Following ⌄

Introduction to Machine Learning world
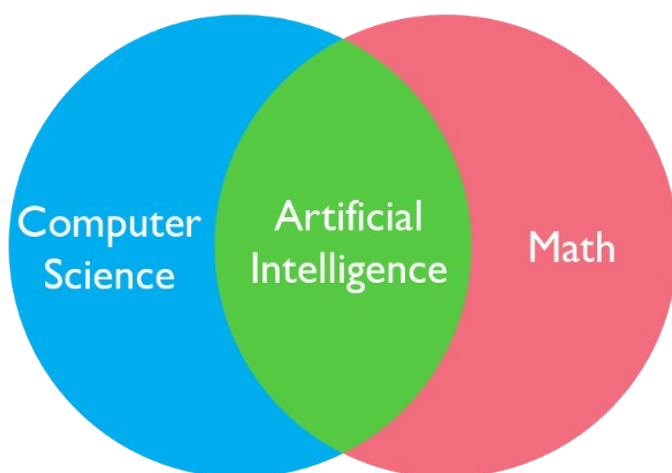
Achraf KHAZRI Follow
Apr 10 · 6 min read

1) What is Artificial intelligence ?

Artificial intelligence is the intersection between math and computer science, it is the theory and development of computer systems able to simulate human actions, and we mean by this making machine act like human being, see like human, speak and understand human language, and do all the stuff that human can do.

Nowadays, there are so many applications of artificial intelligence in different other domain like :

In social media :

- Facebook & Twitter use AI to decide what content to present in their feeds to different audiences.

- Image recognition and sentiment analysis to ensure that content of the

appropriate 'mood' is being served.

Trending AI Articles:

In transportation :

- Google finds the fastest route by processing the traffic data

In daily life:

- Natural language processing : we carry around powerful NLP algorithms in our phones and computers such as google translate.
- Object detection : In our mobile phone we use face detection also in facebook .

In Health:

- Cancer detection, Alzheimer detection, Patient care

In Finance:

- Trading algorithms, Fraud detection, Personal finance

In Energy :

- Oil & Gas exploration

In Government **:**
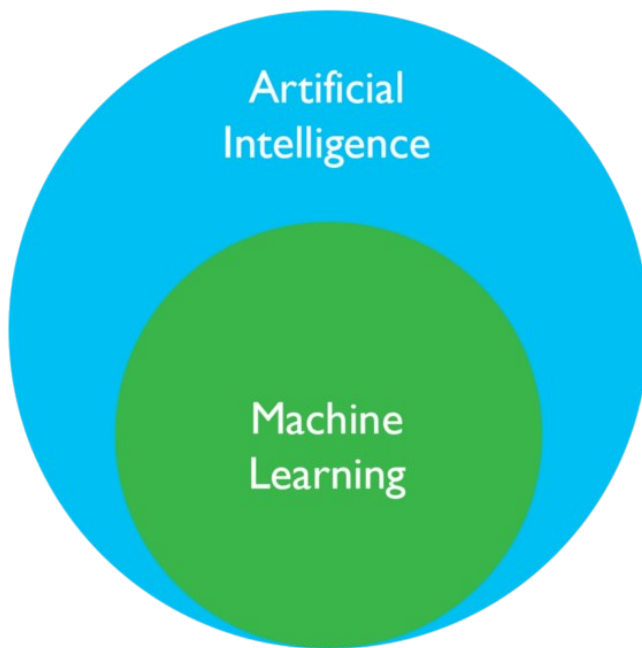
- Defence, Data insights, Safety and security, Smarter cities

In Transport :

- Autonomous cars, Smart parking, Tracking cars for security

As you can see, Artificial intellignce is included in all domains, it's like the electricity of the 21st century, it's important to understand this field and be part of the AI community.

2) What is Machine Learning ?

Machine Learning is a sub-domain of Artificial Intelligence, it is about giving machine the ability to learn without being explicitly programmed and uses algorithms to build a links between informations extracted from data and represent it in some type of trained models.

We have two main parts in every machine learning problem, Dataset and a mathematical model, dataset will be showed to model so it canbe used in the futur in some sort of tasks related to what it learned from the dataset and make predictions.

In the world of Machine learning we have three main tasks :

- Supervised learning

- Unsupervised learning

- Reinforcement learning

3) Supervised Learning

In this task, dataset is labeled, and we mean by this that we have an input data and we have a target that we want the mathemathical model to predict. So to make this model ready for using, we fit those data including target to the model, and each time we fit the data we tell the model "Hey, look, for this input data we have this output (target, label).

Supervised learning have two types :

- **Clasification** : this is when we predict a discrete output or a class or a thing like car, computer, human …
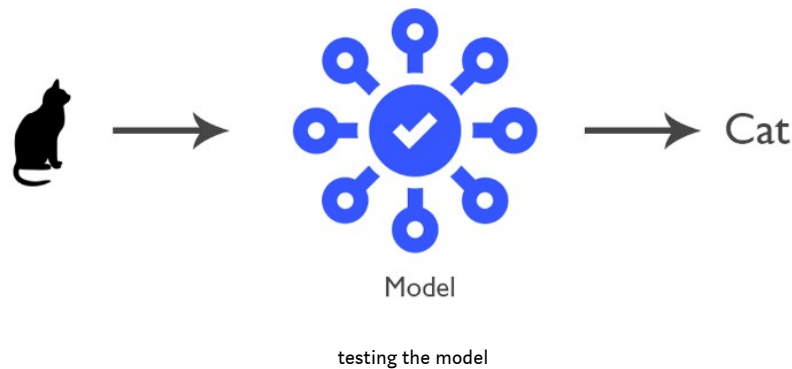
Lest's see this example :

first we fit our dataset the model, the dataset contains animal pictures and a label for each picture (dog, cat and fish).

Fitting data to model

And now after fitting our dataset to the model, we can use it like showed in the picture below.
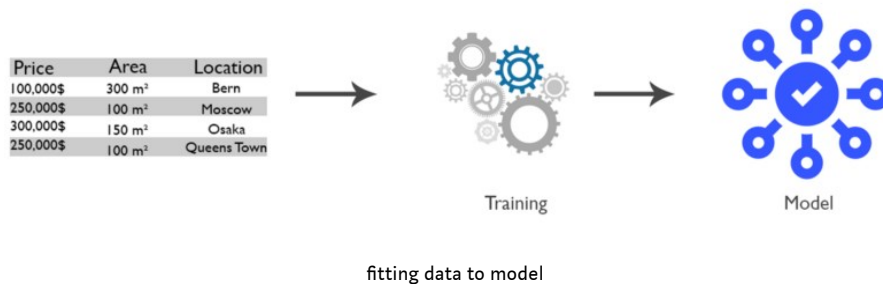

Model

testing the model

As you can see, the model predicts well the category of the picture even it has never shown to it.

- **Regression** : this is when we predict continious output (number)

let's see an example for this task

This time we will fit a dataset that contains as input house area and location and as output or target the house price like shown in the picture below :

Price	Area	Location
100,000$	300 m²	Bern
250,000$	100 m²	Moscow
300,000$	150 m²	Osaka
250,000$	100 m²	Queens Town

Training   Model

fitting data to model

And now after fitting our dataset to the model, we can use it like showed in the picture below.

New York
200 m²   500,000$

Model

testing the model
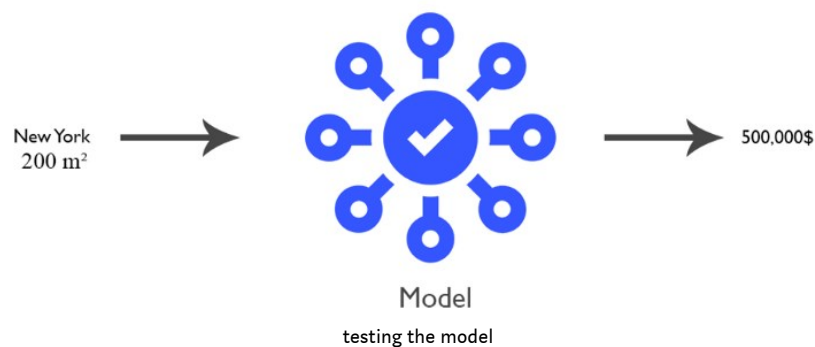
As you can see, the model predicts well the price of the the house even it has never shown to it before.

4) Unsupervised Learning

Unlike supervised learning, in the unsupervised learning task we have unlabeled data, and we mean by this that input data have no targets, we just fit the data to machine learning model to do one of those operation :

Clustering : it consists of getting data without any label and try to group data depending on their features.

the picture below shows an example of clustering.

the picture below shows an example of clustering.

Dimension reduction : it consists of reducing data features without losing their meaning or their most important features, dimension reduction is important when we want to reduce the size of datasets and also we can gain processing time and thats because we have less information.

the picture below shows an example of dimension reduction.
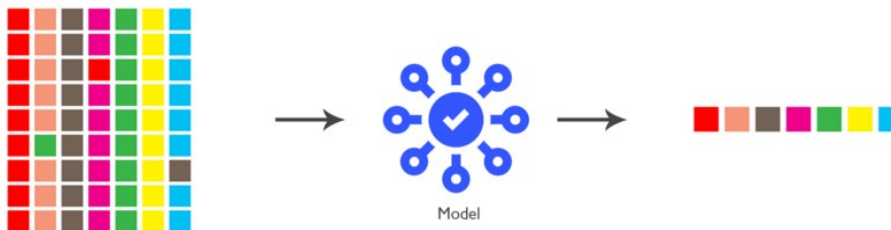
5) Reinforcement Learning

Reinforcment learning is a task of machine learning where we have an Agent in some sort of environment, and we have an objective that we want the agent to acheive. And to acheive this objective, the Agent explore all his capability and do arbitrary actions, and for every action he get a good or bad reward, and considiring that reward, the agent learns wich are the best actions for him to acheive his goal or objective.

We have a mathematical model for the reinforcement learning called Markov Model Process (MDP) and it is constituted of :

- Agent
- Environment
- State
- Action
- Reward

Markov Model Process
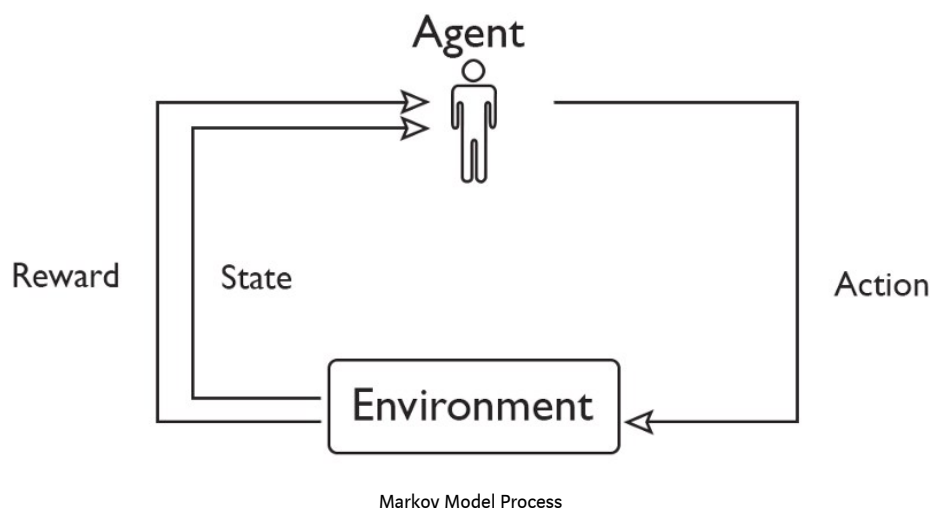
First, the agent get his state in the environment, then, depending on that state he make a move or an action, after that, he gets a reward for that action and finally depending on that reward (positive or negative) he improves his next actions.

Let's see an example to understand more : Cart-Pole problem
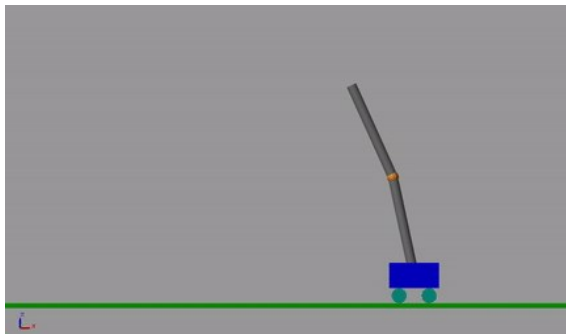
Objective : balance a pole on top of movable chart cart.

State : Angle, angular speed, position, horizontal velocity.

Action : Horizontal force applied on the cart.

Reward : 1 at each time step if the pole is upright.

6) Dataset

To train a machine learning model and see it's performance, we devide our dataset to 3 parts :

- **Training data :** Data used to train machine learning model, about 60% of the original dataset

- **Validation data :** this part of data is used to tune machine learning model hyperparameters to avoid overfitting, underfitting and to improve model performance. This part of dataset generally takes 20% of original data.

- **Testing data :** this the last part of the dataset, and generally takes 20% of the dataset, it is used to test the performance of the model.

7) Metrics

Metrics are used to mesure the performance of trained machine learning models.

Let's some metrics, to make it easy we will present them with an example.

Example : Car-Bus classifier

	Car	Bus
Car	TP = 20	FN = 5
Bus	TN = 3	FP = 18

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precesion = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

TP : True positive

TN : True negative

FN : False negative

FP : False positive

We can assimilate dataset parts and metrics like following :

- Student = Machine learning model
- Training data = Courses
- Validation data = catch-up courses
- Testing data = Exams
- Metrics = Exams notes

Don't forget to give us your   !

Join the Community

Subscribe

Apply To Be A Writer

Machine Learning · Artificial Intelligence · Computer Science · AI · Neural Networks

110 claps

Achraf KHAZRI
Artificial Intelligence
Research Engineer

Follow

Becoming Human: Artificial Intelligence Magazine

Latest News, Info and Tutorials on Artificial Intelligence, Machine Learning, Deep Learning, Big Data and what it means for Humanity.

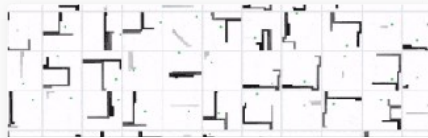Following ⌄

More from Becoming Human: Artificia...

Pipeline Bots Are Insuring Nothing Blows Up

Oliver Mitchell
Jun 24 · 6 min read ★

56

More from Becoming Human: Artificia...

Designing AI: Solving Snake with Evolution

Peter Binggeser
Sep 25, 2017 · 10 min

864

More from Becoming Human: Artificial Intelligence Magazine

Variational AutoEncoders for new fruits with Keras and Pytorch.

Thomas Dehaene
Nov 7, 2018 · 8 min r

995