

Sample results:

1. What was the first large-scale offshore wind project in the US?

System Evaluation Interface

1. Question Input

 Enter your question about environmental news

What was the first large-scale offshore wind project in the US?

2. Run Evaluation

 Run Evaluation

3. Evaluation Results

Base LLM (no RAG)

The first large-scale offshore wind project in the US is the Block Island Wind Farm, located off the coast of Block Island, Rhode Island. It consists of five wind turbines with a total capacity of 30 megawatts and was completed in 2016.

Advanced RAG (Base Model)

The first large-scale offshore wind project in the US is located off the coast of Massachusetts and has started generating power, aligning with President Joe Biden's initiative to expand coastal wind turbines for climate change mitigation.

Basic RAG

The first large-scale offshore wind project in the US was the Block Island Wind Farm off the coast of Rhode Island. It consists of five wind turbines and has a total capacity of 30 megawatts. The project began operations in 2016.

Advanced RAG (LoRA Model)

The first large-scale offshore wind project in the US is located off the coast of Massachusetts. It has recently started producing power, aligning with Joe Biden's goals of increasing coastal wind turbines to address climate change.

The Base LLM (No RAG) and Basic RAG provided the correct answer, identifying the Block Island Wind Farm off the coast of Rhode Island as the first large-scale offshore wind project in the US. Their responses were factually accurate and aligned with general knowledge. However, the Advanced RAG (Base Model) and Advanced RAG (LoRA Model) both incorrectly stated that the first large-scale project was located off the coast of Massachusetts. This suggests that the retrieval process may have surfaced more recent offshore wind projects, leading to a factual error. While the advanced models integrate retrieved information, their responses in this case reflect hallucination rather than accuracy.

2. How did the Zimbabwe elephant deaths happen?

3. 🔍 Evaluation Results

Base LLM (no RAG)

The Zimbabwe elephant deaths occurred in September 2020 when over 30 elephants were found dead in the Pandamasue Forest in Hwange National Park. The cause of death was initially unknown, but it was later determined that the elephants had been poisoned with cyanide. It is believed that the poisoning was carried out by poachers in an attempt to kill the elephants for their tusks. Cyanide poisoning is a common method used by poachers to kill elephants quickly and easily.

Advanced RAG (Base Model)

The Zimbabwe elephant deaths, totaling at least 160, have occurred due to drought conditions in the region. The hot and dry weather has led to these fatalities, with more deaths anticipated as the unfavorable conditions persist.

Basic RAG

The Zimbabwe elephant deaths were caused by a bacterial infection. The elephants were found dead near water sources. The exact cause of the infection is still under investigation.

Advanced RAG (LoRA Model)

The Zimbabwe elephant deaths occurred due to drought conditions that hit the country between August and December last year. Conservationists are concerned that more deaths may happen in the future. The elephants died in a 14,651 sq km area in Zimbabwe.

The Base LLM (No RAG) gave an outdated response, attributing the elephant deaths to cyanide poisoning by poachers—an event that happened in previous years but is not relevant to the latest news reports. Similarly, Basic RAG suggested that the cause was a bacterial infection, which is not supported by recent environmental reports. On the other hand, the Advanced RAG (Base Model) provided a more accurate response, correctly identifying drought as the primary cause of the elephant deaths. The Advanced RAG (LoRA Model) further improved upon this by incorporating additional details, mentioning the specific timeframe (August to December) and highlighting conservationists' concerns about potential future impacts. This demonstrates that the advanced models effectively integrate the latest news data, whereas the basic models struggle to retrieve and synthesize relevant, up-to-date.

3. How have recent climate policies evolved according to environmental news sources?

3. 🔍 Evaluation Results

Base LLM (no RAG)

Recent climate policies have evolved in several ways according to environmental news sources. Some key trends include:

1. Increased focus on renewable energy: Many countries and regions have set ambitious targets for transitioning to renewable energy sources such as wind, solar, and hydropower. This shift is seen as crucial for reducing greenhouse gas emissions and combating climate change.
2. Stricter regulations on fossil fuels: There has been a growing push to phase out coal, oil, and natural gas in favor of cleaner alternatives. This has led to the implementation of stricter regulations on emissions from fossil fuel industries and the promotion of electric vehicles as a more sustainable transportation option.
3. International cooperation: Countries around the world are increasingly working together to address climate change through initiatives such as the Paris Agreement. This global cooperation is seen as essential for achieving meaningful progress in reducing emissions and limiting global warming.
4. Focus on nature-based solutions: There is a growing recognition of the importance of nature-based solutions in combating climate change, such as reforestation, sustainable agriculture, and conservation efforts. These approaches are seen as cost-effective and beneficial for both the environment and local communities.

Overall, recent climate policies have evolved towards more ambitious and comprehensive approaches to addressing climate change, with a greater emphasis on renewable energy, emissions reductions, international cooperation, and nature-based solutions.

Advanced RAG (Base Model)

Recent climate policies have evolved to include commitments to building more clean energy, transitioning from coal to renewables faster, cutting pollution, and providing stronger protections for nature.

Advanced RAG (LoRA Model)

Recent climate policies have evolved to include government pledges to build more clean energy, hasten the switch from coal to renewables, and cut planet-heating pollution faster. The government has promised stronger climate action compared to previous efforts.

Basic RAG

Recent climate policies have evolved to focus more on renewable energy sources and reducing carbon emissions. Environmental news sources report an increasing emphasis on international cooperation and commitments to combat climate change. There is also a growing recognition of the urgency to address the impacts of climate change on a global scale.

The Base LLM (No RAG) provided a broad, well-structured response detailing major trends in climate policy, including the transition to renewable energy, stricter fossil fuel regulations, international cooperation, and nature-based solutions. While informative, this response was based on general pretraining rather than actual recent news data. Basic RAG was slightly more focused but remained vague, summarizing trends without citing specific events or policies. The Advanced RAG (Base Model) and Advanced RAG (LoRA Model) performed significantly better, as they incorporated retrieved information from the dataset. These models mentioned government pledges, commitments to renewables, and regulatory shifts with greater specificity. The LoRA Model further improved on the base model by emphasizing government actions and policy enforcement, making its response the most informative and up-to-date.

4. How does plastic pollution affect marine ecosystems?

3. 🔍 Evaluation Results

Base LLM (no RAG)

Plastic pollution has a significant impact on marine ecosystems in several ways:

1. Physical harm: Marine animals can become entangled in plastic debris, leading to injuries, suffocation, and even death. Animals such as sea turtles, seabirds, and marine mammals often mistake plastic items for food and ingest them, causing internal injuries, blockages, and malnutrition.
2. Chemical pollution: Plastics can release harmful chemicals and toxins into the water, which can contaminate the marine environment and harm marine life. These chemicals can disrupt hormone levels, impair reproduction, and weaken immune systems in marine animals.
3. Habitat destruction: Plastic pollution can smother and damage coral reefs, seagrass beds, and other important marine habitats. This can disrupt the balance of the ecosystem and lead to a decline in biodiversity.
4. Microplastics: Plastic debris can break down into tiny particles known as microplastics, which can be ingested by a wide range of marine organisms, from plankton to fish. These microplastics can accumulate in the food chain and ultimately end up in seafood consumed by humans, posing a potential health risk.

Overall, plastic pollution poses a serious threat to marine ecosystems, disrupting their functioning and causing harm to the diverse species that rely on these habitats for survival. Addressing plastic pollution is crucial to protecting the health and integrity of our oceans.

Advanced RAG (Base Model)

Plastic pollution affects marine ecosystems by leading to the deaths of a million marine animals annually and by contaminating seafood with microplastics, which can enter our bodies when consumed. Additionally, the plastic waste in the ocean poses a significant threat to marine life and ecosystems.

Advanced RAG (LoRA Model)

Plastic pollution in marine ecosystems has a devastating impact, leading to the death of a million marine animals annually. Additionally, plastic waste can enter the food chain, affecting human health as well.

Basic RAG

Plastic pollution in marine ecosystems can harm marine life through ingestion or entanglement. It can also disrupt the food chain and ecosystem balance. Additionally, plastic debris can release harmful chemicals into the water, impacting marine organisms.

The Base LLM (No RAG) provided an extensive answer covering the effects of plastic pollution on marine life, including physical entanglement, chemical contamination, habitat destruction, and microplastics in the food chain. While factually accurate, this response was generic and not sourced from the latest environmental news. The Basic RAG response was more concise but still lacked specificity. In contrast, the Advanced RAG (Base Model) and Advanced RAG (LoRA Model) integrated more precise details, referencing figures (such as "a million marine animals dying annually") and mentioning the impact of plastic pollution on seafood contamination and human health. These additions suggest that the advanced models successfully retrieved and synthesized information from the news dataset, making them more informative than the base models.

5. What are the criticisms of the recent EPA regulations?

3. 🔍 Evaluation Results

Base LLM (no RAG) ↔

1. Industry groups and some politicians argue that the regulations are too strict and will harm businesses, particularly in the energy and manufacturing sectors.
2. Some critics claim that the regulations are based on flawed or incomplete scientific evidence, and that the EPA is overstepping its authority by imposing such stringent rules.
3. Environmental advocates argue that the regulations do not go far enough to protect public health and the environment, and that they should be even stricter to address the growing threats of climate change and pollution.
4. Some critics believe that the regulations will lead to job losses in certain industries, as companies may struggle to comply with the new requirements and may be forced to cut back on production or shut down altogether.
5. There are concerns that the regulations could disproportionately impact low-income and minority communities, as they may be more likely to live near sources of pollution and be exposed to higher levels of environmental harm.

Advanced RAG (Base Model)

The criticisms of the recent EPA regulations include the need for more testing and the recognition that there is still a long way to go in terms of regulation, legislation, and enforcement.

Advanced RAG (LoRA Model)

The criticisms of recent EPA regulations include some pollutants not being tested for, lack of standard testing protocols for many compounds, legislation lagging behind, and inconsistency in approach, particularly on a global level.

Basic RAG

I don't know.

The Base LLM (No RAG) performed well in this case, offering a structured response detailing five major criticisms of EPA regulations, such as claims that they are too strict, based on incomplete science, or disproportionately affect businesses and minority communities. However, this response lacked references to recent news reports, making it less reliable for answering a time-sensitive question. The Basic RAG completely failed here, returning an “*I don't know*” response, likely due to poor retrieval. On the other hand, the Advanced RAG (Base Model) provided a relevant but brief summary, stating that criticisms include the need for more testing and regulation improvements. The Advanced RAG (LoRA Model) outperformed all others by listing specific issues, including untested pollutants, lack of standard testing protocols, outdated legislation, and inconsistencies in regulation. This detailed response suggests that the LoRA model retrieved the most relevant documents and synthesized them effectively, making it the best performer.

General Evaluation Summary

Across five real-world questions, the evaluation of four model variants—Base LLM (No RAG), Basic RAG, Advanced RAG (Base Model), and Advanced RAG (LoRA Model)—revealed clear trends in their performance.

- The Base LLM offered generally informative responses rooted in pretrained knowledge, but lacked temporal relevance and often missed details specific to recent news.

- Basic RAG showed slight improvement by integrating some retrieval capability, but frequently returned vague or incomplete answers and in one case failed to respond at all.
- The Advanced RAG (Base Model) demonstrated a better grasp of news content, offering more targeted responses that incorporated retrieved documents. However, in certain cases it hallucinated or relied on outdated retrievals.
- The Advanced RAG with LoRA Model consistently delivered the most contextually relevant and up-to-date answers. It excelled at synthesizing retrieved summaries, embedding precise facts (e.g., numbers, timeframes, policy names), and addressing queries thoroughly. Particularly on nuanced or time-sensitive topics—like EPA regulations or elephant deaths in Zimbabwe—the LoRA agent stood out for its factual grounding and detailed explanation.

Overall, the LoRA-powered agentic system proved to be the most reliable and accurate, making it the best performer among the four evaluated configurations.