Machine Learning 101 – Whiteboard Friday

Machine learning is only growing in importance for anyone working in the digital world, but it can often feel like an inaccessible subject. It doesn't have to be – and you don't have to miss out on the competitive edge it can give you when it comes to SEO task automation. Put on your technical SEO cap and get ready to take notes, because Britney Muller is walking us through Machine Learning 101 in this week's episode of Whiteboard Friday.

Click on the whiteboard image above to open a high-resolution version in a new tab!

Video Transcription

Hey, Moz fans. Welcome to another edition of Whiteboard Friday. Today I'm talking about all things machine learning, something, as many of you know, I'm super passionate about and love to talk about. So hopefully, this sparks a seed in some of you to explore it a bit further, because it is truly one of the most powerful things to happen in our space in a very long time.

What is machine learning?

So a brief overview, in a nutshell, machine learning is actually a subset of AI, and some would argue we still haven't really reached artificial intelligence. But it's just one facet of the overall AI.

Traditional programming

The best way to think about it is in comparison to traditional programming. So traditional programming, you input data and a program into a computer and out comes the output, whether that be a web page or calculator you built online, whatever that might be.

Machine learning

With machine learning, what you do is you put in the data and the desired output and put this into a computer, and you get a program, otherwise known as a machine learning model. So it’s a bit flipped, and it works extremely well. There are **two primary types of machine learning:**

1. You have **supervised**, which is where you’re basically feeding a model labeled training data,
2. And then **unsupervised**, which is where you’re feeding a program data and letting it create clusters or associations between data points.

The supervised is a bit more common. You’ll see things like classification, linear regression, and image recognition. Things like that are all very common. If you think about machine learning in terms of, okay, there’s all of this data that you’re putting into the model, data is the biggest part of machine learning. A lot of people would argue that if machine learning was a vehicle, data would be the fuel.

It’s a really important part to understand, because unless you have the right types of data to feed a model, you’re not going to get the desired outcome that you would like.

A machine learning model example

So let’s look at an example. If you wanted to build a machine learning model that predicts housing prices, you might have all of this information.

You might have the current price, square foot of these homes, land, the number of bathrooms, the number of bedrooms, you name it. It goes on and on. These are also known as **features**. So what a model is going to try to do, when you put in all of this data, it’s going to try to understand associations between this information and come up with a model that best predicts home prices in the future.

The most basic of these machine learning models is linear regression. So if you think about inputting the data where maybe you just put in the price and the square foot, and you can kind of see the data like this.

You see that as the square foot goes up, so does the price. A model over time, in looking at this data, is going to start to find the smoothest line through the data to have the most accurate predictions in the future.

What you don’t want it to do is to fit every single data point and have a line that looks like that – that’s also known as overfitting – because it doesn’t play nice for new data points. You don’t want a model to get so calculated to your dataset that it doesn’t predict accurately in the future.

A way to look at that is by the loss function. That’s maybe getting a bit deeper in this, but that’s how you would measure how the line is being fit. Let’s see.

What are the machine learning possibilities in SEO?

So what are some of the possibilities in SEO? How can we leverage machine learning in the SEO space?

Automate meta descriptions

So there are couple ways that people are already doing this. You can automate meta descriptions by looking at the page content and using a machine model to summarize the text. So this literally summarizes the content for you and pares it down to a meta description length. Pretty incredible.

Automate titles

You could similarly do this for titles, although I don’t suggest you do this for primary pages. This isn’t going to be perfect. But if you have a huge, huge website, with hundreds of thousands of pages, it gets you halfway there. It’s really interesting to start playing around in that space with these large websites.

Automate image alt text

You can also automate alt text for images. We see these models getting really good at understanding what’s in an image.

Automate 301 redirects

301 redirects, Paul Shapiro has an incredible write-up and basically process for that already.

Automate content creation

Content creation, and if that scares some of you or if you doubt that these models can currently create content that is decent, I challenge you to go check out Talk to Transformer.

It is a pared-back version of OpenAI, which was founded by Elon Musk. It’s pretty incredible and a little scary as to how good the content is just from that pared back model. So that is for sure possible in the future and even today.

Automate product/page suggestions

In addition to product and page suggestions.

So this is just going to get better. Imagine us providing content and UX specifically for the unique users that come to our site, highly personalized content, highly personalized experiences. Really exciting stuff moving forward.

Resources

I’ve got some resources I highly suggest you check out.

Google Codelabs is one of my favorites, just because it walks you through the steps. So if you go to Google Codelabs, filter by TensorFlow or machine learning, you can see the possible examples there. Colab notebooks or Jupyter notebooks are where you’ll likely be doing any of the machine learning that you want to do on your own.

Kaggle.com is the number one resource for data science competitions. So you get to really see what are the examples, how are people using machine learning today. You’ll see things like TSA has put up over $1 million for a data science team to come up with a model that predicts potential threats from security footage.

This stuff gets really interesting really fast. It’s also so

important to have diversity and inclusion in this space to avoid really dangerous models in the future. So it's something to definitely think about.

TensorFlow is a great resource. It's what Google put out, and it's what a lot of their machine learning models is built off of. They've got a really great JavaScript platform that you can play around with.

Andrew Ng has an incredible machine learning course. I highly suggest you check that out.

Then **Algorithmia** is sort of a one-stop shop for models. So if you don't care to dip your toes into machine learning and you just want say a summarizer model or a particular type of model, you could potentially find one there and do a plug-and-play of sorts.

So that's pretty interesting and fun to explore. The last thing is a machine learning model is only as good as the data. I can't express that enough. So a lot of machine learning and data scientists, it's all data cleaning and parsing, and that's the bulk of the work in this field.

It's important to be aware of that. So that's it for Machine Learning 101. Thank you so much for joining me, and I hope to see you all again soon. Thanks.

Video transcription by Speechpad.com

If you enjoyed this episode of Whiteboard Friday, you'll be delighted by all the cutting-edge SEO knowledge you'll get from our newly released MozCon 2019 video bundle. Catch more useful technical tips in Britney's talk, plus 26 additional future-focused topics from our top-notch speakers:

Grab the sessions now!

We suggest scheduling a good old-fashioned knowledge share with your colleagues to educate the whole team – after all, who didn’t love movie day in school? ;-)

http://tracking.feedpress.it/link/9375/12923731?utm_medium=feed&utm_source=feedpress.me&utm_campaign=Feed%3A+mozblog