

Insights of Big Data and Artificial Intelligence

Author, Ajit Singh

A Data Science Foundation Blog

April 2019

www.datascience.foundation

Copyright 2016 - 2017 Data Science Foundation

The Big Data vs. AI compare and contrast it, in fact, a comparison of two very closely related data technologies.

Is Big Data vs. artificial intelligence even a fair comparison? To some degree it is, but first let's cut through the confusion.

Artificial Intelligence. Big Data. Those are two buzzwords you are hearing an awful lot lately, perhaps to the point of confusion. What are the similarities and differences between artificial intelligence and Big Data? Do they have anything in common? Are they similar? Can a valid comparison even be made?

The one thing the two technologies do have in common is interest. A survey about Big Data and AI by NewVantage Partners of c-level executives found 97.2% of executives stated that their companies are investing in, building, or launching Big Data and AI initiatives.

More significantly, 76.5% of executives feel AI and Big Data are becoming closely interconnected and that the greater availability of data is empowering AI and cognitive initiatives within their organizations.

Pitting artificial intelligence against Big Data is a natural mistake to be made, partly because the two actually do go together. But they are different tools for achieving the same task. But first thing's first: defining the two. A lot of people don't even know that much.

"I find many people don't really know a lot about what true big data or big data analytics is, or what 'AI' is beyond a few prominent examples," said Alan Morrison, senior research fellow with consulting giant PriceWaterhouse Coopers.

AI vs. Big Data: the Differences

He said a major differentiator is that Big Data is the raw input that needs to be cleaned, structured and integrated before it becomes useful, while artificial intelligence is the output, the intelligence that results from the processed data. That makes the two inherently different.



Artificial intelligence is a form of computing that allows machines to perform cognitive functions, such as acting or reacting to input, similar to the way humans do. Traditional computing apps also react to data but the reactions and responses all have to be hand-coded. If any kind of curve ball is thrown, like an unexpected result, the app can't react. So AI systems are constantly changing their behavior to accommodate changes in findings and modifying their reactions.

An AI-enabled machine is designed to analyze and interpret data and then solve the problem or address the issue based on those interpretations. With machine learning, the computer learns once how to act or react to a certain result and knows in the future to act in the same way.

Big Data is old style computing. It doesn't act on results, it merely looks for them. It defines very large sets of data, but also data that can be extremely varied. In Big Data sets there can be structured data, such as transactional data in a relational database, and less structured or unstructured data, such as images, email data, sensor data, and so on.

They also have differences in use. Big Data is primarily about gaining insight. How does Netflix know what movies or TV shows to suggest to you based on what you watch? Because it looks at the habits of other customers and what they like and deduces you might feel the same.

AI is about decision making, and learning to make better decisions. Whether it is self-tuning software, self-driving cars or examining medical samples, AI is doing tasks previously done by humans but faster and with reduced errors.

Artificial Intelligence and Big Data Together

Although they are very different, AI and Big Data still do work well together. That's because AI needs data to build its intelligence, particularly machine learning. A machine learning image recognition app, for instance, looks at thousands of images of an airplane to learn what constitutes an airplane so it can recognize them in the future.



Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3670 Email: contact@datascience.foundation Web: www.datascience.foundation
Registered in England and Wales 4th June 2015, Registered Number 9624670

Of course, there is the important step of data preparation, which Morrison noted. “The data you start with is Big Data, but to train the model, that data needs to be structured and integrated well enough that machines are able to reliably identify useful patterns in the data,” he said.

Big Data hoovers up massive amounts of data and the wheat has to be separated from the chafe first before anything can be done with it. Data used in AI and ML is already “cleaned,” with extraneous, duplicate and unnecessary data already removed. So there is that big first step.

After that, AI can thrive. Big Data can provide the data needed to train the learning algorithms. There are two types of data learning: the initial training, which is a sort of priming the pump, and routinely gathered data. AI apps never stop learning once the initial training is done. They continue to take in new data and adjust their actions along the way as the data changes. So data is needed initially and continuously.

The two styles of computing both use pattern recognition, but differently. Big Data analytics finds patterns through sequential analysis, sometimes of cold data, or data that is not freshly gathered. Hadoop, the basic framework for Big Data analysis, is a batch process originally designed to run at night during low server utilization.

Machine learning learns from collected data and keeps collecting. Your self-driving car never stops gathering data, and it keeps learning and honing its processes. Data is always coming in fresh and always acted upon.

Big Data’s Role in AI

AI has been talked about forever. It was a plot point of “The Matrix,” a film that came out in 1999. The humans were fighting machines that had gotten too smart. But in execution, AI remained a fringe technology until recently.

The big leap has been the advent of massively parallel processors, particularly GPUs, which are massive parallel processing units with thousands of cores, vs. the dozens in a CPU. This has greatly sped up the existing AI algorithms and has now made them viable.

With Big Data to feed these processors, machine learning algorithms can learn how to reproduce a certain behavior, including collecting the data to in turn speed up the machine. AI doesn’t deduce conclusions like humans do. It learns through trial and error, and that requires massive amounts of data to teach the AI.

The more data an AI app has, the more accurate the outcome it can achieve. In the past, AI didn’t work well because of slow processors and small data sets. There were no sensors like today, where a car can have dozens of sensors built in. And there was no real-time data because the Internet wasn’t widely available.

Today, we have everything we need; the fast processors, the input devices, the network, and the massive amounts of data sets. It’s safe to say there is no Artificial Intelligence without Big Data.

Data Science Foundation

About the Data Science Foundation

The Data Science Foundation is a professional body representing the interests of the Data Science Industry. Its membership consists of suppliers who offer a range of big data analytical and technical services and companies and individuals with an interest in the commercial advantages that can be gained from big data. The organisation aims to raise the profile of this developing industry, to educate people about the benefits of knowledge based decision making and to encourage firms to start using big data techniques.

Contact Data Science Foundation

Email: contact@datascience.foundation
Telephone: 0161 926 3641
Atlantic Business Centre
Atlantic Street
Altrincham
WA14 5NQ
web: www.datascience.foundation

Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3670 Email: contact@datascience.foundation Web: www.datascience.foundation
Registered in England and Wales 4th June 2015, Registered Number 9624670