MACHINE LEARNING: HYPE VS REALITY

There is a gap between popular belief in machine learning and what the tools of machine learning can actually accomplish. Pandora's box opens and dreams follow suit.

A combination of boredom and new gun technology in post-Civil War America led Jules Verne to write a book about creating a bullet large enough to house a person, dig a very deep hole in the Earth, and load it with pounds of gunpowder to launch a person to the moon—about a month-long journey for the "Spam in a Can". Jules carefully provides the calculations for the dimensions of the Earth barrel, and the number of barrels of gunpowder needed to accomplish such a feat. We now know that's not the best way to attend space, but, at the time, dreams…

Machine learning is in a similar place today. There're lots of ideas, plenty of movies, even A.I. treatises and frameworks—everyone reaches to be a diplomat for the incoming A.I. species. Is all this buzz merely unfulfilled hype, or is there some real potential here? Let's take a look.

Machine learning in 2020

Machine learning is a revolutionary technology, like the car was to the horse. And even though all the information is free, right there on the internet, the knowledge is still inaccessible—like an alphabet spelling out the words right in front of someone, wishing to be read, but the audience has yet to acquire the reading skill.

The technology behind machine learning is largely inaccessible. Some people get to engineer; others get to gossip. Engineers often have their heads down, rarely looking up to see what others say about them; their voice absent from the sewing circles. Gossipers, however, are only bound by the limits of their imaginations. The people on the "in" can get paid huge sums of money. They're modern wizards. Such a dynamic breeds fascination, unreal stories, and mistrust.

The gap in understanding allows Machine Learning practitioners to appear as one of two kinds of wizards. They can choose to be like Criss Angel or John Edwards, who perform only camera tricks and announce to the world they perform real magic. Or they can produce magicians like Penn & Teller or James Randi, who do what they do best—engineer fantastic, real tricks to delight an audience, never crossing the line of passing off their tricks as if they had real wizardry powers.

Today, machine learning is an iceberg of massive proportions sitting at the top of the Hype Cycle, the Peak of Inflated Expectations. Slowly, bits and pieces of it get chipped away by the population, falling through a Trough of Disillusionment to be refined into a usable product.

Compared to expectations, ML is slow to take off

Business leaders are skeptical and, thus, slow to implement. It is a lot easier to speak the words "machine learning" than it is to implement it. They could hear that space is the next big frontier, but we already know how difficult that can be; they'll be hesitant, or downright opposed to, to change their businesses' direction.

Organizations are right not to jump on board immediately. Machine learning has a lot of hype, and many people jump in not knowing what is needed. After all, an estimated 85% of AI projects won't ship.

Machine learning is not intrinsically valuable

Lots of gossip paints machine learning as inherently important, a silver bullet for societies and businesses. What's often missing, though, is that machine learning has to be used for something. It doesn't throw money at the bottom line by posting up on a fence.

A general consensus in the industry, The Economist wrote in February, is that data is like oil or sunshine. Like oil, it needs to be refined. Like sunshine, it needs to be collected, and a percent of what is collected can be used.

Business leaders shouldn't buy into the shininess of machine learning. Instead, they must consider any ML strategy in light of the value of their data, using machine learning to extract the data's potential.

Data is not just data

To use machine learning, your company will need data. And not just any data. You will need to match the data with the intent of the machine learning models. This requires intention and design.

A couple of problems arise, either:

- You have lots of data already, but it doesn't align with the intent of the ML model.
- The data does not exist. It will take your company some efforts to get the data you need.

Not the right data

Companies have been collecting data for a long time, but because they haven't had machine learning in mind, they haven't been asking the right questions for their datasets to be valuable under the machine learning lens.

A storage facility company may have all kinds of data about their user and their user's behavior. They know everything from their user's eye color to how often their user enters and exits their storage facility. But how can they use this information to create a machine learning product?

Machine learning models can only predict items based on the data given. The storage organization might be able to predict how frequently a user comes in and out of the facility based on their storage unit size, personal income, career type, etc. That may be useful in some way for the organization. But it must be asked:

How valuable is that as a product feature to the user? Can it attract new users?

Starting from scratch

That brings us to having to collect data from scratch. Organizations are failing to deploy their ML products because they jump in without having the appropriate data. If they have a really cool feature in mind, they will need to design a way for user's behaviors to provide the kind of data necessary to collect the data.

Starting from zero means two things:

- The system design will have to incorporate data collection.
- Time must pass in order for data to pile up for ML analysis.

Coding is not just coding

Finally, a note on coding. Coding here is not the same coding there. Organizations with developers who have worked there for 20 years might know tremendous amounts about software development and systems, as one might after putting in their 10,000 hours several times over. But, because machine learning is new and complex, they may not have a clue how to deploy a model or build the ML infrastructure—even if they are senior engineers or managers.

New things always require a learning curve, and learning curves are barriers to entry.

The future of machine learning

Machine learning can do a lot of things. It can even fulfil the hype. But it can only fit into the right hype. Computer literacy will help bridge the gap between hype and reality; the engineers and the dreamers.

Additional resources

For more on machine learning and AI technologies, check out the BMC Machine Learning & Big Data Blog and these articles:

- How Machine Learning Benefits Businesses
- AIOps Machine Learning: Supervised vs Unsupervised
- Machine Learning with TensorFlow and Keras, with 10+ articles
- How to Create a Machine Learning Pipeline
- AWS Sagemaker vs Amazon Machine Learning