

0 5 . 0 3 . 1 7

WorldQuant Perspectives

WELCOME TO THE MACHINE

Unlike the science of artificial intelligence, which has yet to live up to the promise of replicating the human brain, machine learning is changing the way we do everything from face recognition and social media advertising to medical diagnosis and stock trading.

WORLDQUANT[®]

WorldQuant, LLC
1700 East Putnam Ave.
Third Floor
Old Greenwich, CT 06870
www.weareworldquant.com

AS SO OFTEN HAPPENS WITH SOPHISTICATED BUT FASHIONABLE ideas, there is some confusion around the concepts of artificial intelligence and machine learning. In 1950 — five full years before the term “artificial intelligence” (AI) was coined — British mathematician Alan Turing introduced a test for determining an intelligent computer: A machine is truly intelligent, he wrote, if a human asking it questions can’t tell the difference between the computer’s answers and those of a human.

Since then, AI has become an element of popular culture, represented as something that can easily satisfy the Turing test. In Hollywood movies like *2001: A Space Odyssey*, the *Terminator* saga and, most recently, *Her*, AI appears almost humanlike. However, Turing’s definition of an intelligent machine is more philosophical than practical. What nowadays is referred to as AI usually is mostly based on machine learning technologies. The recent overwhelming achievements of machine learning — spam detection; text, speech and face recognition; medical diagnosis; stock trading; and social media advertising — have nothing to do with the popular perception of AI. The common misunderstanding of the technologies behind these successes has blurred what are purely engineering problems. In this article we’ll try to dispel the misconceptions and briefly explain some of the most advanced methods of machine learning.

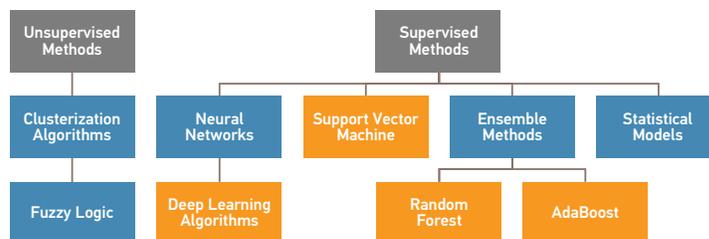
First, it’s important to understand that there is no generally accepted definition of artificial intelligence in modern science. You will hardly ever hear the acronym AI used by the scientists and engineers responsible for the technologies behind automated decision making, image recognition or web searching. AI is an extremely wide family of methods and algorithms with one common attribute: All of them were inspired by attempts to understand and reproduce human decision making, human vision and human searching. Indeed, it has always been painful for information scientists that any three-year-old child can immediately distinguish between pictures of a dog and a cat with 100 percent precision, while for computer algorithms this task remains a challenge. The natural reaction to such situations has been, “Let’s just do the same as humans.” But doing the same has appeared to encourage so many different approaches and interpretations that it becomes difficult to describe AI as a well-determined area of knowledge.

Machine learning traces its roots to Frank Rosenblatt, an American psychologist who in 1957 invented the perceptron while working at Cornell Aeronautical Laboratory in Buffalo, New York. The idea behind the perceptron was both obvious and

promising: Using the assumption that the human brain is nothing but a collection of neurons, Rosenblatt came up with the notion of creating an artificial neuron that would make it “feasible to construct an electronic or electromechanical system which will learn to recognize similarities or identities between patterns of optical, electrical, or tonal information, in a manner which may be closely analogous to the perceptual processes of a biological brain.” An artificial neuron is the simplest linear classifier of input signals and has parameters that can be fit (taught) on a set of training examples. A perceptron is an assembly of artificial neurons that may represent stronger and more complex classifiers. The disadvantage is that having too many parameters may lead to what is called overfitting, when good training results do not imply good results on new data.

METHOD ACTING

Machine learning typically follows one of two paths (the most popular methods are in orange).



The field of machine learning encompasses several families of decision-making, optimization and classification algorithms that rely on theory and produce proven results if applied properly. Machine learning can resolve classification tasks (whether the picture shows a dog or a cat) or prediction tasks (trying to determine the price of stocks tomorrow). Its algorithms can usually be trained on a set of examples, applying the knowledge obtained to new data. Here we give a brief overview of the most developed directions and methods in machine learning.

In supervised learning we expect the machine to train itself from the set of examples provided. For instance, in the case of the classification of cats and dogs, we prepare a set of pictures with the correct labels and let the computer teach itself to distinguish between them. By contrast, unsupervised algorithms are used for clustering tasks when we don’t know the correct answers in advance. In this case the machine search is directed with some predefined quality criteria. For investors, the task of predicting stock prices can be a good application for supervised learning, whereas a machine algorithm that selects stocks for a portfolio will likely be unsupervised.

FUZZY LOGIC METHODS

In classical logic every statement is either true or false. In real life this is not always enough: We can think of some statements that are likely true, indefinite or unlikely true. In other words, there is a grayscale between yes and no. Allowing machines to operate in such an underdefined world, we can produce a whole line of reasoning called fuzzy logic, which has imprecise inference rules and provides mechanisms to make decisions with a lack of information. Fuzzy logic is used in many applications, including face recognition, air conditioners, washing machines, antilock braking systems, subway and unmanned-vehicle controllers, weather forecasting systems, models for new-product pricing or project risk assessment, medical diagnosis and treatment plans, and stock trading. Digital camera autofocus systems are a good example of a simple technical device that is usually based on fuzzy logic.

STATISTICAL MODELS

Statistical models — hidden Markov, linear discriminant analysis, logistic regression, naive Bayes — are a group of algorithms that are good for solving relatively simple problems that do not need highly precise classification or prediction. These methods are easy to implement and not very sensitive to missing data. The disadvantage is that each of these approaches presumes some specific data model. For example, the naive Bayes method is sometimes used in medical diagnostics to derive the diagnosis from the results of some measurements; it's critical that all the measurements are independent.

Statistical models constitute the basis of many systems. They are used for classification in various fields, from physics to linguistics to financial forecasting. Naive Bayes models are widely applied in medical diagnostics; hidden Markov models are typically used for voice recognition and natural language processing.

“ THE RECENT OVERWHELMING ACHIEVEMENTS OF MACHINE LEARNING HAVE NOTHING TO DO WITH THE POPULAR PERCEPTION OF AI. ”

“ YOU WILL HARDLY EVER HEAR THE ACRONYM AI USED BY THE SCIENTISTS AND ENGINEERS RESPONSIBLE FOR THE TECHNOLOGIES BEHIND AUTOMATED DECISION MAKING, IMAGE RECOGNITION OR WEB SEARCHING. ”

SUPPORT VECTOR MACHINE

In the 1960s, Soviet scientists Vladimir Vapnik and Alexey Chervonenkis invented the generalized portrait algorithm, which led to the family of classification algorithms known as support vector machine (SVM). The generalized portrait and SVM algorithms were new steps in machine learning. In contrast to Rosenblatt's perceptron, in SVM the solution presumes the best generalization of the results. In other words, the accuracy achieved during training is preserved off-sample and prevents overfitting.

Unlike many other machine learning algorithms, SVM has a strong theoretical base. It is very robust and has several modifications for various types of problems, but it usually requires a lot of training time and is not ideal for parallel computations — a disadvantage for modern algorithms. SVM is very common in a lot of applications, including handwritten text recognition, classification of images, classification of proteins and genes in biology, and identification of diseases according to symptoms and physical measures.

NEURAL NETWORKS

The perceptron, which attempted to represent the structure of a human brain with artificial neurons, was a neural network (NN). Scientists later learned that the key to replicating the function of the brain is not the neurons but how we connect them to one another and how we train them. Although there is no theory that can suggest how to build a neural network for any specific task, there is a well-known backpropagation algorithm that allows neural networks to be trained. (Think of a neural network not as a specific algorithm but as a specific way to represent algorithms.)

Neural networks are used in a lot of applications today — most notably, image recognition. But the combination of a lack of theory and the lack of internal transparency of neural networks (their internal logic is hard to interpret) has somewhat held back their development. Nevertheless, neural networks are very efficient for anyone who has access to a supercomputer. They play an important role in many artificial intelligence systems, including computers that beat human players in chess, determine credit ratings and detect fraudulent Internet activity.

ENSEMBLE METHODS

Ensemble methods like the AdaBoost and random forest algorithms aggregate the solutions of multiple weak classifiers or predictors with poor accuracy to obtain a stronger classification or more precise prediction. The Random Forest algorithm constructs a set of simple decision trees by relying on training examples and then makes these trees vote to obtain the final decision for new cases. AdaBoost (which stands for “adaptive boosting”) aggregates various weak classifiers to obtain a final decision as a weighted combination of weak decisions. These algorithms are very scalable and good for parallel computations, making them a perfect fit for computer vision applications.

DEEP LEARNING

Few areas of machine learning have received more hype than deep learning. Some data scientists think deep learning is just a buzzword or a rebranding of neural networks. The name comes from Canadian scientist Geoffrey Hinton, who created an unsupervised method known as restricted Boltzmann machine (RBM) for pretraining neural networks with a large number of neuron layers. RBM was meant to improve backpropagation training, but there is no strong evidence that it is really an improvement.

“MACHINE LEARNING TRACES ITS ROOTS TO FRANK ROSENBLATT, AN AMERICAN PSYCHOLOGIST WHO IN 1957 INVENTED THE PERCEPTRON.”

“THE COMBINATION OF A LACK OF THEORY AND THE LACK OF INTERNAL TRANSPARENCY OF NEURAL NETWORKS HAS SOMEWHAT HELD BACK THEIR DEVELOPMENT.”

Another alternative direction of deep learning is recurrent neural networks (RNN), which are widely used in voice recognition (“Hello, Siri!”) and natural language processing. More frequently, deep learning is used to refer to the convolutional neural network (CNN). The architecture of CNN was introduced originally by Japanese research scientist Kunihiko Fukushima, who developed the neocognitron model (feed-forward neural network) in 1980, and longtime New York University computer scientist Yann LeCun, who modified the backpropagation algorithm for neocognitron training. CNN is ideally suited for computer vision and image recognition applications. It can easily detect shifts within an image, so if we have to recognize, for example, pictures of the same face, CNN doesn’t really care if the face appears in the corner of the image or if it is upside down. Although CNN requires a lot of computational power for training, it can be easily paralleled. Unfortunately, it inherits all the disadvantages of neural networks.

When Alan Turing came up with his test more than 60 years ago to determine whether a machine can think, UNIVAC had just introduced the first tape drive for a commercial computer. Since then computational power and data storage have increased exponentially, but what is called artificial intelligence is still far — and distinct — from human intelligence. Moreover, although it was originally inspired by human thinking patterns, AI keeps going in a different direction, formalizing problems and employing mathematical logic that is unnatural for human beings. Of course, there is a chance that some future breakthrough in neuropsychology will bring us closer to understanding the internal mechanics and logic of intelligence so we’ll be able to engineer it, but for now the human brain continues to be one of life’s biggest mysteries. Perhaps machine learning systems — which talk to us, beat us in chess and recognize our faces — can help us tackle the problem. ◀

REFERENCES

Yoshua Bengio, "Learning Deep Architectures for AI",
Foundations and Trends in Machine Learning 2, no. 1 (2009): 1–127.

Leo Breiman, "Random Forests," *Machine Learning* 45,
no. 1 (2001): 5–32.

Corinna Cortes and Vladimir Vapnik, "Support-Vector Networks,"
Machine Learning 20, no. 3 (1995): 273–297.

Yoav Freund and Robert E. Schapire, "A Decision-Theoretic
Generalization of On-Line Learning and an Application to
Boosting," *Journal of Computer and System Sciences* 55, no. 1
(August 1997): 119–139.

Geoffrey Hinton, "Learning Multiple Layers of Representation,"
Trends in Cognitive Sciences 11, no. 10 (2007): 428–434.

Yann LeCun, Koray Kavukcuoglu and Clément Farabet,
"Convolutional Networks and Applications in Vision," Proceedings
IEEE International Symposium on Circuits and Systems, 2010.

Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models
and Selected Applications in Speech Recognition,"
Proceedings of the IEEE (1989): 257–286.

Lotfi A. Zadeh, "Fuzzy Logic = Computing with Words,"
IEEE Transactions on Fuzzy Systems 4, no. 2 (May 1996): 103–111.

Thought Leadership articles are prepared by and are the property of WorldQuant, LLC and are circulated for informational and educational purposes only. This article is not intended to relate specifically to any investment strategy or product that WorldQuant offers, nor does this article constitute investment advice or convey an offer to sell, or the solicitation of an offer to buy, any securities or other financial products. In addition, the above information is not intended to provide, and should not be relied upon for, investment, accounting, legal or tax advice. Past performance should not be considered indicative of future performance. WorldQuant makes no representations, express or implied, regarding the accuracy or adequacy of this information, and you accept all risks in relying on the above information for any purposes whatsoever. The views expressed herein are solely those of WorldQuant as of the date of this article and are subject to change without notice. No assurances can be given that any aims, assumptions, expectations and/or goals described in this article will be realized or that the activities described in the article did or will continue at all or in the same manner as they were conducted during the period covered by this article. WorldQuant does not undertake to advise you of any changes in the views expressed herein. WorldQuant may have a significant financial interest in one or more of any positions and/or securities or derivatives discussed.