SmartDataCollective

The World's Best Thinkers on Data

Predictive Analytics on Big Data – What Does the Future Hold?

The future of targeting and online marketing begins with predictive analytics on big data, according to today's Predictive Analytics World Session presented by Dr. Usama Fayyad.

Known as the industry's first chief data officer (his former position at Yahoo!), Fayyad is the current chairman & CTO of ChoozOn Corporation, a consumer deals search engine.

Fayyad really knows big data because he's been developing technologies and research to harness, process and elicit insights from it for the past 20 years. More specifically, he has helped organizations including NASA, Audience Science, Microsoft and Yahoo! develop data mining technologies and data strategies.

In his 11:35 a.m. ET session, Fayyad will discuss how online marketing has progressed from brand advertising to search marketing to a sophisticated landscape of methods backed by data and predictive analytics. Known as targeting in the online marketing world, predictions are made based on consumer behavior, context and application.

Targeting cannot exist without predictive analytics and this session will reveal where online marketing is headed when backed by strategies that use big data to offer up more relevant advertising.

The session explores a topic that has implications that are almost as big as the data researchers use to help marketers target advertising to consumers. It offers up a huge helping of privacy and ethical issues as well as a need for truth in research. As Fayyad said on Twitter recently, "I agree w/@kdnuggets and I say, "With any data comes great responsibility, not just with Big Data as http://bit.ly/oaNsvm says."

Gregory Piatetsky, editor of KD Nuggets (a newsletter and data mining online community) said, "With Big Data Come Big Responsibilities says [Technology Review] I say any data – big, small or medium – needs careful look at assumptions." And that's what we're getting at with the truth in research. For instance, the article that Fayyad and Piatetsky referenced noted that researchers often use Facebook to garner impressions of people's behavior, but if you don't combine the human element (interviews, observation, etc.), you can only get a cursory view of the truth.

The article gave even more examples of how big data can raise the privacy shield for consumers. For instance, Kate Crawford, a researcher involved in the paper "Six Provocations for Big Data" and an associate professor at the University of New South Wales, says that aggregating data from multiple sources can reveal a person's identity from social media and search engine data.

She says that the reason behind this is that the companies involved in collecting the data may “have no obligation to support scientific inquiry.” This includes requiring that companies pay for the data or manipulate the data by eliminating certain study methods. An example of this is that data samples can be nominal (or not representing a random review of all the data). Crawford and Danah Boyd (her research partner) reference consumer sentiments garnered from Twitter activity as an example. Their premise is that if you dig into the data about Twitter usage, you can see a large portion of users are not saying anything. About 40% of Twitter users are there just to listen. This could signify a “certain type of person” and not the unbiased story.

To wrap up, we can see how Fayyad approaches integrity in handling and analyzing data in a recent interview. His advice to data miners is that “an ounce of knowledge is worth a ton of data,” so it’s important to “seek and model what the experts know or your results will look silly.” Additionally, he says that data analysts should “incorporate the business and legal constraints into their mining.” This protection can help alleviate the implications associated with big data or any data for that matter.

Next Steps: Be sure to follow the conversation throughout PAWCON on Twitter. We also have a complimentary webcast on Spotfire’s predictive analytics tool if you’re interested.