

DEMYSTIFYING RESEARCH PAPERS WITH GEMINI, FROM COMPLEX TO CONCISE

Applications of Gen AI
(Google Gemini)

**GENERATIVE AI
FOR ALL**

 **DINESHLAL**



DINESH LAL
(DIRECTOR, DATA SCIENCE)



WHAT IS COVERED IN THE DOCUMENT

- In this document we will discuss Generative AI use case in **Research Paper Summarization**
- Research Paper can be sometime very complex to understand
- We will see how to use **Gemini** can make this task simple
- We will refer to Research Paper "**Llama 2: Open Foundation and Fine-Tuned Chat Models**" for this Article



Open Gemini Advanced - <https://gemini.google.com/app>

Gemini Advanced

Gemini has just been updated. [See update](#)

Hello, Dinesh

How can I help you today?

Generate a fantastical image

Generate a futuristic image

Role play as a character from a novel

Goodness, what a delightful day! The sun fairly beams upon us - I should say it's an exceptional day for a walk in the countryside, wouldn't you agree? Though I'm not certain what sort of weather you are enjoying

Generate a stunning, playful image

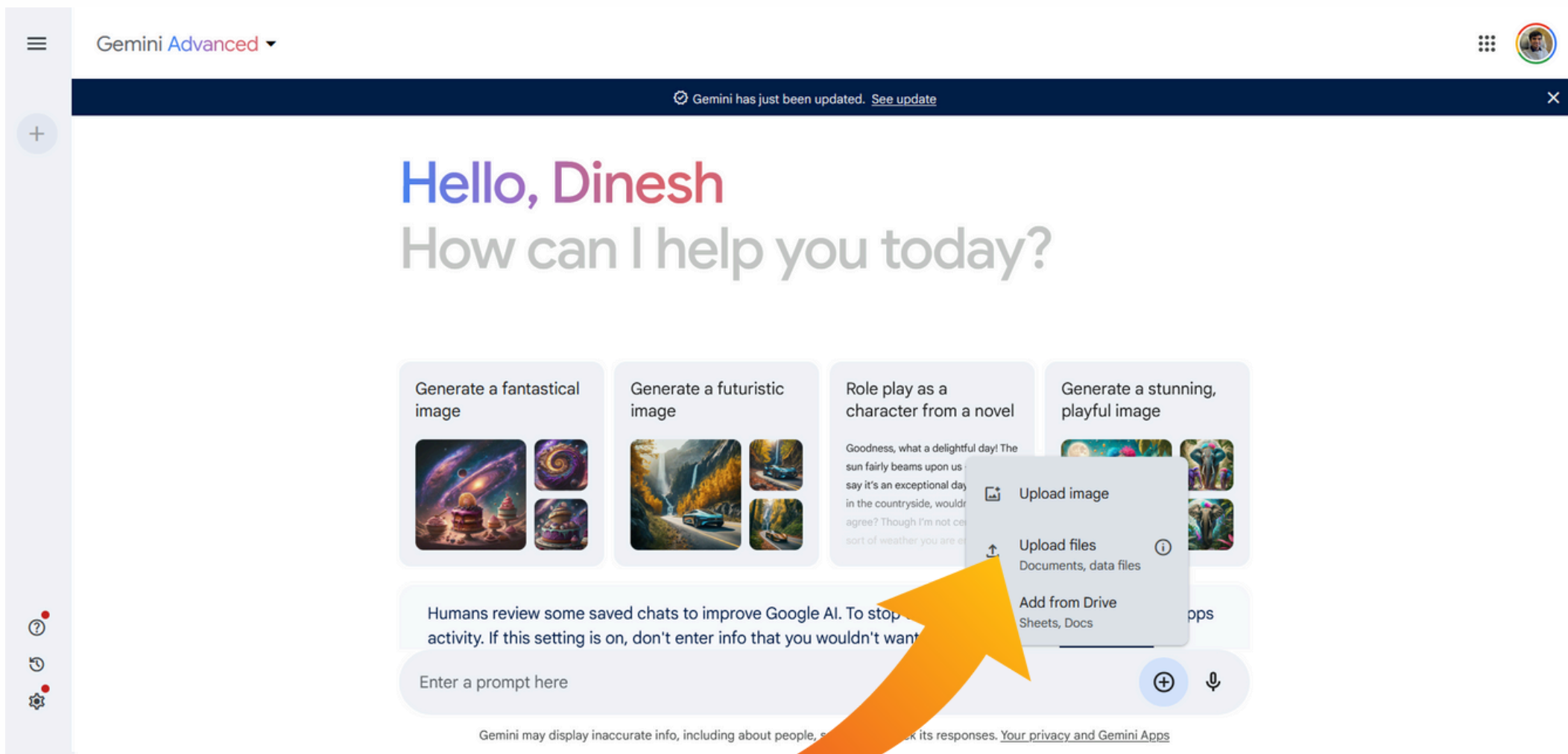
Humans review some saved chats to improve Google AI. To stop this for future chats, turn off Gemini Apps activity. If this setting is on, don't enter info that you wouldn't want reviewed or used. [How it works](#)

Enter a prompt here

Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)

Click Here

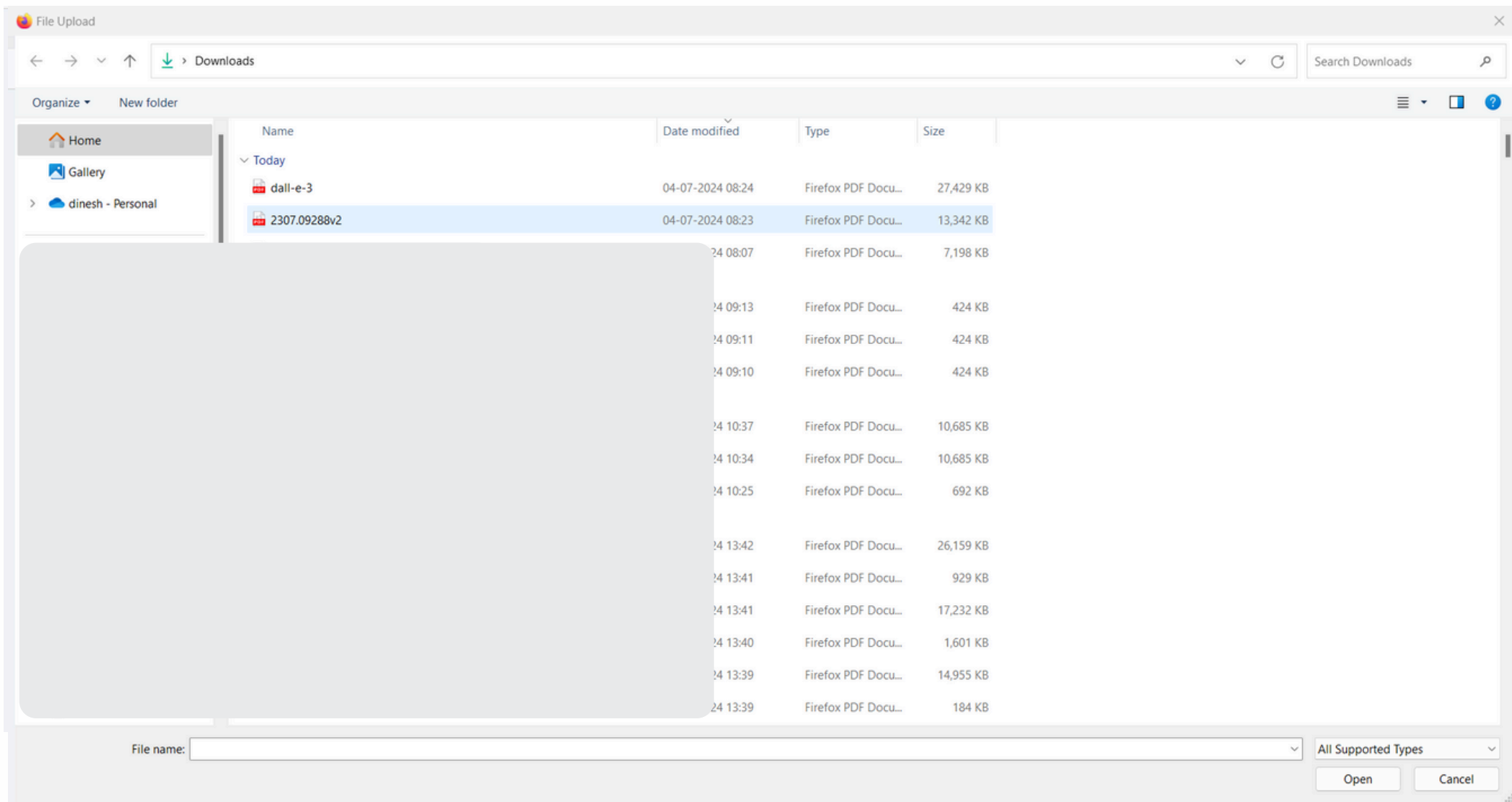
Click on Upload Files Option



The screenshot displays the Gemini Advanced web interface. At the top, it says "Gemini Advanced" with a dropdown arrow. A notification bar indicates "Gemini has just been updated. See update". The main heading reads "Hello, Dinesh" followed by "How can I help you today?". Below this are four interactive cards: "Generate a fantastical image", "Generate a futuristic image", "Role play as a character from a novel", and "Generate a stunning, playful image". A text box below the cards contains a disclaimer: "Humans review some saved chats to improve Google AI. To stop this activity. If this setting is on, don't enter info that you wouldn't want to be shared." Below the text box is a prompt input field labeled "Enter a prompt here" and a plus icon. A dropdown menu is open over the plus icon, showing three options: "Upload image", "Upload files" (which is highlighted with an orange arrow), and "Add from Drive". The "Upload files" option includes the subtext "Documents, data files".

Click on upload files

Select the Desired File to Upload

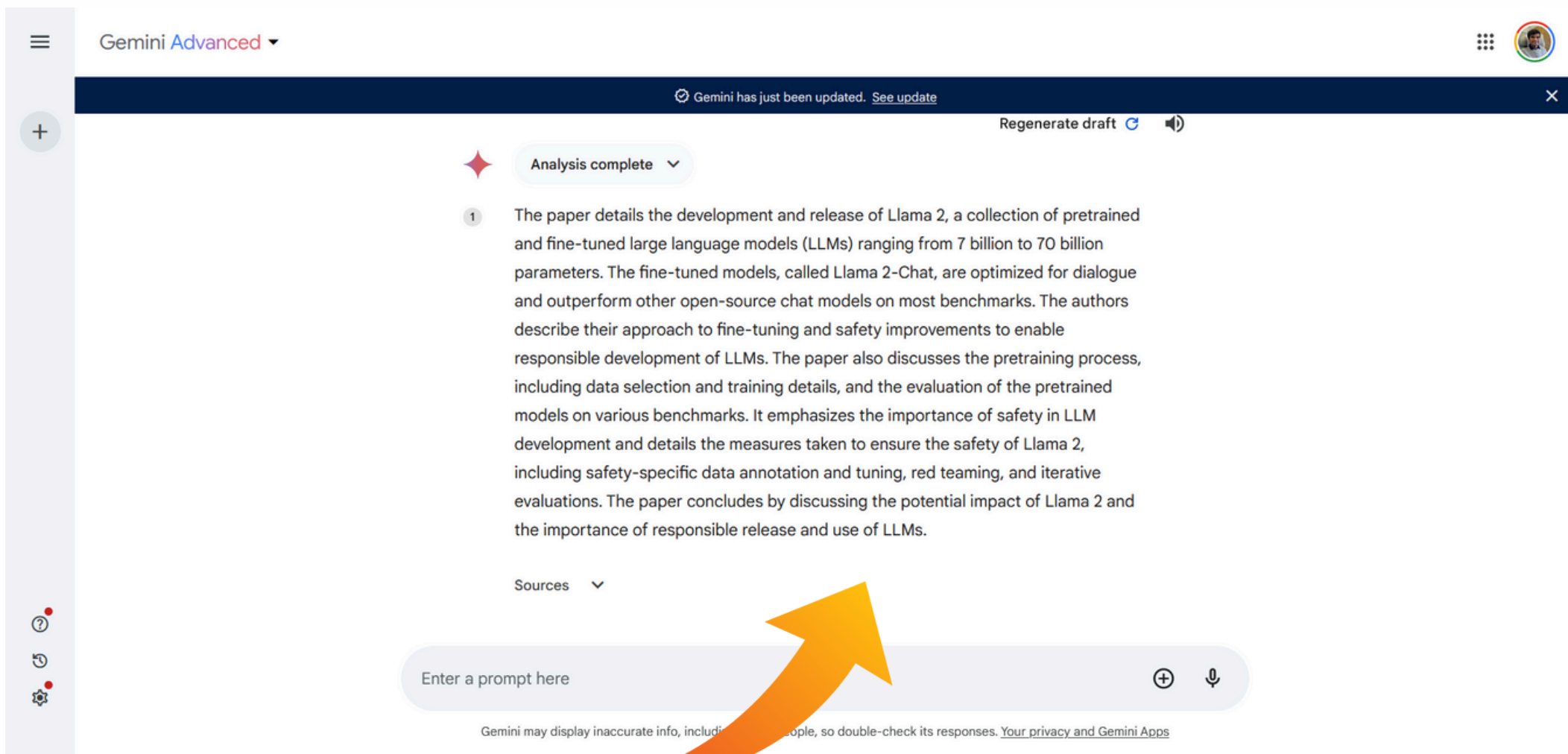


Summarize the Research Paper

The screenshot shows the Gemini Advanced web interface. At the top, there is a navigation bar with the Gemini logo and a notification: "Gemini has just been updated. See update". The main heading asks "How can I help you today?". Below this are four suggested actions: "Help me to understand Shakespeare", "Generate an image in a specific style", "Walk me through solving a problem", and "Generate a futuristic image". A text input field contains the prompt: "Help me summarize the Research Paper. Keep the response concise, I want to know at high level what is covered in the Paper." Below the prompt, a PDF file named "2307.09288v2" is attached. The interface also includes a sidebar with navigation icons and a footer with a disclaimer: "Gemini may display inaccurate info, in particular about people, so double-check its responses. Your privacy and Gemini Apps".

Write the Prompt as mentioned above

Gemini Generates A Concise Summary Output



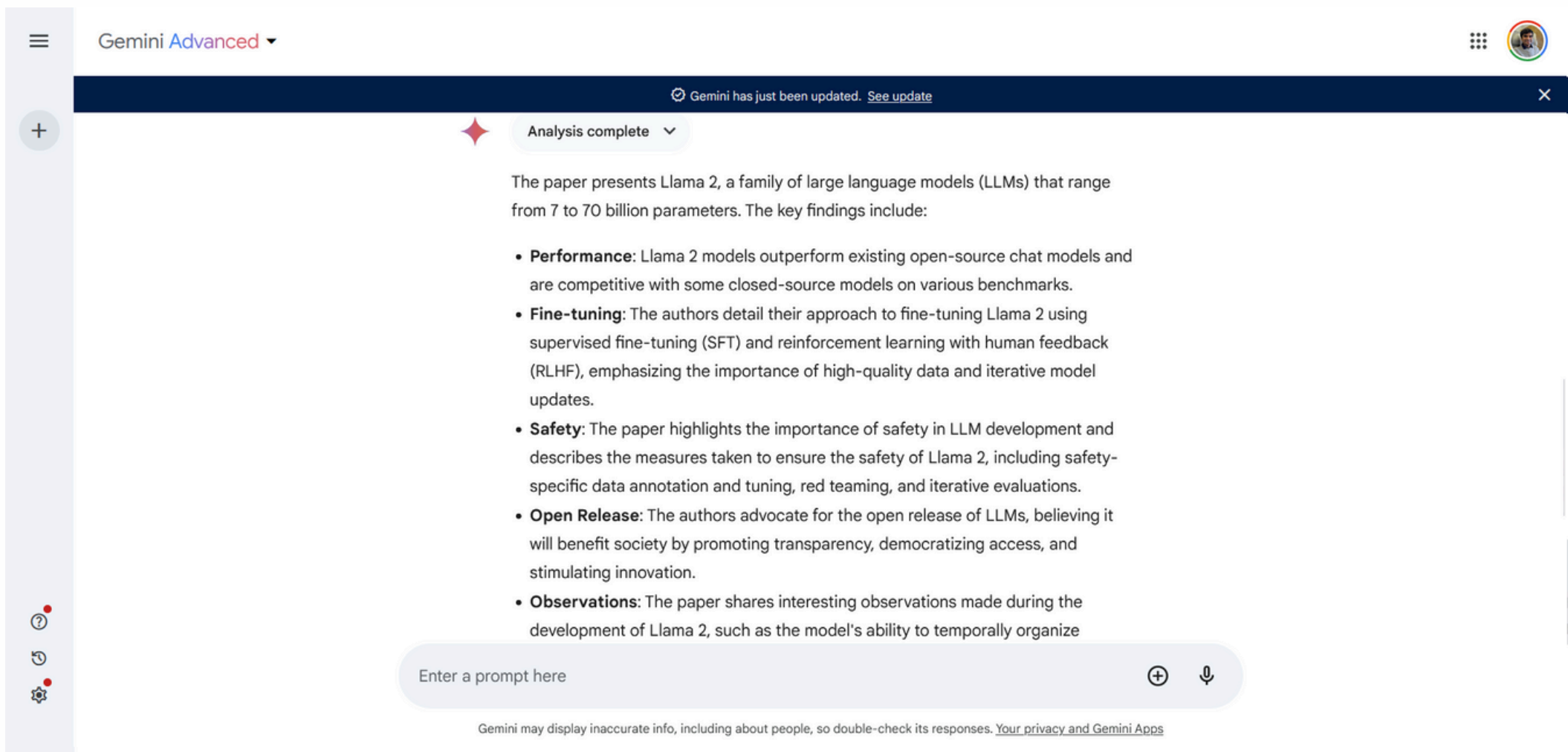
The screenshot displays the Gemini Advanced web interface. At the top, the text "Gemini Advanced" is visible on the left, and a notification bar states "Gemini has just been updated. See update". On the right, there are options for "Regenerate draft" and a speaker icon. The main content area shows a status "Analysis complete" with a dropdown arrow. Below this, a numbered list item (1) contains a concise summary of a paper about Llama 2. The summary text is: "The paper details the development and release of Llama 2, a collection of pretrained and fine-tuned large language models (LLMs) ranging from 7 billion to 70 billion parameters. The fine-tuned models, called Llama 2-Chat, are optimized for dialogue and outperform other open-source chat models on most benchmarks. The authors describe their approach to fine-tuning and safety improvements to enable responsible development of LLMs. The paper also discusses the pretraining process, including data selection and training details, and the evaluation of the pretrained models on various benchmarks. It emphasizes the importance of safety in LLM development and details the measures taken to ensure the safety of Llama 2, including safety-specific data annotation and tuning, red teaming, and iterative evaluations. The paper concludes by discussing the potential impact of Llama 2 and the importance of responsible release and use of LLMs." Below the summary is a "Sources" dropdown menu. At the bottom of the interface is a text input field with the placeholder "Enter a prompt here" and icons for adding attachments and voice input. A large orange arrow points from the text "Output from Gemini" below to the summary text in the screenshot.

Output from Gemini

Ask for the Key Findings

The screenshot displays the Gemini Advanced chat interface. At the top left, there is a hamburger menu icon and the text "Gemini Advanced" with a dropdown arrow. At the top right, there is a grid icon and a circular profile picture. A dark blue notification bar at the top center contains the text "Gemini has just been updated. [See update](#)" with a close button on the right. The main chat area shows a user's prompt: "What are the key findings of the paper?". Below the prompt, the AI's response is in progress, indicated by a red star icon and the text "Analyzing... Working on my response". Three horizontal progress bars of varying lengths and colors (pink, purple, and red) are shown below the AI's text. At the bottom of the chat area, there is a text input field with the placeholder "Enter a prompt here" and three icons: a plus sign, a microphone, and a square. Below the input field, a small disclaimer reads: "Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)".

Gemini Shares The Key Findings of The Paper



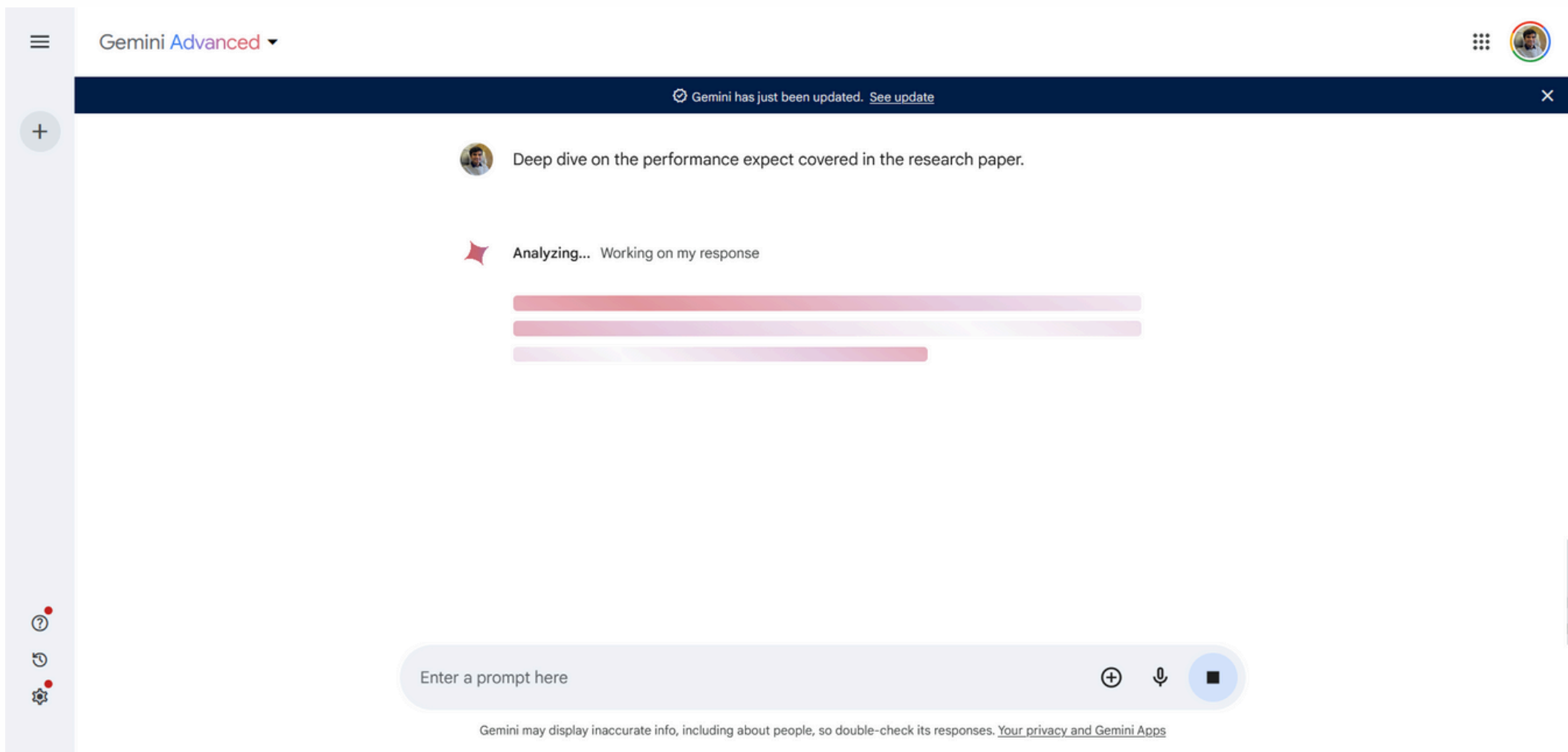
The screenshot displays the Gemini AI interface. At the top, it says "Gemini Advanced" with a dropdown arrow. A notification bar indicates "Gemini has just been updated. See update". Below this, a status bar shows "Analysis complete" with a dropdown arrow. The main content area contains the following text and list:

The paper presents Llama 2, a family of large language models (LLMs) that range from 7 to 70 billion parameters. The key findings include:

- **Performance:** Llama 2 models outperform existing open-source chat models and are competitive with some closed-source models on various benchmarks.
- **Fine-tuning:** The authors detail their approach to fine-tuning Llama 2 using supervised fine-tuning (SFT) and reinforcement learning with human feedback (RLHF), emphasizing the importance of high-quality data and iterative model updates.
- **Safety:** The paper highlights the importance of safety in LLM development and describes the measures taken to ensure the safety of Llama 2, including safety-specific data annotation and tuning, red teaming, and iterative evaluations.
- **Open Release:** The authors advocate for the open release of LLMs, believing it will benefit society by promoting transparency, democratizing access, and stimulating innovation.
- **Observations:** The paper shares interesting observations made during the development of Llama 2, such as the model's ability to temporally organize

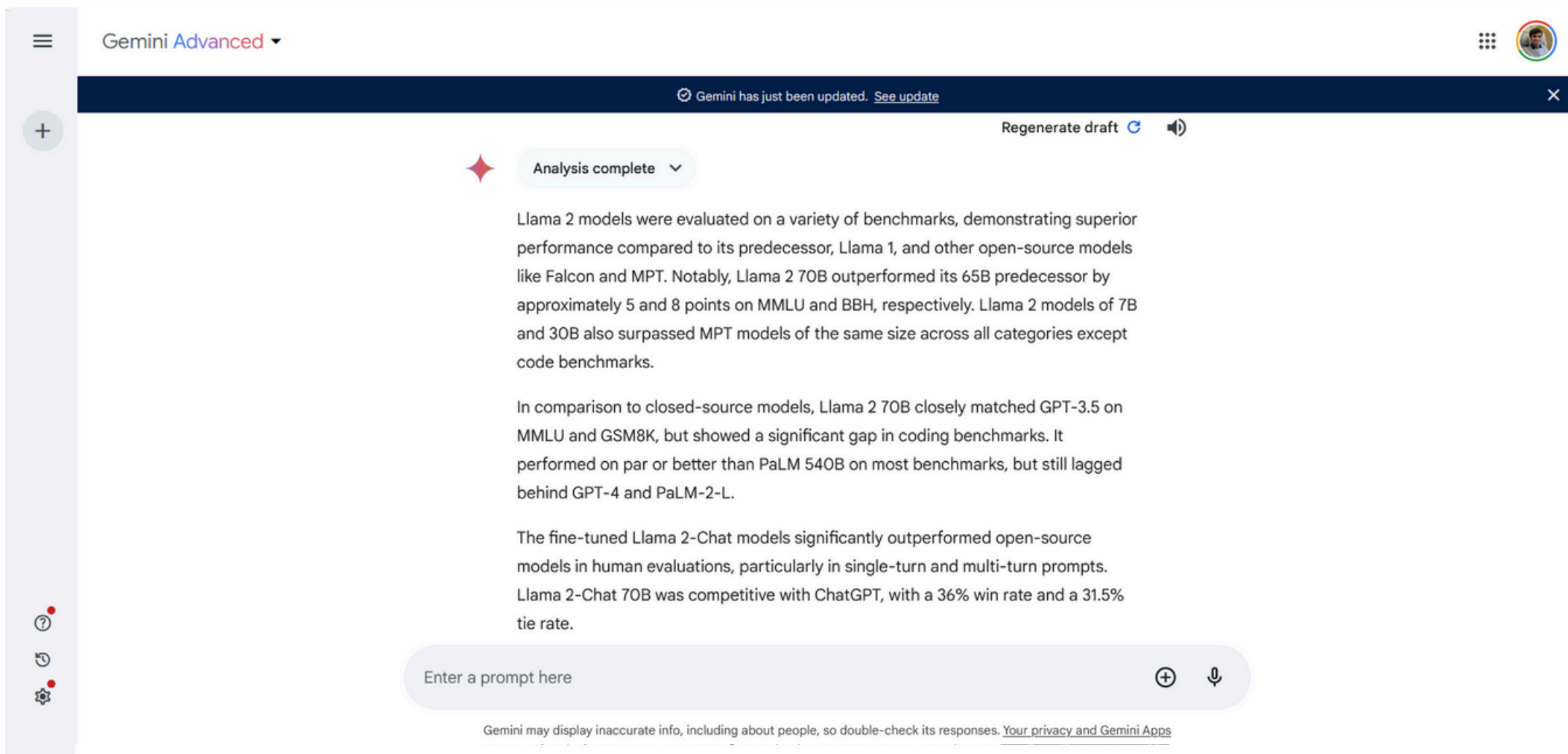
At the bottom of the interface, there is a text input field labeled "Enter a prompt here" with a plus icon and a microphone icon. Below the input field, a disclaimer reads: "Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)"

Deep Dive on a Specific Area of Paper



The screenshot displays the Gemini Advanced web interface. At the top left, there is a hamburger menu icon and the text "Gemini Advanced" with a dropdown arrow. At the top right, there is a grid icon and a circular profile picture. A dark blue notification bar at the top center contains the text "Gemini has just been updated. [See update](#)" with a close button on the right. The main chat area shows a user prompt: "Deep dive on the performance expect covered in the research paper." Below the prompt, a red star icon indicates the AI is "Analyzing... Working on my response". Three horizontal bars of varying lengths and colors (red, purple, and pink) represent the AI's response. At the bottom, there is a text input field with the placeholder "Enter a prompt here" and icons for adding attachments, voice input, and a stop button. A small disclaimer at the bottom center reads: "Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)".

Gemini Does the Deep Dive and Share Output



The screenshot displays the Gemini Advanced chat interface. At the top, there is a notification bar stating "Gemini has just been updated. See update". Below this, a "Regenerate draft" button is visible. The main content area features a notification "Analysis complete" with a red star icon. The analysis text is as follows:

Llama 2 models were evaluated on a variety of benchmarks, demonstrating superior performance compared to its predecessor, Llama 1, and other open-source models like Falcon and MPT. Notably, Llama 2 70B outperformed its 65B predecessor by approximately 5 and 8 points on MMLU and BBH, respectively. Llama 2 models of 7B and 30B also surpassed MPT models of the same size across all categories except code benchmarks.

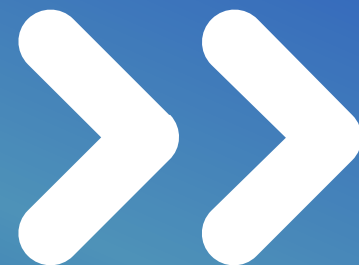
In comparison to closed-source models, Llama 2 70B closely matched GPT-3.5 on MMLU and GSM8K, but showed a significant gap in coding benchmarks. It performed on par or better than PaLM 540B on most benchmarks, but still lagged behind GPT-4 and PaLM-2-L.

The fine-tuned Llama 2-Chat models significantly outperformed open-source models in human evaluations, particularly in single-turn and multi-turn prompts. Llama 2-Chat 70B was competitive with ChatGPT, with a 36% win rate and a 31.5% tie rate.

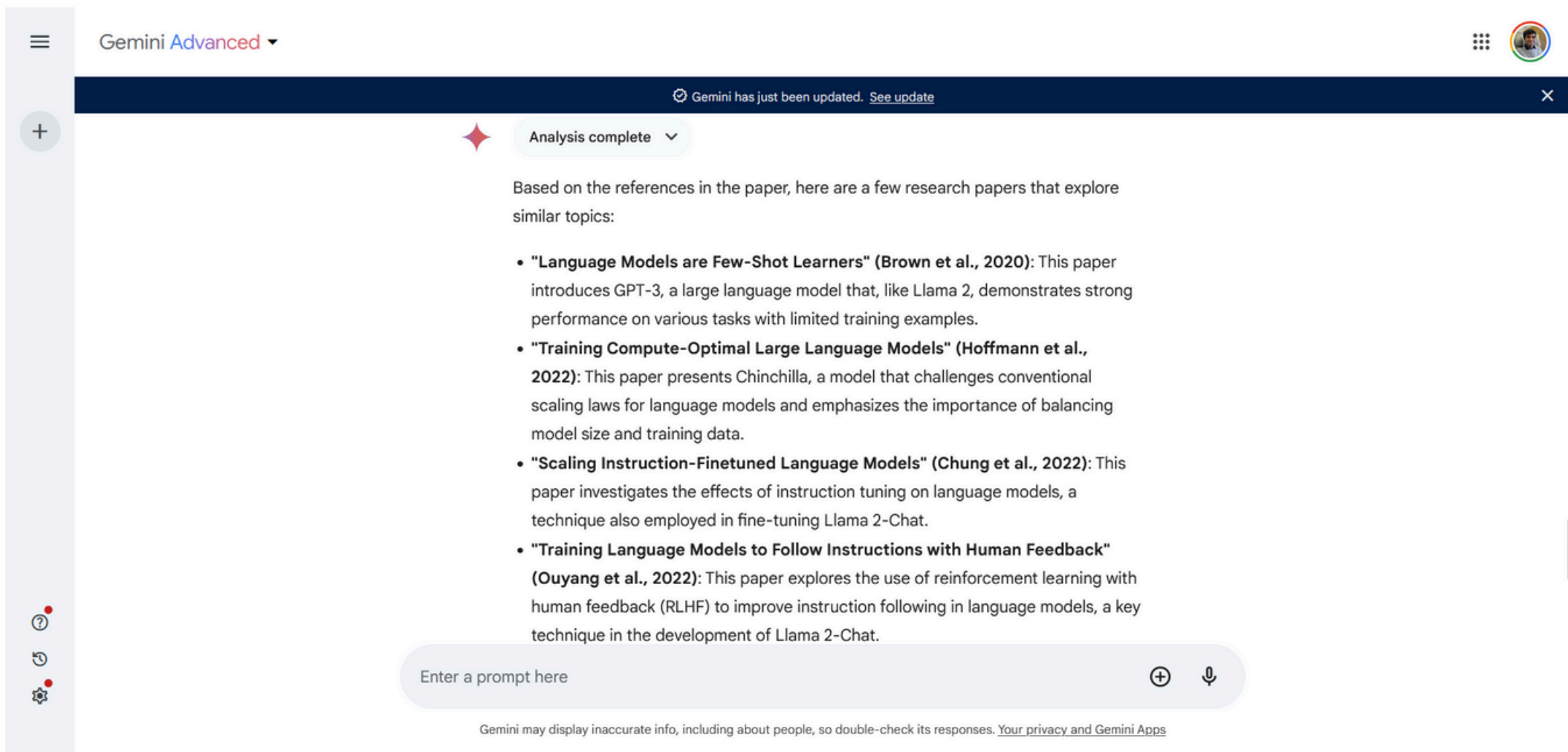
At the bottom of the interface, there is a text input field labeled "Enter a prompt here" with a plus icon and a microphone icon. A footer note states: "Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)".

Ask for Similar Research Paper for More Deep Dives

The screenshot displays the Gemini Advanced chat interface. At the top left, there is a hamburger menu icon and the text 'Gemini Advanced'. At the top right, there is a grid icon and a user profile picture. A dark blue notification bar at the top center reads 'Gemini has just been updated. See update'. The main chat area shows a user's message: 'Suggest me some alternative research papers similar to this one.' Below it, the AI's response is 'Analyzing... Working on my response', accompanied by three horizontal pink progress bars. At the bottom, there is a text input field with the placeholder 'Enter a prompt here' and icons for adding attachments, voice input, and a stop button. A small disclaimer at the bottom center states: 'Gemini may display inaccurate info, including about people, so double-check its responses. Your privacy and Gemini Apps'.



Gemini Shares a List of Papers that can be Read Further



The screenshot shows the Gemini Advanced chat interface. At the top, there is a navigation bar with the Gemini logo and a user profile icon. A notification banner states "Gemini has just been updated. See update". Below this, a status indicator shows "Analysis complete". The main content area displays a list of research papers based on references in a paper. The list includes:

- **"Language Models are Few-Shot Learners" (Brown et al., 2020):** This paper introduces GPT-3, a large language model that, like Llama 2, demonstrates strong performance on various tasks with limited training examples.
- **"Training Compute-Optimal Large Language Models" (Hoffmann et al., 2022):** This paper presents Chinchilla, a model that challenges conventional scaling laws for language models and emphasizes the importance of balancing model size and training data.
- **"Scaling Instruction-Finetuned Language Models" (Chung et al., 2022):** This paper investigates the effects of instruction tuning on language models, a technique also employed in fine-tuning Llama 2-Chat.
- **"Training Language Models to Follow Instructions with Human Feedback" (Ouyang et al., 2022):** This paper explores the use of reinforcement learning with human feedback (RLHF) to improve instruction following in language models, a key technique in the development of Llama 2-Chat.

At the bottom of the chat area, there is a text input field with the placeholder "Enter a prompt here" and icons for adding attachments and voice input. A disclaimer at the very bottom reads: "Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)"



Thank You

**SPECIAL THANKS TO GEMINI
FOR THE SUPPORT ON CONTENT**