Google Uses Machine Learning for Crawling, Indexing & Ranking

There has been plenty of speculation about just how much Google uses machine learning when it comes to the search algorithm, especially outside of RankBrain itself. During a recent Google Webmaster Office Hours, the role of machine learning in the algo came up.

The question was specifically about Panda and whether it uses a machine learning algo classifier on sample pages to train

high quality versus low quality content.

John Mueller doesn't specifically talk about Panda and machine learning, but instead he discusses Google's use of machine learning overall in their search algos. And while we know Google does use machine learning for RankBrain, which enables Google to serve high quality results for the type of search queries that are never or rarely seen, they haven't talked much about it from outside of RankBrain.

First, Mueller talks about the Panda aspect of it.

> *I don't actually know what the Panda algorithm uses so I can't really answer that question for you.*

He does clarify later that it wouldn't be out of the question for Google Panda to have a machine learning component, but he doesn't know.

Then he goes on to talk about how Google is using machine learning, and he says Google is using it to understand better "how we should crawl, index and rank pages."

> *In general, we do use a lot of machine learning to try and better understand how we should crawl, index and rank pages, so that's something that wouldn't be totally out of the question, but as far as I know, or actually I don't know what the Panda algorithm does there.*

> *So that's kind of something where I think machine learning has a lot of potential to try and understand pages a little bit better. It's not like an automatic solution, though. It's not that we can just like feed it a bunch of pages and say these are good, these are bad and then it'll figure out the whole rest of the web on its own. It does take quite a bit of work to actually make machine learning work well enough.*

Ever since Google first announced RankBrain, and the use of machine learning in algo, many SEOs have been speculating about just how much it is used, especially beyond the scope of RankBrain.

The use of machine learning to help determine crawl budget could mean that what a larger scope of sites do could impact an individual site's crawl budget as well, despite the fact many SEOs manipulate how large sites are crawled, such as through robots.txt, noindex and nofollow, to see their site crawled optimally. But if Google is using machine learning, whether for training or directly, to help determine how they should crawl a site, SEO's own crawl budgeting could have less of an impact, especially in the future as we know Google use using machine learning more and more.

As for indexing, it is harder to know for certain how Google is using it, or again if it is merely for training purposes.

Here are the full comments:

[embedded content]

The following two tabs change content below.
Jennifer Slegg is a longtime speaker and expert in search engine marketing, working in the industry for almost 20 years. When she isn't sitting at her desk writing and working, she can be found grabbing a latte at her local Starbucks or planning her next trip to Disneyland. She regularly speaks at Pubcon, SMX, State of Search, Brighton SEO and more, and has been presenting at conferences for over a decade.

Latest posts by Jennifer Slegg (see all)

http://www.thesempost.com/google-uses-machine-learning-crawling-indexing-ranking/