

Self-Supervised Learning's Impact on AI and NLP



Although humans can learn from observing a few examples of a given task, ML algorithms cannot. They can, however, learn through self-supervised learning. We explain how self-supervised learning works.

Many organizations understand the power of artificial intelligence (AI) systems and the countless benefits they offer to enhance the way companies do business. However, as more organizations implement AI systems, there's a growing need for these systems to mirror human language and intelligence to solve modern use cases. It might seem simple, but mimicking human language and mastering its unique complexities continue to be among AI's biggest challenges.

Although AI systems have been around for decades and organizations see their value, the limitations are becoming more visible -- particularly around the amount of data required to train machine learning (ML) algorithms and the flexibility of these algorithms to understand human language.

ML algorithms have proven to be essential, especially in AI applications in customer services. They have the capability to process information and automate con-

versations, increasing the ability of businesses to communicate with customers at anytime from anywhere. As consumer behavior continues to shift, businesses are also transitioning away from high-frequency, one-way communications and toward two-way conversations. These algorithms will play an important role in the customer journey, but it will be critical for organizations to gain a deeper understanding of human language as they try to improve customer interactions.

AI systems have the power to gain a deeper understanding beyond the traditional means of analyzing data. If done successfully, they will exceed human performance in language tasks, bringing AI one step closer to human-level intelligence and transforming how we engage with brands, businesses, and organizations on a global scale. Fortunately, ML can now make this a reality through self-supervised learning.

Self-Supervised Learning Is the Future of AI

As babies, we learn about the world mainly through observation and trial and error. This paves the way for us to develop the ability to learn complex tasks such as driving a car. Unfortunately, although humans can

learn from just observing a few examples of a given task, ML algorithms cannot. They can, however, learn through self-supervised learning.

Self-supervised learning allows ML algorithms to train on low-quality, unlabeled data -- a raw form of data not associated with any tag or label. This type of data is usually plentiful and more readily available in most organizations. The technique typically involves taking an input data set and concealing part of it. The self-supervised learning algorithm must then analyze the visible data to create rules and methods that enable it to predict the remaining hidden data.

As a result, this process creates the labels that will allow the system to learn. With self-supervised learning, there is no need to have a person label large amounts of data. Self-supervised learning creates a data-efficient AI system that can analyze and process data without human intervention, eliminating the need for full "supervision." It also helps organizations better utilize unlabeled data and streamline data processes.

Human brains, especially those of young children, are constantly active and making sense of the world by predicting what will happen next. They might be caught off guard when their predictions do not match reality but these are learning opportunities. Along the same lines, ML algorithms learn to fill in the gaps using semi supervised learning. ML algorithms trained using self-supervised learning can better understand common human cues and beat human performance in language tasks.

The Impact on Natural Language Processing

Although self-supervised learning is still relatively new, it has already made a significant impact on natural language processing (NLP). In 2018, NLP was thrust into the spotlight when Google introduced the BERT model. After recycling an architecture typically used for machine translation, engineers made the model learn the meaning of a word in relation to its context in a sentence, giving it the ability to complete a wide range

of language tasks.

In the last two years, there have been more breakthroughs in NLP than in the past four decades. These AI algorithms now beat human performance in understanding the topic of a text and finding the answer to a random question, doing so in more than 100 languages at once. Many chatbots and virtual assistants utilize NLP technology and with more companies turning to such technologies to answer customer questions in real time, this technology is playing a critical role in enhancing the customer experience.

Other practical applications of NLP and ML include:

- **Blocking phishing attempts.** Phishing attacks are on the rise and businesses need to be able to detect proactively these threats. Different algorithms in the NLP field can identify and analyze a message that may be fraudulent, which can allow organizations to weed out spam messages before they are sent to consumers.

- **Data analytics.** Most databases require structured data, which means that users must enter the data in a specific format, usually numerical. However, converting text data to structured data is a tedious job that costs significant time and effort to do well. NLP algorithms can do this work, performing data analytics on raw, unstructured text without a database.

- **Business analytics.** These algorithms understand what is said and can decipher underlying emotions and intentions. Organizations can now use all the interactions they have with their customers as valuable input to better analyze and understand how well they are doing.

Reducing AI's data dependency and streamlining processes will require the capabilities of self-supervised learning to be successful. Self-supervised learning will allow AI systems to automatically act more human-like and better understand language. The quest to unlock the true potential of ML is well within reach. Once we achieve this, we will uncover even more possibilities in the world of ML