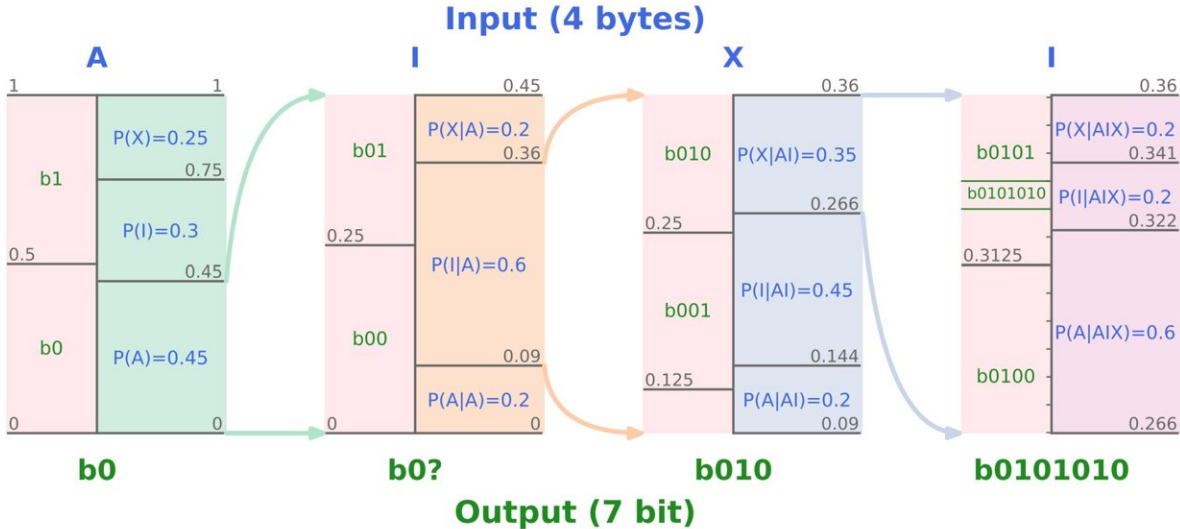


AI model beats PNG and FLAC at compression

October 3 2023, by Peter Grad



Arithmetic encoding of the sequence ‘AIXI’ with a probabilistic (language) model ? (both in blue) resulting in the binary code ‘0101001’ (in green). Arithmetic coding compresses data by assigning unique intervals to symbols based on the probabilities assigned by ?. It progressively refines these intervals to output compressed bits, which represent the original message. To decode, arithmetic coding initializes an interval based on the received compressed bits. It iteratively matches intervals with symbols using the probabilities given by ? to reconstruct the original message. Credit: *arXiv* (2023). DOI: 10.48550/arxiv.2309.10668

What would we do without compression?

Those music libraries and personal photo and video collections that would force us to purchase one [hard drive](#) after another can instead be squeezed into portions of a single drive.

Compression allows us to pull up volumes of data from the Internet virtually instantaneously.

Interruptions and irritating lag times would mar cellphone conversations without compression.

It allows us to improve [digital security](#), stream our favorite movies, speed up data analysis and save significant costs through more efficient digital performance.

Some observers wax poetic about compression. The popular science author Tor Nørretranders once said, "Compression of large amounts of information into a few exformation-rich macrostates with small quantities of nominal information are not only intelligent: they are very beautiful. Yes, even sexy. Seeing a jumble of confused data and shreds of rote learning compressed into a concise, clear message can be a real turn-on."

An anonymous author described compression as "a symphony for the modern age, transforming the cacophony of data into an elegant and efficient melody."

And futurist Jason Luis Silva Mishkin put it succinctly: "In the digital age, compression is akin to magic; it lets us fit the vastness of the world into our pockets."

Ever since the earliest days of digital compression when acronyms such as PKZIP, ARC and RAR became a part of computer users' routine vocabulary, researchers have continued to explore the most efficient

means of squeezing data into smaller and smaller packets. And when it can be done without loss of data, it is that much more valuable.

Researchers at DeepMind recently announced they have discovered that [large language models](#) can take [data compression](#) to new levels.

In a paper, "Language Modeling Is Compression," published on the preprint server *arXiv*, Grégoire Delétang said DeepMind's large language model Chinchilla 70B achieved remarkable lossless compression rates with image and audio data.

Images were compressed to 43.4% of original size, and [audio data](#) were reduced to 16.4% of original size. In contrast, standard image compression algorithm PNG squeezes images to 58.5% of original size, and FLAC compressors reduce audio files to 30.3%.

The results were particularly impressive because unlike PNG and FLAC, which were designed specifically for image and audio media, Chinchilla was trained to work with text, not other media.

Their research also brought to light a different view on scaling laws, that is, how compression quality changes as the size of compressed data changes.

"We provide a novel view on scaling laws," Delétang said, "showing that the dataset size provides a hard limit on model size in terms of compression performance."

In other words, there are upper limits to advantages achieved with large language model compressors the larger their dataset is.

"Scaling is not a silver bullet," Delétang said.

"Classical compressors like gzip aren't going away anytime soon since their [compression](#) vs. speed and size trade-off is currently far better than anything else," Anian Ruoss, a DeepMind research engineer and co-author of the paper, said in a recent interview.

More information: Grégoire Delétang et al, Language Modeling Is Compression, *arXiv* (2023). [DOI: 10.48550/arxiv.2309.10668](https://doi.org/10.48550/arxiv.2309.10668)

© 2023 Science X Network

Citation: AI model beats PNG and FLAC at compression (2023, October 3) retrieved 29 February 2024 from <https://techxplore.com/news/2023-10-ai-png-flac-compression.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.