Get started Open in app

Follow 574K Followers

The Limitations of Machine Learning

Machine learning is now seen as a silver bullet for solving all problems, but sometimes it is not the answer.

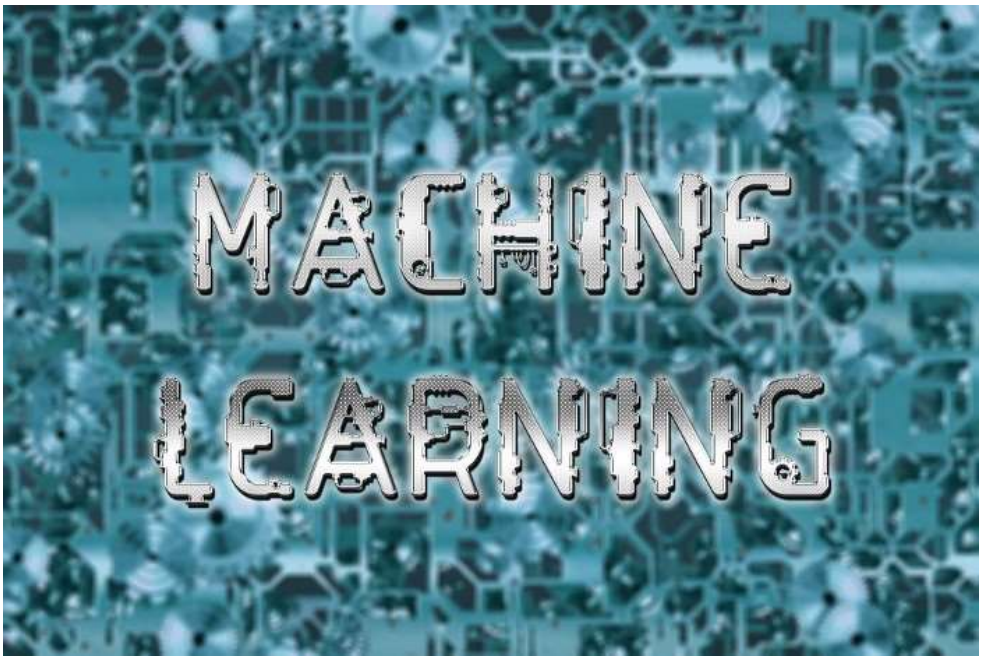
Matthew Stewart, PhD Researcher Jul 29, 2019 · 12 min read

> *"If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future."*
>
> **— Andrew Ng**

Twitter
Facebook
LinkedIn
Copy link

Most people reading this are likely familiar with machine learning and the relevant algorithms used to classify or predict outcomes based on data. However, it is important to understand that machine learning is not the answer to all problems. Given the usefulness of machine learning, it can be hard to accept that sometimes it is not the best solution to a problem.

In this article, I aim to convince the reader that there are times when machine learning is the right solution, and times when it is the wrong solution.



Get started Open in app

massive amounts of data, especially by large companies such as Facebook and Google. This amount of data, coupled with the rapid development of processor power and computer parallelization, has now made it possible to obtain and study huge amounts of data with relative ease.

Nowadays, hyperbole about machine learning and artificial intelligence is ubiquitous. This is perhaps rightly so, given the potential for this field is massive. The number of AI consulting agencies has soared in the past few years, and, according to a report from Indeed, the number of jobs related to AI ballooned by 100% between 2015 and 2018.

As of December 2018, Forbes found that 47% of business had at least one AI capability in their business process, and a report by Deloitte projects that a penetration rate of enterprise software with AI built-in, and cloud-based AI development services, will reach an estimated 87 and 83 percent respectively. These numbers are impressive — if you are planning to change careers anytime soon, AI seems like a pretty good bet.

So it all seems great right? Companies are happy and, presumably, consumers are also happy — otherwise, the companies would not be using AI.

It is great, and I am a huge fan of machine learning and AI. However, there are times when using machine learning is just unnecessary, does not make sense, and other times when its implementation can get you into difficulties.

Limitation 1 — Ethics

Machine learning, a subset of artificial intelligence, has revolutionized the world as we know it in the past decade. The information explosion has resulted in the collection of massive amounts of data, especially by large companies such as Facebook and Google. This amount of data, coupled with the rapid development of processor power and computer parallelization, has now made it possible to obtain and study huge amounts of data with relative ease.

It is easy to understand why machine learning has had such a profound impact on the world, what is less clear is exactly what its capabilities are, and perhaps more importantly, what its limitations are. Yuval Noah Harari famously coined the term ‘dataism’, which refers to a putative new stage of civilization we are entering in which we trust algorithms and data more than our own judgment and logic.

Whilst you may find this idea laughable, remember the last time you went on vacation and followed the instructions of a GPS rather than your own judgment on a map — do you question the judgment of the GPS? People have literally driven into lakes because they blindly followed the instructions from their GPS.

The idea of trusting data and algorithms more than our own judgment has its pros and cons. Obviously, we benefit from these algorithms, otherwise, we wouldn’t be using them in the first place. These algorithms allow us to automate processes by making informed judgments using available data. Sometimes, however, this means replacing

Get started Open in app

The most commonly discussed case currently is self-driving cars — how do we choose how the vehicle should react in the event of a fatal collision? In the future will we have to select which ethical framework we want our self-driving car to follow when we are purchasing the vehicle?

If my self-driving car kills someone on the road, whose fault is it?

Whilst these are all fascinating questions, they are not the main purpose of this article. Clearly, however, machine learning cannot tell us anything about what normative values we should accept, i.e. how we should act in the world in a given situation. As David Hume famously said, one cannot 'derive an ought from an is'.

Limitation 2 — Deterministic Problems

This is a limitation I personally have had to deal with. My field of expertise is environmental science, which relies heavily on computational modeling and using sensors/IoT devices.

Machine learning is incredibly powerful for sensors and can be used to help calibrate and correct sensors when connected to other sensors measuring environmental variables such as temperature, pressure, and humidity. The correlations between the signals from these sensors can be used to develop self-calibration procedures and this is a hot research topic in my research field of atmospheric chemistry.

However, things get a bit more interesting when it comes to computational modeling.

Running computer models that simulate global weather, emissions from the planet, and transport of these emissions is very computationally expensive. In fact, it is so computationally expensive, that a research-level simulation can take weeks even when running on a supercomputer.

Good examples of this are MM5 and WRF, which are numerical weather prediction models that are used for climate research and for giving you weather forecasts on the morning news. Wonder what weather forecasters do all day? Run and study these models.

Running weather models is fine, but now that we have machine learning, can we just use this instead to obtain our weather forecasts? Can we leverage data from satellites, weather stations, and use an elementary predictive algorithm to discern whether it is going to rain tomorrow?

The answer is, surprisingly, yes. If we have knowledge of the air pressures around a certain region, the levels of moisture in the air, wind speeds, and information about neighboring points and their own variables, it becomes possible to train, for example, a neural network. But at what cost?

Get started Open in app

physics of the weather system.

Machine learning is stochastic, not deterministic.

A neural network does not understand Newton's second law, or that density cannot be negative — there are no physical constraints.

However, this may not be a limitation for long. There are multiple researchers looking at adding physical constraints to neural networks and other algorithms so that they can be used for purposes such as this.

Limitation 3 — Data

This is the most obvious limitation. If you feed a model poorly, then it will only give you poor results. This can manifest itself in two ways: lack of data, and lack of **good** data.

Lack of Data

Many machine learning algorithms require large amounts of data before they begin to give useful results. A good example of this is a neural network. Neural networks are data-eating machines that require copious amounts of training data. The larger the architecture, the more data is needed to produce viable results. Reusing data is a bad idea, and data augmentation is useful to some extent, but having more data is always the preferred solution.

If you can get the data, then use it.

Lack of Good Data

Despite the appearance, this is not the same as the above comment. Let's imagine you think you can cheat by generating ten thousand fake data points to put in your neural network. What happens when you put it in?

It will train itself, and then when you come to test it on an unseen data set, it will not perform well. You had the data but the quality of the data was not up to scratch.

In the same way that having a lack of good features can cause your algorithm to perform poorly, having a lack of good ground truth data can also limit the capabilities of your model. No company is going to implement a machine learning model that performs worse than human-level error.

Similarly, applying a model that was trained on a set of data in one situation may not necessarily apply as well to a second situation. The best example of this I have found so far is in breast cancer prediction.

Mammography databases have a lot of images in them, but they suffer from one problem that has caused significant issues in recent years — almost all of the x-rays are from white women. This may not sound like a big deal, but actually, black women have

Get started Open in app

training an algorithm primarily on white women adversely impacts black women in this case.

What is needed in this specific case is a larger number of x-rays of black patients in the training database, more features relevant to the cause of this 42 percent increased likelihood, and for the algorithm to be more equitable by stratifying the dataset along the relevant axes.

If you are skeptical of this or would like to know more, I recommend you look at this article.

Limitation 4 — Misapplication

Related to the second limitation discussed previously, there is purported to be a "*crisis of machine learning in academic research*" whereby people blindly use machine learning to try and analyze systems that are either deterministic or stochastic in nature.

For reasons discussed in limitation two, applying machine learning on deterministic systems will succeed, but the algorithm which not be learning the relationship between the two variables, and will not know when it is violating physical laws. We simply gave some inputs and outputs to the system and told it to learn the relationship — like someone translating word for word out of a dictionary, the algorithm will only appear to have a facile grasp of the underlying physics.

For stochastic (random) systems, things are a little less obvious. The crisis of machine learning for random systems manifests itself in two ways:

- P-hacking
- Scope of the analysis

P-hacking

When one has access to large data, which may have hundreds, thousands, or even millions of variables, it is not too difficult to find a statistically significant result (given that the level of statistical significance needed for most scientific research is $p < 0.05$). This often leads to spurious correlations being found that are usually obtained by p-hacking (looking through mountains of data until a correlation showing statistically significant results is found). These are not true correlations and are just responding to the noise in the measurements.

This has resulted in individuals 'fishing' for statistically significant correlations through large data sets, and masquerading these as true correlations. Sometimes, this is an innocent mistake (in which case the scientist should be better trained), but other times, it is done to increase the number of papers a researcher has published — even in the world of academia, competition is strong and people will do anything to improve their metrics.

Get started Open in app

There are inherent differences in the scope of the analysis for machine learning as compared with statistical modeling — statistical modeling is inherently confirmatory, and machine learning is inherently exploratory.

We can consider confirmatory analysis and models to be the kind of thing that someone does in a Ph.D. program or in a research field. Imagine you are working with an advisor and trying to develop a theoretical framework to study some real-world system. This system has a set of pre-defined features that it is influenced by, and, after carefully designing experiments and developing hypotheses you are able to run tests to determine the validity of your hypotheses.

Exploratory, on the other hand, lacks a number of qualities associated with the confirmatory analysis. In fact, in the case of truly massive amounts of data and information, the confirmatory approaches completely break down due to the sheer volume of data. In other words, it simply is not possible to carefully lay out a finite set of testable hypotheses in the presence of hundreds, much less thousands, much less millions of features.

Therefore and, again, broadly speaking, machine learning algorithms and approaches are best suited for exploratory predictive modeling and classification with massive amounts of data and computationally complex features. Some will contend that they can be used on "small" data but why would one do so when classic, multivariate statistical methods are so much more informative?

ML is a field which, in large part, addresses issues derived from information technology, computer science, and so on, these can be both theoretical and applied problems. As such, it is related to fields such as physics, mathematics, probability, and statistics but ML is really a field unto itself, a field which is unencumbered by the concerns raised in the other disciplines. Many of the solutions ML experts and practitioners come up with are painfully mistaken…but they get the job done.

Limitation 5 — Interpretability

Interpretability is one of the primary problems with machine learning. An AI consultancy firm trying to pitch to a firm that only uses traditional statistical methods can be stopped dead if they do not see the model as interpretable. If you cannot convince your client that you understand how the algorithm came to the decision it did, how likely are they to trust you and your expertise?

As bluntly stated in "*Business Data Mining — a machine learning perspective*":

"A business manager is more likely to accept the [machine learning method] recommendations if the results are explained in business terms"

These models as such can be rendered powerless unless they can be interpreted, and the process of human interpretation follows rules that go well beyond technical prowess. For

Get started Open in app

The blossoming -omics sciences (genomics, proteomics, metabolomics and the like), in particular, have become the main target for machine learning researchers precisely because of their dependence on large and non-trivial databases. However, they suffer from the lack of interpretability of their methods, despite their apparent success.

Summary and Peter Voss' List

While it is undeniable that AI has opened up a wealth of promising opportunities, it has also led to the emergence of a mindset that can be best described as "AI solutionism". This is the philosophy that, given enough data, machine learning algorithms can solve all of humanity's problems.

As I hope I have made clear in this article, there are limitations that, at least for the time being, prevent that from being the case. A neural network can never tell us how to be a good person, and, at least for now, do not understand Newton's laws of motion or Einstein's theory of relativity. There are also fundamental limitations grounded in the underlying theory of machine learning, called computational learning theory, which are primarily statistical limitations. We have also discussed issues associated with the scope of the analysis and the dangers of p-hacking, which can lead to spurious conclusions. There are also issues with the interpretability of results, which can negatively impact businesses that are unable to convince clients and investors that their methods are accurate and reliable.

Whilst in this article I have covered very broadly some of the most important limitations of AI, to finish, I will outline a list published in an article by Peter Voss in October 2016, outlining a more comprehensive list on the limitations of AI. Whilst current mainstream techniques can be very powerful in narrow domains, they will *typically* have some or all of a list of constraints that he sets out and which I'll quote in full here:

- Each narrow application needs to be specially trained
- Require large amounts of *hand-crafted, structured* training data
- Learning must generally be supervised: Training data must be tagged
- Require lengthy offline/ batch training
- Do not learn incrementally or interactively, in real-time
- Poor transfer learning ability, reusability of modules, and integration
- Systems are opaque, making them very hard to debug
- Performance cannot be audited or guaranteed at the 'long tail'
- They encode correlation, not causation or ontological relationships
- Do not encode entities or spatial relationships between entities

Get started Open in app

Not well suited for high level, symbolic reasoning or planning

All that being said, machine learning and artificial intelligence will continue to revolutionize industry and will only become more prevalent in the coming years. Whilst I recommend you utilize machine learning and AI to their fullest extent, I also recommend that you remember the limitations of the tools you use — after all, nothing is perfect.

Newsletter

For updates on new blog posts and extra content, sign up for my newsletter.

Newsletter Subscription

Enrich your academic journey by joining a community of scientists, researchers, and industry professionals to obtain...

mailchi.mp

Sign up for The Variable

By Towards Data Science

Every Thursday, the Variable delivers the very best of Towards Data Science: from hands-on tutorials and cutting-edge research to original features you don't want to miss. Take a look.

Get this newsletter

Machine Learning Data Science Artificial Intelligence Towards Data Science Deep Learning

About Write Help Legal

Get the Medium app

Download on the App Store GET IT ON Google Play