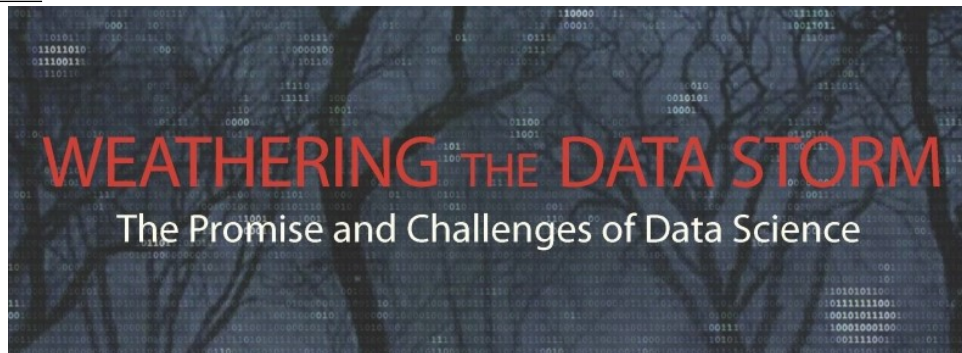


Science Dryad

Data science blog

Grad student descent

Posted on 2014/01/25 by sciencedryad



(<https://sciencedryad.files.wordpress.com/2014/01/picdatastorm.png>) On January 24, I attended a 1-day data science symposium at Harvard University with the fun title ‘Weathering the Data Storm (<http://compute.fest.seas.harvard.edu/data-storm/>)’. I imagine being in a tiny boat on the endless beautiful sea of data, and then a big data storm comes up! Numbers and pieces of text fly through the air... they hit me hard in the face like hail, pile up in my boat... and I’m in dire need of some clever algorithms to take care of all that data, so that I won’t get hurt, my boat won’t sink!

In line with the fun title, there were lots of fun talks. The funniest quote of the day clearly goes to Ryan Adams (<http://www.seas.harvard.edu/directory/rpa>) from Harvard University, when he introduced a new name for a common machine learning ‘method’: grad student descent. He talked about a ‘meta-problem’ of machine learning: Most machine learning algorithms are sufficiently complex to give great results – if they are run with parameters that are adapted to the problem at hand. For example, to work with a neural

network you have to choose the number of layers, the weight regularization, the layer size, which non-linearity, the batch size, the learning rate schedule, the stopping conditions... How do people choose these parameters? Mostly with ad hoc, black magic methods. One method, common in academia, is 'grad student descent' (a pun on gradient descent), in which a graduate student fiddles around with the parameters until it works. It's kind of sad, but it's so true! Of course, Ryan Adams then went on to discuss better solutions ('meta-algorithms' that automatically find the parameters), but it was the 'grad student descent' that stuck to everyone's mind.

Rachel Schutt (<http://idse.columbia.edu/rachel-schutt-0>) from News Corps mused on the perennial question 'What is a data scientist?' She cited the well-known definition by Josh Wills from Cloudera, which I really like:

Data scientist = "Person who is better at statistics than any software engineer and better at software engineering than any statistician."

But I hadn't yet heard the clever rephrasing by Will Cukierski of Kaggle:

Data scientist = "Person who is worse at statistics than any statistician and worse at software engineering than any software engineer."

Both quotes nail down the interdisciplinary nature of the field of data science (and are really funny). This interdisciplinarity is something that I really like. Whenever I go to data science meetings, I meet people from so many different backgrounds – it is very enriching, and the melting pot of so many different ideas and ways of thinking is enticing. It also matches my own diverse background, with lots of math, physics, statistics, biology, programming thrown together...

It was also great to see some data science tools celebrities. Fernando Perez, who started iPython in 2001, talked about the great features of iPython – for example, I didn't know that it also supports other languages like R, Julia, or SQL. And Jeff Heer, creator of D3, showed some awesome D3 visualizations, including the most funny alternative-visualizations sequence I have ever seen (the first 15 seconds of this [video](http://vimeo.com/29862153) (<http://vimeo.com/29862153>) by Mike Bostock).

Advertisements

[Junior Data Engineer](#)

Business & Decision

Woluwé-Saint-
APPLY[Data Engineer](#)

This entry was posted in [Data science](#), [Meetings](#) and tagged [conference](#), [data science](#), [data scientist](#), [Harvard University](#), [machine learning](#). Bookmark the [permalink](#).

One thought on “Grad student descent”

[Angelika](#) says:

on 2015/10/26 at 11:01

Hi, after reading this amazing post i am also cheerful to share my knowledge here with friends.

[Reply](#)

[Create a free website or blog at WordPress.com.](#)