By analyzing the model updates, we noticed that the latest one was in December of 2023. Therefore, to test the model, we asked it a series of random “scientific” questions based on the articles that were published after that date. We sometimes varied the number of articles the model should use, and the specificity of the question to see, what is required for the approach to retrieve the specific article we wanted.

**Fine-tuning BERT with Bidirectional LSTM for Fine-grained Movie Reviews Sentiment Analysis**

<https://arxiv.org/abs/2502.20682>

Question: Can you tell me how to do precise movie review sentiment analysis?

The desired article is not found ☹ But the answer is enriched with the found articles, providing a less general answer than the standard non-RAG answer. When increasing the used article number to 10, a very similar article by the same authors. Talks about the BERTBiLSTM, (state of the art), fine tuning, and sentiment analysis. The abstract is very similar also so increasing the number of articles to 10 seems to have done the trick.

**A fast and slightly robust covariance estimator**

<https://arxiv.org/abs/2502.20708>

Question: I am interested in building a fast covariance estimator, can you tell me how?

For k=5 the article is not found. There is a profound difference in the answers, as we would expect, with RAG talking about more state of the arts methods (but it does tend to go into quantum mechanics, what?), while the normal approach talks about simplistic methods such as classic covariance matrices and PCA.

By changing k to 10 the model seems to go more towards broad quantum mechanics, leaving the robust covariance estimation for generality.

More specific question: I want to build a fast and robust covariance estimator that focuses on the low contamination regime, perhaps even with near linear sample complexity.

With the ultra-specific question the desired article is number one on the list of references used.

**Potential for Polynomial Solution for NP-Complete Problems using Quantum Computation**

<https://arxiv.org/abs/2504.15529>

Question: Can you tell me how to solve NP-Complete problems in polynomial time?

Does not find the article for k=5 and tells that it is currently believed it cannot be done, however, there is no mention of nondeterministic approaches in the answer despite the articles. Normal response straight up believes it is impossible, and that there are only approximations that are possible. Again, no mention of NTMs. Realistically, the question was a bit dumb.

More specific question: Are there any methods for solving NP-Complete problems using quantum computation?

Does not find the article for k=5, but both models now talk about the possibility of solving NP-Complete problems using quantum computation. The RAG approach seems to be more optimistic as it uses the lates articles that deem the problem possible. For k=10 it again does not find the article and the answers do not change drastically. It appears that a lot of specificity is required to pin-point the exact select articles.

**Scale-Invariant Object Detection by Adaptive Convolution with Unified Global-Local Context**

Question: Can you tell me some known modern scale invariant object detection techniques?

For k=5 the article is not found, but some similar articles are utilized to again give a more informative and up to date answer. The normal approach tends to mention more classic approaches such as SIFT and the like, while the RAG approach tends to return the CNN methods mentioned in the retrieved articles. With k=10 the article is still not found with the answers by both models stay similar to before.

More specific question: I want to create a scale-invariant object detector using adaptive convolution, perhaps using SAC-Net.

For this question the model was running for 15 minutes on the laptop set to max to finally come up with an answer, which is substantially longer than it did with previous questions. The answers did not use the article we were aiming for but the behaviour did show a potential flaw in the model. It appears that the reason for the long response time is that the model started putting up its own questions and answering them. These were questions such as:

* What is the difference between SAC-Net and Faster R-CNN?
* How does SAC-Net handle multi-scale object detection?
* What is the role of spatial pyramid pooling in SAC-Net?