Twitter Thread by Santiago



Santiago
@svpino

A surprising situation.

A machine learning model predicting 3 classes ends up with these results:

- **Class A: 90% accuracy**
- **Class B: 80% accuracy**
- **Class C: 70% accuracy**

Class C seems to be the worst-performing.

But you know already that I'll say it isn't. ↓

Here is the correct answer: We don't know yet which class is the worst.

This is important to understand, so keep reading.

Many people make this same mistake every day and try to fix things that ain't broken.

We are missing something, so let's go and get it.

How about if we ask a friend to look at the dataset manually?

Let's say these are the results of our friend doing the job:

- Class A: 95% accuracy
- Class B: 100% accuracy
- Class C: 71% accuracy

Do you see what just happened?

The model is actually really accurate when predicting Class C! 70% versus 71% when done by a human.

Class B, however, is a disaster: 80% versus 100% when done manually.

Class B seems to be the worse performing class!

The work that our friend did is called a "baseline."

Without having it, it's hard to determine how good are the results of a model.

I've seen tons of money down a drain just because teams didn't create a baseline beforehand.

But now, that won't happen to you ■.

Every week, I post 2 or 3 threads like this, breaking down machine learning concepts and trying to give you ideas on applying them in real-life situations.

If you find this helpful, follow me @svpino so we can do this thing together!