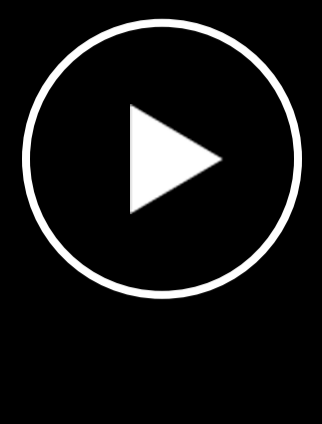


SPONSOR CONTENT

How deep learning is poised to transform the life sciences and healthcare



BURLINGTON, Mass. — Deep-learning algorithms have begun to make their mark on the life sciences and healthcare space, but we have a long way to go to realize their seemingly limitless potential.

Appearing at the [MilliporeSigma Scientific Symposium](#) in October, Dr. Christopher Bouton, Vyasa Analytics CEO, said the ability to leverage big data through these algorithms will be nothing less than transformative.

“Many are saying we’re in a period that will prove to be as impactful and valuable as the introduction of the Internet,” he said. “But we’re at the point where you first saw Netscape Navigator. There’s all of this incredible impact that will be the result of these technologies, much of which we don’t even understand yet.”

The following excerpts from Bouton’s presentation (watch the video here) have been edited for conciseness and clarity.

Clarifying the terms

- **Artificial intelligence (AI)** broadly encompasses the various approaches for getting machines to operate more like humans.
- **Machine learning**, a subset of AI, is a set of algorithms or capabilities we have to handle data, find patterns, etc.
- **Deep learning**, a subset of machine learning, is a type of algorithm with some really interesting properties that we’ll explore below.

Getting at the “dark data”

Big data is great. But if I can’t gain insights and knowledge from it, then what’s it worth to me? How can I take all this information being generated and actually create new insights that will drive value, new products and new cures?

This is especially important because most of the data in enterprises today is what’s called “dark data.” An estimated 80% to 90% of the data in any given organization is dark, or siloed, meaning it’s really hard to access and do any sort of analysis with it. The danger is that people are making decisions without all of the necessary information because they don’t know how to get at it. AI, particularly deep learning, lets us better utilize “dark data” within an organization to become more efficient and productive.

Practically speaking

“Big data” is a vague term, but one that broadly describes the fact that for the past few decades it has become possible to generate and store more data than we can easily interpret. Experts like Dr. Bouton have dedicated their careers to figuring out how to make these data useful.

“Machine learning can be thought of as a response to this kind of problem. Selecting the right and relevant set of features to create an accurate set of parameters requires specialized domain knowledge and usually takes a lot of experimentation to get right. This is the reason blending data experts and life science expertise together in collaborative teams is critical to our business,” said Udit Batra, CEO, MilliporeSigma, the life science business of Merck KGaA, Darmstadt, Germany

The next step in neural networks

What we’re doing with deep learning is creating neural networks — basically, a series of interconnected mathematical formulas. People have built neural networks since the 1950s. What’s the difference now? We’ve realized that if you put a bunch of layers in between the input into the network and the output from the network, these networks do a really good job at figuring out how to get to the right answer. These deep-learning algorithms are different than anything else we’ve done before.

Applying deep learning in the life sciences

Deep-learning algorithms have proven to be powerful across many aspects of the life sciences and healthcare. Here are eight examples of how they’re being applied throughout the pipeline:

1. **Deep-learning-powered lead optimization:** Lead optimization is a key problem. Once you’ve gotten your lead, how do you explore a 10⁸⁰ space to find a better molecule based on your lead? According to a recent publication, a type of deep-learning engine called an “encoder” can encode structures with a discrete library into a continuous gradient space. That allows you to explore a much larger area of compound space than you have in your discrete library. You can then drop novel molecules out of the gradient space, design small compounds and test them to see how they perform, compared with all of the molecules in your existing library.
2. **Predicting compound activity:** One variant of deep learning is known as “one-shot deep learning,” in which you train the algorithm to identify differences, rather than similarities, in data. That requires a lot less training data, which is a critical advantage. A recent paper shows that one-shot deep-learning approaches excel in predicting compound bioactivity based on training with a small set of data. So, if you’re in a new space and don’t know much about the activity of the compounds you’re using, you can still apply these approaches to get a reading of a novel molecule.
3. **Cell assay imaging analytics:** Applying deep learning to cell assay imaging is an obvious path to pursue. All of the papers that have been published in this space so far have shown that deep-learning algorithms can do as well as or better than humans in detecting things like phenotypes. Better yet, they’re more efficient. You still need a human who has to be trained initially and can understand the outcomes. But overall, people are freed up to think about higher-level matters, rather than sitting there looking at images.
4. **Toxicity prediction:** Deep learning has been shown to be significantly more effective than existing methods in predicting the toxicity of any given molecule. Over time, it learns to look for the specific elements or substructures that are causing the toxicity.
5. **Counterfeit scanning:** Deep-learning systems can be trained to detect counterfeit drugs on the web and other sources by examining the package labeling or the pills themselves. They can pick up on little differences between real and counterfeit with a pretty high degree of accuracy. For example, logos or lettering printed on packaging can be slightly off because the presses aren’t exactly the same as the original.
6. **Electronic health record (EHR) analysis:** Deep-learning approaches have been shown to significantly outperform traditional methods in doing things like patient cohort identification, readmissions analysis, clinical trial recruitment and clinical predictive modeling from EHR data stores.
7. **Language translation:** Clinical trial protocol translation is a critical issue. If you provide a protocol in one language and then have someone translate it into another language, how do you know if they have translated all of the important phrases accurately? You can use deep learning to translate back into the original language and figure out how to normalize all those key criteria in your clinical trial.
8. **Electronic laboratory notebook (ELN) analysis:** Many life sciences companies know they have a huge amount of information in their ELNs, but no good way of getting it out. You can have deep-learning systems go in and pull out this information; they can literally read the text in an ELN and figure out what it means.

A massive market opportunity

In 2016, the healthcare sector accounted for the highest revenue share of the global AI market, about 15%. Growing adoption of deep-learning technology will help fuel the global market at an estimated compound annual growth rate of 53% or more through 2020, with the single largest push coming from the life sciences and healthcare space. That’s an estimated \$968 billion [in 2021 value] worth of potential incremental growth, [according to Technaxiomarket research company Technavio](#).