



MACHINE LEARNING **BOOTCAMP**

WHOAMI?



Hi, 🖐️

- Name ? Aymen Berriche
- Third year Student at ESI Algiers
- GitHub Campus Expert
- GDG Algiers Development Co-Manager

Plan of Attack?



- What Is Machine Learning?
- Why Use Machine Learning?
- Types of Machine Learning Systems
- Supervised/Unsupervised learning
- Batch/Online learning

Plan of Attack?



- Instance-Based Versus Model-Based Learning
- Career Paths

What Is Machine Learning?

Machine Learning is the science (and art) of programming computers so they can learn from data.

What Is Machine Learning?

Here is a slightly more general definition:

[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.

—Arthur Samuel, 1959

What Is Machine Learning?

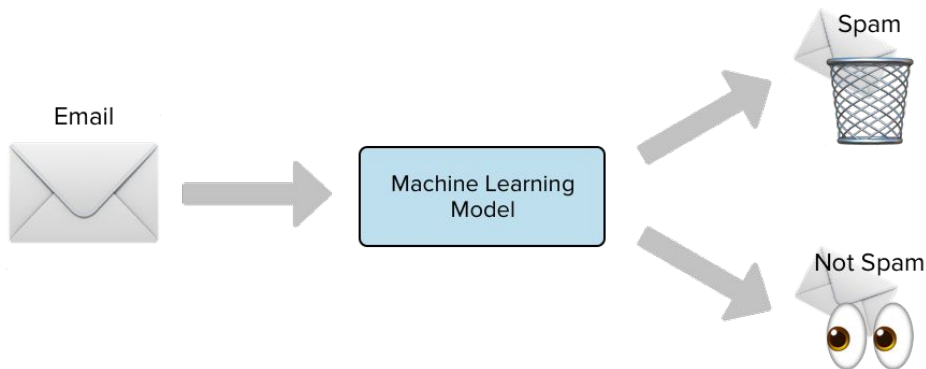
And a more engineering-oriented one:

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**.

—Tom Mitchell, 1997

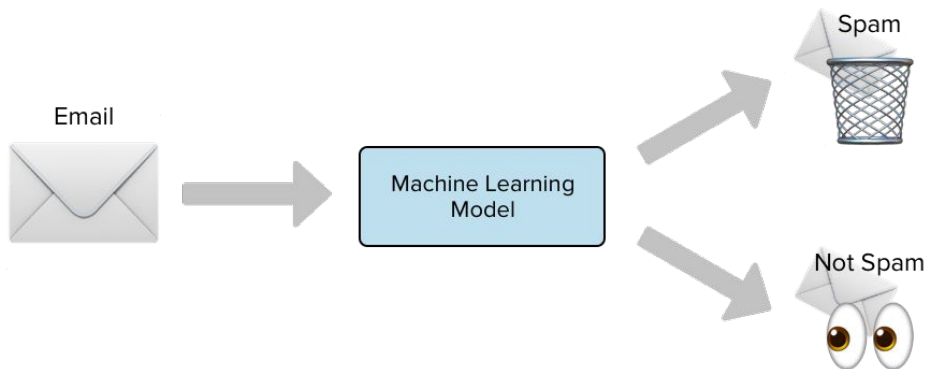
What Is Machine Learning?

- Experience **E** ?
- Task **T** ?
- Performance measure **P** ?



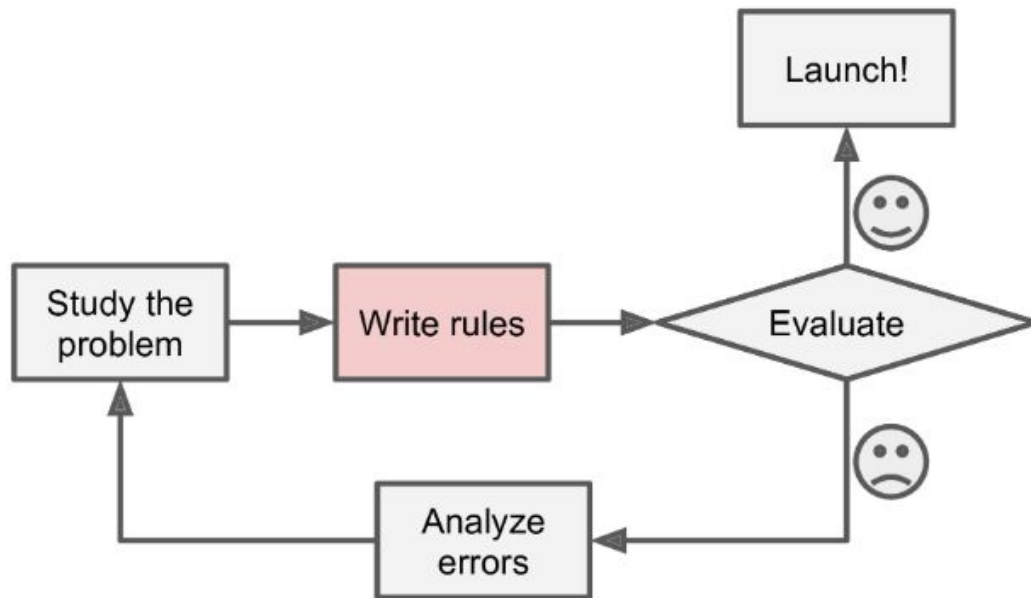
Why Use Machine Learning?

Consider how you would write a spam filter using traditional programming techniques ?



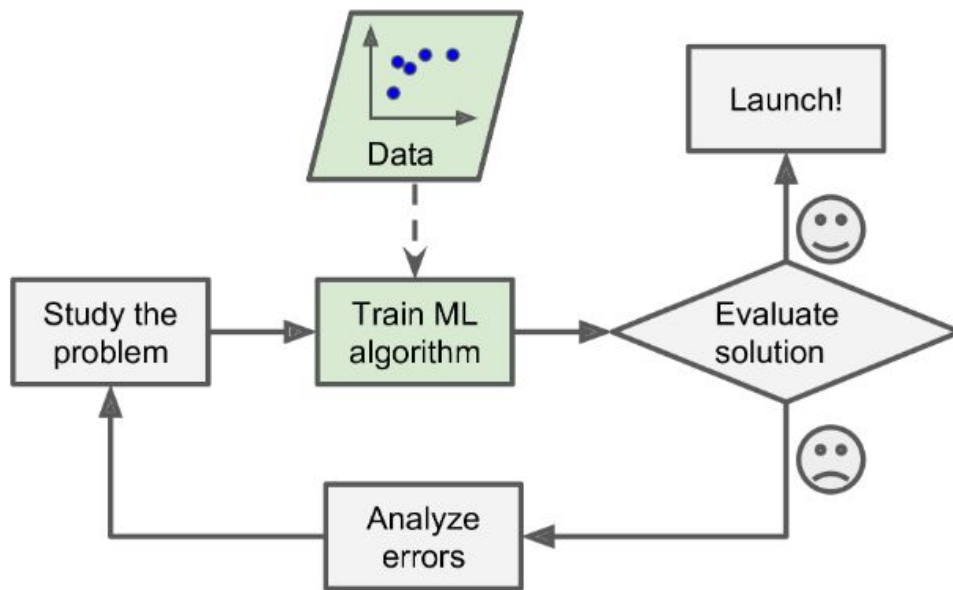
Why Use Machine Learning?

How would you do it ? The traditional approach



Why Use Machine Learning?

How would you do it ? Machine Learning approach



Why Use Machine Learning?

How would you do it ? Machine Learning approach

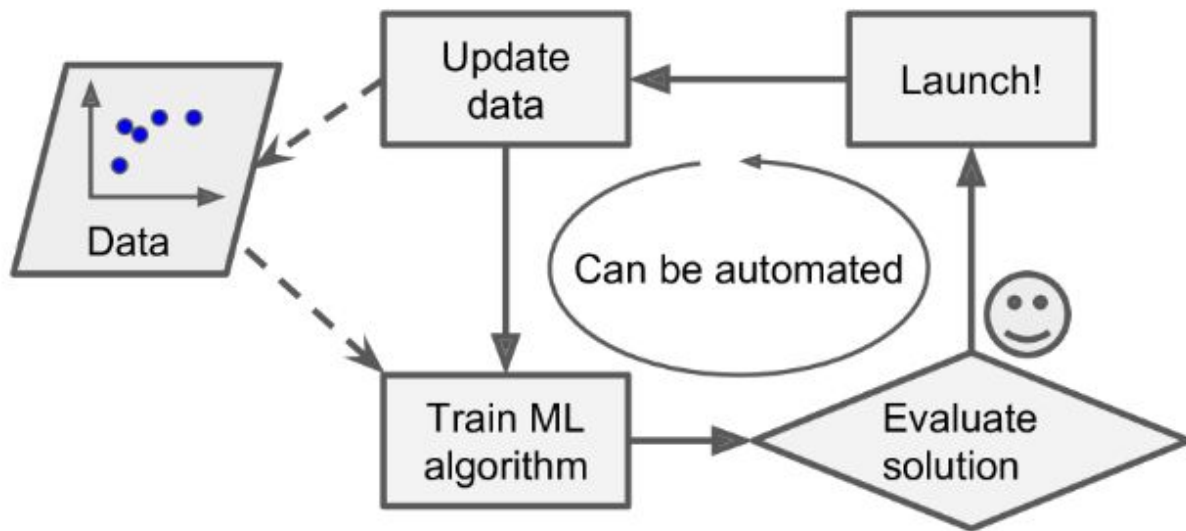
What if spammers notice that all their emails containing “4U” are blocked?, they might start writing “For U” instead

Consequence? your model isn't performing very well

What is your solution for this issue ?

Why Use Machine Learning?

How would you do it ? Automatically adapting to change



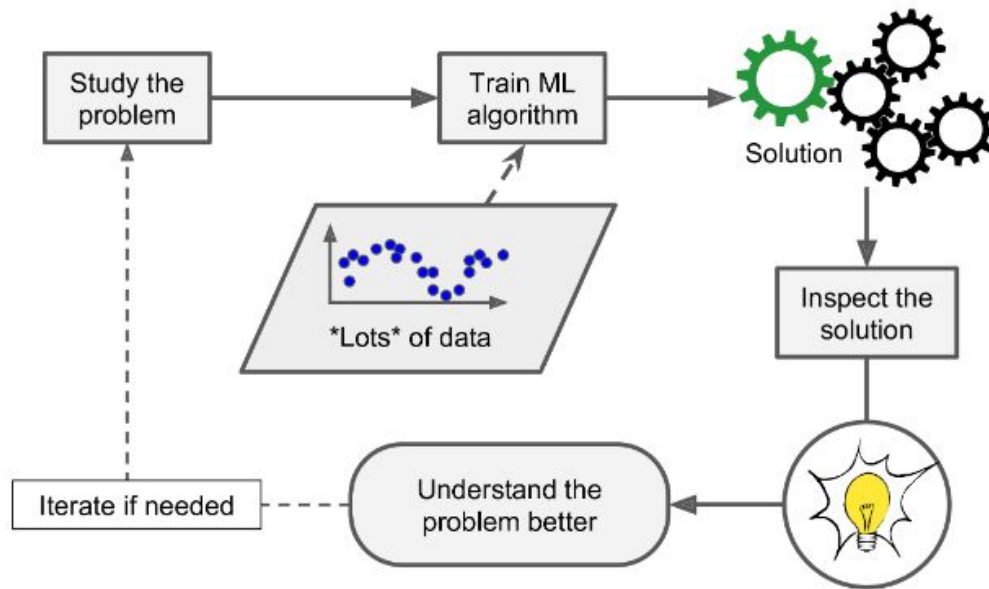
Why Use Machine Learning?

Machine Learning can help humans learn, How ?

Applying ML techniques to dig into large amounts of data can help discover patterns that were not immediately apparent. This is called ***data mining***.

Why Use Machine Learning?

Machine Learning can help humans learn, How ?



Why Use Machine Learning?

To summarize, Machine Learning is great for:

- Problems for which existing solutions require a lot of hand-tuning or long lists of rules
- Complex problems for which there is no good solution at all using a traditional approach
- Fluctuating environments: a Machine Learning system can adapt to new data.
- Getting insights about complex problems and large amounts of data.

Types of Machine Learning Systems

Let's classify them in broad categories based on:

Whether or not they are trained with human supervision

- Supervised
- Unsupervised

Types of Machine Learning Systems

Let's classify them in broad categories based on:

Whether or not they can learn incrementally on the fly

- Online learning
- Batch learning

Types of Machine Learning Systems

Let's classify them in broad categories based on:

Whether they work by simply comparing new data points to known data points

or instead

detect patterns in the training data and build a predictive model

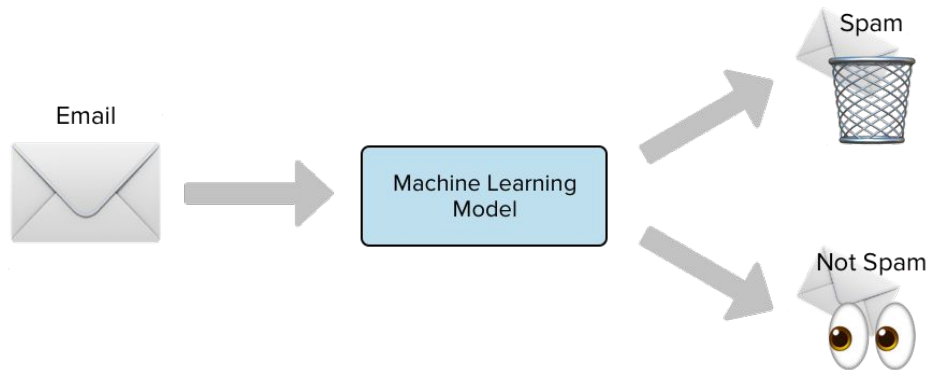
- Model-based learning
- Instance-based learning

Types of Machine Learning Systems

We can combine between criterias

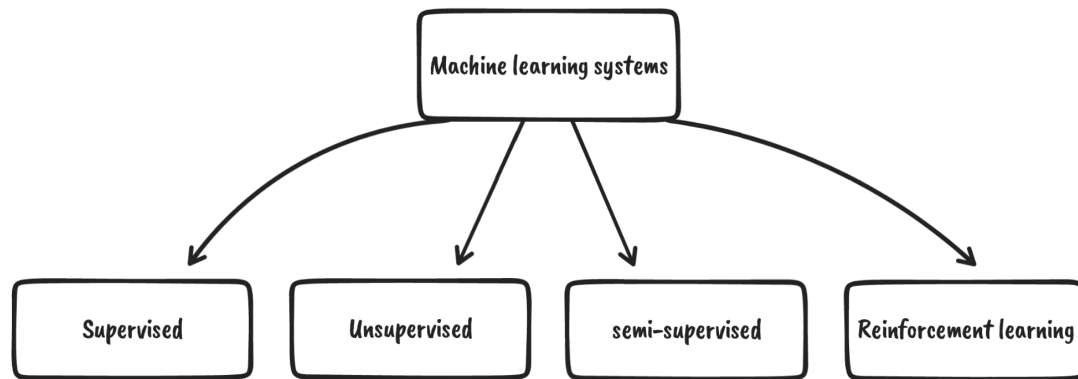
a **state-of-the-art spam filter** may learn on the fly using a deep neural network model trained using examples of spam and ham, this makes it a :

- an online
 - model-based
 - supervised
- machine learning system



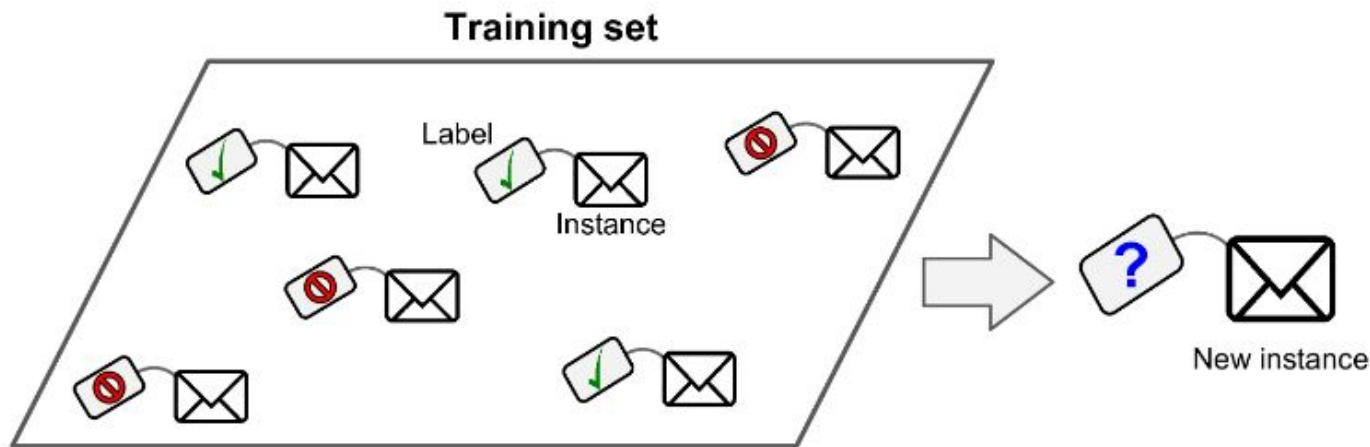
Supervised/Unsupervised learning

Machine Learning systems can be classified according to the **amount** and **type** of supervision they get during training



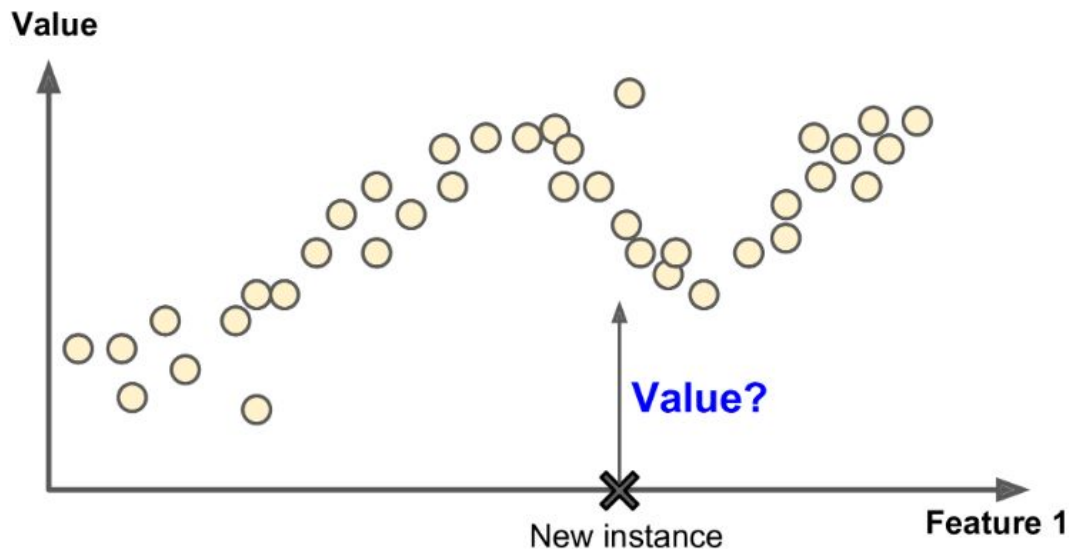
Supervised learning

the training data you feed to the algorithm includes the desired solutions, called **labels**



Supervised learning

Another typical task is to predict a value, such as the price of a car, given a set of features (mileage, age, brand, etc.) This sort of task is called **regression**



Supervised learning

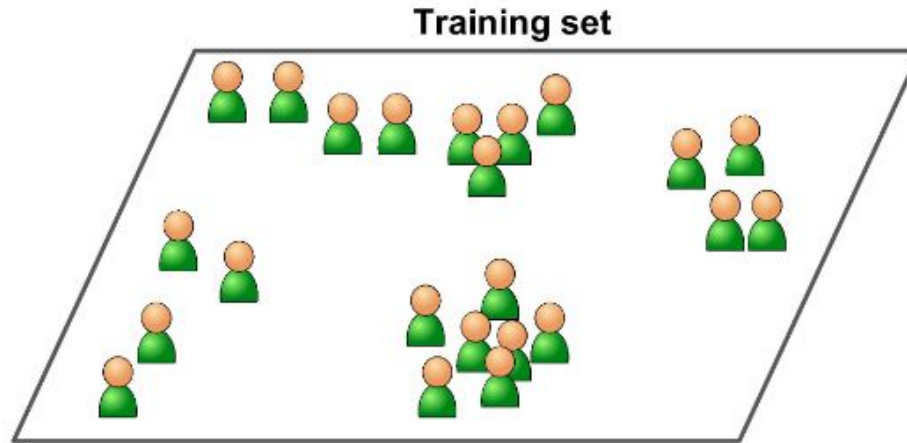
Here are some of the most important supervised learning algorithms

- k-Nearest Neighbors
- Linear Regression
- Logistic Regression
- Support Vector Machines (SVMs)
- Decision Trees and Random Forests
- Neural networks

Unsupervised learning

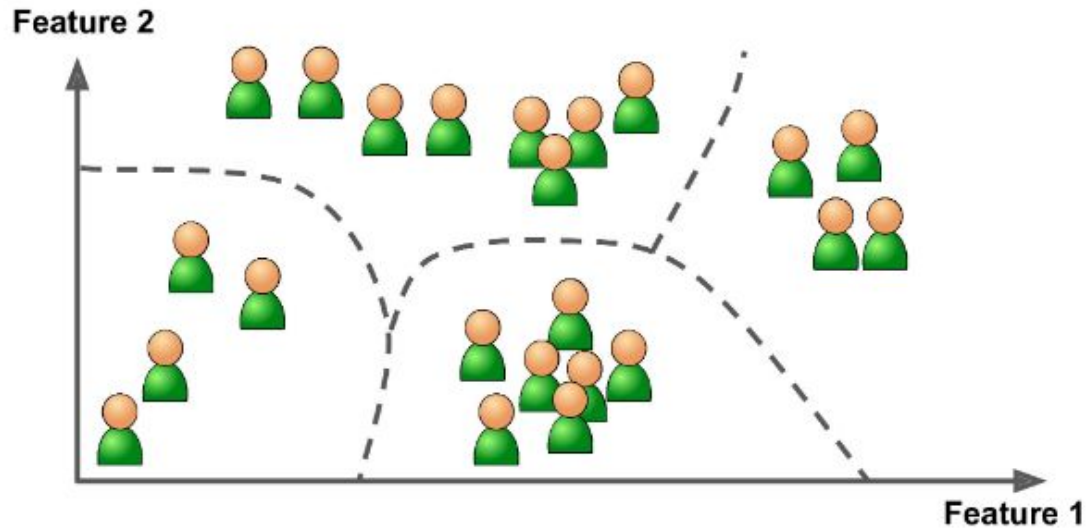
as you might guess, the training data is unlabeled

The system tries to learn without a teacher.



Unsupervised learning

Clustering



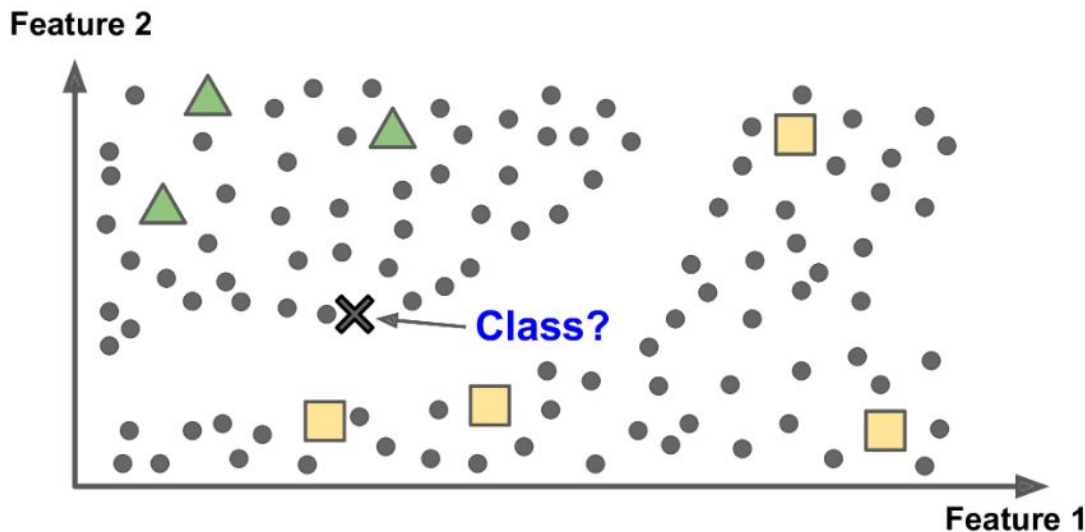
Unsupervised learning

an important unsupervised task is anomaly detection—for example, detecting unusual credit card transactions to prevent fraud



Semi-supervised learning

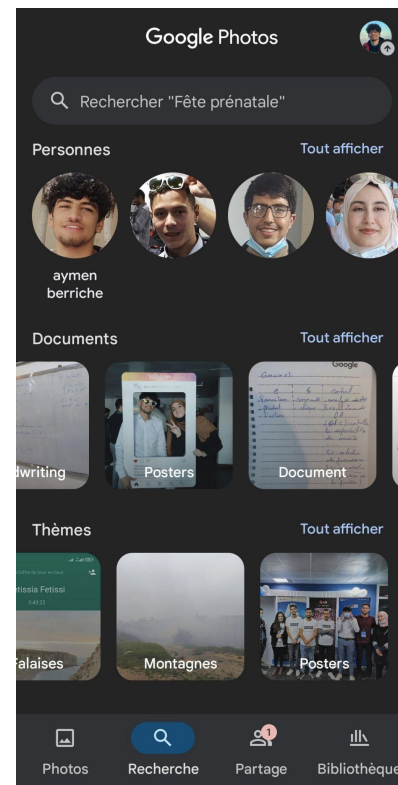
Some algorithms can deal with **partially** labeled training data, usually a lot of **unlabeled** data and a little bit of **labeled** data. T



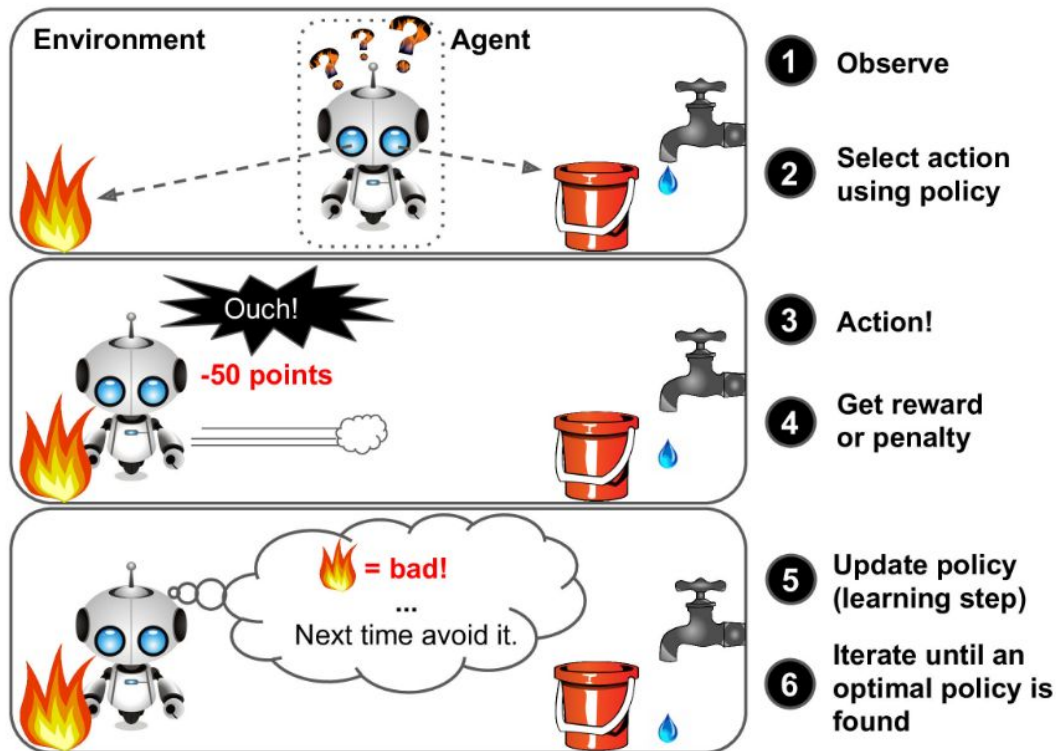
Semi-supervised learning

Google Photos is a good example of this. Once you upload all your family photos to the service, it automatically recognizes that the same person A shows up in photos 1, 5, and 11, while another person B shows up in photos 2, 5, and 7.

This is the unsupervised part of the algorithm (clustering). Now all the system needs is for you to tell it who these people are.



Reinforcement learning



Reinforcement learning

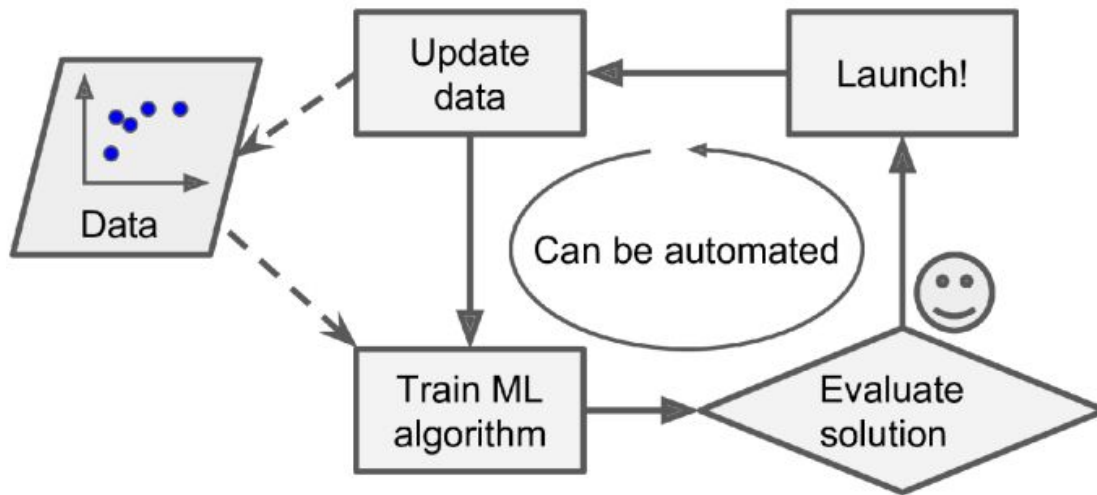


Batch/Online learning

Another criterion used to classify Machine Learning systems is whether or not the system can **learn incrementally** from a stream of incoming data.

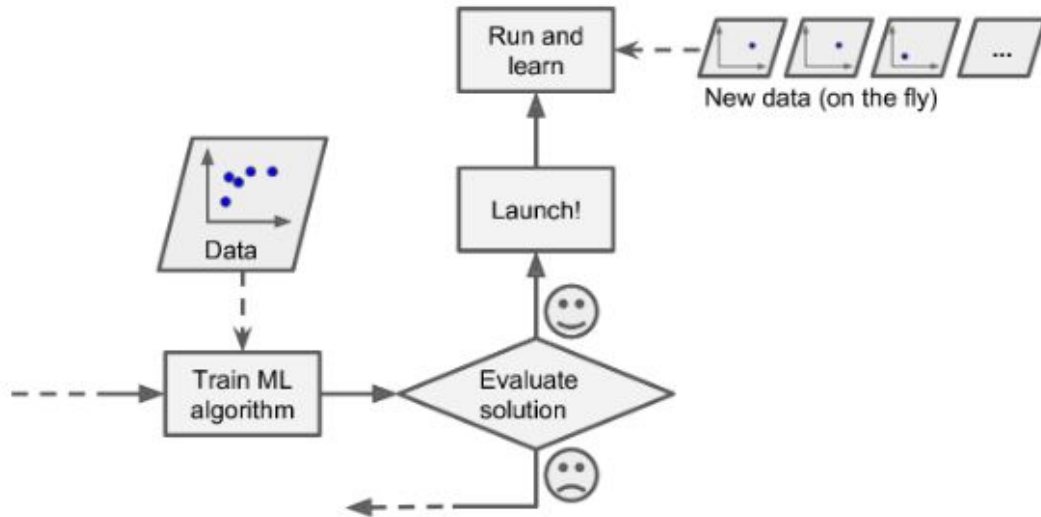
Batch learning

In batch learning, the system is incapable of learning incrementally: it must be trained using **all the available data**



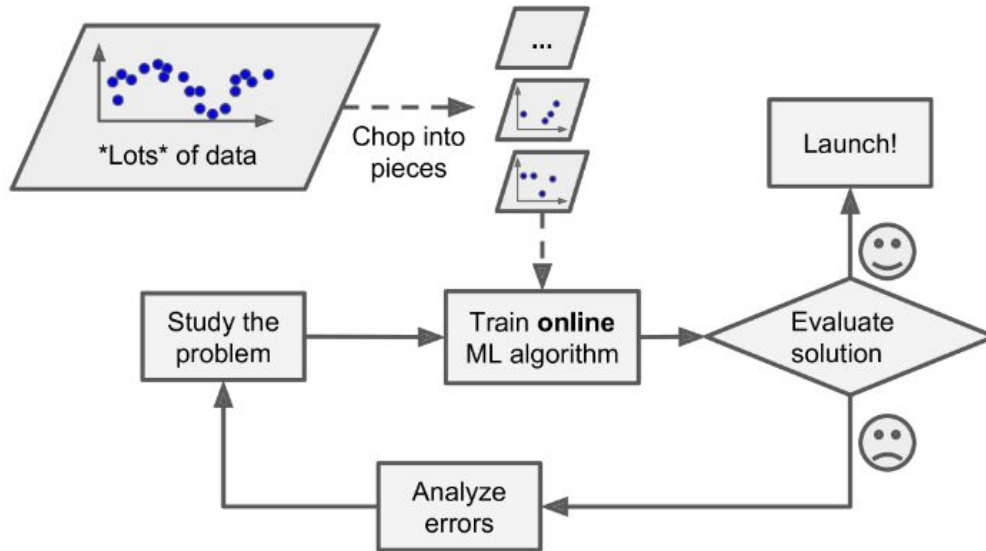
Online learning

you train the system incrementally by feeding it data instances sequentially, either individually or by small groups called **mini-batches**



Online learning

Online learning algorithms can also be used to train systems on huge datasets that cannot fit in one machine's main memory



Instance-Based Versus Model-Based Learning

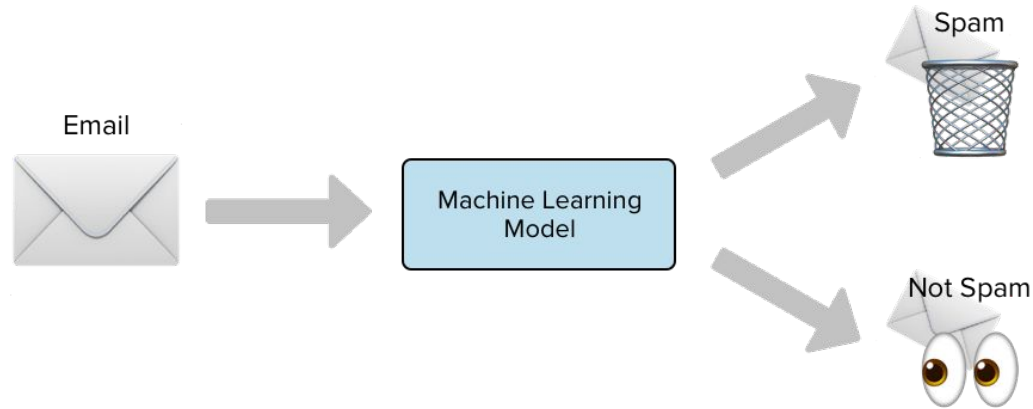
One more way to categorize Machine Learning systems is by how they generalize.

There are two main approaches to generalization:

- instance-based learning
- model-based learning

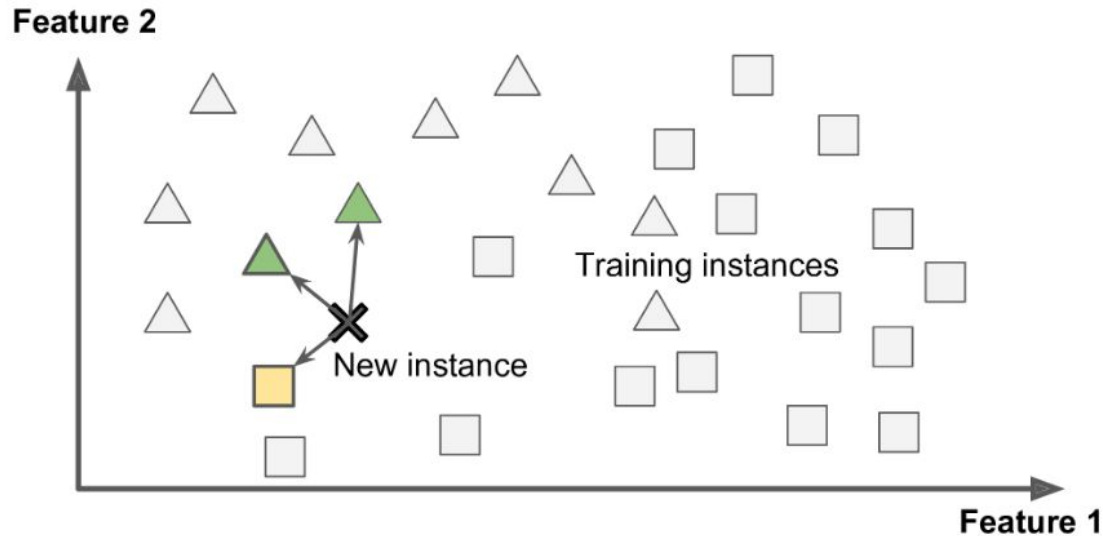
Instance-Based learning

How would we do it using instance-based learning ?



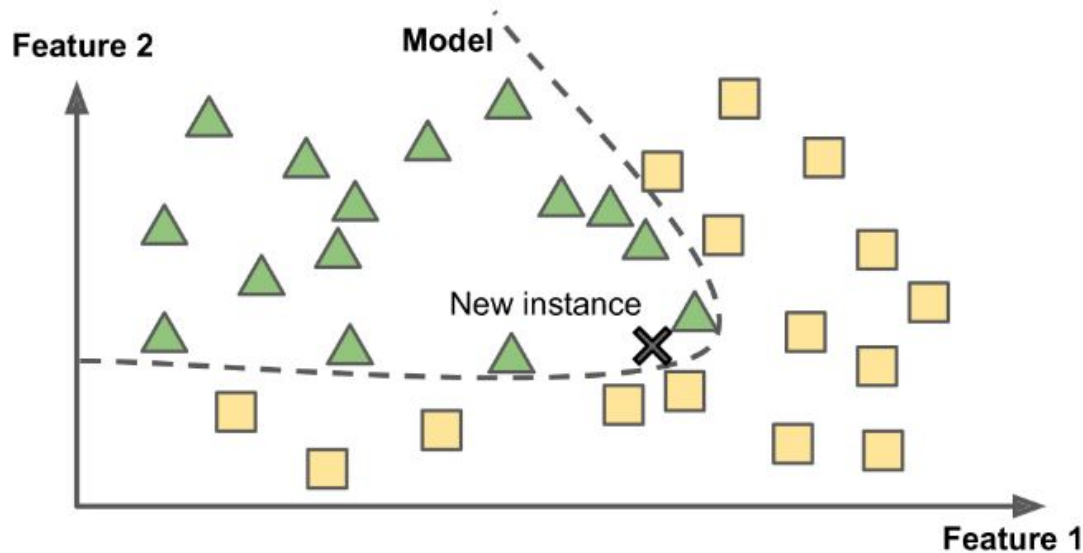
Instance-Based learning

requires a **measure of similarity** between two emails, for example :the number of words they have in common.



Model-Based learning

build a model of these examples, then use that model to make predictions



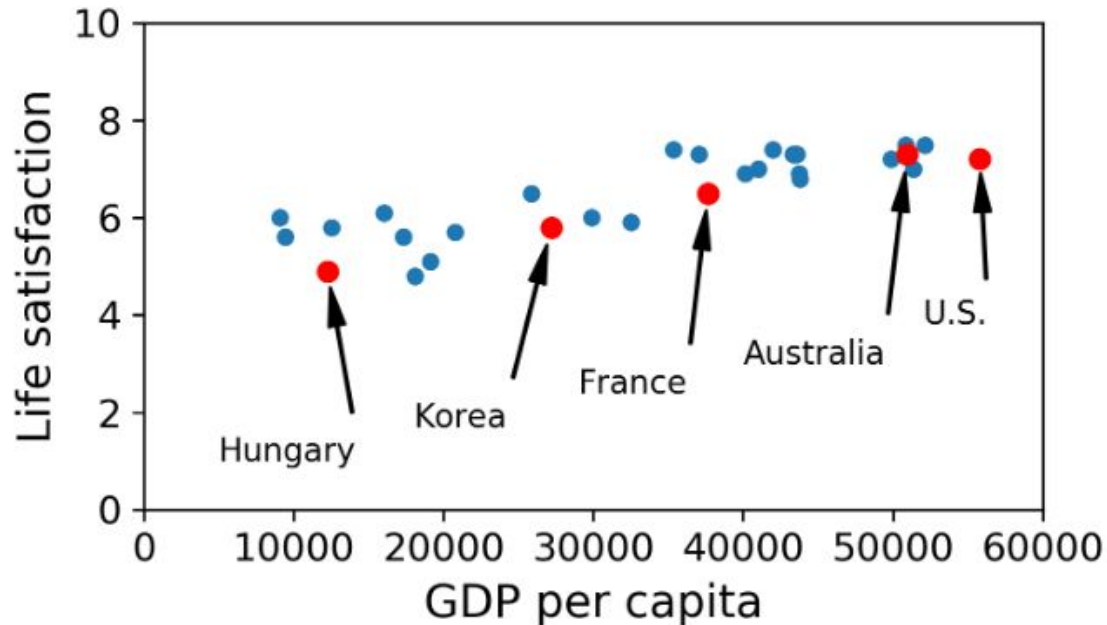
Model-Based learning

For example, suppose you want to know if money makes people happy, so you download the Better Life data

Country	GDP per capita (USD)	Life satisfaction
Hungary	12,240	4.9
Korea	27,195	5.8
France	37,675	6.5
Australia	50,962	7.3
United States	55,805	7.2

Model-Based learning

Let's plot the data for a few random countries



Model-Based learning

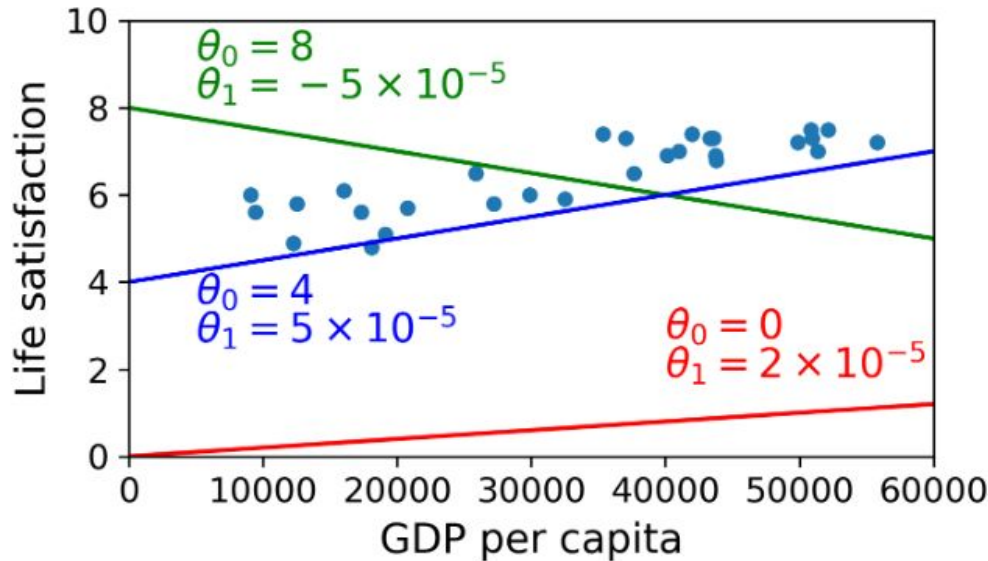
model selection: you selected a linear model of life satisfaction with just one attribute: GDP per capita.

A simple linear model

$$\text{life_satisfaction} = \theta_0 + \theta_1 \times \text{GDP_per_capita}$$

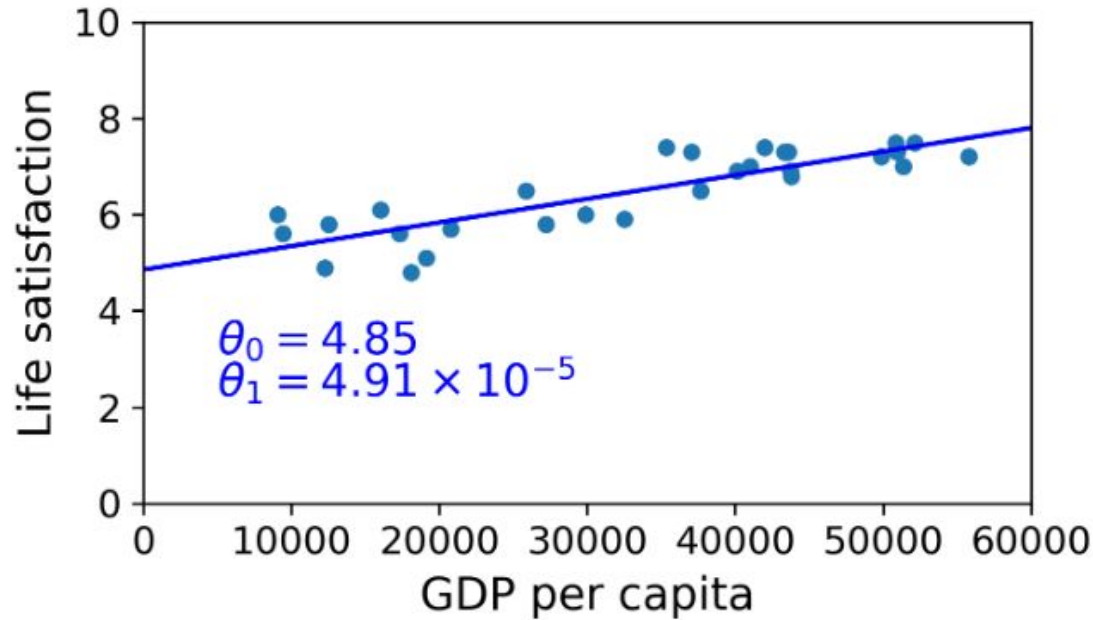
Model-Based learning

This model has two model parameters, θ_0 and θ_1 , By tweaking these parameters, you can make your model represent any linear function



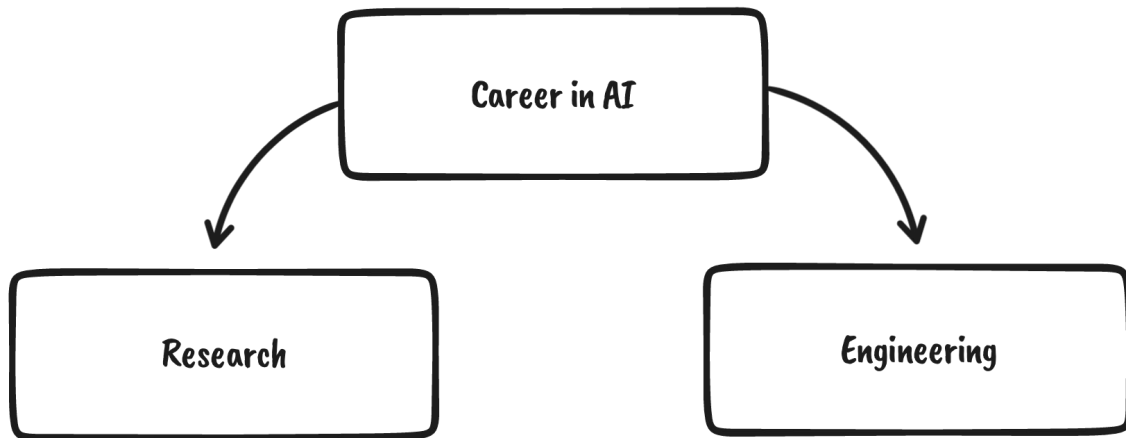
Model-Based learning

The linear model that fits the training data best



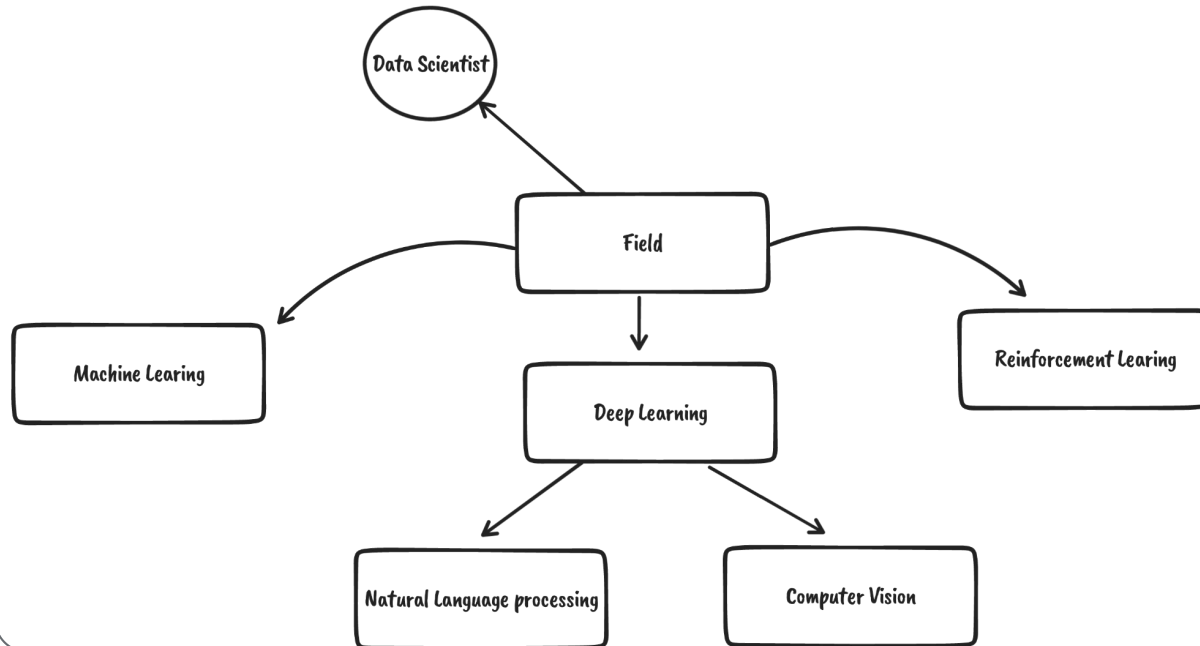
Career Paths

Design your path for a career in AI



Career Paths

Design your path for a career in AI



Career Paths

Design your path for a career in AI

Highly required skills as an Engineer

- Degree in Computer Science, Maths, Machine Learning, or related field.
- Experience in programming and debugging.
- Experience using deep learning frameworks such as PyTorch, Tensorflow, Jax, Keras...

=> Open-source contribution is a nice way to demonstrate your skills.

Career Paths

Design your path for a career in AI

Highly required skills as an **Researcher**

- Phd Degree in Computer Science, Maths, Machine Learning, or related field or equivalent proven experience.
- Experience using deep learning frameworks such as PyTorch, Tensorflow, Jax, Keras...
- Publishing scientific papers at conferences (JMLR, ICLR, NeurIPS, ICML, GECCO, etc.).

=> Scientific publication is a nice way to demonstrate your skills.

Career Paths

Design your path for a career in AI

Researcher

- Uses scientific knowledge to understand the world around us or uncover new things
- Seeks knowledge, sometimes just for the sake of knowing

Engineer

- Applies scientific knowledge to create something new, solve a problem or improve something
- Designing new devices and systems that serve a useful purpose that is not met by existing technology

Thank you for listening, any questions?



MACHINE LEARNING
BOOTCAMP