3Qs with Dr. Kamil Grajski

As Managing Director of Data Science at TROVE, Dr. Kamil Grajski is a senior leader of the Science Squad™, TROVE's industry-leading data science team, charged with delivering timely, high quality commercial-grade predictive analytics products and services. In his work, Dr. Grajski leverages Trove's open-source Big Data platform to run predictive Solvers that leverage decision trees, deep-learning neural networks, proportional hazards models, random forests, regression, and time series forecasting, among many other advanced techniques optimized for complex business analytics challenges.

1

What is deep learning?

Deep learning is an artificial intelligence (AI) technique based on neural networks. Looking at challenges the human brain is uniquely good at accomplishing, such as speech, face, and object recognition, deep learning mimics the brain's approach. That approach has to do with solving tasks iteratively and at successively increasing levels of abstraction. With object recognition, for example, the brain first pulls together basic information from the visual field, then passes that information "up a level" where it is combined and further analyzed to determine whether the object is a face, a cat, or car. I'm oversimplifying here for sure, but the idea of iteration and successive levels of abstraction are key. And whether you know it or not, you are interacting with deep-learning applications every day when you talk to your smart phone or virtual assistant.

2

Why is deep learning important to TROVE and the field of predictive data science?

Every advance in AI and learning techniques gets us closer to doing better things with data, including making better predictions. Our mission at TROVE, and what we truly excel at here, is making data **useful** to our clients, and that means helping them take all the data available to them and distilling it into something that helps them move the needle in some way, whether predicting high probability points of failure on the grid, understanding which customers are most likely to buy an electric vehicle, or pinpointing the best candidate to receive a direct-mail piece.

So deep learning is important to TROVE and to all companies interested in improving the outputs of data, because it improves what can be inferred from it. The best way to understand this improvement is to look at it vs. an established analytics approach, such as linear regression. In linear regression, data points are plotted, a statistical distribution is established, and a line can be drawn between them to predict one variable from another. Deep learning is able to discover and make use of more complex relationships that may exist amongst variables. The outcome, of course, is more finely tuned – and sometimes even unexpected – results.

3

How is TROVE currently using deep learning in its predictive data science work for clients?

At TROVE, we've done some really interesting work in computer-vision applications. In one use case, as part of an asset and capital management program, we developed a deep-learning model to help an electric utility analyze images of transmission towers taken by drones and satellites. The model was used to generate equipment inventories and to assess evidence of wear and tear on the most frequent and most vulnerable tower elements. This capability has the potential alternatively to generate cost savings through the more efficient use of analysts' time or to increase the frequency of inspections while managing cost.

Another promising area for deep learning is direct marketing. Deep learning has the potential to increase the financial return of marketing campaigns. For example, think of a pilot direct-mail program seeking to target a subset of tens of millions of prospects. A deep-learning application can examine hundreds of variables and their complex interactions that may have impacted historical response rates. The result is the ability to generate improved estimates of responders amongst the prospects.

Even the smallest percentage gain in predicting responders translates directly into increased returns for a given fixed marketing budget. Today, we are validating the use of such techniques in the marketing efforts of direct-mail, financial services, and insurance companies.

The results look promising!