

Title: Knowledge cutoff

URL: https://en.wikipedia.org/wiki/Knowledge_cutoff

PageID: 80445288

Categories: Category:Artificial intelligence, Category:Machine learning algorithms, Category:Natural language processing

Source: Wikipedia (CC BY-SA 4.0).

In machine learning , a knowledge cutoff (or data cutoff) is the point in time beyond which a model has not been trained on new data. The term is mostly used in reference to a large language model (LLM). [1] Any information about events after this date is absent from the model's internal knowledge base. [1] It cannot access information about later events without a system for real-time data access like RAG . [2] While useful for training and tuning LLMs , knowledge cutoffs introduce new limitations like hallucinations , information gaps, and temporal bias . [1]

Overview

A model with a fixed knowledge cutoff is unable to provide information on facts or developments that have emerged since that time, since the model is not connected to the internet. [1] Therefore, it may occasionally produce incorrect answers. [1] This is caused by the fact that training on newer data would cause a major price concern, given that training the most powerful large language models may soon cost over a billion dollars according to Time . [3]

Notable AI model cutoff dates include:

The GPT-4 model has a knowledge cutoff of September 2021. [4]

The GPT-4 Turbo model has an updated knowledge cutoff of December 2023. [4]

The Llama 4 models have a knowledge cutoff of August 2024. [5]

Effects of knowledge cutoffs

Knowledge gaps

Knowledge cutoffs create information gaps. The model lacks any knowledge of events or discoveries that are not included in its training data. [1] This can lead to hallucinations , where the model generates plausible but verifiably false statements. Such inaccuracies occur because LLMs choose words from an internal dictionary and select the most plausible words, which may or may not be accurate. [6]

Effective vs. reported cutoffs

A research paper on arXiv indicates that a model's functional knowledge may not be uniformly limited by its stated cutoff date. This effective cutoff often differs for various subjects and is influenced by the distribution of information within the training data itself. [7] Due to the high cost of retraining large language models, these models are rarely completely retrained to increase their knowledge cutoff. [8] Some models can also use integrated search tools to access more recent information, which blurs the line of their inherent knowledge base. For example, GPT-4 , can access its search tool and give real-time info. [4]

Attempts to overcome knowledge cutoffs

Retrieval-augmented generation

Retrieval-augmented generation (RAG) is a common technique used to overcome the limitations of a knowledge cutoff. [2] In a RAG system, the language model is connected to an external knowledge base or search engine to retrieve live data. This architecture allows the model to find current information relevant to a query and incorporate it into its response, often with citations. [2] Grounding a model in external data helps reduce the frequency of hallucinations and improves output accuracy. However, the external knowledge base might be outdated or contain biases, which

may also lead to incorrect information or hallucinations . [9] For example, Google AI Overviews have created false claims and the results are sometimes unreliable, since it either fail at understanding the source, or at generating the actual response properly. [9] However, a method to mitigate this is to apply techniques like reinforcement learning from human feedback, which can improve the quality of a large language model 's responses. [9]

Continual learning

Another approach is continual learning , which involves methods like adapters and LoRA . [10] These fine-tuning techniques permit efficient, incremental updates to a model without the high cost of a full retraining cycle. However, this does not give real-time awareness, since adding modules to the system may result in algorithmic bias and catastrophic forgetting, as the weights in the model become biased towards the new set of data. [10]

See also

Retrieval-augmented generation

Continual learning

Language model

Large language model

Hallucination (artificial intelligence)

Algorithmic bias

References