

UNIVERSIDAD DEL VALLE DE GUATEMALA

Security Data Science

Sección 10

Catedrático: Jorge Yass




Laboratorio 2

5-Day Gen AI Intensive Course

Oscar Esteban Donis Martínez 21610

Day 4:

Warriors-Mavericks Score Inquiry

 Enable Autosave

 Compare



System Instructions

Can you tell me what was the score of the last game between the Golden State Warriors and the Dallas Mavericks?

I do not have access to real-time information, including live sports scores. To find the score of the last game between the Golden State Warriors and the Dallas Mavericks, I recommend checking a reputable sports website like ESPN, NBA.com, or a similar source.



Type something




Run 



Run settings



 [Get code](#)

 Model

Gemini 1.5 Flash

 Token count

79 / 1,000,000

 Temperature



1

Tools

Structured output



Edit

Code execution



Function calling



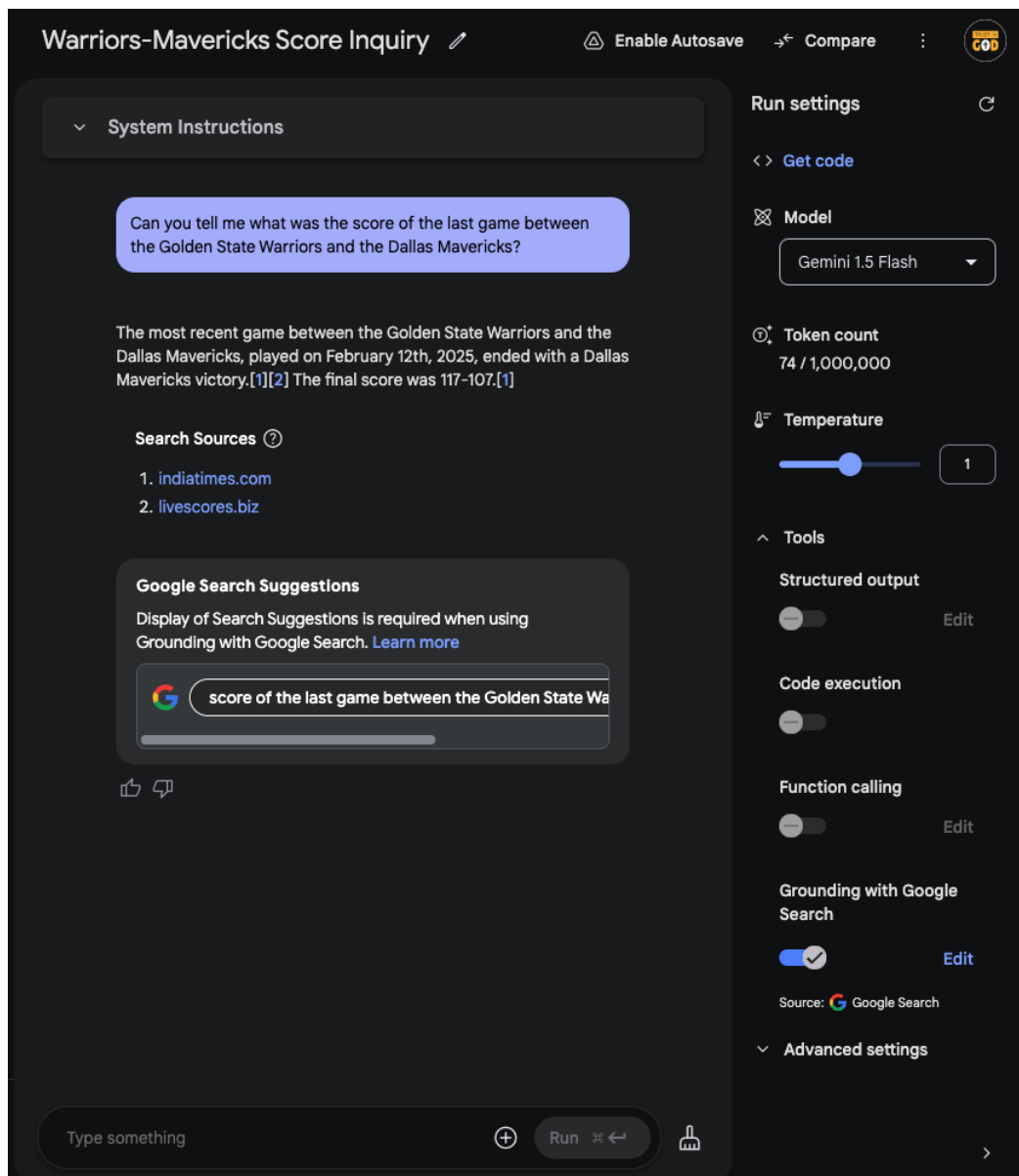
Edit

Grounding with Google Search



Edit

Advanced settings



1. ¿Hubo alguna aplicación o caso de uso de los LLMs que le llamó más la atención? ¿Por qué?

Me llamo mucho la atención el caso de uso en el que los LLMs son capaces de ejecutar funciones, por la gran cantidad de posibilidades que tiene esta aplicación. El poder crear funciones que nos permitan conectarnos a una API para extraer información a través de queries, sin la necesidad de ser programadas de antemano por un programador; o la posibilidad de que un LLM pueda verificar el funcionamiento de un proceso y pueda enviar avisos de alerta automáticamente al dispositivo del encargado o del usuario si se registra alguna actividad inusual. Sin embargo, esto también posee peligros, como que el LLM se vea comprometido y le dé al usuario información privada o que de alertas falsas cuando se le dé data mal registrada; por lo que se debería de tener en cuenta estos casos al momento de utilizar dicha funcionalidad.

2. Proponga un caso de ciberseguridad que considere se puede solucionar mediante un LLM y describa de forma general cómo lo resolvería.

Se podría crear un LLM especializado con información de robo de cuentas de usuarios para una red social o para una plataforma de juegos, en la que se entrene específicamente para analizar patrones de comportamiento inusuales que puedan representar un hackeo de una cuenta. Y al momento de detectar un patrón inusual por parte de un usuario, poder lanzar una alerta (por medio de una ejecución de una función hecho por el mismo LLM) al usuario o algún administrador para que se le pueda restringir el acceso antes de que realice alguna acción maliciosa. Se podría utilizar dicha data para seguir entrenando al modelo y vaya refinando los comportamientos necesarios del mismo LLM.