

Machine Learning

INTRODUCTION

Machine Learning is probably the most well known and widely-used form of Artificial Intelligence. The term was popularized by Arthur Samuel in 1959 and highlighted the notion that it is possible for a machine to solve problems “without being explicitly programmed.” The main idea of machine learning is that it is possible to derive a statistical model capable of making predictions (a predictive model) from a dataset (training dataset) in an automated way. This approach has many applications and advantages when compared to human-written code, such as the speed and ease by which these models can be changed and optimized and the ability to find patterns in the dataset that would be very hard for human eyes to detect. There are many current applications of machine learning today—it is almost ubiquitous—from text suggestion on our phones and spam filtering of our email to face recognition and self driving cars.

(See our Quick Thing on [ChatGPT & Generative AI](#) for more about artificial intelligence.)

EXAMPLES

1. **Virtual Assistants:** Alexa, Siri, Google Now: these voice-activated programs running in our phones, TVs, smart speakers, and other devices help us with a variety of tasks in our everyday lives, things like fact-checking during a conversation, creating appointments, getting directions, or texting a friend our ETA (estimated time of arrival). They rely heavily on machine learning for speech recognition but also collect data from every interaction we make with the software, data that is used in later interactions to customize the experience.
2. **Product Recommendations:** Most online shopping sites, such as Alibaba, use machine learning to choose which items to show you when purchasing on the site. This is done by analysing the behaviour of other customers with a similar profile as yourself and determining which items have a higher probability of being sold and showing them higher on the results page.
3. **Moderating comments:** Instagram uses machine learning to implement an anti-bullying filter for comments. The model takes a user-generated comment and determines if it follows a similar pattern of others that have been previously flagged for abuse.

HOW DO MACHINES LEARN?

There are three main ways machines can create models based on data: Supervised Learning, Unsupervised Learning and Reinforcement Learning.

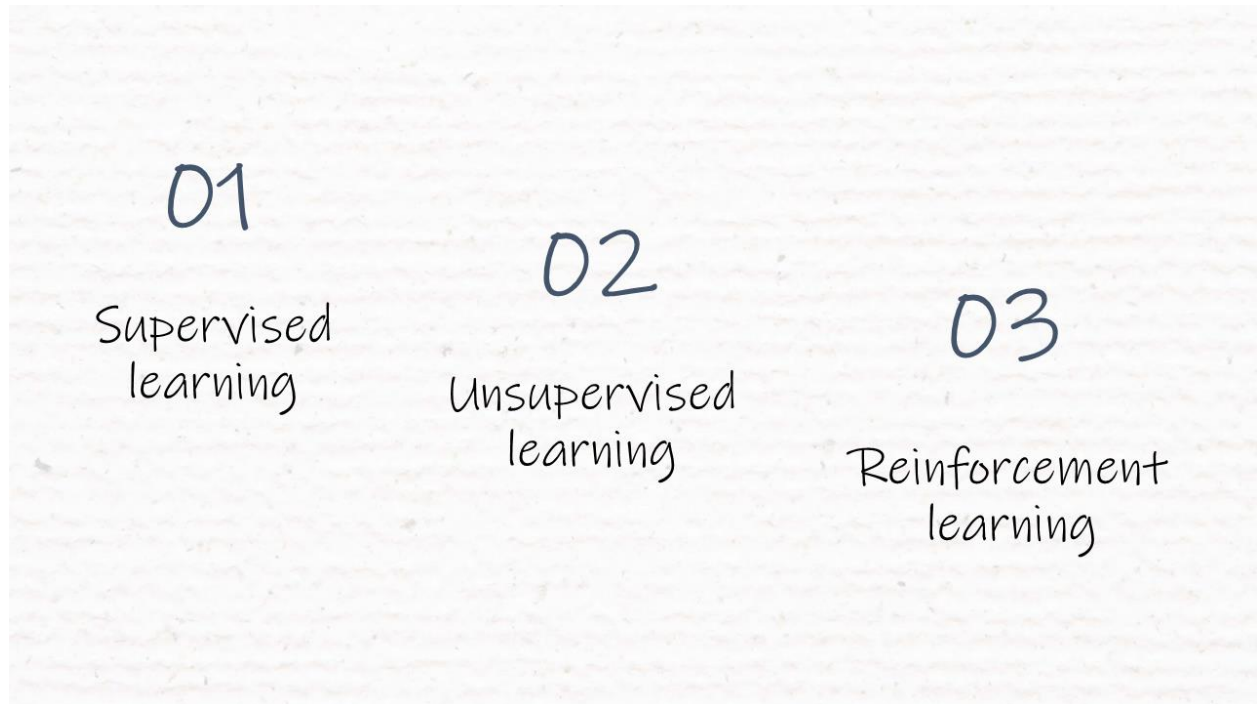


Image description: 3 types of machine learning, supervised learning, unsupervised learning, reinforcement learning.

Supervised Learning: The machine is given a pre-tagged dataset to generate a predictive model. For example, if we wanted to create a model to distinguish a good hotel review from a bad one (what is called sentiment analysis) we will feed the model generator a dataset consisting of the comment and whether the comment is bad or good. After processing, the resulting model will be able to classify a new comment as good or bad with a certain level of accuracy, but it will never be 100% correct.

Supervised learning

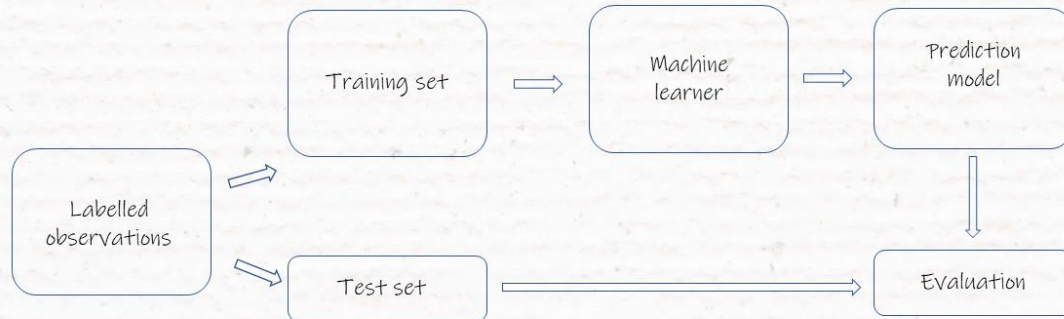


Image description: the process of supervised learning. Start with a dataset that has already been labeled being split into a training set and a test set. The training set then is used to train the machine. The test set is used when the machine has been trained, and the output will be evaluated.

Unsupervised Learning

Here the data is fed to the algorithm without any intervention, and the machine will find patterns and group similar data according to these patterns. For example, given a dataset of news articles, an unsupervised learning algorithm could produce a series of clusters that correspond with the nature of the news article, say, sports or economy. This is called topic modelling. However, it is up to a human to make sense and give meaning to those clusters.

Unsupervised learning

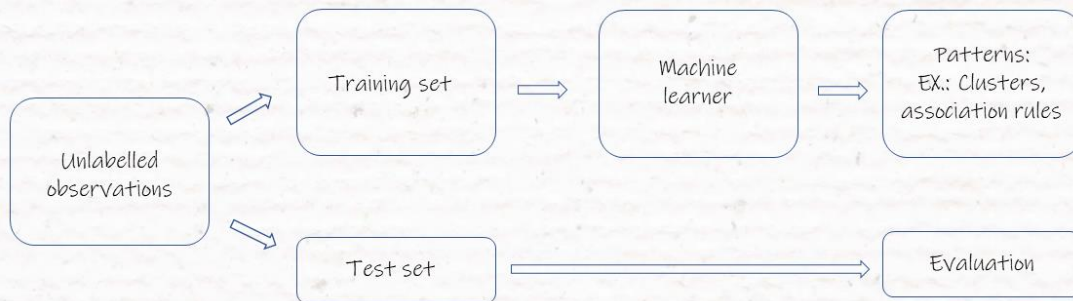


Image description: the process of unsupervised learning. Start with a dataset with unlabelled observations. This is split into a training set and a test set. The machine learns by finding patterns such as clusters and association rules in the data. The test set output is evaluated.

Reinforcement Learning

Unlike the other two main forms of machine learning, reinforcement learning doesn't require a dataset to be provided. Instead, the designers give the computer a set of rules, like the rules of game, and a goal but no hints on how to achieve the goal. The machine will be rewarded or penalized depending on the outcome of the action taken, with the idea of maximizing the rewards while achieving the goal. A very well-known example is AlphaGo Zero. Using reinforcement learning, the computer learned to play the strategy game Go from scratch, by playing against itself. After 40 days of training, it was able to defeat the human champion.

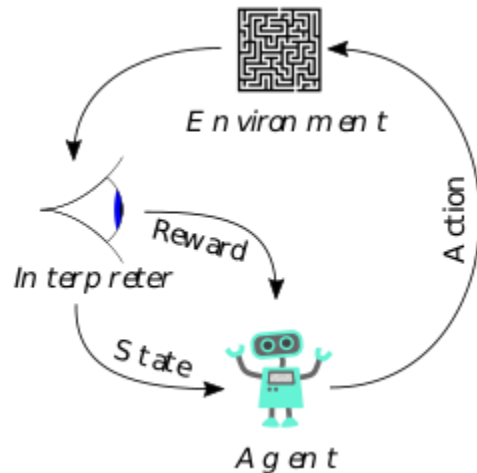


Image description: diagram for reinforcement learning. A computer takes actions based on the input and the interpreter decides if the action taken is appropriate based on the given stimulus.

QUIZ

1. To improve the spam filter, your email provider asks you to flag whether certain emails are spam or legitimate (ham). What kind of machine learning technique are they using?
 - a. **Supervised Learning**
 - b. Unsupervised Learning
 - c. Reinforcement Learning

Feedback: The correct answer is “supervised learning”. The email provider is using your input to construct a tagged dataset (spam vs ham) that will be used re-train the predictive model.

2. The term “machine learning” means that computers are capable of solving some problems without being explicitly programmed.
 - a. **Yes**
 - b. No

Feedback: Yes, that is the expression Arthur Samuel used when coining the term “machine learning” in 1959.

3. Can machine learning be used to fight misinformation on social media?
 - a. **Yes**
 - b. No

Feedback: Yes, for example, Facebook has been using machine learning techniques in conjunction with human reviewers to fight the dissemination of fake news in their platform.

ACTIVITY

Purpose: This activity gives you an opportunity to create and test your own predictive model.

Task: Go to the Teachable Machine site and follow the instructions to create a model that can recognize two (or more) objects. For example: if you have a favourite coffee mug, you can gather a few others and teach the machine to tell them apart.

Time: This activity should take about 20 minutes to complete.

Next: Now that you know how to make a predictive model. Can you think of a possible application in your own studies or research?

FURTHER RESOURCES

1. Udemy Course

- a. **Title:** Machine Learning for Absolute Beginners - Level 1
- b. **Description:** Learn the Fundamental Concepts of Artificial Intelligence and Machine Learning as the Next Game-Changing Technology
- c. **URL:** <https://concordia.udemy.com/course/machine-learning-for-absolute-beginners-level-1/>

2. Book

- a. **Title:** Artificial intelligence in practice: how 50 successful companies used artificial intelligence to solve problems
- b. **Description:** Cyber-solutions to real-world business problems. Artificial Intelligence in Practice is a fascinating look into how companies use AI and machine learning to solve problems. Presenting 50 case studies of actual situations, this book demonstrates practical applications to issues faced by businesses around the globe
- c. **URL:** <https://concordiauniversity.on.worldcat.org/oclc/1086409401>

3. Article:

- a. **Title:** The Imitation Game

- b. **Description:** Alan Turing ‘s original paper from 1950 where he considers, perhaps for the first time, the question, ‘Can machines think?’
- c. **URL:** <https://doi.org/10.1093/mind/LIX.236.433>

REFERENCES

- Silver, D., Schrittwieser, J., Simonyan, K. ... Hassabis, D. Mastering the Game of Go Without Human Knowledge. *Nature*. 2017. 550(7676): 354-359.
- Wikipedia. Arthur Samuel. 2023
- Wikipedia. Machine Learning. 2023
- Blei, D. Probabilistic Topic Models. *Communications of the ACM*. 2012; 55(4): 77-84.
- Lyons, T. Increasing Our Efforts to Fight False News. *Meta*, 2018 June 21.
- Bernard, M. & Ward, M. Artificial Intelligence in Practice: How 50 Successful Companies Used Artificial Intelligence to Solve Problems. Chichester, West Sussex: Wiley, 2019.

VIDEO TRANSCRIPT

Things for Digital Knowledge. Concordia University Library.

Francisco Berrizbeitia, Developer

Intro slide: Concordia library logo

Interviewer: Can you give us a quick introduction on machine learning? What is it? [question displayed on screen]

Francisco Berrizbeitia: Yes, of course. Machine learning is probably the most popular and most widely-used form of artificial intelligence. The term was coined in 1959 by Arthur Samuel. And the basic concept behind it is that computers are capable of resolving problems without being explicitly programed by a human.

Visual slide text: Use a computer to solve problems, with the phrase “a computer” circled

Francisco Berrizbeitia: So basically, what they do is design these algorithms, the output is an algorithm or a predictive model, and they give you an example. It's if you want to perform an algorithm that adds two numbers, like a very simple one. You have two inputs and say, two and two, and the output will be, you know, the addition of those two numbers in this case, is four.

Visual image description: A diagram shows numbers (such as A & B) were input into an algorithm and the output is a number

Visual image description: Diagram shows a table being labelled as input, a table representing many data being input into a machine learner, and output.

Francisco Berrizbeitia: In the case for machine learning scenario, the input will be a dataset, it would be a ton of numbers and the output won't be a number. Well, it won't give you a result. And what it will give you is a statistical model that can then be re-used in all sorts of applications in the real world.

Interviewer: Can you give us a few examples in real life? [question displayed on screen]

Francisco Berrizbeitia: Oh, AI these days is pretty much ubiquitous. You can find it like everywhere from your phone every time you say, Hey Siri or Hey Google or stuff like that.

Visual image description: Virtual assistant picture, slide text: voice activated

Visual video description: A person laughing while holding a phone

Francisco Berrizbeitia: Those virtual assistants use AI intensively from understanding the way you talk, recognizing your speech, what you mean when you talk to them, and they also learn from you; your tastes, where you live, your location. The idea will be to try to customize the experience you have with them to you. And another example will be like online shopping, for example.

Visual image description: Product recommendations

Visual video description: A person's hand swiping a device screen and then typing on a keyboard

Francisco Berrizbeitia: They use extensively AI, sites like Alibaba or Amazon. They do analyze all the data from people who are navigating their sites and shopping on their sites. And based on that, they will be able to give you recommendations of products.

Visual image description: A robot reading a laptop. Slide text: hate speech is crossed out

Francisco Berrizbeitia: And it's also used on social media to do some form of moderation. Instagram, for example, has an interesting use of it because they used it to detect online bullying on the posts. So it has many, many applications in the real world.

Interviewer: Oh, I see. But how does a machine learn? Can you tell us a bit? [question displayed on screen]

Francisco Berrizbeitia: Yes. There are different ways machines can learn. They are divided by the type of data they receive, and the three main forms of doing it are called supervised learning, unsupervised learning, and reinforcement learning.

Visual slide text: 1 is supervised learning, 2 is unsupervised learning, 3 is reinforcement learning

Visual image description: Diagram of supervised learning. Start with labelled observations. These are split into a training set and test set. The training set then is used to train machine learners and a prediction model will be generated. A test set is then used as the evaluation.

Francisco Berrizbeitia: In supervised learning, the input data is a dataset that has been pre-labeled by humans. Say, for example, if you want to build a model that is able to predict whether an email is legitimate or not, you will need a dataset that has a lot of email examples labeled, such as legitimate or illegitimate, ham or spam.

Visual image description: Diagram of unsupervised learning. Start with unlabeled observations, which are split into a training set and test set. A training set is set aside to train a machine learner, then patterns will be generated, such as clusters, associations, rules. A test set is then used as evaluation.

Francisco Berrizbeitia: In unsupervised learning, the algorithm is fed with data that is untouched, that is just raw data. And then the algorithm will find patterns in the dataset and output these patterns. Then it is up to the human to make sense of what this pattern means, and an example would be, say, you feed this algorithm a big dataset of news articles. It will cluster those articles based on similarity, say for example, similar words used for example in different clusters. And then it's up to the human to make sense of those clusters, like, there will be semantic similarities within those clusters if, for example, that they all belong to the same category. So like they will cluster all the sports articles together, all the economics articles, and so on and so forth. In reinforcement learning, it's a bit different in a sense. You don't feed the machine data, what you feed the machine is a set of rules like the rules of the game.

Visual image description: Diagram of reinforcement learning. The computer takes actions based on the input, the interpreter decides if the action taken is appropriate based on the given stimulus.

Francisco Berrizbeitia: Say, for example, if you want to teach a machine how to play chess, you will input all the rules, how you can move all the pieces of chess, then you would feed them what the goal is of this game, and also their reward and the punishment, you know, of how you play that while playing the game. And then by repetition, by running thousands of times, the repetition of the game is basically playing against itself. The machine will then learn how to play chess. There are examples of these, like the AlphaGo machine, which was amazing, it beat the grandmaster.

Visual image description: Thank you and credit. CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon and infographics & images by Freepik. Music: Lofi Summer Background by Vladislav Kurnikov on Pixabay.