Surpassing human-level performance

(SPEECH)
[inaudible]

(DESCRIPTION)
Logo, Deep Learning.ai. Title, Comparing to Human-level performance, Surpassing human-level performance

(SPEECH)
teams often find it exciting to surpass human-level performance on the specific recreational classification task.

Let's talk over some of the things you see if you try to accomplish this yourself.

We've discussed before how machine learning progress gets harder as you approach or even surpass human-level performance.

Let's talk over one more example of why that's the case.

(DESCRIPTION)
Surpassing human-level performance. Team of humans, One Human, Training Error, Dev Error.

(SPEECH)
Let's say you have a problem where a team of humans discussing and debating achieves 0.5% error, a single human 1% error, and you have an algorithm of 0.6% training error and 0.8% dev error.

So in this case, what is the avoidable bias?

(DESCRIPTION)
Arrow from team of humans, 0.5% down to training error, 0.6%

(SPEECH)
So this one is relatively easier to answer, 0.5% is your estimate of base error, so your avoidable bias is, you're not going to use this 1% number as reference, you can use this difference, so maybe you estimate your avoidable bias is at least 0.1% and your variance as 0.2%.

(DESCRIPTION)
Arrow between Training error, 0.6% down to Dev error, 0.5%

(SPEECH)
So there's maybe more to do to reduce your variance than your avoidable bias perhaps.

But now let's take a harder example, let's say, a team of humans and single human performance, the same as before, but

(DESCRIPTION)
0.5% and 1%

(SPEECH)
your algorithm gets 0.3% training error, and 0.4% dev error.

Now, what is the avoidable bias?

It's now actually much harder to answer that.

Is the fact that your training error, 0.3%, does this mean you've over-fitted by 0.2%, or is base error, actually 0.1%, or maybe is base error 0.2%, or maybe base error is 0.3%?

You don't really know, but based on the information given in this example, you actually don't have enough information to tell if you should focus on reducing bias or reducing variance in your algorithm.

So that slows down the efficiency where you should make progress.

Moreover, if your error is already better than even a team of humans looking at and discussing and debating the right label, for an example, then it's just also harder to rely on human intuition to tell your algorithm what are ways that your algorithm could still improve the performance?

(DESCRIPTION)
Circles 0.5% and 0.3%

(SPEECH)
So in this example, once you've surpassed this 0.5% threshold, your options, your ways of making progress on the machine learning problem are just less clear.

It doesn't mean you can't make progress, you might still be able to make significant progress, but some of the tools you have for pointing you in a clear direction just don't work as well.

(DESCRIPTION)
Problems where ML significantly surpasses human-level performance

(SPEECH)
Now, there are many problems where machine learning significantly surpasses human-level performance.

For example, I think, online advertising, estimating how likely someone is to click on that.

Probably, learning algorithms do that much better today than any human could, or making product recommendations, recommending movies or books to you.

I think that web sites today can do that much better than maybe even your closest friends can.

All logistics predicting how long will take you to drive from A to B or predicting how long to take a delivery vehicle to drive from A to B, or trying to predict whether someone will repay a loan, and therefore, whether or not you should approve a loan offer.

All of these are problems where I think today machine learning far surpasses a single human's performance.

Notice something about these four examples.

All

(DESCRIPTION)
Structured Data

(SPEECH)
four of these examples are actually learning from structured data, where you might have a database of what has users clicked on, database of proper support for, databases of how long it takes to get from A to B, database of previous loan applications and their outcomes.

(DESCRIPTION)
Not Natural Perception

(SPEECH)
And these are not natural perception problems, so these are not computer vision, or speech recognition, or natural language processing task.

Humans tend to be very good in natural perception task.

So it is possible, but it's just a bit harder for computers to surpass human-level performance on natural perception task.

(DESCRIPTION)
Lots of data

(SPEECH)
And finally, all of these are problems where there are teams that have access to huge amounts of data.

So for example, the best systems for all four of these applications have probably looked at far more data of that application than any human could possibly look at.

And so, that's also made it relatively easy for a computer to surpass human-level performance.

Now, the fact that there's so much data that computer could examine, so it can petrifies that's called patterns than even the human mind.

(DESCRIPTION)
Speech Recognition

(SPEECH)
Other than these problems, today there are speech recognition systems that can surpass human-level performance.

(DESCRIPTION)
Some image recognition

(SPEECH)
And there are also some computer vision, some image recognition tasks, where computers have surpassed human-level performance.

But because humans are very good at this natural perception task, I think it was harder for computers to get there.

And

(DESCRIPTION)
Medical

(SPEECH)
then there are some medical tasks, for example, reading ECGs or diagnosing skin cancer, or certain narrow radiology task, where computers are getting really good and maybe surpassing a single human-level performance.

And I guess one of the exciting things about recent advances in deep learning is that even for these tasks we can now surpass human-level performance in some cases, but it has been a bit harder because humans tend to be very good at this natural perception task.

So surpassing human-level performance is often not easy, but given enough data there've been lots of deep learning systems have surpassed human-level performance on a single supervisory problem.

So that makes sense for an application you're working on.

I hope that maybe someday you manage to get your deep learning system to also surpass human-level performance.