Very Short, Did Read

A new summarization model boils down AI research papers to a single sentence.
What's new: TLDR from Allen Institute for AI creates at-a-glance summaries of scientific research papers. It's up and running at Semantic Scholar, a research database, where searches now return its pithy precis.
How it works: The researchers trained BART, a pretrained language model, using a multitask learning strategy. Because the dataset of summaries was small, the authors also trained the model to generate titles, a task for which far more data was available.

- The researchers compiled SciTLDR, a dataset that comprises 5,411 one-sentence summaries of 3,229 papers. All papers are paired with at least one one-sentence summary written by the author of the paper. One-third of them also come with a one-sentence summary written by students based on peer-review comments.
- To this dataset, they added 20,000 scientific papers and their titles.
- They trained the model to generate either a summary or a title depending on a control code.
- To save on computation, TLDR analyzes only a paper's abstract, intro, and conclusion.

Results: TLDR was able to summarize research articles that averaged 5,000 words long using around 20 words. Human experts ranked the output of TLDR, BART model trained only on SciTLDR, the author-generated summaries, and student-generated summaries of 51 papers chosen at random. TLDF outperformed BART trained on SciTLDR, achieving a mean reciprocal rank, where 1 is highest, of 0.54 versus 0.42. Its output ranked on par with the author generated summaries (0.53) but worse than the student generated summaries (0.60).
Behind the news: Most summarizers produce summaries that average between 150 and 200 words.
Why it matters: At least 3 million scientific papers are published annually, Semantic Scholar estimates, and a growing portion of them describe innovations in AI, according to the AI Index from Stanford Human-Centered Artificial Intelligence. This model, along with the excellent Arxiv Sanity Preserver, promises a measure of relief to weary engineers and students. (To learn more about the Allen Institute for AI's research, watch our Heroes of NLP interview with AI2 CEO Oren Etzioni here.)
We're thinking: Some papers can be summed up in a couple of dozen words, but many are so complex that no single sentence can do them justice. We look forward to *n*-sentence summarizers.