Machine Learning for Data Analysis

In the past few years, We have created more data in the past few years than we have in all the previous millennia of humankind before then human history combined. This data offers a treasure trove of value for companies and other decision-making entities. However, that value remains untapped — or worse, misinterpreted — when we don't have the tooling required to process it this unprecedented amount of information. In this article, we'll look at how machine learning can help humans gain insight into patterns and critical information buried in our repository of data these data troves.

What is Machine Learning?

The heart of machine learning consists of Sself-learning algorithms that evolve by continuous improvement upon their assigned task comprise the heart of machine learning. When structured appropriately and fed the proper data, these algorithms eventually produce results like pattern identification and predictive modelingmodelling.

Data is to machine learning as learning is what practicing is to us: the more the better. Machine-learning algorithms fine-tune themselves based on the data they train with, just as an Olympic-level athlete perfects a sport by practicing it every day.

A variety of languages work with machine learning, including Python, R, Java, JavaScript, and Scala. Python is a primary choice for machine learning developers because of its TensorFlow library, which offers a comprehensive ecosystem of machine learning tools. If you're interested in coding out an actual example algorithm, check out our article on machine learning with Python.

What is Big Data?

Data includes numbers, words, measurements, and observations available in formats that computers can process. Big data refers to vast sets of such information, — whether structured or unstructured — of such information.

Big data can be summarized by the three Vs: volume, velocity, and variety. *Volume* refers to the huge amount of data available. *Velocity* is the speed at which data is accumulated. *Variety* refers to the different sources that the data it comes from. The advent of the digital era presents a challenge for traditional data processing software: information is becoming available with such volume, velocity, and variety as to outpace human-centered computation. Two other Vs often serve as an addendum to the three aforementioned: *veracity* refers to the inconsistency

and uncertainty in the sourced data, and *value* measures the usefulness of the data extracted from the received data.

Proper analysis of data requires someone with business skills, programming knowledge, and a thorough repertoire of math and data analysis techniques. But how can a professional using traditional techniques sort through the millions of credit card scores or billions of social media interactions now available? That's where machine learning comes in.

How Big Data and Machine Learning Intersect

Machine-learning algorithms become more effective when the size of the training dataset grows. So when big data and machine learning combine, we have a win-win: the algorithms help us keep up with a continuous influx of information while also automatically improving themselves with a huge volume and variety of data. Let's look at how this integration process might work.

By feeding big data to a machine-learning algorithm, we first expect to see defined and analyzed results—think hidden patterns and analytics that assist in predictive modeling. For some companies, this output might automate specific aspects of the decision-making process. More likely than not, however, organizations will review the algorithm's findings, looking for valuable insights that can guide business operations.

This is where people come back into the picture. While AI and data analytics run on computers that outperform us by a vast margin, they lack the skills to replace human decision-making. Expertise, critical thinking, understanding contextintention, and the ability to use holistic approaches comprise human characteristics that computers have yet to replicate. Without an expert providing it with the right data, the value of the algorithm's results diminishes. Without an expert interpreting its output, the algorithm's suggestions may negatively impact compromise company decisions.

Machine Learning Applications for Big Data

Let's look at some real-life examples that demonstrate how big data and machine learning can complement one another.

Cloud Networks

Suppose Aa research firm possesses a huge amount of medical data it wants to study, but doing so would require expensivecalls for servers, online storage, networking, and security assets— all of which would be an unreasonable expense. Instead, the firm decides to invest in Amazon EMR, a cloud service that offers data analysis models within a managed framework.