



INTELLIGENT INNOVATION

Enter the library of semantic fingerprints

Innovations in natural language understand mimic how our brains catalog words in a new way.

April 29, 2022 | 4 min read

By [Rodika Tollefson](#) | Illustration by [Chris Gash](#)

Today's businesses are drowning in the volumes of data they create and collect daily. Many data-related jobs are too complex or time-consuming, which is why organizations are turning to artificial intelligence (AI) and its branches of machine learning (ML) and natural language processing (NLP).

Most commonly, the ML models that steer those tasks rely on deep learning: neural networks trained with large data sets—sometimes billions of data points. While these models are proving more adept than people at conquering the data mountain, the result is high computational cost and a growing data center footprint.

But what if, instead of using processing-heavy neural networks, the AI could be trained to “learn” the same way the human brain does?

Austria-based [Cortical.io](#) believes it has the answer. The 10-year-old company has developed a novel alternative to the massive data sets and statistical computations that neural networks use for deep learning.

“These small startups are saying that it doesn't make sense to always use that sort of sledgehammer approach of deep learning. By taking a very narrow approach, they're able to train on a much smaller data set, be up and running much more quickly and, frankly, be much more accurate.

—Alan Pelz-Sharpe, founder, Deep Analysis

Share



YOU MAY ALSO LIKE

WORKFORCE AND CULTURE

[Building a company culture to believe in](#)
Apr 29, 2022 | 8 min read

SUSTAINABILITY

[Calculating \(and cutting\) your carbon footprint at the grocery store](#)
Apr 22, 2022 | 5 min read

RETAIL AND MARKETING

[AI gets emotional: how perceptive analytics could revolutionize marketing](#)
Jul 6, 2021 | 7 min read

Natural language understanding (NLU) underpins Cortical.io's patented system. An evolution as well as a subbranch of NLP, NLU helps computers understand language context. The Cortical.io platform trains the models by mimicking biology—the way the human brain learns, processes and understands language.

The training can be achieved with as few as a hundred documents and is as simple as having a person click through them to categorize bits of text. The process is anywhere from 50 to 1,000 times more energy-efficient and less expensive to implement than traditional models, according to company co-founder and CEO Francisco Webber.

This solution works well for specific use cases because not all the processes that employ NLP need the broad capabilities of deep learning. New companies like Cortical.io are entering the market to challenge that status quo—and provide niche solutions that require a fraction of the computing power.

“These small startups are saying that it doesn't make sense to always use that sort of sledgehammer approach of deep learning,” says Alan Pelz-Sharpe, founder of [Deep Analysis](#), an analyst firm focused on emerging technologies. “By taking a very narrow approach, they're able to train on a much smaller data set, be up and running much more quickly and, frankly, be much more accurate.”

Speed and accuracy

Take the example of a global accounting firm that needed to help customers disclose their leases. Most customers had thousands of lease agreements to comb through, and those agreements didn't contain standard language that could be used effectively in keyword search. Cortical.io's contract intelligence solution allowed the firm to train a model by annotating about 50 documents and then to automatically extract and classify data from the rest of the leases—reducing project completion time by about 80%.

Webber, who developed the technology, says those are typical results for the company's contract intelligence solution. “And you don't need a large team of support clerks—only two or three people monitoring the engine so it works properly,” he says.

Speed and cost savings are not the only advantages. In general terms, contract interpretation is a good use case for AI. [Research suggests](#) that machine-based analysis is 98% accurate, compared to 92% for humans, according to World Commerce & Contracting, a nonprofit association that promotes commercial practice standards.

In the case of Cortical.io, accuracy and consistent quality result from the model training. For example, some employees may have less experience or produce lower quality work when fatigued at the end of the day. AI solves these problems when the customer's most experienced subject matter experts train the model. And it doesn't require an in-house AI expert, so anyone within the company can take the pre-trained model and train it further by adding new classifiers for their use case.

The Cortical.io platform eliminates another challenge inherent with traditional ML approaches—what the industry calls a “black box” or transparency problem.

“These complex [ML] models are making the decisions once they're up and running, but it's not possible to know how they're making them,” explains Pelz-Sharpe. “So you have to trust the technology enormously because it can never explain itself to you.”

In contrast, the Cortical.io platform provides an audit track. The customer can trace all of the system's decisions and inspect every semantic step to understand why a certain document was classified in a particular way.

“What caught our attention with Cortical.io is the ability to recognize context,” adds Pelz-Sharpe, whose company wrote a “vendor vignette” on Cortical.io. “And what makes it interesting is that it's language agnostic. It can take a commercial lease, for example, whether written in English, Japanese or Spanish, and read it equally well.”

Peeking under the hood

Webber's design uses a neuroscience-grounded theory he developed called [semantic folding](#). In simple terms, here's how this complex theory works: Cortical.io converts words into so-called “semantic fingerprints,” or pixel patterns, the same way a brain's neocortex turns words into representations and creates context, or meanings.

“The fingerprints, or the patterns for words in different languages, are about 80% the same.

—Francisco Webber, cofounder and CEO, Cortical.io

Share

The result is a library of these semantic fingerprints that represents relationships among unstructured data sets (words/text), along with context. Based on the trained model, the platform takes the content that needs to be analyzed—whether that's a large volume of contracts, emails or other documents—and compares it to those digital fingerprints to find similar meanings.

The representation of each word as a fingerprint is why the Cortical.io system is language agnostic. “It turns out that the fingerprints, or the patterns for words in different languages, are about 80% the same,” Webber says.

Growing demand

[International Data Corporation](#) estimates that the amount of data created between 2020 and 2025 will be “greater than twice the amount of data created since the advent of digital storage.” The intelligent document processing market (the broad category into which Cortical.io fits) is trying to keep pace. Research firm [Everest Group](#) estimated this market grew 55–65% in 2021, with deep learning, NLP and ML among the core technologies powering the capabilities.

Pelz-Sharpe says that tools like those offered by Cortical.io could have a tremendous impact on business, particularly in sectors like insurance, healthcare and government.

“If these solutions can take 20–30% of the burden of processing documents like claims and automate it, it's absolutely huge,” he says. “There's a lot of opportunity here to solve business problems more efficiently, quickly and affordably—as well as with greater accuracy and transparency.”

Want to read more stories like this? Check out [Realize magazine](#).

Topics in this article:

- Artificial Intelligence
- Intelligent Innovation
- Machine Learning
- Workforce and Culture

YOU MAY ALSO LIKE



INTELLIGENT INNOVATION

[Will blockchain disrupt the video streaming industry?](#)

Feb 10, 2022 | 7 min read



ARTS AND SCIENCE

[How this museum brings art to life with a historic hologram](#)

Aug 11, 2021 | 5 min read



RETAIL AND MARKETING

[AI-generated facial images promise new markets—and new risk](#)

Jul 26, 2021 | 7 min read



EDUCATION

[This AI-powered app helps non-native English speakers ace pronunciation](#)

Jul 16, 2021 | 6 min read

We use cookies to personalise and enhance your experience on our site as described in our Cookie Consent Tool. You may agree to all cookies on our site by clicking [Accept All](#), or decline non-essential cookies by clicking [Decline All](#). If you decline all or some non-essential cookies, this may impact your site experience and our ability to provide certain services, features or offers. You may no longer see ads that interest you. For more information on our data practices, see our [Privacy Statement](#).

Accept All
Decline All