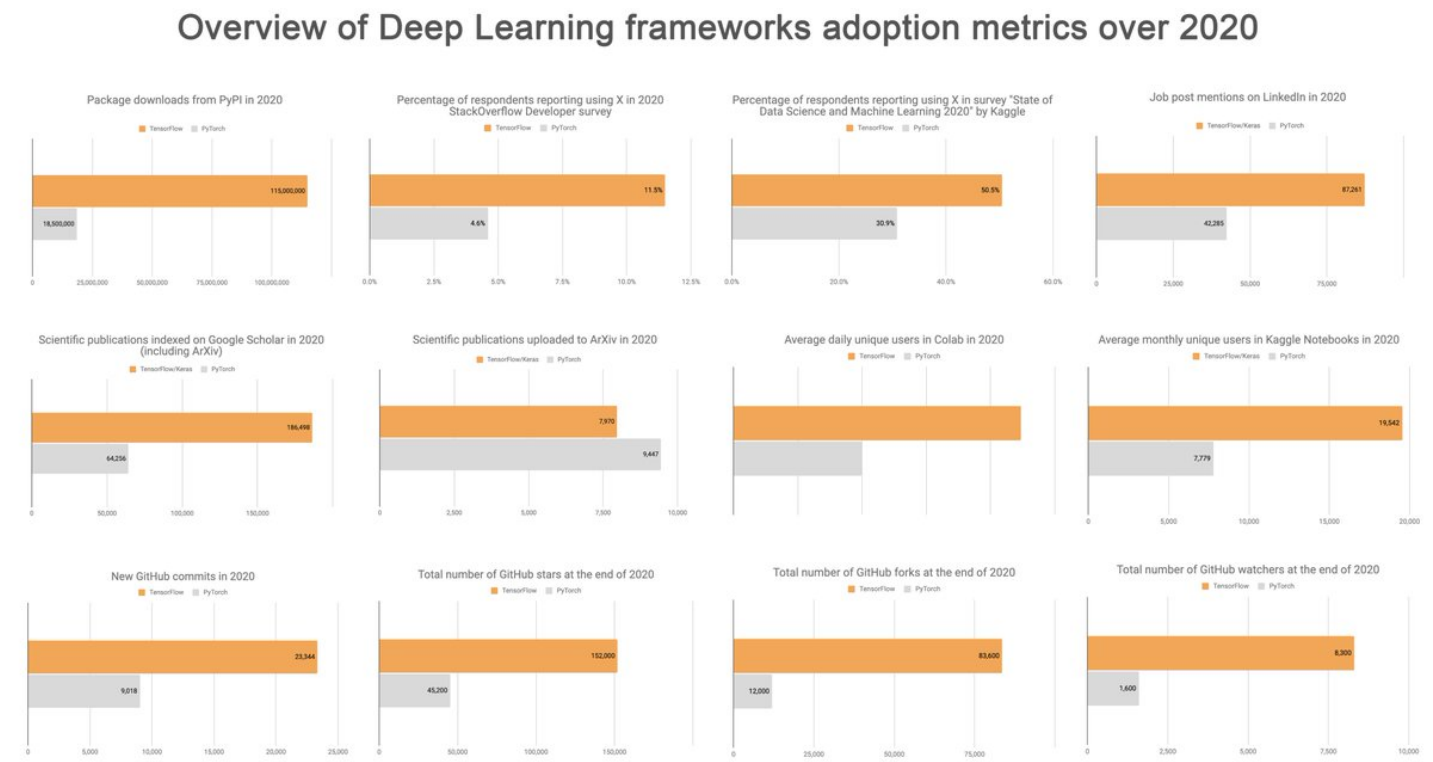
Twitter Thread by François Chollet

François Chollet
@fchollet

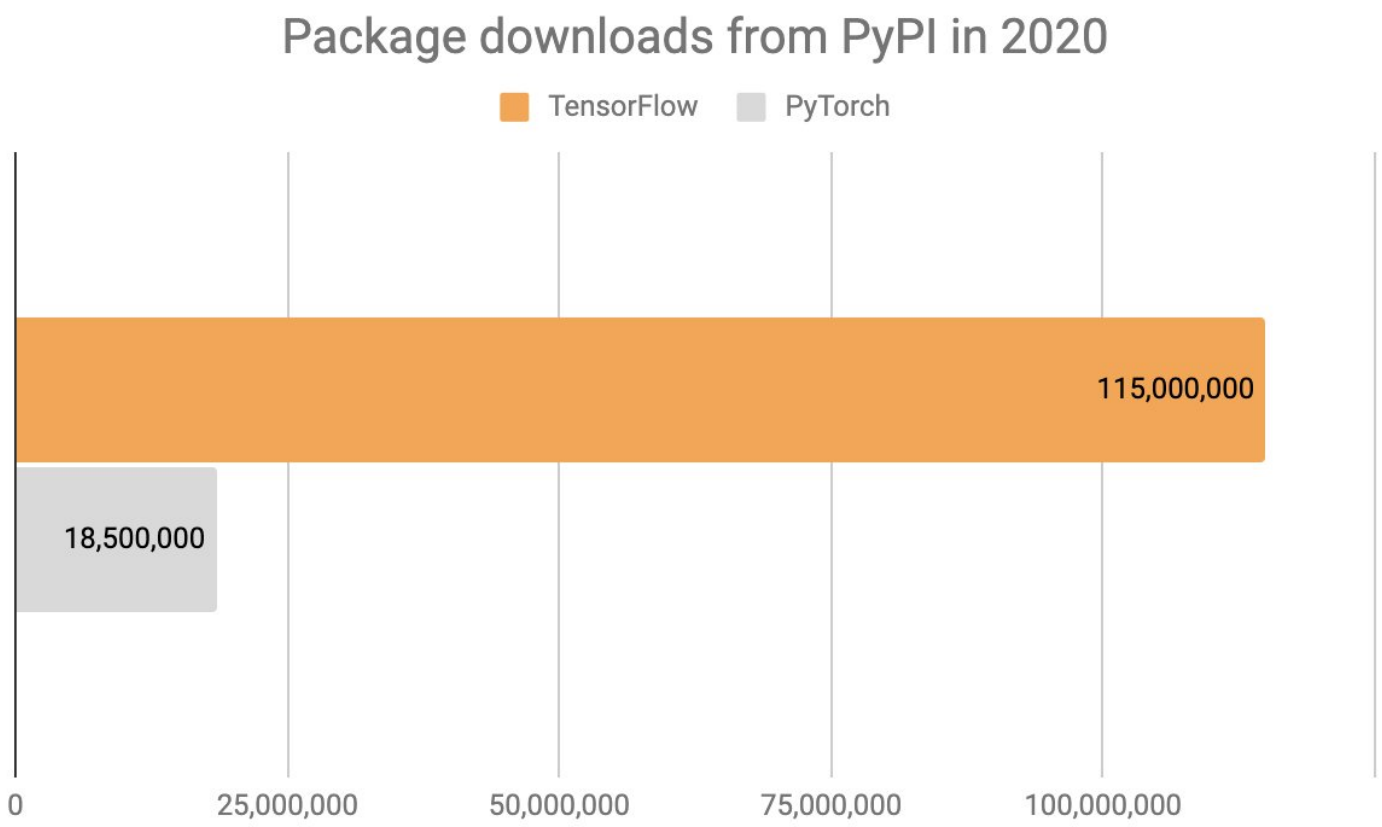
Here's an overview of key adoption metrics for deep learning frameworks over 2020: downloads, developer surveys, job posts, scientific publications, Colab usage, Kaggle notebooks usage, GitHub data.

TensorFlow/Keras = #1 deep learning solution.



Note that we benchmark adoption vs Facebook's PyTorch because it is the only TF alternative that registers on the scale. Another option would have been sklearn, which has massive adoption, but it isn't really a TF alternative. In the future, I hope we can add JAX.

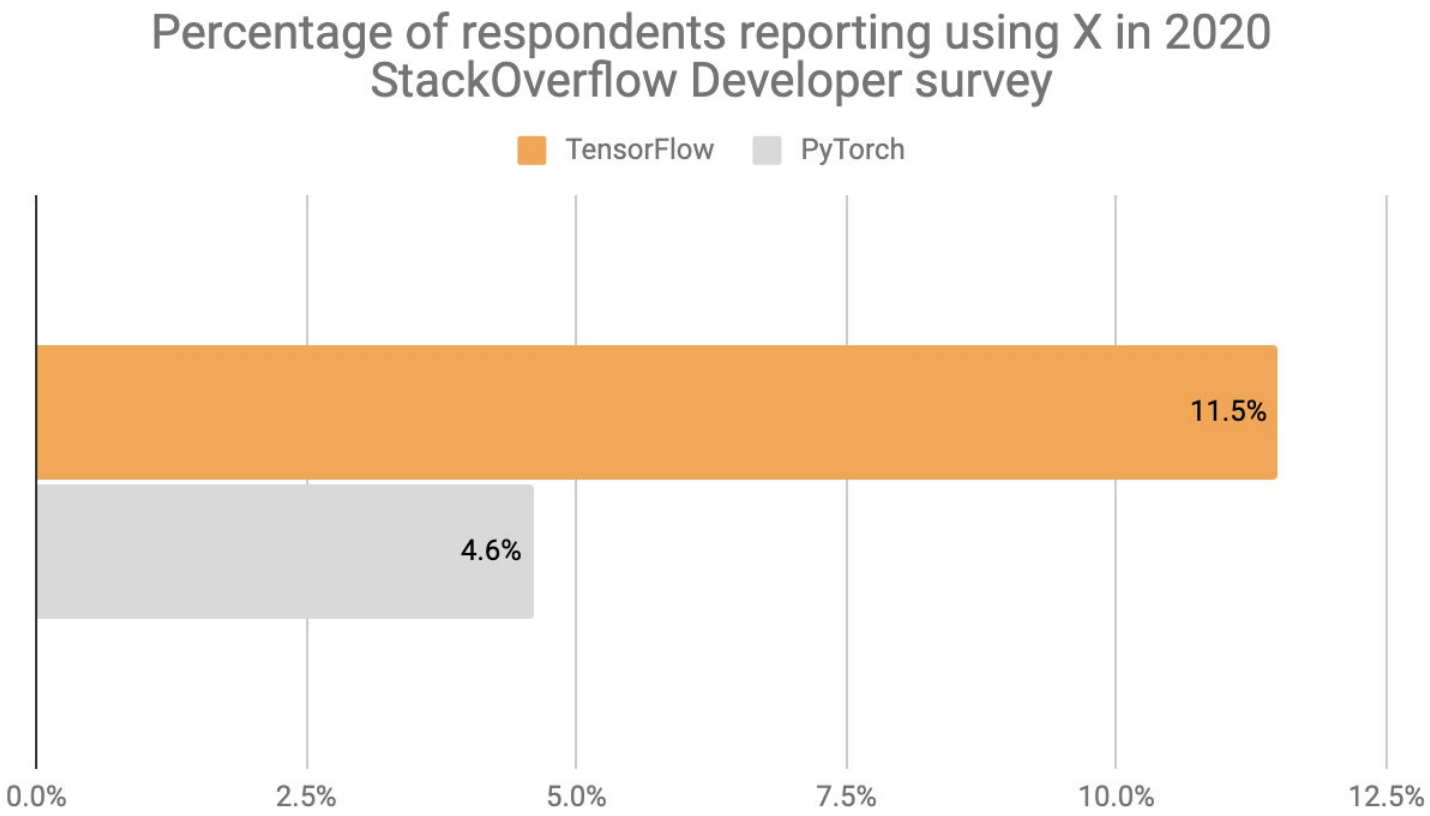
TensorFlow has seen 115M downloads in 2020, which nearly doubles its lifetime downloads. Note that this does *not* include downloads for all TF-adjacent packages, like tf-nightly, the old tensorflow-gpu, etc.



Also note that most of these downloads aren't from humans, but are automated downloads from CI systems (but none are from Google's systems, as Google doesn't use PyPI).

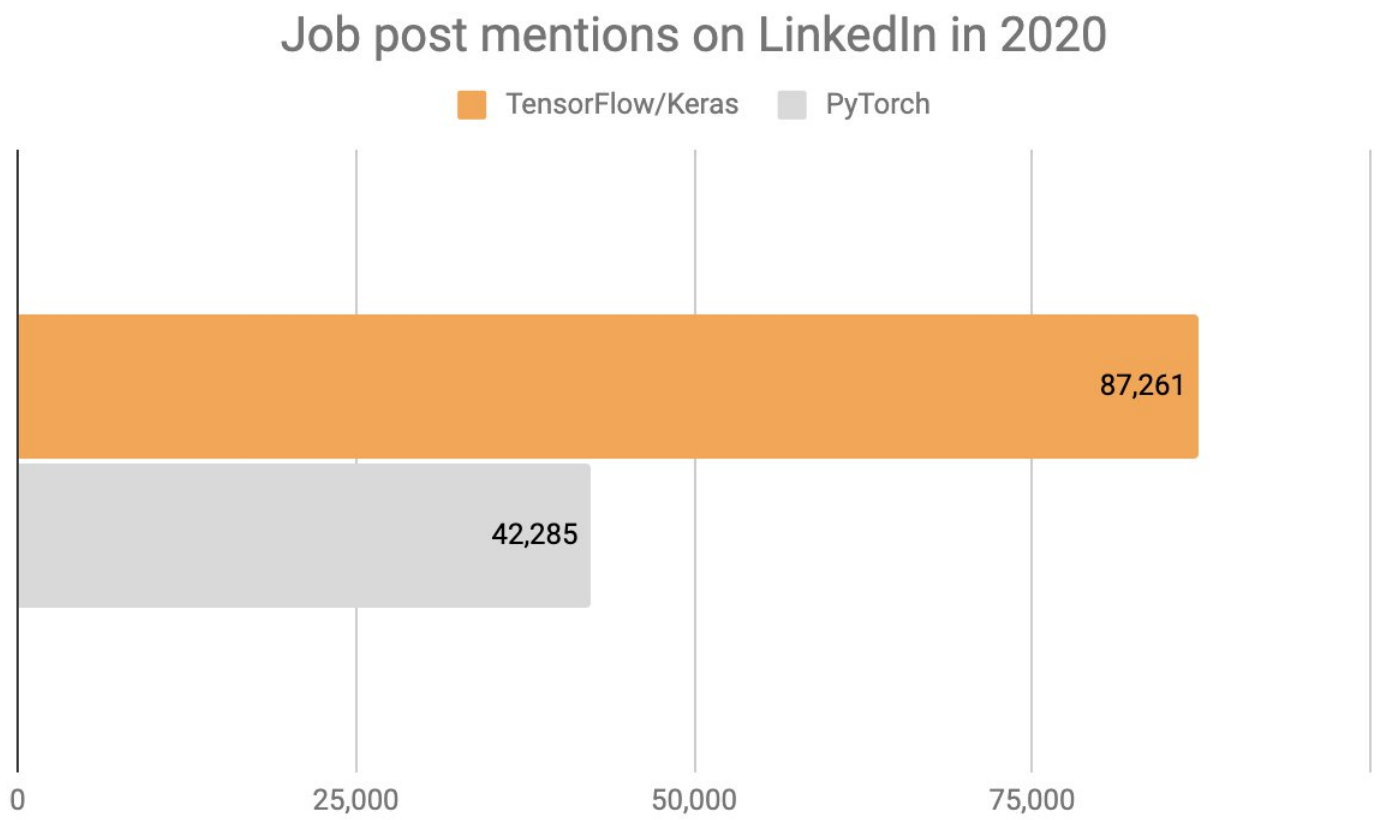
In a way, this metric reflects usage in production.

There were two worldwide developer surveys in 2020 that measured adoption of various frameworks: the one from StackOverflow, targeting all developers, and the one from Kaggle.

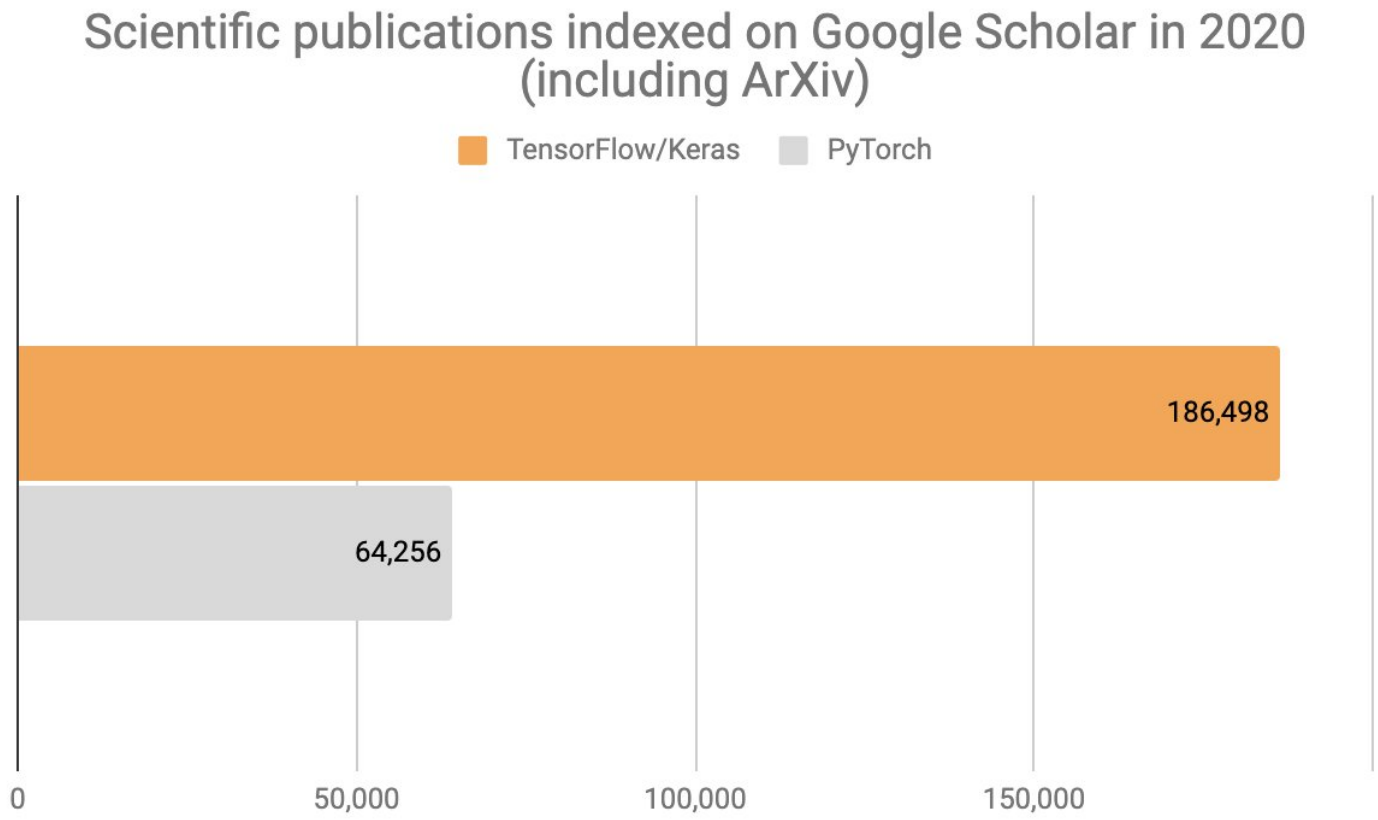


Note that the StackOverflow survey listed both TF and Keras; Keras had very strong metrics, and I suspect many people checked Keras without checking TF. So if "TF/Keras" was a choice, it would have significantly higher numbers here (probably around 15% overall usage).

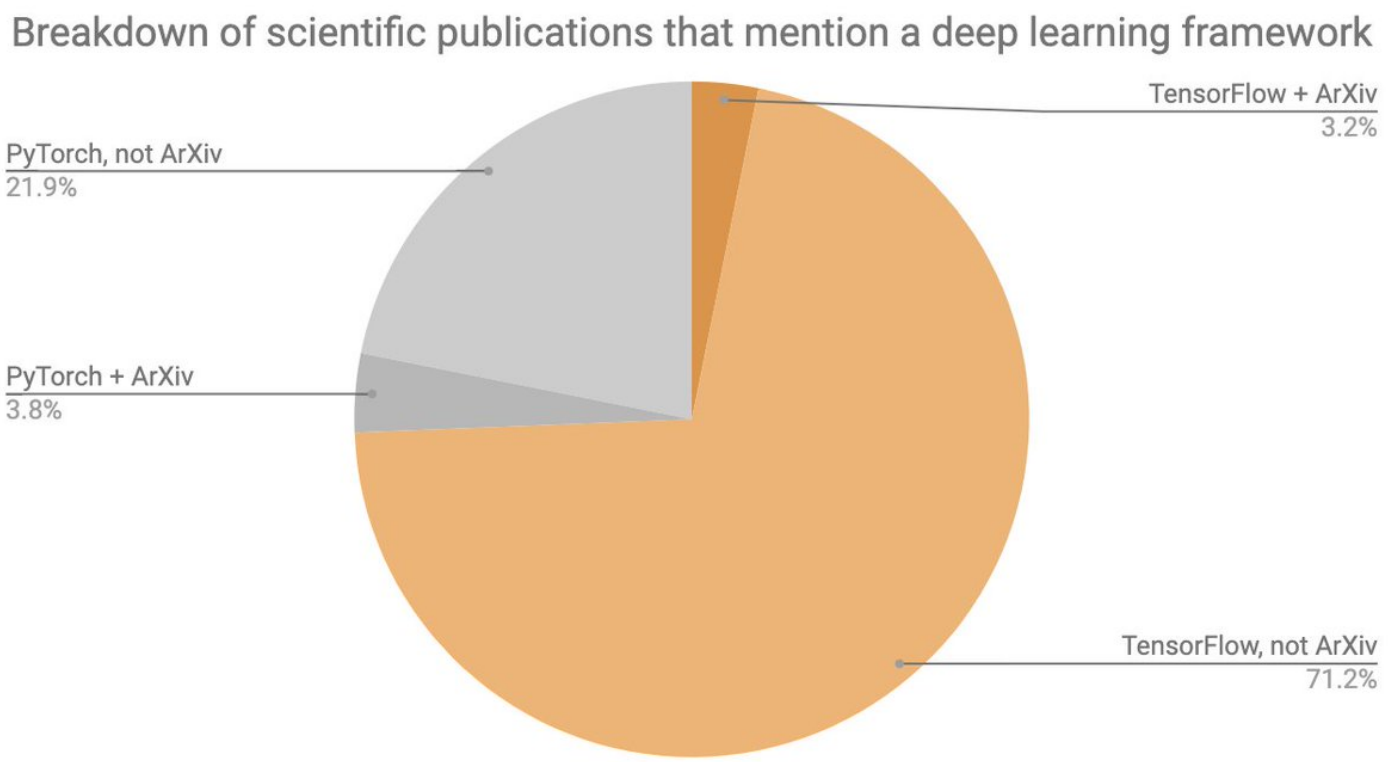
Mentions in LinkedIn job posts is a metric that I'm not quite sure is meaningful, unfortunately. It doesn't reflect the stack of companies that hire, only the keywords tracked by recruiters.



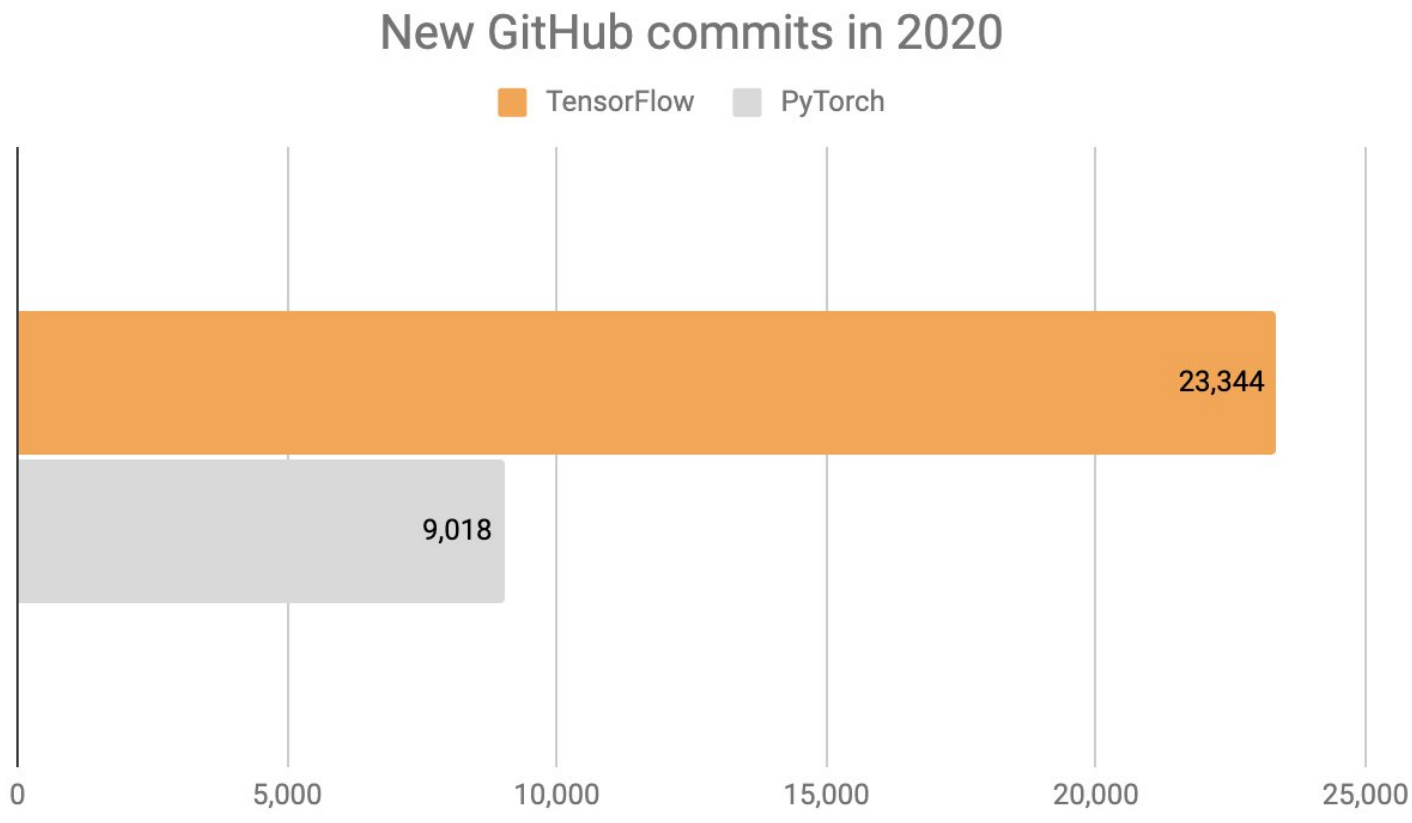
We can track usage in the research community in two categories: ArXiv, which represents "pure deep learning" research, and Google Scholar, which represents all publications, including applications of deep learning to biology, medicine, etc.



Deep learning research is an important but small niche (~20k users of deep learning out of several millions in total) and it is the only niche where PyTorch is neck-to-neck with TensorFlow.



Finally, GitHub metrics. GitHub makes it possible to track new commits over the last year, but doesn't make it possible to track new stars/forks/watchers, hence why I'm displaying total numbers for these rather than 2020 increases.



Note that the GitHub metrics are only for the TensorFlow repo, not the dozens of large TensorFlow adjacent repos (like the Keras repo, etc).

Overall: 2020 has been a difficult year, in particular one during which many businesses have cut their exploratory investments in deep learning because of Covid, causing a slump from March to November. However, on balance, TF/Keras has still seen modest growth over the year.

Our current growth rate is solid, and our prospects for 2021 are looking bright! I'll post an update to these metrics in 2021. Here's to another year full of improvement, growth, and focusing on delighting our users :)