

Machine Learning: A Primer

 medium.com/@lizziedotdev/lets-talk-about-machine-learning-ddca914e9dd1

June 27, 2018



Lizzie Turner

May 26, 2018

.

12 min read

.

an introduction for both technical and non-technical readers



Image via

Do you use a personal assistant client like Siri or Alexa? Do you rely on a spam filter to keep your email inbox clean? Do you subscribe to Netflix and rely on its scarily-accurate suggestions to discover new movies to watch? If you said ‘yes’ to any of these questions, congratulations! You’ve made fine use of machine learning!

Although it sounds like a complicated idea that requires a lot of technical background, machine learning is actually a fairly simple concept. To better understand it, let’s go over the **what, who, when, where, how, and why**.

What is machine learning?

One day ladies will take their computers for walks in the park and tell each other, “My little computer said such a funny thing this morning”. — Alan Turing

At its core, machine learning is “the practice of using algorithms to parse data, learn from it, and then make a determination or prediction about something in the world.” ([Nvidia](#)) This means that, rather than explicitly programming a computer to perform some task, you teach the computer how to develop an algorithm to complete the task. There are three major types of machine learning, all with specific advantages and disadvantages: **supervised learning**, **unsupervised learning**, and **reinforcement learning**.

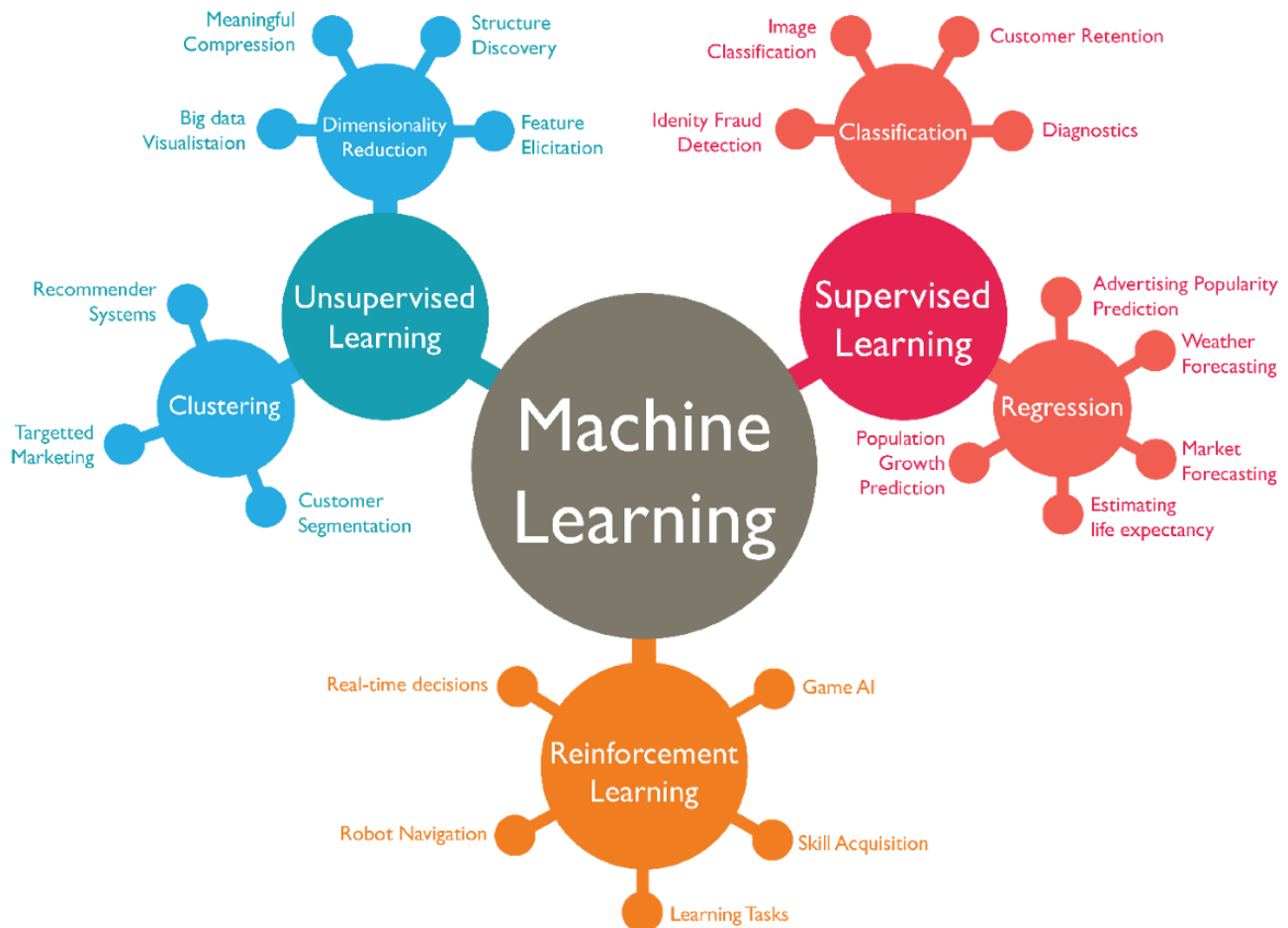


Image via

Supervised learning involves sets of labeled data. The computer can use the provided data to recognize new instances of each labeled type using certain patterns. The two main types of supervised learning are **classification** and **regression**. In **classification**, a machine is trained to separate a group into specific classes. A simple example of classification is a spam filter on your email account. The filter analyzes the emails you've previously marked as spam and compares them to new emails. If they match to a certain percentage, these new emails are tagged as spam and sent to the appropriate folders. Emails are less similar are classified as normal and sent to your inbox. The second type of supervised learning is **regression**. In regression, a machine uses previous (labeled) data to make predictions about the future. Weather apps are good examples of regression. Using historical data about weather events (i.e. average temperature, humidity, and precipitation amounts), your phone's weather app can look at the current weather and make predictions about the weather in a future time frame.

In **unsupervised learning**, the data is unlabeled. As most real-world data is unlabeled, these algorithms are particularly useful. Unsupervised learning is divided into **clustering** and **dimensionality reduction**. Clustering is used to group objects based on properties and behaviors. This is different from classification because these groups are not provided to you. An example of clustering is dividing a group into different subgroups (say, based on age

and marital status) to then send targeted marketing to. **Dimensionality reduction**, on the other hand, involves reducing the variables of a data set by finding commonalities. Most big data visualization uses dimensionality reduction to identify trends and rules.

Finally, **reinforcement learning** uses a machine's personal history and experiences to make decisions. The classic application of reinforced learning is game playing. As opposed to supervised and unsupervised learning, reinforcement learning is not concerned with providing "correct" answers or outputs. Instead, it focuses exclusively on performance. This mirrors how humans learn based on positive and negative consequences. If a child touches a hot stove, he quickly learns not to repeat that action. In this same vein, a chess-playing computer can learn not to move its King to a space accessible to an opponent's piece. This fundamental lesson of chess can then be expanded and extrapolated on until the machine is capable of playing (and eventually beating) top human players.

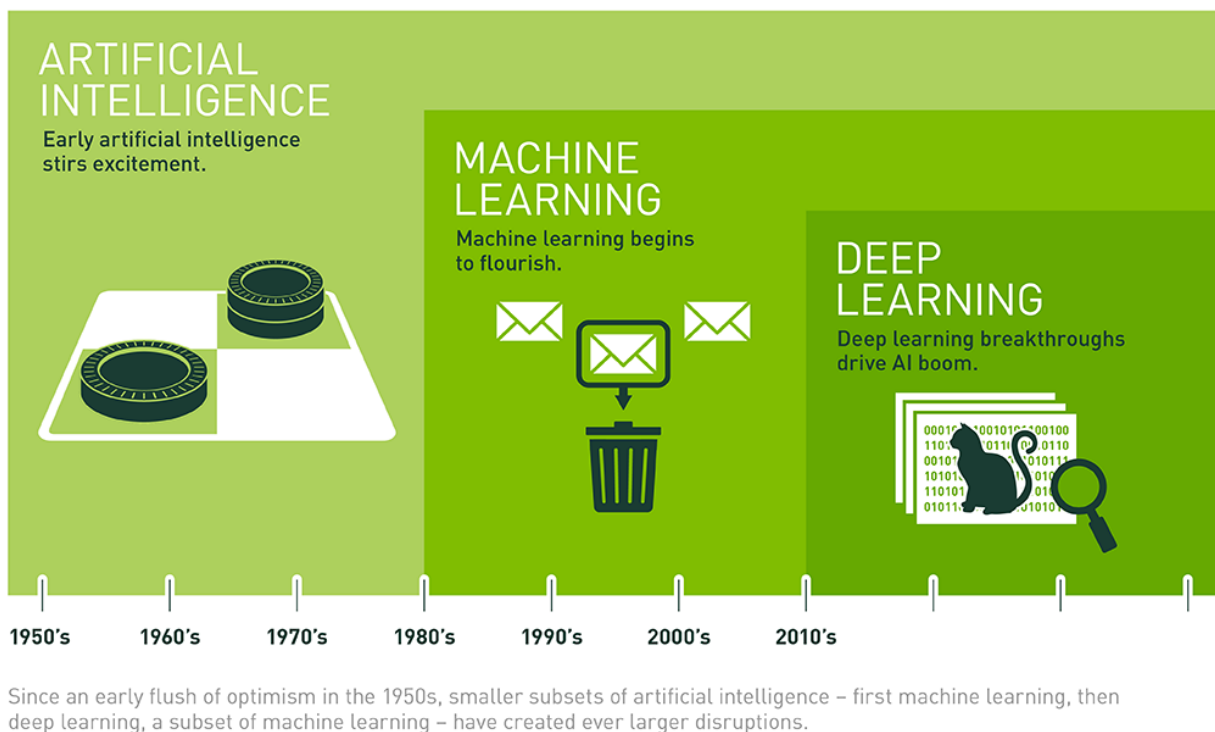


Image via

But wait, you might be saying. *Are we talking about artificial intelligence?* Sort of. Machine learning is a branch of artificial intelligence. **Artificial intelligence** is concerned with creating machines that can perform complex tasks as well (or better) than humans. These tasks often involve judgement, strategy, and cognitive reasoning, skills originally considered “off limits” to machines. While this sounds simple, the scope of these skills is immense — think language processing, image recognition, planning, and much more. Machine learning uses specific algorithms and programming approaches to achieve artificial intelligence. Without machine learning, our previously mentioned chess program would require millions

and millions of lines of code, accounting for all edge cases and containing all possible moves from its opponents. With machine learning, we can reduce the code base into a fraction of its former self. Pretty nifty, huh?

There's just one missing piece: **deep learning** and **neural networks**. We'll go into them in more detail later, but for now, be aware that deep learning is a subset of machine learning focused on specifically mimicing the biology and process of the human brain.

Who developed machine learning? Where and when?

| A breakthrough in machine learning would be worth ten Microsofts. — Bill Gates

In my opinion, the earliest development in machine learning was Thomas Bayes' 1783 publication of his eponymous theorem. **Bayes' theorem** finds the probability of an event given historical data about similar events. It is, unsurprisingly, the basis of the Bayesian branch of machine learning which seeks to find the most likely occurrence based on previous information. In other words, Bayes' theorem is just a fancy schmancy math way of *learning from experience*, the fundamental idea of machine learning.

Centuries later, in 1950, computer scientist (and all-around badass) Alan Turing created the so-called **Turing Test**, where a computer must fool a human through text conversation into thinking she is speaking to another human. Turing argued that a machine could only be considered "intelligent" if it passed this test. Shortly after this, in 1952, Arthur Samuel created the first true machine learning program — a simple checkers game where the computer was able to learn strategy from previous plays and improve future performance. This was followed up with Donald Michie's 1963 reinforcement learning-based tic-tac-toe program. For the next several decades, advances in machine learning followed the same general pattern — a technological breakthrough led to newer, more sophisticated computers, often tested by playing strategy games against professional human players. It hit its peak in 1997 when the IBM chess computer *Deep Blue* beat world champion Garry Kasparov in a chess match. More recently, Google developed *AlphaGo*



Image via

focusing on the ancient Chinese board game Go, widely considered to be the hardest game in the world. Although Go was assumed to be too complex for a computer to master, *AlphaGo* finally prevailed in 2016, beating Lee Sedol in a five-game match.

The largest breakthrough in machine learning was the 2006 development of deep learning. **Deep learning** is a class of machine learning that aims to mimic the thought process of the human brain and is often used in image and speech recognition. The advent of deep learning led to many of the technologies we use (and possibly take for granted) today. Have you ever uploaded a picture to your Facebook account, only for it to suggest tagging the people in the picture? Facebook is using a neural network to recognize the faces in your photo. Or what about Siri? When you ask your iPhone about today's baseball scores, your speech is analyzed by a sophisticated speech-parsing algorithm. None of this would be possible without deep learning.

For a more thorough machine learning timeline, make sure to check out by the Google cloud team!

How does machine learning work?

Attention all math shy readers: I regret to inform you that a basic understanding of some key mathematics concepts is required to fully comprehend most machine learning algorithms. But fear not! The required concepts are simple and draw on classes you've probably already taken. Machine learning uses **linear algebra**, **calculus**, **probability** and **statistics**.

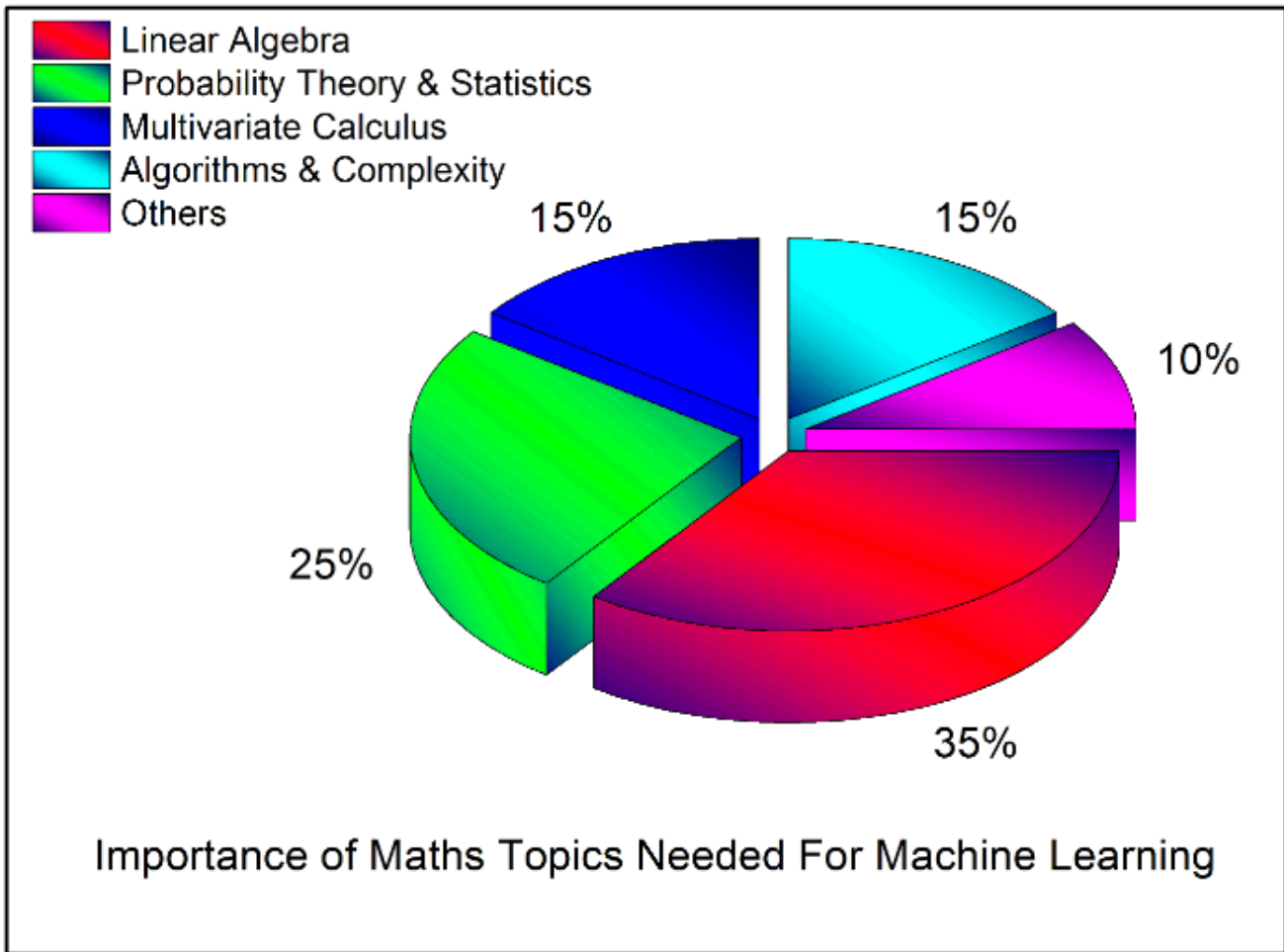


Image via

Top 3 linear algebra concepts:

1. matrix operations
2. eigenvalues/eigenvectors
3. vector spaces and norms

Top 3 calculus concepts:

1. partial derivatives
2. vector-valued functions
3. directional gradients

Top 3 statistics concepts:

1. Bayes' theorem
2. combinatorics
3. sampling methods

For specific math resources, I highly recommend from MetaDesignIdeas.

Once you have a basic mathematic understanding, it's time to start thinking about the entire machine learning process. There are 5 main steps.

The Machine Learning Process



Image via

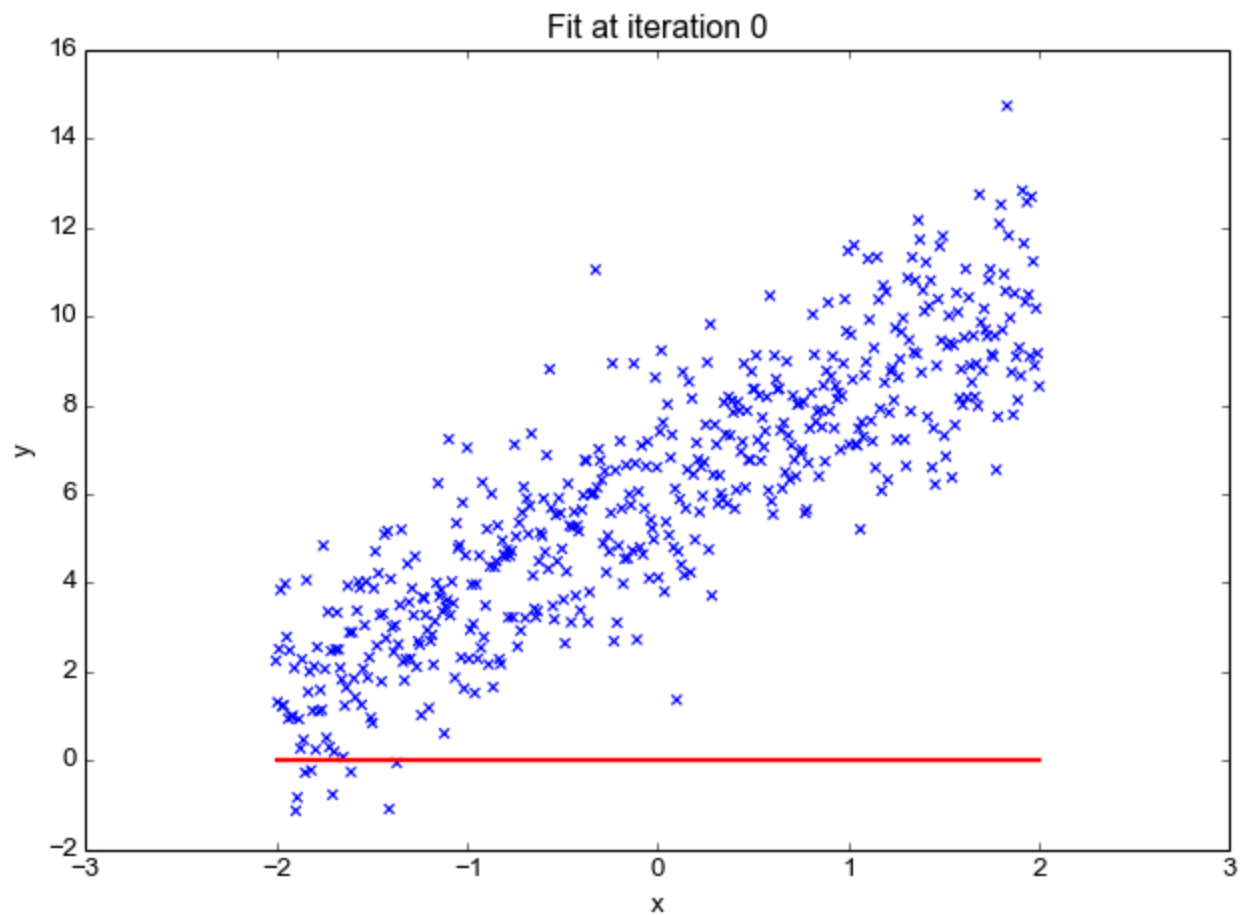
The above diagram explains the steps in a much clearer way than I could, so take a minute to study it before we focus on to the most crucial part: choosing the correct algorithm for the data and situation.

┆ We don't have better algorithms, we just have more data. — Peter Norvig

Let's review some common groupings of algorithms:

Regression Algorithms

Possibly the most popular machine learning algorithms, linear regression algorithms are supervised learning algorithms that predict a specific outcome based on continuous variables. Logistic regression, on the other hand, are used specifically to predict discrete values. Both these (and all other regression algorithms) are known for their speed; they are consistently ranked among the very fastest machine learning algorithms.



GIF via

Instance-Based Algorithms

Instance-based analysis uses specific instances of provided data to predict an outcome. The most famous instance-based algorithm is k-Nearest Neighbor, also known as kNN. Used for classification, kNN compares the distance of data points and assigns each point to the group it is closest to.

kNN Algorithm

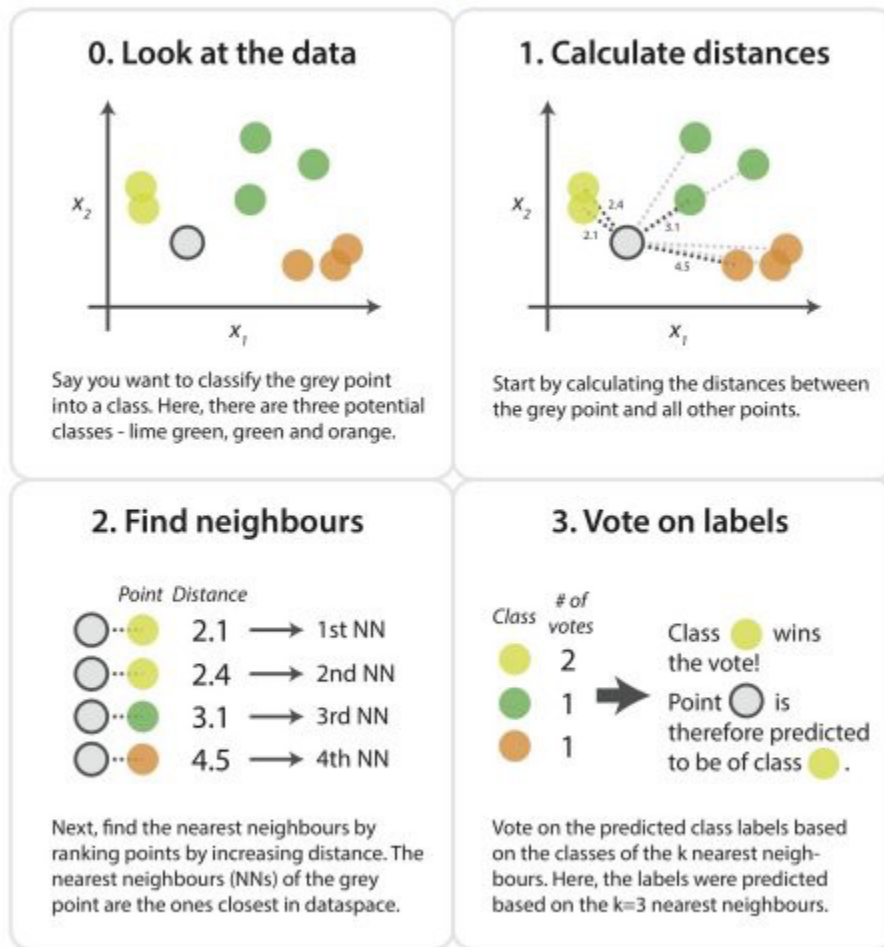


Image via

Decision Tree Algorithms

Decision tree algorithms take groups of “weak” learners and have them work together to form one strong algorithm. These learners are organized in a tree-like structure, branching off one another. A popular decision tree algorithm is the **Random Forest Algorithm**. In this algorithm, the weak learners are randomly chosen. This tends to lead to a strong predictor. In the example below, we can discover numerous common traits (like eyes that are or are not blue), none of which would be enough on their own to identify the animal. When we put all these observations together, however, we are able to form a more complete picture and make a much more accurate prediction.

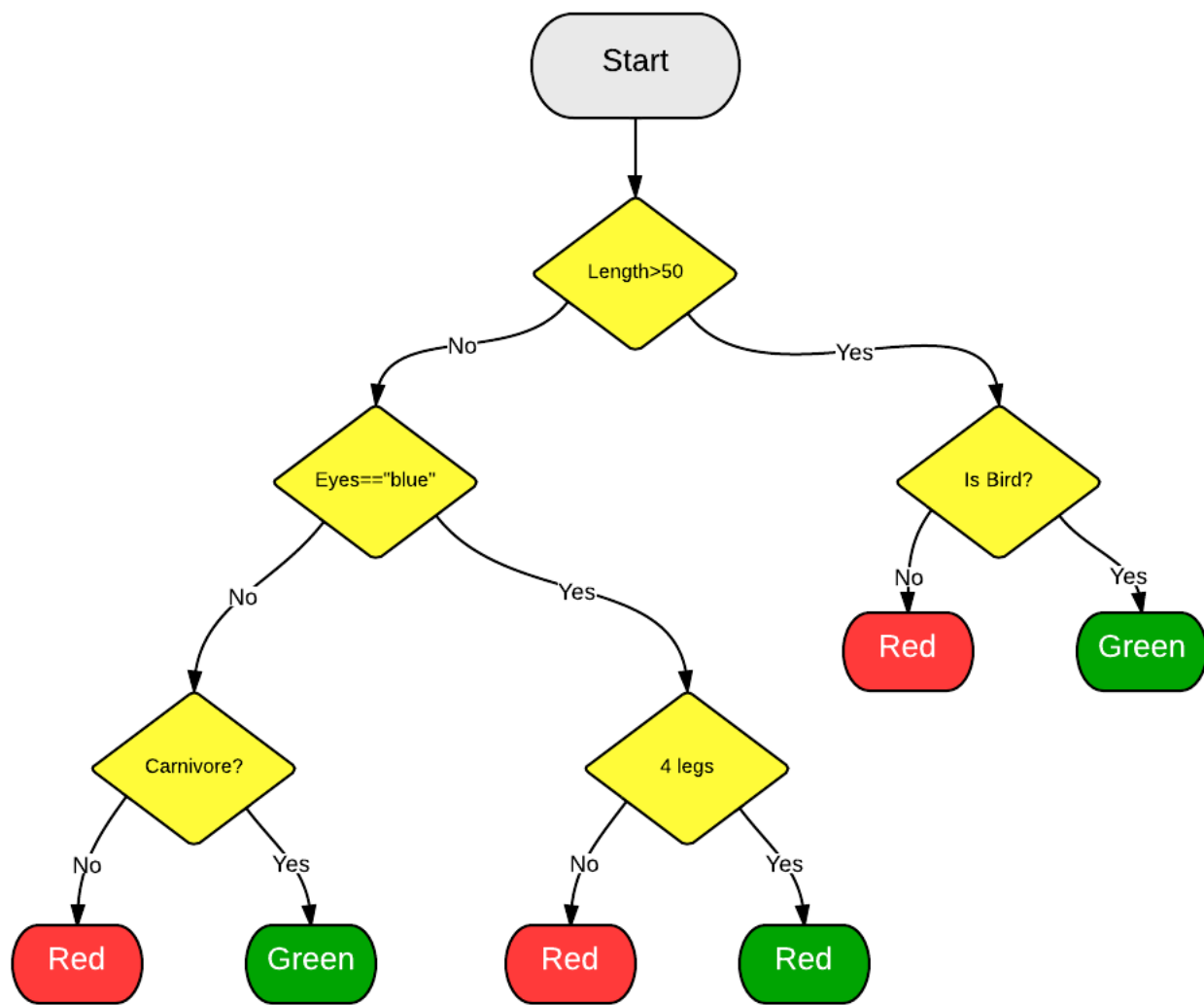


Image via

Bayesian Algorithms

Unsurprisingly, these algorithms are explicitly based on Bayes' theorem. The most popular is **Naive Bayes**, which is often used in text analysis. Most spam filters, for example, use Bayesian algorithms. They use user-inputted data labeled by class to compare new data against and categorize appropriately.

Bayes Rule

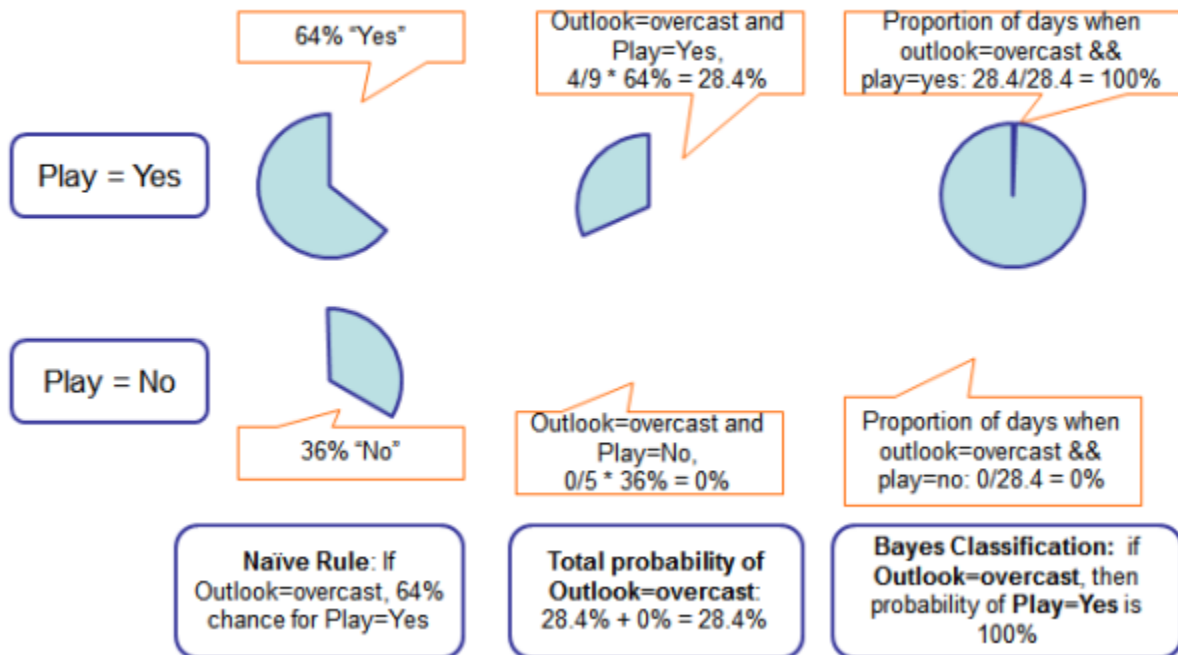


Image via

Clustering Algorithms

Clustering algorithms focus on finding commonalities between elements and grouping them accordingly. A common clustering algorithm is **K-Means Clustering**. In K-Means, an analyst selects the number of clusters (denoted by the variable K) and the algorithm groups the elements by physical distance into appropriate clusters.

Demonstration of the standard algorithm

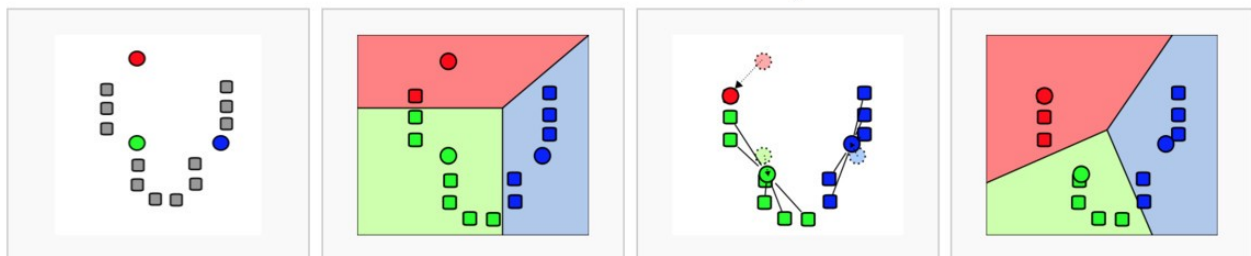
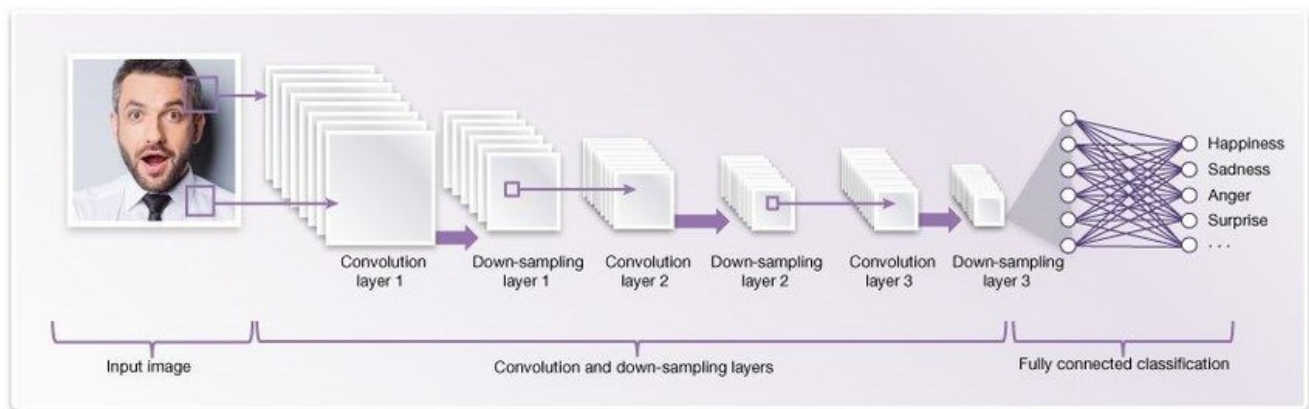


Image via

Deep Learning and Neural Network Algorithms

Artificial neural network algorithms are based on the structure of biological neural networks. Deep learning takes the neural network model and updates it. They are large, extremely complex neural networks that use small amounts of labeled data and much larger amounts of unlabeled data. Neural networks and deep learning have many inputs that go through several hidden layers before resulting in one or more outputs. These connections form a specific cycle that mimics the way the human brain processes information and makes logical connections. In addition, the hidden layers often get smaller and more nuanced as the algorithm runs.



Other Algorithms

The diagram below is the best one I've found to show the major machine learning algorithms, their categories, and their relationships with each other.

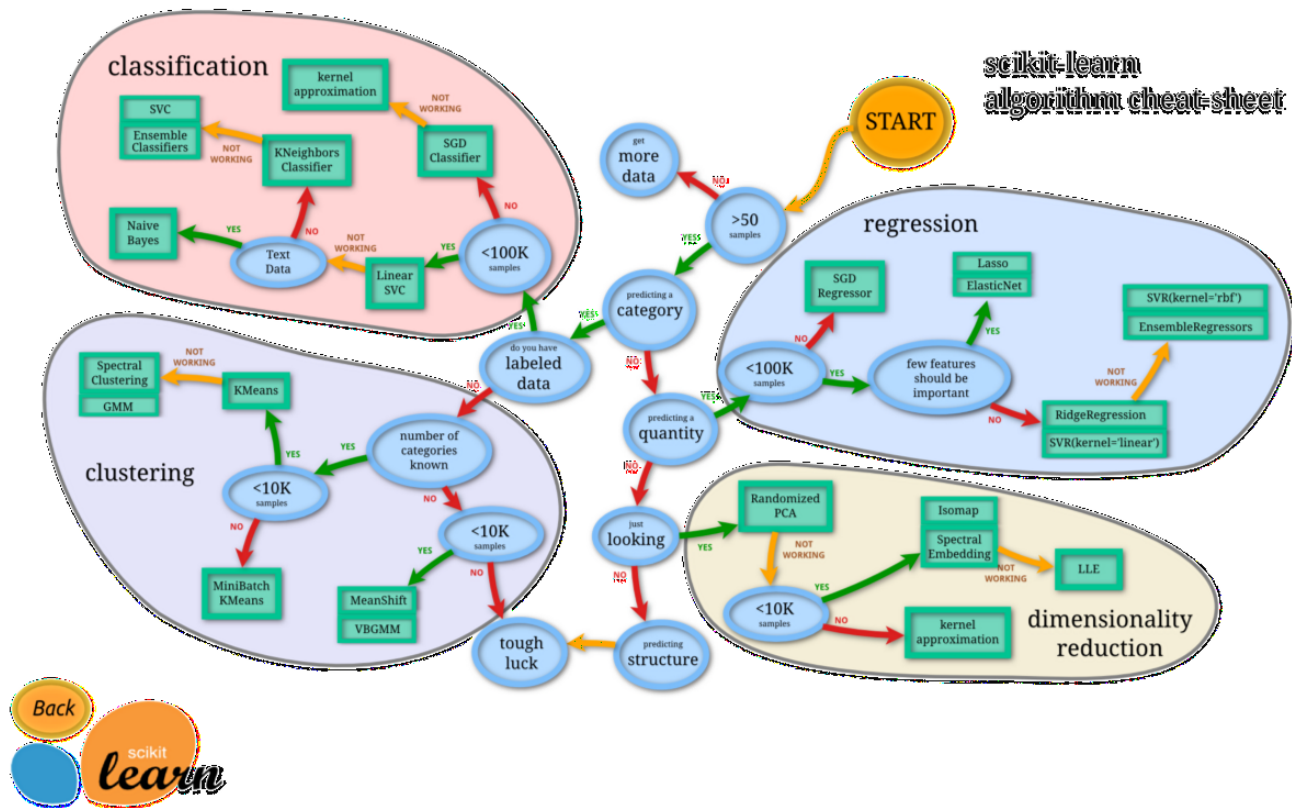


Image via

The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning....**Before we demand more of our data, we need to demand more of ourselves.** — Nate Silver

Once you've chosen and run your algorithm, there is one extremely important step left: visualizing and communicating the results. Although this can seem silly and superficial compared to the nitty gritty of algorithmic programming, good visualization is a key separator of good data scientists and great scientists. What good are amazing insights if no one can understand them?

Why is machine learning important?

Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don't think AI will transform in the next several years.— Andrew Ng

It should now be clear that machine learning has huge potential to change and improve the world. Through research teams like Google Brain and the Stanford Machine Learning Group, we are making huge strides towards true artificial intelligence. But what, exactly, are the next major areas where machine learning can make an impact?

Internet of Things

The term internet of things, or **IOT**, refers to the network-connected physical devices in your home and office. A popular IOT device is the smart lightbulb, sales of which skyrocketed over the last few years. With advances in machine learning, IOT devices are smarter and more sophisticated than ever. Machine learning has two main applications pertaining to IOT: making your devices better and gathering your data. Making the devices better is very straightforward: using machine learning to personalize your environment, i.e. using facial recognition software to sense who is the room and adjust the heat and AC accordingly. Gathering data is even more straightforward; by keeping network-connected devices (like an Amazon echo) powered on and listening in your home, companies like Amazon gather key demographic information to pass onto advertisers, like what television shows you watch, what time you wake up and go to sleep, and how many people live in your home.

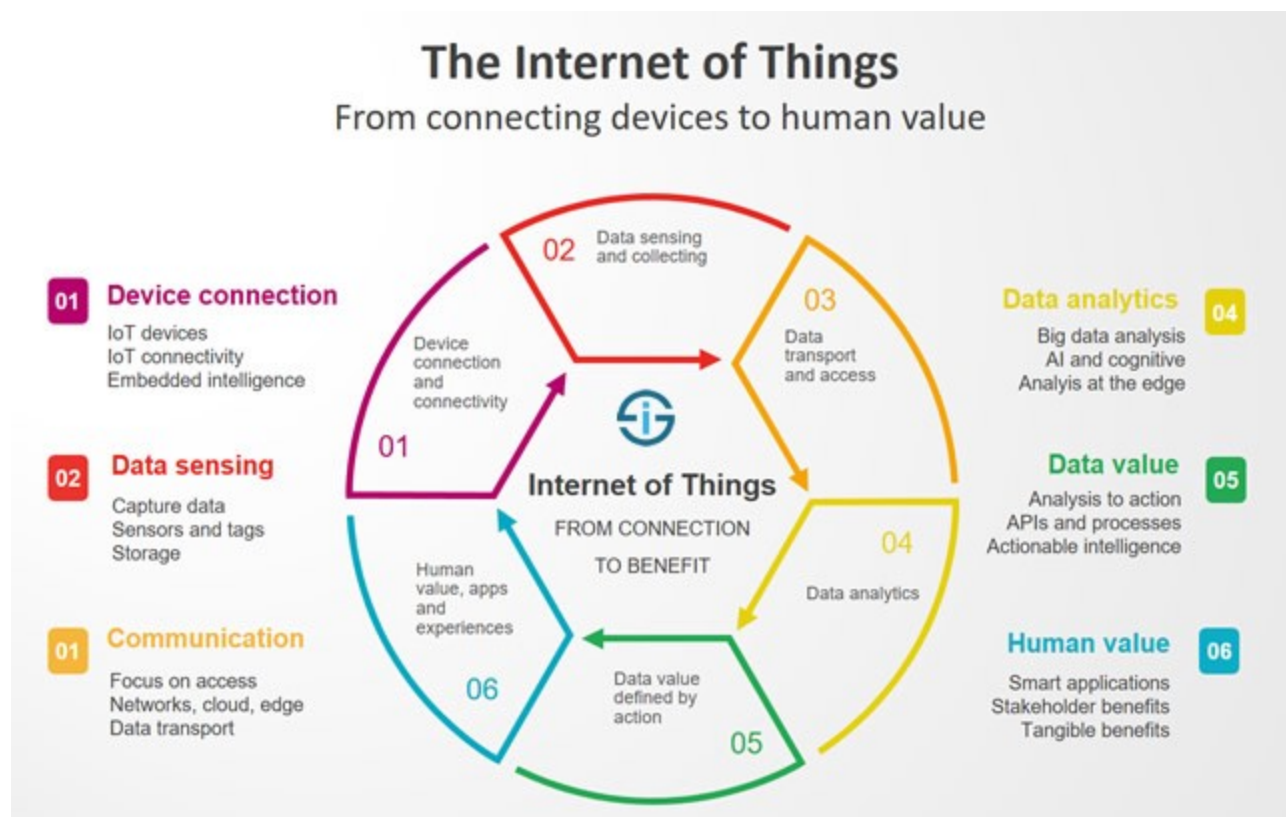


Image via

Chatbots

We have seen a proliferation of **chatbots** in the last few years and sophisticated language processing algorithms are improving them every day. Chatbots are used by companies, on both their own mobile apps and third-party apps like Slack, to provide virtual customer service that is both faster and more efficient than a traditional (human) representative. To order a shirt from the clothing company H&M for example, you can now tell their chat bot in natural language what you'd like and what size you need and order the item without ever leaving the chat screen.

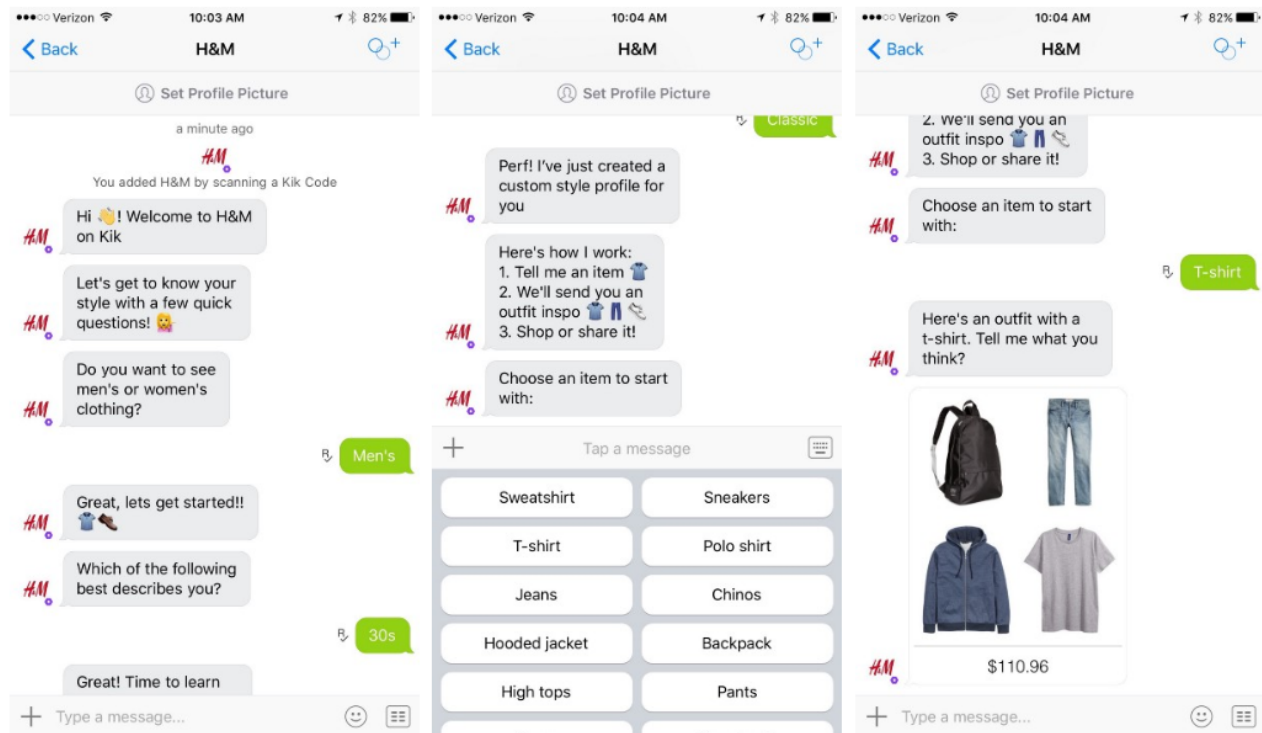


Image via z

Autonomous (a.k.a. self driving) Cars

My personal favorite of the next big machine learning projects is one of the furthest from widespread production. Nevertheless, self driving cars are currently in development at several huge companies like Chevrolet (through their Cruise brand), Uber, and Tesla. These cars use technologies made possible through machine learning for navigation, maintenance, and safety procedures. One example is the traffic sign sensors, which use supervised learning algorithms to identify and parse traffic signs and compare them to a labeled data set of standard signs. Thus, the car sees a stop sign and recognizes that it does, in fact, signify stop, rather than yield or one way or pedestrian crossing.

HOW UBER'S FIRST SELF-DRIVING CAR WORKS

Top mounted **LIDAR** beams 1.4 million laser points per second to create a 3D map of the car's surroundings.

There are **20 cameras** looking for braking vehicles, pedestrians, and other obstacles.

A **colored camera** puts LIDAR map into color so the car can see traffic light changes.

Antennae on the roof rack let the car position itself via GPS.

LIDAR modules on the front, rear, and sides help detect obstacles in blind spots.

A **cooling system** in the car makes sure everything runs without overheating.

SOURCE: Uber

BUSINESS INSIDER

Image via

So that concludes our extremely brief journey into the world of machine learning. I encourage you to read all the articles I've linked below and, if you have a fundamental understanding of Python, play around with your own simple machine learning projects. Happy coding!