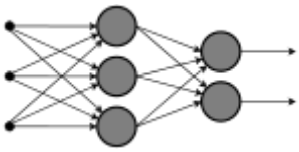
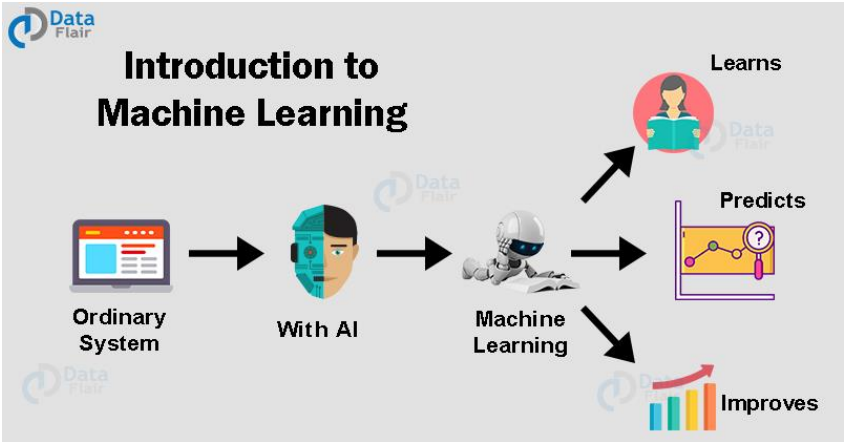
Machine Learning

Practical-1: Introduction to Machine Learning



What is Machine Learning?

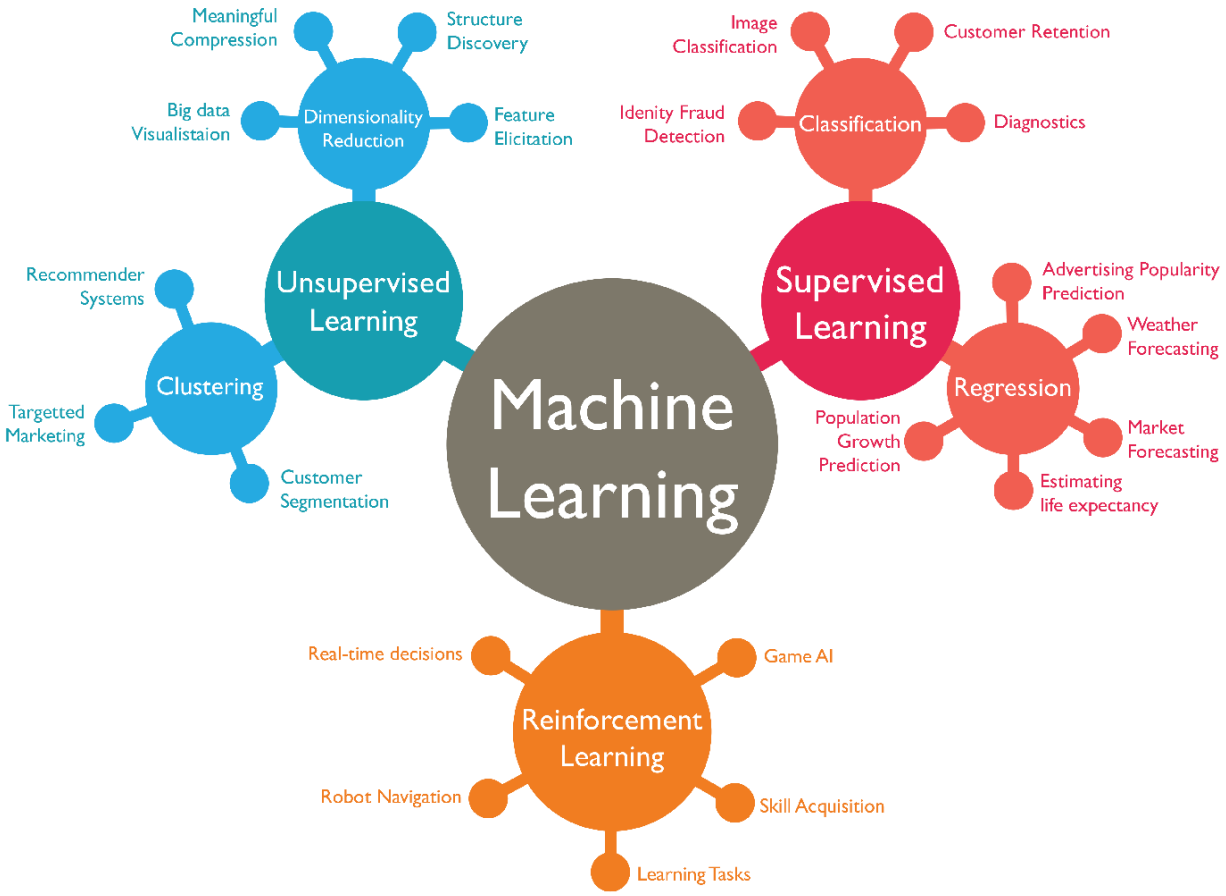
- Machine learning is the study of computer algorithms that improve automatically through experience and by the use of data.
- Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.
- Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, uncovering key insights within data mining projects.



- These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics.
- As big data continues to expand and grow, the market demand for data scientists will increase, requiring them to assist in the identification of the most relevant business questions and subsequently the data to answer them.

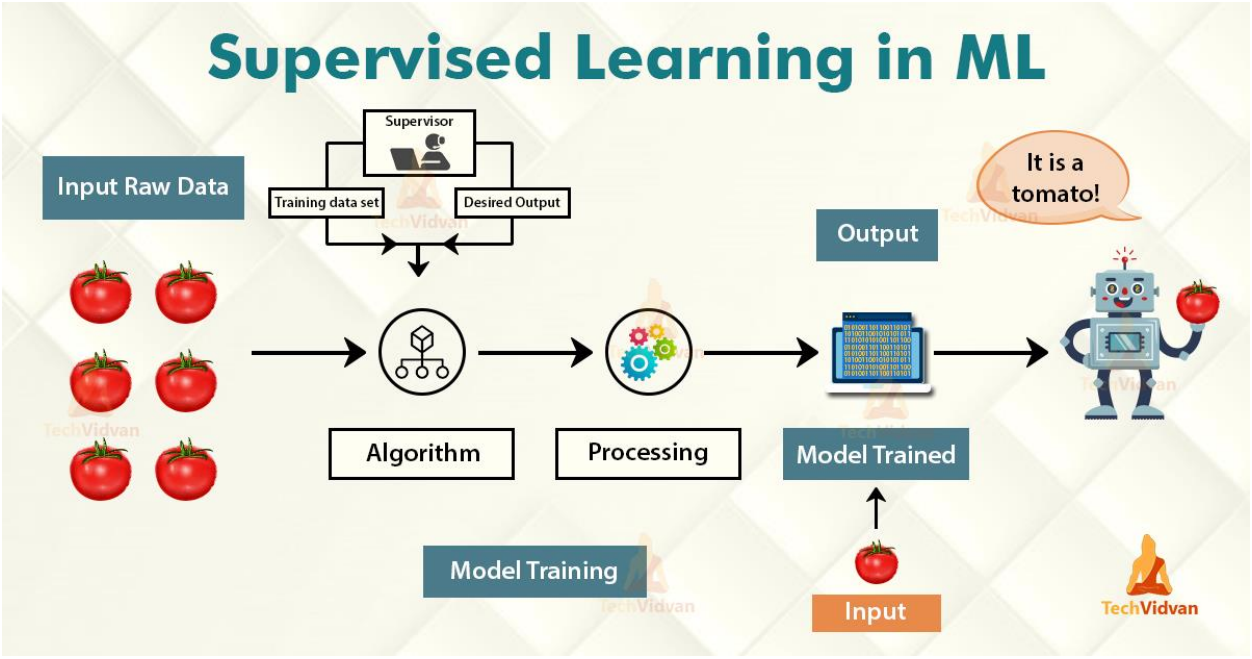
Types of Machine Learning

1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning



1. Supervised Learning

- Supervised learning is the most popular paradigm for machine learning. It is the easiest to understand and the simplest to implement. It is very similar to teaching a child with the use of flash cards.



- Supervised learning is often described as task-oriented because of this. It is highly focused on a singular task, feeding more and more examples to the algorithm until it can accurately perform on that task. This is the learning type that you will most likely encounter, as it is exhibited in many of the following common applications:

1) Spam Classification:

If you use a modern email system, chances are you've encountered a spam filter. That spam filter is a supervised learning system. Fed email examples and labels spam/not spam, these systems learn how to preemptively filter out malicious emails so that their user is not harassed by them.

2) Advertisement Popularity:

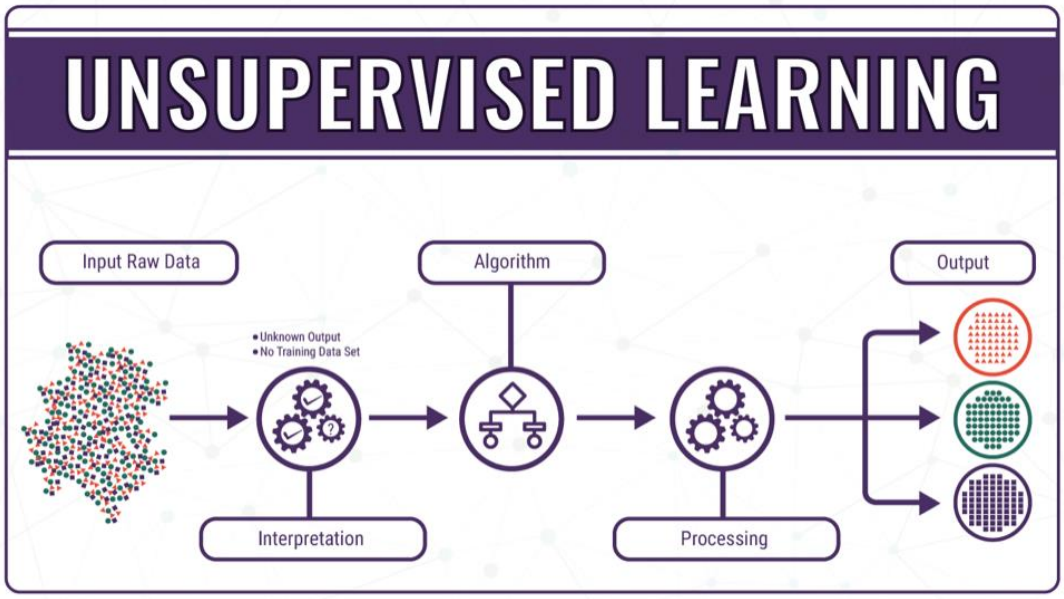
Selecting advertisements that will perform well is often a supervised learning task. Many of the ads you see as you browse the internet are placed there because a learning algorithm said that they were of reasonable popularity.

3) Face Recognition:

Most likely your face has been used in a supervised learning algorithm that is trained to recognize your face. Having a system that takes a photo, finds faces, and guesses who that is in the photo suggesting a tag is a supervised process. It has multiple layers to it faces and then identifying them, but is still supervised nonetheless.

2. Unsupervised Learning

- Unsupervised learning is very much the opposite of supervised learning. It features no labels.
- Instead, our algorithm would be fed a lot of data and given the tools to understand the properties of the data.
- From there, it can learn to group, cluster, and/or organize the data in a way such that a human can come in and make sense of the newly organized data.



- Unsupervised learning is based upon the data and its properties, we can say that unsupervised learning is data-driven. The outcomes from an unsupervised learning task are controlled by the data and the way its formatted. Some areas you might see unsupervised learning crop up are:

1) Recommender Systems:

If you've ever used YouTube or Netflix, you've most likely encountered a video recommendation system. These systems are often times placed in the unsupervised domain. o Buying Habits: It is likely that your buying habits are contained in a database somewhere and that data is being bought and sold actively at this time. These buying habits can be used in unsupervised learning algorithms to group customers into similar purchasing segments.

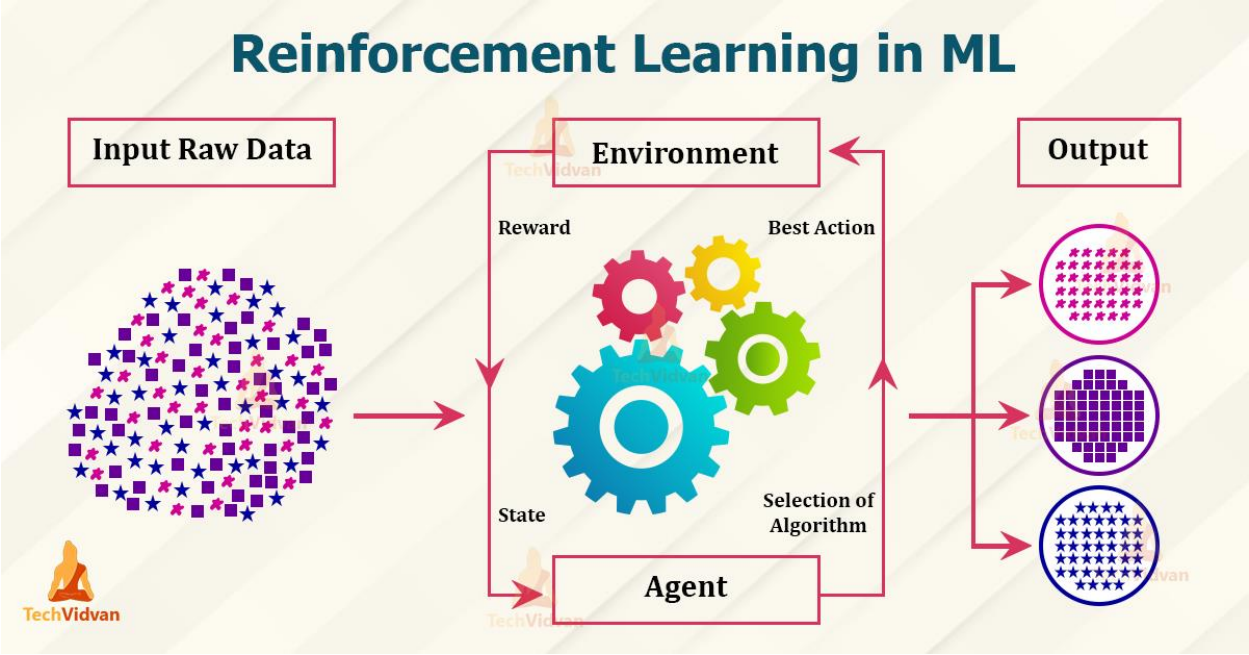
2) Grouping User Logs:

Less user facing, but still very relevant, we can use unsupervised learning to group user logs and issues. This can help companies

identify central themes to issues their customers face and rectify these issues, through improving a product or designing an FAQ to handle common issues.

3. Reinforcement Learning

- Reinforcement learning is fairly different when compared to supervised and unsupervised learning.
- Where we can easily see the relationship between supervised and unsupervised (the presence or absence of labels), the relationship to reinforcement learning is a bit murkier.



- Some people try to tie reinforcement learning closer to the two by describing it as a type of learning that relies on a time-dependent sequence of labels, however, my opinion is that that simply makes things more confusing.

1) Video Games:

One of the most common places to look at reinforcement learning is in learning to play games. Look at Google's reinforcement learning application, AlphaZero and AlphaGo which learned to play the game Go.

2) Industrial Simulation:

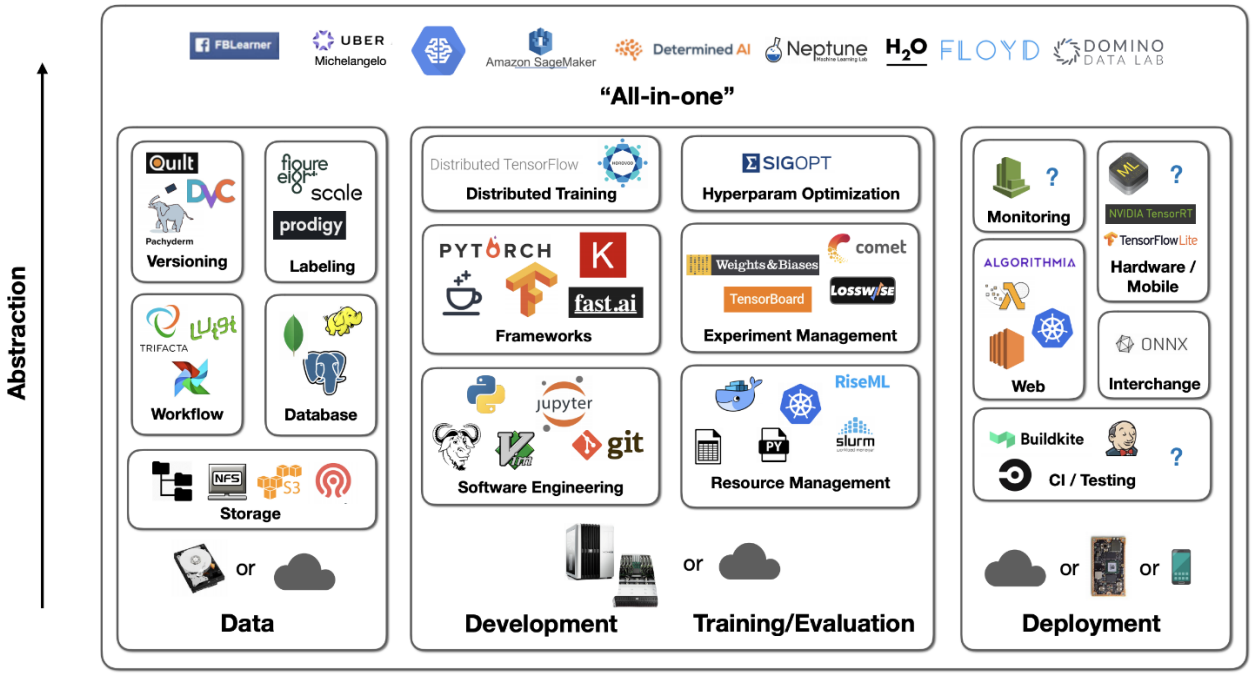
For many robotic applications (think assembly lines), it is useful to have our machines learn to complete their tasks without having to hardcode their processes.

3)Resource Management:

Reinforcement learning is good for navigating complex environments. It can handle the need to balance certain requirements. Take, for example, Google's data centers.

Tools and Technology for Machine Learning

1) Python
2) TensorFlow
3) Amazon Machine Learning (AML)
4) Google Cloud AutoML
5) Scikit-Learn
6) IBM Watson Studio
7) Microsoft Azure ML Studio


Abstraction
FBLearner
UBER Michelangelo
Amazon SageMaker
Determined AI
Neptune
H2O
FLOYD
DOMINO DATA LAB
"All-in-one"
Quilt
DVC
Pachyderm
Versioning
pachyderm scale
prodigy
Labeling
TRIFACTA
Workflow
Database
NFS
S3
Storage
or
Data
Distributed TensorFlow
Distributed Training
SIGOPT
Hyperparam Optimization
PYTORCH
K
fast.ai
Frameworks
Weights & Biases
comet
TensorBoard
Losswise
Experiment Management
jupyter
git
Software Engineering
RiseML
PY
slurm
Resource Management
or
Development
Training/Evaluation
Monitoring
ALGORITHMIA
Web
NVIDIA TensorRT
TensorFlow Lite
Hardware / Mobile
ONNX
Interchange
Buildkite
CI / Testing
or
or
Deployment