting a chart



Machines follow instructions

What is Machine Learning?

[We want a machine to act like a human]

What is Machine Learning?



[to identify this object.]

What is Machine Learning?



I ~~made~~ **met** him yesterday

[Natural Language understand, and correct grammar]

What is Machine Learning?



recognize face

[Recognize Faces]

What is Machine Learning?



[What do we do?

Just like, what we did to human,

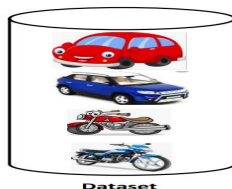
we need to provide experience
to the machine.

]

What is Machine Learning?



+

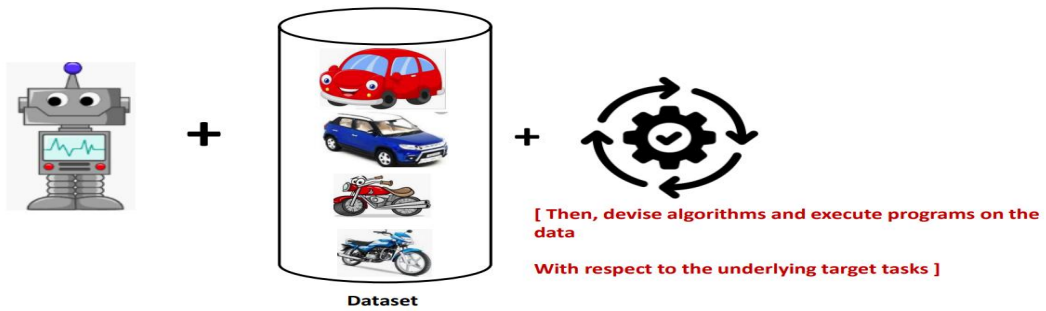


[This what we called as Data
or Training dataset

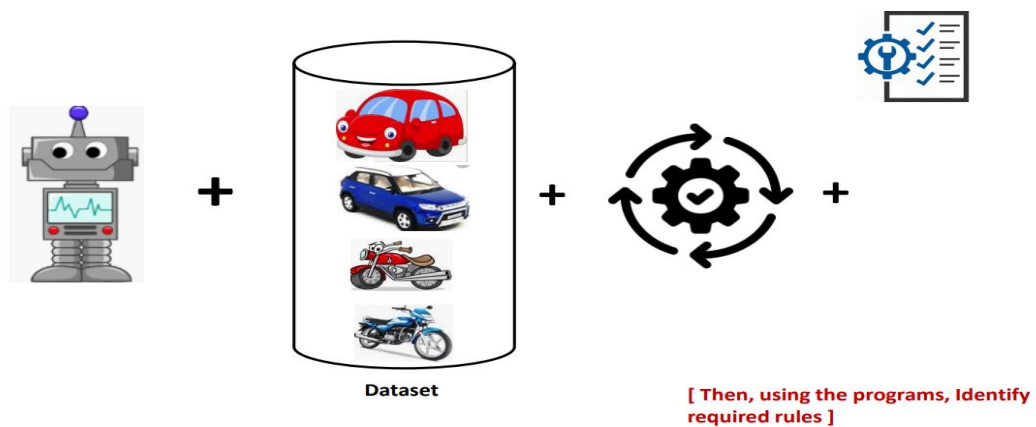
So, we first need to provide
training dataset to the
machine

]

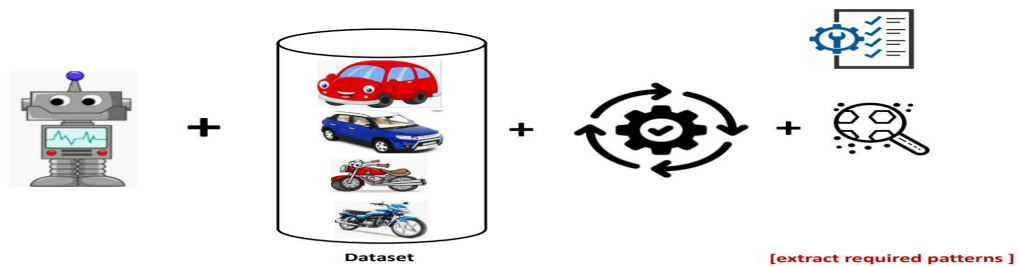
What is Machine Learning?



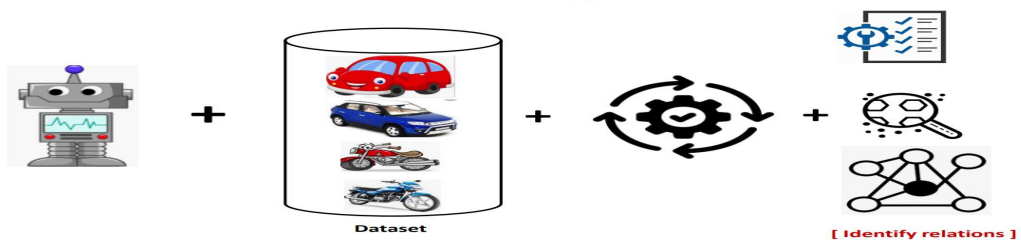
What is Machine Learning?



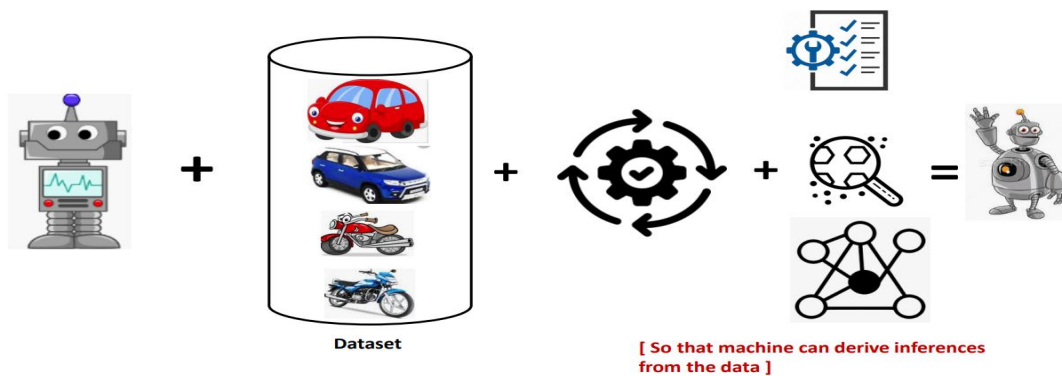
What is Machine Learning?



What is Machine Learning?



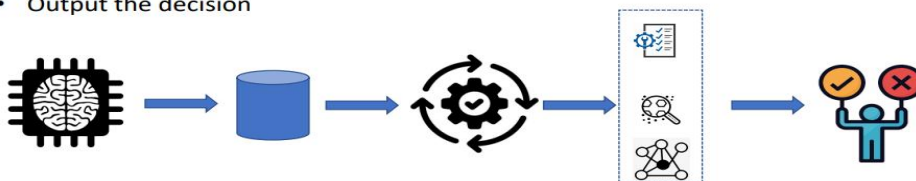
What is Machine Learning?



In summary, what is machine learning?

Given a machine learning problem

- Identify and create the appropriate dataset
- Perform computation to learn
 - Required rules, pattern and relations
- Output the decision



Definitions of Machine Learning

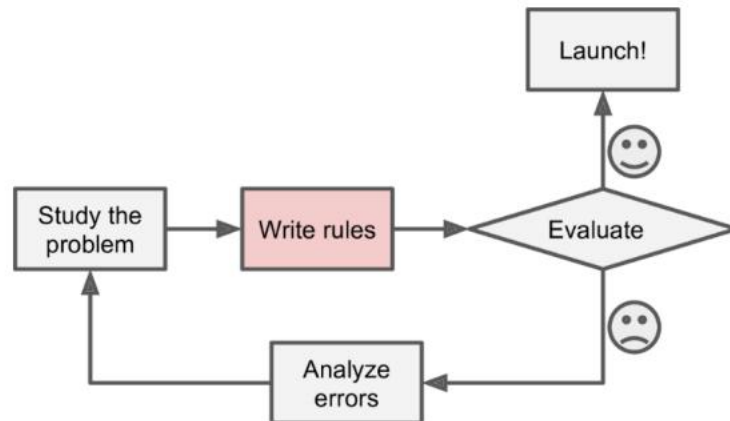
- Machine Learning is the science (and art) of programming computers so they can learn from data.
- Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.
- A computer program is said to learn from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E .

Need to use Machine Learning

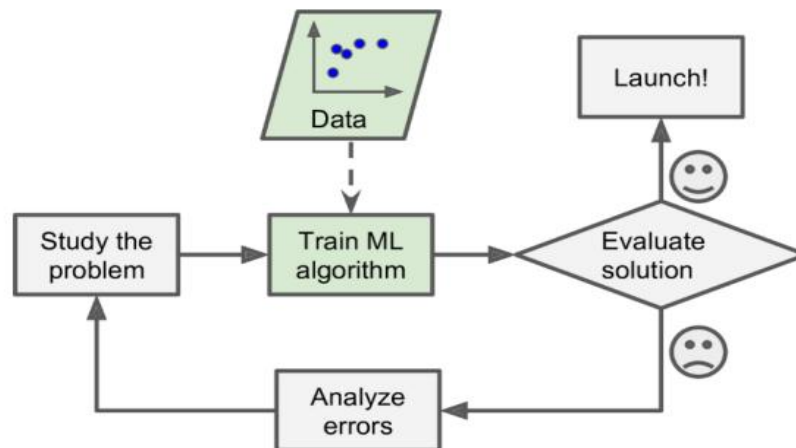
Machine Learning is great for:

- Problems for which existing solutions require a lot of hand-tuning or long lists of rules: one Machine Learning algorithm can often simplify code and perform

better. Eg: writing traditional programming vs machine learning approach for a spam filter:



Traditional programming for spam filter detection

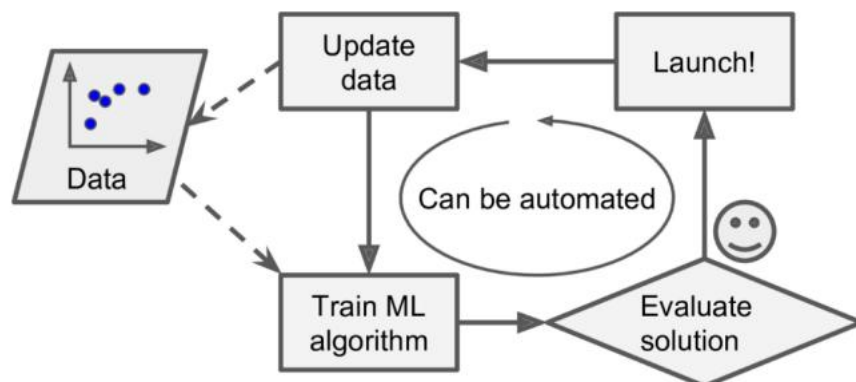


Machine Learning approach for spam filter detection

In traditional programming, one writes a detection algorithm for each of the patterns that you noticed, and the program would flag emails as spam if a number of these patterns are detected. Since the problem is not trivial, your program will likely become a long list of complex rules—pretty hard to maintain. In contrast, a spam filter based on Machine Learning techniques automatically learns which words and phrases are good predictors of spam by detecting unusually frequent patterns of words in the spam examples compared to the ham examples. The program is much shorter, easier to maintain, and most likely more accurate.

- Complex problems for which there is no good solution at all using a traditional approach: the best Machine Learning techniques can find a solution.

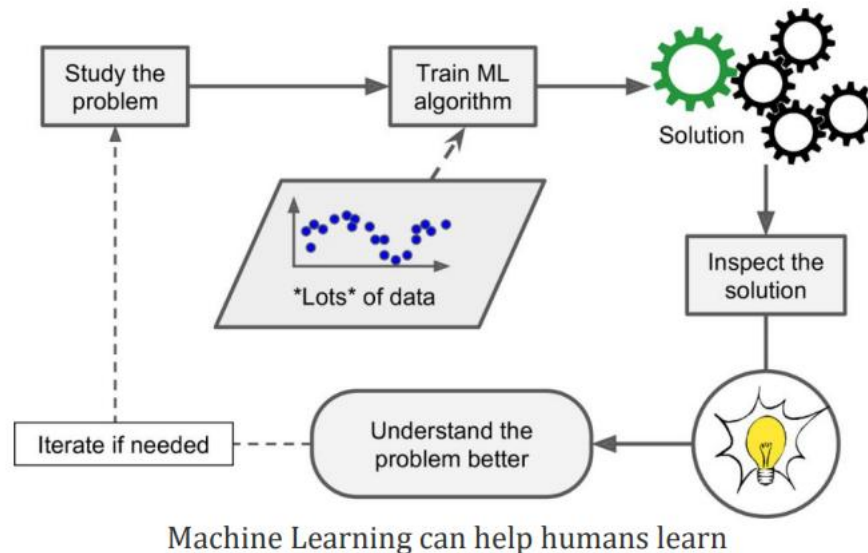
- Machine Learning shines is for problems that either are too complex for traditional approaches or have no known algorithm. For example, consider speech recognition: say you want to start simple and write a program capable of distinguishing the words “one” and “two.” You might notice that the word “two” starts with a high-pitch sound (“T”), so you could hardcode an algorithm that measures high-pitch sound intensity and use that to distinguish ones and twos. Obviously this technique will not scale to thousands of words spoken by millions of very different people in noisy environments and in dozens of languages. The best solution (at least today) is to write an algorithm that learns by itself, given many example recordings for each word.
- Fluctuating environments: a Machine Learning system can adapt to new data. Eg: In the spam filter detection example, if spammers notice that all their emails containing “4U” are blocked, they might start writing “For U” instead. A spam filter using traditional programming techniques would need to be updated to flag “For U” emails. If spammers keep working around your spam filter, you will need to keep writing new rules forever. In contrast, a spam filter based on Machine Learning techniques automatically notices that “For U” has become unusually frequent in spam flagged by users, and it starts flagging them without your intervention.



Machine Learning solutions adapt to change

- Getting insights about complex problems and large amounts of data. ML algorithms can be inspected to see what they have learned (although for some algorithms this can be tricky). For instance, once the spam filter has been trained on enough spam, it can easily be inspected to reveal the list of words and combinations of words that it believes are the best predictors of spam. Sometimes this will reveal unsuspected correlations or new trends, and

thereby lead to a better understanding of the problem. Applying ML techniques to dig into large amounts of data can help discover patterns that were not immediately apparent. This is called data mining.



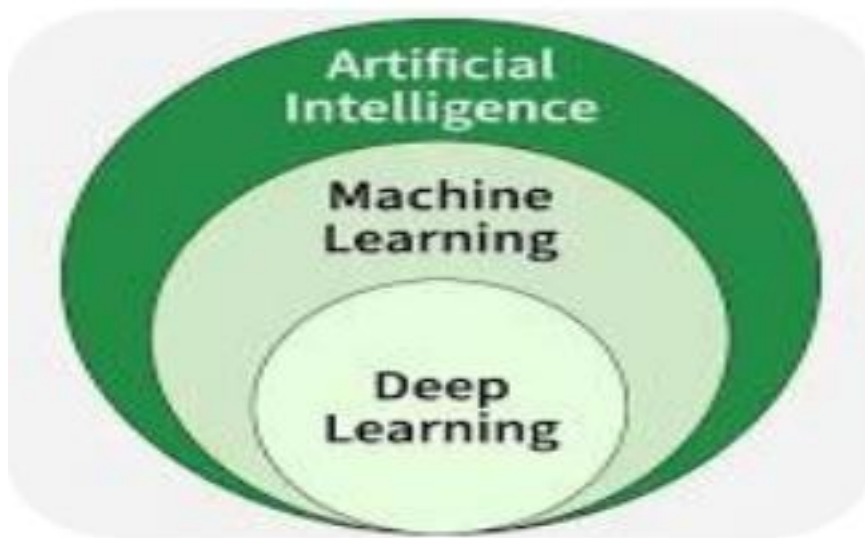
AI VS ML VS DEEP LEARNING

What Should be your First Step?

- When deciding between AI, ML, and Deep Learning, it is important to follow a logical sequence instead of choosing one randomly.
- A structured progression helps build strong fundamentals and ensures long-term career growth, as explained below:
- 1. Start With ML: Machine Learning builds strong foundational knowledge for advanced AI and Deep Learning concepts.
 2. Understand AI: Learning AI concepts provides a broader understanding of intelligent systems and applications.
 3. Move To DL: Deep Learning should follow after mastering Machine Learning fundamentals and core mathematical concepts.
 4. Learn Basics: Knowledge of Python, statistics, and linear algebra is essential before starting Machine Learning.
 5. Follow Sequence: A structured learning path ensures steady progress and avoids unnecessary confusion.

Key Differences Between AI, ML, and Deep Learning

- Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) are often used interchangeably, but they are a set of nested concepts. Think of them as a set of Russian dolls: Deep Learning is a specialized subset of Machine Learning, which is a core discipline of Artificial Intelligence



Scope and Definition

- Artificial Intelligence: Broad field focused on creating intelligent machines capable of simulating human thinking and decision-making processes.
- Machine Learning: Subset of AI that enables systems to learn from data patterns and improve performance without explicit programming.
- Deep Learning: Advanced subset of ML using neural networks with multiple layers to process complex data and solve sophisticated problems.

Learning Approach

- Artificial Intelligence: Uses rule-based systems, logic, and algorithms to perform tasks requiring human-like intelligence and reasoning.
- Machine Learning: Applies statistical models and algorithms to identify patterns in data and make predictions or decisions automatically.
- Deep Learning: Utilizes layered neural networks to automatically learn hierarchical data representations without requiring manual feature engineering processes

Data Dependency

- Artificial Intelligence: Can operate with predefined rules and limited datasets for decision-making in controlled environments.
- Machine Learning: Requires structured and labeled datasets to train models effectively and generate accurate predictive outcomes.
- Deep Learning: Depends on large volumes of unstructured data like images, audio, and text for optimal performance.

Complexity Level

- Artificial Intelligence: Includes both simple rule-based systems and complex intelligent systems depending on application requirements.
- Machine Learning: Moderately complex, involving algorithm selection, feature engineering, and model evaluation for effective performance.
- Deep Learning: Highly complex, requiring advanced architectures, large datasets, and significant computational power for model training.

Real-World Applications

- Artificial Intelligence: Used in automation, chatbots, robotics, and intelligent systems across multiple industries and business functions.
- Machine Learning: Applied in recommendation systems, fraud detection, predictive analytics, and customer behavior analysis solutions.
- Deep Learning: Powers advanced applications like image recognition, speech processing, autonomous vehicles, and natural language understanding systems.