Machine Learning for Anyone Who Took Math in Eighth Grade

Explaining modern A.I. with easy math and oversimplified analogies

I usually see artificial intelligence explained in one of two ways: through the increasingly sensationalist perspective of the media or through dense scientific literature riddled with superfluous language and field-specific terms.

There’s a less publicized area in between these extremes where I think literature needs to step up a bit. News about “breakthroughs” like that stupid [robot Sophia](https://www.hansonrobotics.com/sophia/) hype up A.I. to be something akin to human consciousness while in reality, Sophia is about as sophisticated as AOL Instant Messenger’s [SmarterChild](https://en.wikipedia.org/wiki/SmarterChild).

Scientific literature can be even worse, causing even the most driven researcher’s eyes to glaze over after a few paragraphs of gratuitous pseudo-intellectual trash. The public can be left in between the hype and the mystery when all it really takes to understand fundamental A.I. is some middle school math. In order to accurately assess A.I., the general population needs to know what it really is.

I may be prone to oversimplification — and I’ll ask all my math, data science, and engineering colleagues to bear with me as I do it — but sometimes that’s what pretentious science needs.

## Basic A.I. and machine learning in two seconds

Quintessential, classic A.I. is anything that mimics human intelligence. This could be anything from video game bots to Sophia to sophisticated platforms like [Deepmind’s Alphago](https://deepmind.com/research/alphago/).

#### [Source](https://geospatialmedia.s3.amazonaws.com/wp-content/uploads/2017/05/AAEAAQAAAAAAAAhPAAAAJDlkMWMwNTA1LTZkZjUtNDA5MS1hYT.jpg) (ignore deep learning — in this context it’s the same)

Machine learning is a subset of A.I. — and more interesting, I’d argue — that allows machines to “learn” from real-world data instead of acting on a set of predefined rules.

But what does “learn” mean? It might not be as wild as you’d think it is.

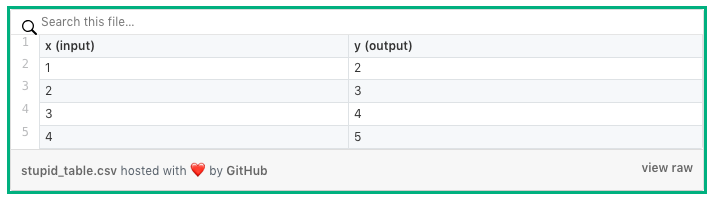
## Machine learning is just y=mx+b on crack

This is my favorite go-to explanation. If you’re watching *Black Mirror*, it’s pretty easy to start to visualize modern A.I. as a conscious entity — something that thinks, feels, and makes complex decisions. This is even more prevalent in the media where practicallyevery reference to A.I. doing anything is personified and then likened to *Terminator*’s Skynet or *The* *Matrix.*

In reality, that’s not true at all. In its current state, A.I. is just math. Sometimes it’s difficult math, and sometimes it requires extensive knowledge of computer science, statistics, and other fields, but a modern A.I. is, at its core, just a mathematical function.

No worries if you’re unfamiliar with math functions because you don’t remember them or don’t use them. For grasping this, we only need the easy stuff: There is input (x), and there is output (y), and the function is what happens between the input and output — it’s the relationship between the two.

Really simplified A.I. is a function expressed as y=mx+b. We know xand y; we just need to find mand b to understand what the relationship between x and y is. For example, in the table below, xis the input, and yis the output.



For this pattern, what we have to do to xto get to y is multiply x by 1 (giving us the m value) and add 1 (giving us the b value). And so, the function is y=1x+1.

In essence, this is all machine learning is. Using input x, we made a predictionof what ywould probably be for all examples. The fancy part is how you teach a machine to learn what function best describes the data — but when you’re done, what you’re left with is generally some form of y=mx+b. Once we have that function, we can also plot it on a graph:

#### [Math](https://i.ytimg.com/vi/m9-_sxSU_WQ/maxresdefault.jpg)

For more explanation of functions, [Math Is Fun](https://www.mathsisfun.com/sets/function.html) has an intuitive and straightforward site (even if the name is a potential red flag and the site looks like their web designer quit sometime in the early 2000s).

## Humans can’t do the math machines can

Obviously, y=1x+1is a really simple example. The whole reason we have machine learning is because humans can’t look at millions of data points and come up with a complex function to describe the output. Instead, we can have a computer look at the input (x) and the output (y) and figure out what ties them together.

Since figuring out the function is based on data, there must be enough of it that a correct function can be found. If we only have one example for xand y, neither we nor a machine could predict only one accurate function. In the original example where x=1 and y=2, the function could be y=2x, y=x+1, y=((x+1)\*5–9)⁵ + 1, or any number of possibilities. Without enough data, too many errors are likely when we try to apply the function to more data. In machine learning models, however, A.I. can process a lot of data and return a function for all — or maybe at least many — of the data points.

Another benefit to machines doing this work is that data in the real world isn’t always perfect. In the example below, a machine has determined several functions that fit most of the data, but sometimes, the lines don’t pass through every point. Real-world data is unpredictable and can rarely be described perfectly. This imperfection and complexity can trip up or even stump humans at times, but finding a “best fit” for the data is much faster with machine learning.

#### [This is a basic example of a machine learning a function that best represents the dataImage result for linear regression gif](https://cdn-images-1.medium.com/max/1600/1*eeIvlwkMNG1wSmj3FR6M2g.gif)

One more thing humans can’t generally do that machines can is factor in a bunch of variables. Finding a function can be easy with just one input (x)and one output (y), but what if there isn’t just one input variable? There might be x1 and x2 or a hundred things that all work together to lead to y. Very quickly, functions can become complex. A.I. is more capable of analyzing the relationship among all those variables.

## Machine learning and A.I. in the real world

Let’s look at a real-world example. I work in pharmaceuticals, so let’s say we have a cancer-related dataset that has two input variables on tumor sizes — radius and perimeter — and two potential outputs for whether or not the tumor is benign or metastatic (potentially life-threatening). It may seem complicated, but we need only apply the familiar y=mx+b concept:

* y is the diagnosis and can be 0 (benign) or 1 (metastatic).
* x¹is the radius.
* x² is the perimeter.
* Each x has an unknown m; let’s call them “something.”
* b stays the same as the unknown constant.

The function, then, looks like y=(m1x1)+(m2x2)+b. Or, by substituting words: the diagnosis equals (one something times the radius) plus (another something times the perimeter) plus a constant. It’s not that different from a basic function, but the complexity of managing multiple variables is getting out of the realm of human capability. So instead of looking at the data and trying to find out what the “somethings” are that affect the input factors to give an accurate estimate of diagnosis, we have A.I. do it for us. That’s machine learning.

Of course, even the most detailed, multi-factored data isn’t perfect, and therefore our machine learning model won’t be either. But we don’t need it to be right 100 percent of the time (just like humans); we simply need it to come up with the best possible function it can that’s right most of the time.

This was definitely a scratch on the surface of all the incredible math and computer science that goes into machine learning. But even at complex levels, the concept is the same. No matter how wildly impressive or ridiculous machine learning and A.I. may seem, it all stems from functions a machine has learned to best describe data.