

Guatemala 16 de febrero de 2025 Josue Argueta - 211024

Security Data Science

Laboratorio 2, Curso intensivo de Gen AI

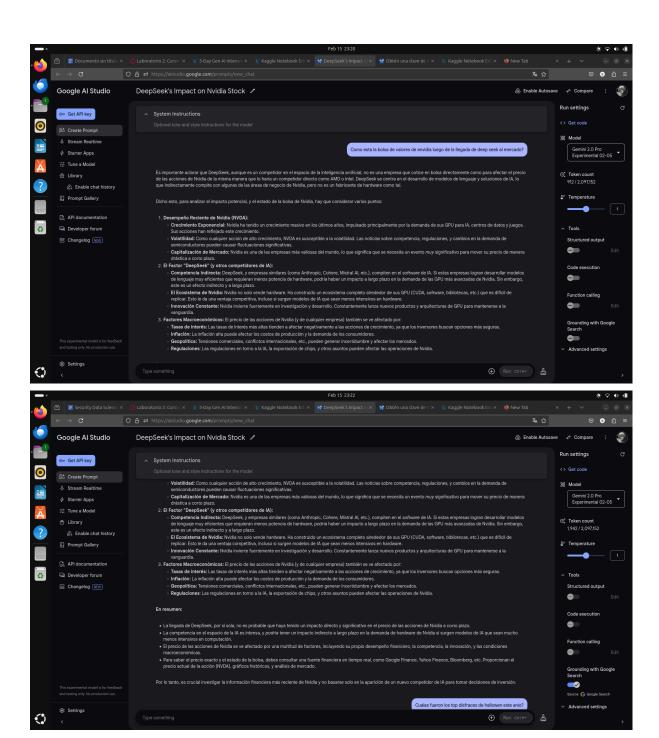
PREGUNTAS

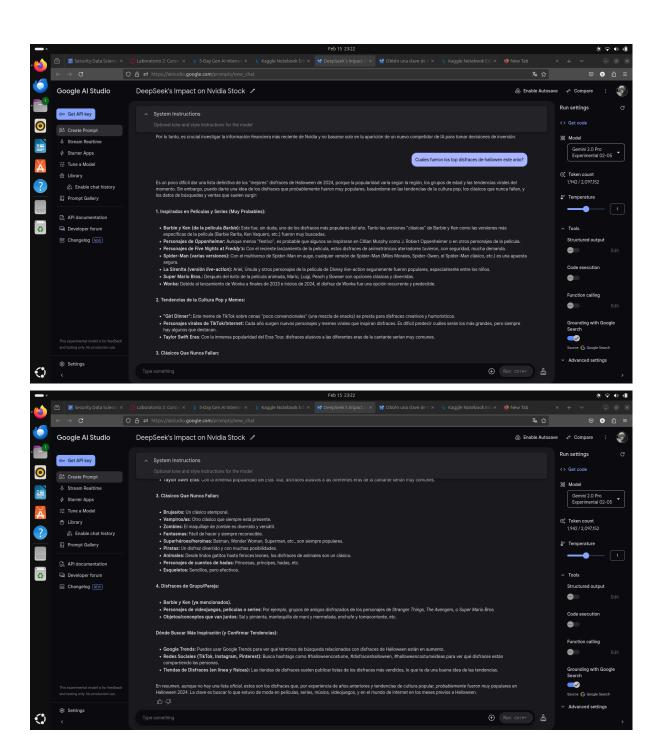
- 1. ¿Hubo alguna aplicación o caso de uso de los LLMs que le llamó más la atención? ¿Por qué? Entre lo que estaba viendo, comprendí que mayormente era para clasificar información, poder determinar qué tipo de elemento perteneces a esto u otro. Claro, modificando ciertos parámetros que tiene el modelo, para ser más preciso. El que más me gustó fue el día 2, parte 2. Donde se tenía un texto base y se determina cual de los otros textos era más cercano a ello. También, la parte 3 tiene más relación con el día 4, donde se clasificaba información, según el contexto que este tenía. Esos casos eran muy interesantes, ya que simplifican muchas de las cosas que podrían ser complicadas a la vista humana. Claro, cada texto o contexto tiene un patrón, desde el tipo de letras o palabras que utiliza para expresarlo y cómo es que se desenvuelve. Pero como mencione, el primero me interesó, porque me recordó a un tema de entropía que se realizó antes.
- 2. Proponga un caso de ciberseguridad que considere se puede solucionar mediante un LLM y describa de forma general cómo lo resolvería

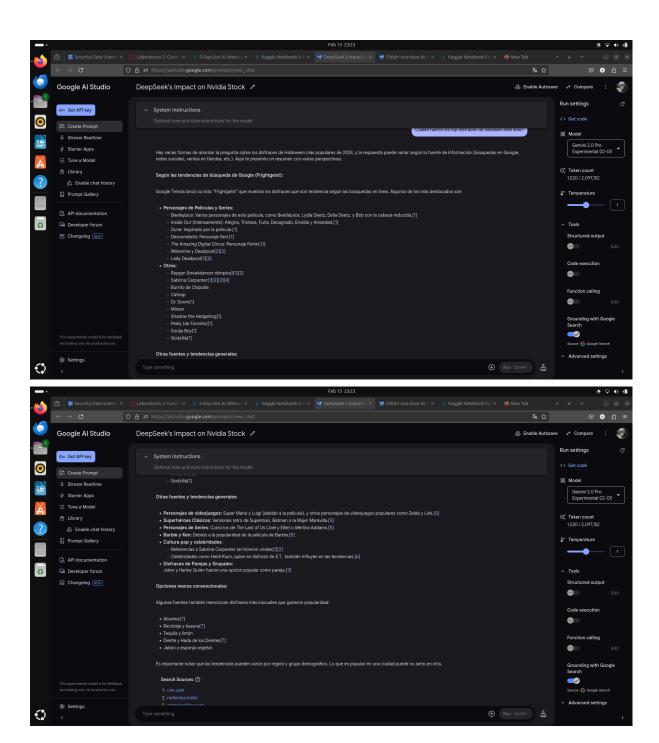
 En este caso, la clasificación de sitios maliciosos. Poder determinar si es o no un sitio malicioso, también, el phishing en general, pues hay mensajes que pueden tener contexto real, pero un enlace falso. Poder utilizar un LLM para poder determinar qué tipo de enlace es más propenso o tiende a ser phishing o no. En el laboratorio 1, halle una manera para determinar si es o no igual al sitio original. levenshtein distance, ayuda a determinar cuántas diferencias tiene entre uno u otro carácter o cambios que tuvo el sitio como tal, por ejemplo google -> g00gle, se determina que tiene 2, pues 00 no son oo. No solo es por el cambio de letras, sino que, también por falta de texto o exceso del mismo. En fin, me gusto, porque es para poder determinar entre uno y otro, cual es el más cercano, pero, tomando en cuenta que se clasifique el tipo de phishing, donde el más cercano, pero no tan exacto, seria el crítico, pues es el tipo de phishing realmente imperceptible al ojo humano.

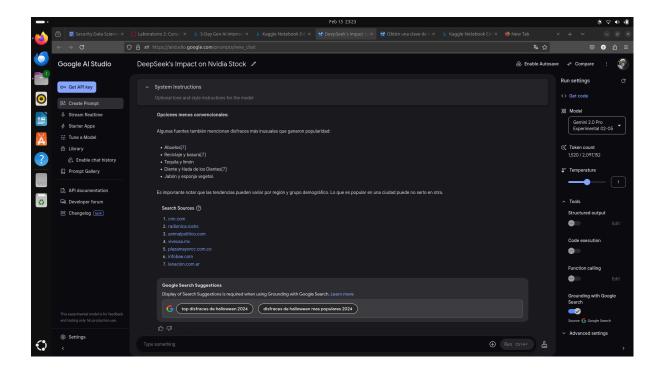
ANEXO

Este es del dia 4, sobre taylor swift prediction









Taylor Swift

