**MACHINE LEARNING**

**Introduction:**

The scientific field of machine learning (ML) is a branch of artificial intelligence, as defined by Computer Scientist and machine learning pioneer Tom M. Mitchell “Machine learning is the study of computer algorithms that allow computer programs to automatically improve through experience .”.

Machine-learning algorithms use statistics to find patterns in massive amounts of data. And data, here, encompasses a lot of things numbers, words, images, clicks, what have you. If it can be digitally stored, it can be fed into a machine-learning algorithm.

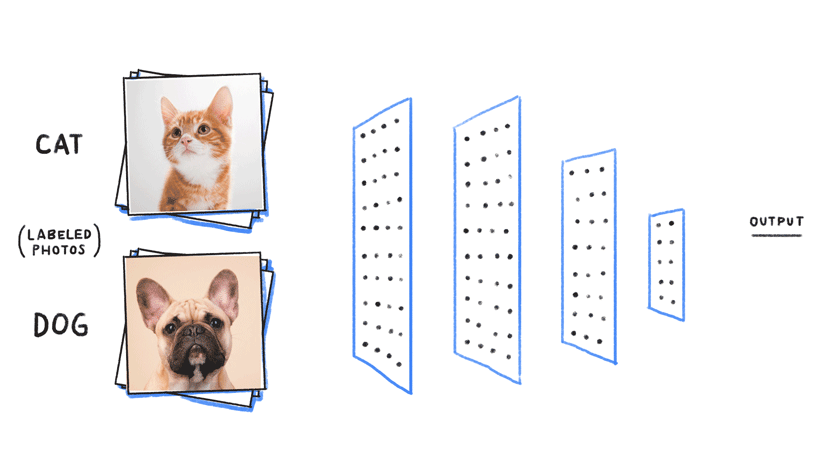
Machine learning is the process that powers many of the services we use today recommendation systems like those on Netflix, YouTube, and Spotify; search engines like Google and Baidu; social-media feeds like Facebook and Twitter; voice assistants like Siri and Alexa.

**Working of Machine Learning**

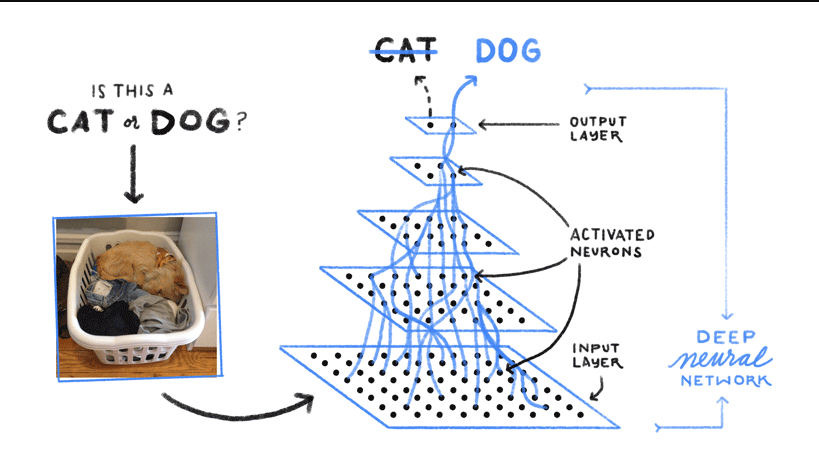
In practice, the patterns that a computer (machine learning system) learns can be very complicated and difficult to explain. Consider searching for dog images on Google search Google is incredibly good at bringing relevant results, yet how does Google search achieve this task? In simple terms, Google search first gets a large number of examples (image dataset) of photos labeled “dog” — then the computer (machine learning system) looks for patterns of pixels and patterns of colors that help it guess (predict) if the image queried it is indeed a dog.

At first, Google’s computer makes a random guess of what patterns are reasonable to identify an image of a dog. If it makes a mistake, then a set of adjustments are made in order for the computer to get it right. In the end, such collection of patterns learned by a large computer system modeled after the human brain (deep neural network), that once is trained can correctly identify and bring accurate results of dog images on Google search, along with anything else that you could possibly think of such process is called the training phase of a machine learning system.

At first, Google’s computer makes a random guess of what patterns are reasonable to identify an image of a dog. If it makes a mistake, then a set of adjustments are made in order for the computer to get it right. In the end, such collection of patterns learned by a large computer system modeled after the human brain (deep neural network), that once is trained can correctly identify and bring accurate results of dog images on Google search, along with anything else that you could possibly think of such process is called the training phase of a machine learning system.



Imagine that you were in charge of building a machine learning prediction system to try and identify images between dogs and cats. The first step, as we explained above, would be to gather a large number of labeled images with “dog” for dogs and “cat” for cats. Second, we would train the computer to look for patterns on the images to identify dogs and cats, respectively.



Once the machine learning model has been trained .we can throw at it (input) different images to see if it can correctly identify dogs and cats. As seen in the image above, a trained machine learning model can (most of the time) correctly identify such queries.

**Methodologies:**

Machine learning algorithms are often categorized as supervised or unsupervised.

* **Supervised machine learning** algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.
* **Unsupervised machine learning** algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn’t figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.
* **Semi-supervised machine learning** algorithms fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabeled data generally doesn’t require additional resources.
* **Reinforcement machine learning algorithms** is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

Machine learning enables analysis of massive quantities of data. While it generally delivers faster, more accurate results in order to identify profitable opportunities or dangerous risks, it may also require additional time and resources to train it properly. Combining machine learning with AI and cognitive technologies can make it even more effective in processing large volumes of information.

**Software for Machine Learning**

Free and Open Source software

* scikit-learn
* Shogun
* Spark MLlib
* SystemML
* TensorFlow
* Torch / PyTorch

Proprietary Software

* Amazon Machine Learning
* Angoss KnowledgeSTUDIO
* Azure Machine Learning
* Ayasdi
* IBM Data Science Experience
* Google Prediction API

**Applications:-**

* Virtual Personal Assistants:- Siri, Alexa, Google Now are some of the popular examples of virtual personal assistants.
* Predictions while Commuting:- When booking a cab, the app estimates the price of the ride,.
* Videos Surveillance:- The video surveillance system nowadays are powered by AI that makes it possible to detect crime before they happen. They track unusual behaviour of people like standing motionless for a long time, stumbling, or napping on benches etc.
* Social Media Services:- From personalizing your news feed to better ads targeting, social media platforms are utilizing machine learning for their own and user benefits.
* Email Spam and Malware Filtering:- There are a number of spam filtering approaches that email clients use. To ascertain that these spam filters are continuously updated, they are powered by machine learning.
* Search Engine Result Refining:- Google and other search engines use machine learning to improve the search results for you.

**Limitations:**

Although machine learning has been transformative in some fields, machine-learning programs often fail to deliver expected results. Reasons for this are numerous: lack of (suitable) data, lack of access to the data, data bias, privacy problems, badly chosen tasks and algorithms, wrong tools and people, lack of resources, and evaluation problems.

In 2018, a self-driving car from Uber failed to detect a pedestrian, who was killed after a collision. Attempts to use machine learning in healthcare with the IBM Watson system failed to deliver even after years of time and billions of investment.

Machine learning approaches in particular can suffer from different data biases. A machine learning system trained on current customers only may not be able to predict the needs of new customer groups that are not represented in the training data. When trained on man-made data, machine learning is likely to pick up the same constitutional and unconscious biases already present in society.

Machine learning systems used for criminal risk assessment have been found to be biased against black people.

In 2015, Google photos would often tag black people as gorillas, and in 2018 this still was not well resolved, but Google reportedly was still using the workaround to remove all gorillas from the training data, and thus was not able to recognize real gorillas at all .

**References**

* <https://expertsystem.com/machine-learning-definition/>
* <https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/>
* <https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0>
* <https://medium.com/towards-artificial-intelligence/what-is-machine-learning-ml-b58162f97ec7>
* <https://en.wikipedia.org/wiki/Machine_learning>