

LMSYS Chatbot Arena: Benchmarking LLMs in the Wild


[Blog](#) | [GitHub](#) | [Paper](#) | [Dataset](#) | [Twitter](#) | [Discord](#) | [Kaggle Competition](#)

 **NEWS: We got a shorter URL! Reach us via lmarena.ai**

Rules

- Ask any question to two chosen models (e.g., ChatGPT, Gemini, Claude, Llama) and vote for the better one!
- You can chat for multiple turns until you identify a winner.

Choose two models to compare

claude-3-5-sonnet-20240620	claude-3-opus-20240229
 Expand to see the descriptions of 70 models	
<div>Model A</div> <div><p>Fix and handle edge cases where conversions between units might produce incorrect results due to rounding errors.</p><p>Certainly! I'll modify the unit converter to handle edge cases and minimize rounding errors. We'll use the <code>decimal</code> module for more precise calculations and implement proper rounding. Here's an updated version of the script with these improvements:</p><pre>import tkinter as tk from tkinter import ttk, filedialog import requests import json import csv import statistics from decimal import Decimal, get</pre></div>	<div>Model B</div> <div><p>Fix and handle edge cases where conversions between units might produce incorrect results due to rounding errors.</p><p>Certainly! Let's update the script to handle edge cases and fix potential rounding errors during unit conversions. We'll use the <code>decimal</code> module for precise decimal arithmetic.</p><p>Here's the updated script:</p><pre>import tkinter as tk from tkinter import ttk from tkinter import filedialog import requests import statistics from decimal import Decimal, get def convert_single():</pre></div>

👉 A is better

👉 B is better

🤝 Tie

👎 Both are bad

Send

🗑️ Clear history

🔄 Regenerate

📷 Share

Parameters

Terms of Service

Users are required to agree to the following terms before using the service:

The service is a research preview. It only provides limited safety measures and may generate offensive content. It must not be used for any illegal, harmful, violent, racist, or sexual purposes. Please do not upload any private information. The service collects user dialogue data, including both text and images, and reserves the right to distribute it under a Creative Commons Attribution (CC-BY) or a similar license.

Please report any bug or issue to our [Discord](#)/arena-feedback.

Acknowledgment

We thank [UC Berkeley SkyLab](#), [Kaggle](#), [MBZUAI](#), [a16z](#), [Together AI](#), [Hyperbolic](#), [RunPod](#), [Anyscale](#), [HuggingFace](#) for their generous [sponsorship](#).



Built with Gradio