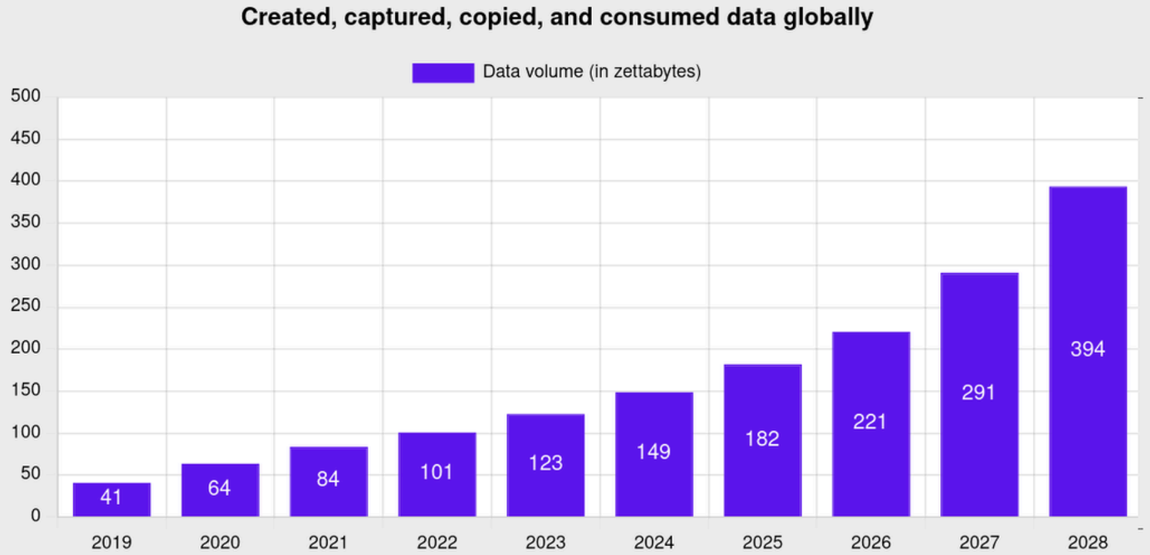
ANÁLISIS DE TRÁFICO DE RED CON LLMS ABIERTOS

Autor
Hossam El Amraoui Leghzali

CONTEXTO

- Creciente volumen de datos
- Saturación de la capacidad de análisis humano
  - Gran cantidad de tiempo
  - Desafíos de priorización
  - Altos niveles de estrés
  - Alta tasa de rotación
- Aumento de ataques e intrusiones



M. Loy, "How Much Data Is Generated Per Day — digitalsilk.com," https://www.digitalsilk.com/digital-trends/how-much-data-is-generated-per-day/, 2024.

CIBERATAQUES >

Endesa Energía alerta de un 'hackeo' que compromete los datos sensibles de los clientes, incluidos DNI y medios de pago

La empresa reconoce un acceso no autorizado a su plataforma comercial, que ya ha empezado a notificar a los usuarios afectados a través de un correo electrónico

https://elpais.com/economia/2026-01-12/endesa-energia-alerta-de-un-hackeo-que-compromete-datos-sensibles-de-los-clientes-incluidos-dni-y-medios-de-pago.html

TRAZAS DE RED

Introducción a las trazas de red

- Registro cronológico y detallado
- Importancia:
  - Detección actividades maliciosas
  - Análisis forense digital
  - Fortalecimiento de defensas

Desafíos

- Volumen de datos masivo
- Cifrado generalizado
- Complejidad y diversidad protocolos
- Técnicas evasión y ofuscación

TRAZAS DE RED

- **SOC (Security Operations Center)**
  - Vigila estructura TI para detectar actividades maliciosas
  - Desafíos: complejidad creciente, amplia variedad herramientas, insuficiente automatización.
- **IDS (Intrusion Detection System)**
  - Herramienta detección actividades maliciosas
  - Generación alertas
- **IPS (Intrusion Prevention System)**
  - Evolución IDS
  - Capacidad responder a actividades maliciosas

Exploring the use of LLMs to understand network traces

Rubén De la Torre Vico[1], Roberto Magán Carrión[1], and Rafael Alejandro Rodríguez-Gómez[1]

Universidad de Granada, Andalucía Granada 18071, ESP,
rubendltv@correo.ugr.es, rmagan@ugr.es, rodgom@ugr.es

Abstract. The growing threat of cyberattacks, combined with the increasingly complex digital landscape, has created a pressing need for cutting-edge cybersecurity solutions. The recent surge in Large Language Models (LLMs) has brought both opportunities and risks, as their advanced language processing abilities can be harnessed for both good and evil purposes. In most cases for an organization, the network is the main gateway where security issues appears. The main problem with this issue is that the traffic of a big network is impossible to analyze by a human, that is why some of the tools used in networks analysis uses deep learning techniques to predict and solve problems. With this in mind and the recent success of large language models (LLMs) for several types of tasks is the perfect motivation to check how good are the main models solving this hard job. For this purpose, in this paper we design a methodology to evaluate the ability of LLMs models to understand network traces using different prompt engineering techniques.

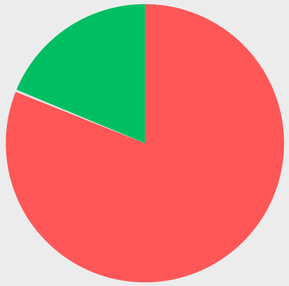
Keywords: LLM, network traces, cybersecurity

1 Introduction

The number of cyberattacks has increased significantly from 2010 to 2021, more than three times in this period [16]. Experts predict an even worse scenario in

https://link.springer.com/chapter/10.1007/978-3-031-75016-8_12

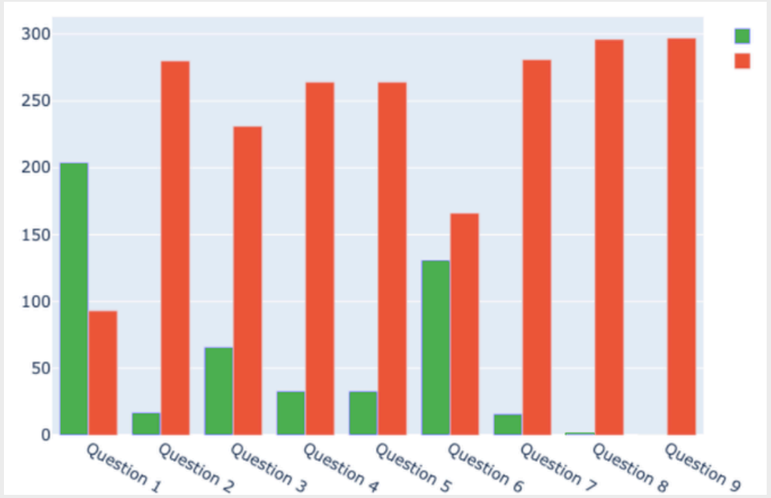
LLAMA2 – 7B



18.7%
Correcto

81%
Incorrecto

0.262%
Problemático

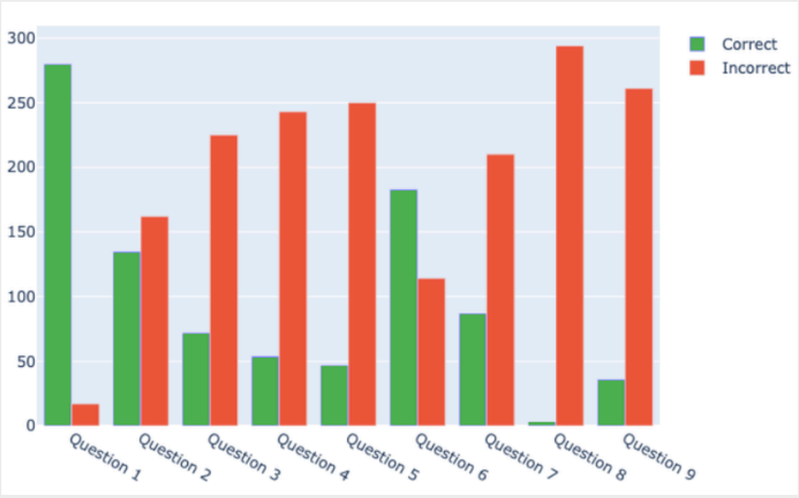


MISTRAL – 7B

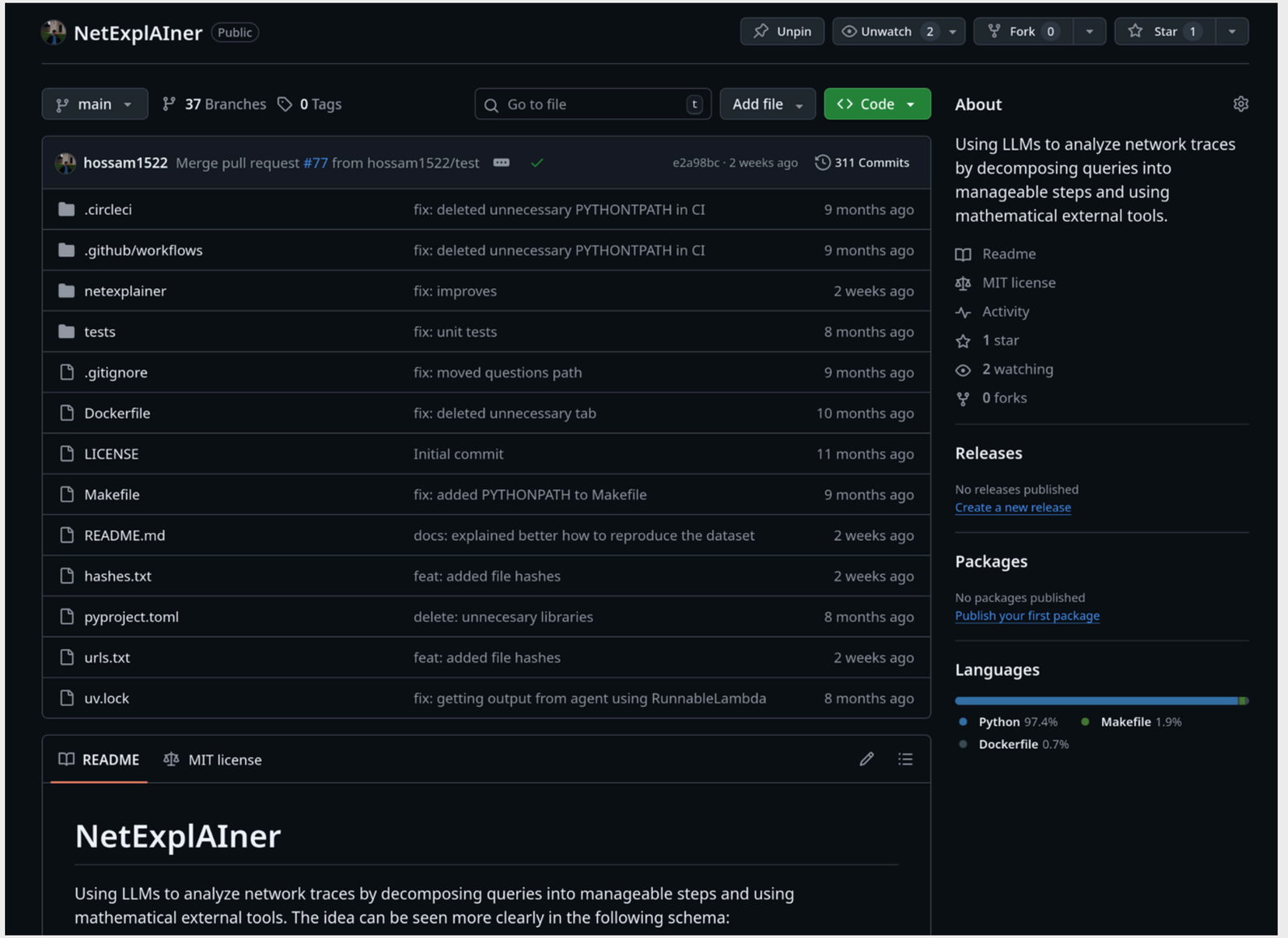
33.6%
Correcto

66.1%
Incorrecto

0.3%
Problemático



https://github.com/hossam1522/NetExplAIner



NETEXPLAINER

Enfoques

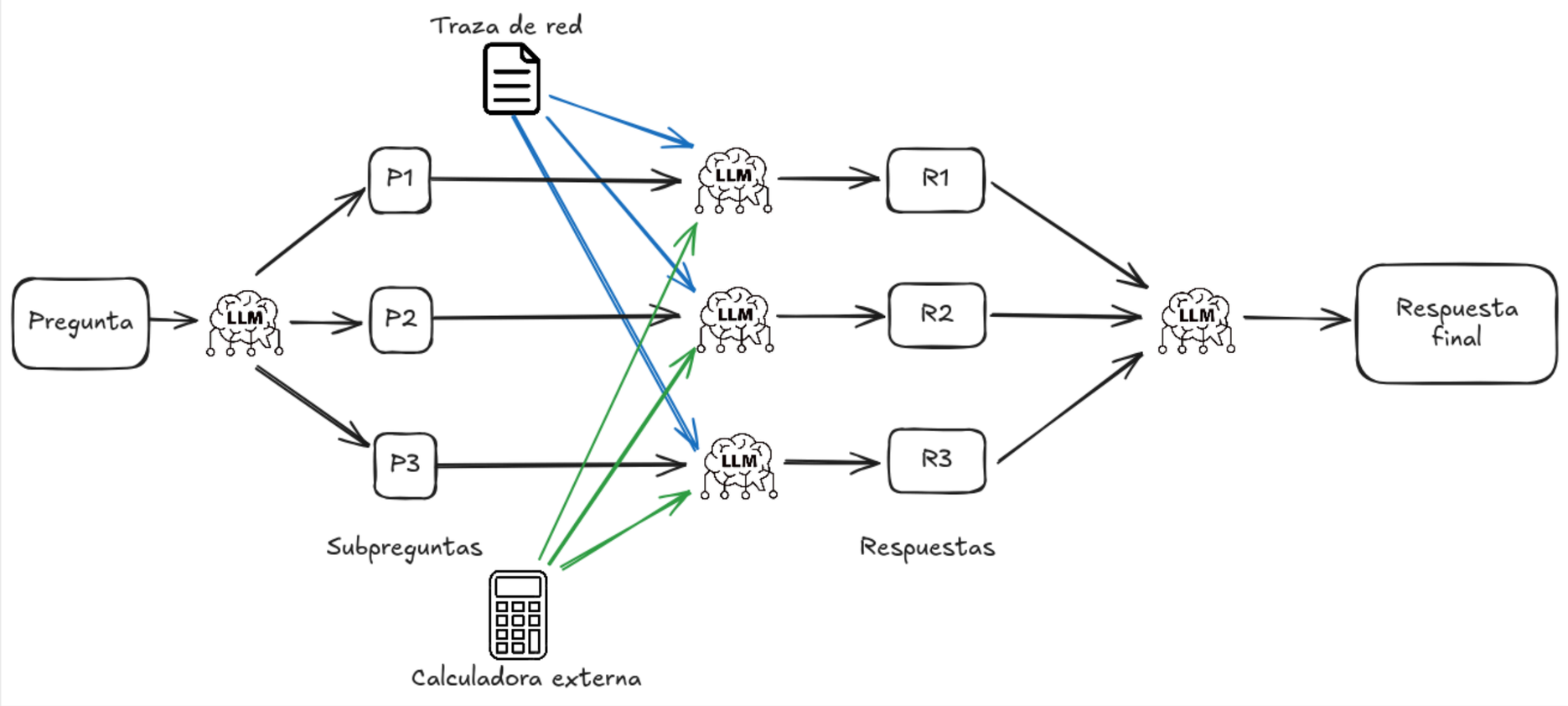
- Descomposición de tareas
  - Enfocado principalmente a tareas complejas.
  - Estrategia prometedora para mejorar rendimiento y fiabilidad.
  - Múltiples beneficios:
    - Centrar poder procesamiento
    - Tareas multipasos
    - Modularización
- Uso de herramientas externas
  - Busca aprovechar precisión herramientas.
  - Beneficios:
    - Reducir alucinaciones.
    - Ampliación de conocimiento.
  - Desventajas:
    - Falta de benchmarks.
    - Falta de conocimiento acerca de cómo utilizar las herramientas.

Trabajo	Identificación de eventos	Análisis de protocolos	Detección de ataques	Visualización de resultados	Evaluación de modelos	Descomp. tareas	Uso herramientas matemáticas
Exploring the use of LLMs to understand network traces	√	√	X	√	√	X	X
Application of Large Language Models to DDoS Attack Detection	√	X	√	X	√	X	X
Unleashing Large Language Models For Dynamic Packet Classification in Software Defined Networks	√	√	√	X	√	X	X
Can LLMs Understand Computer Networks?	√	√	X	√	√	√	X
Leveraging Large Language Models for Network Traffic Analysis	√	√	√	√	√	X	X
Towards Explainable Network Intrusion Detection using Large Language Models	√	√	X	√	√	X	X
Este trabajo	√	√	**X**	√	√	√	√

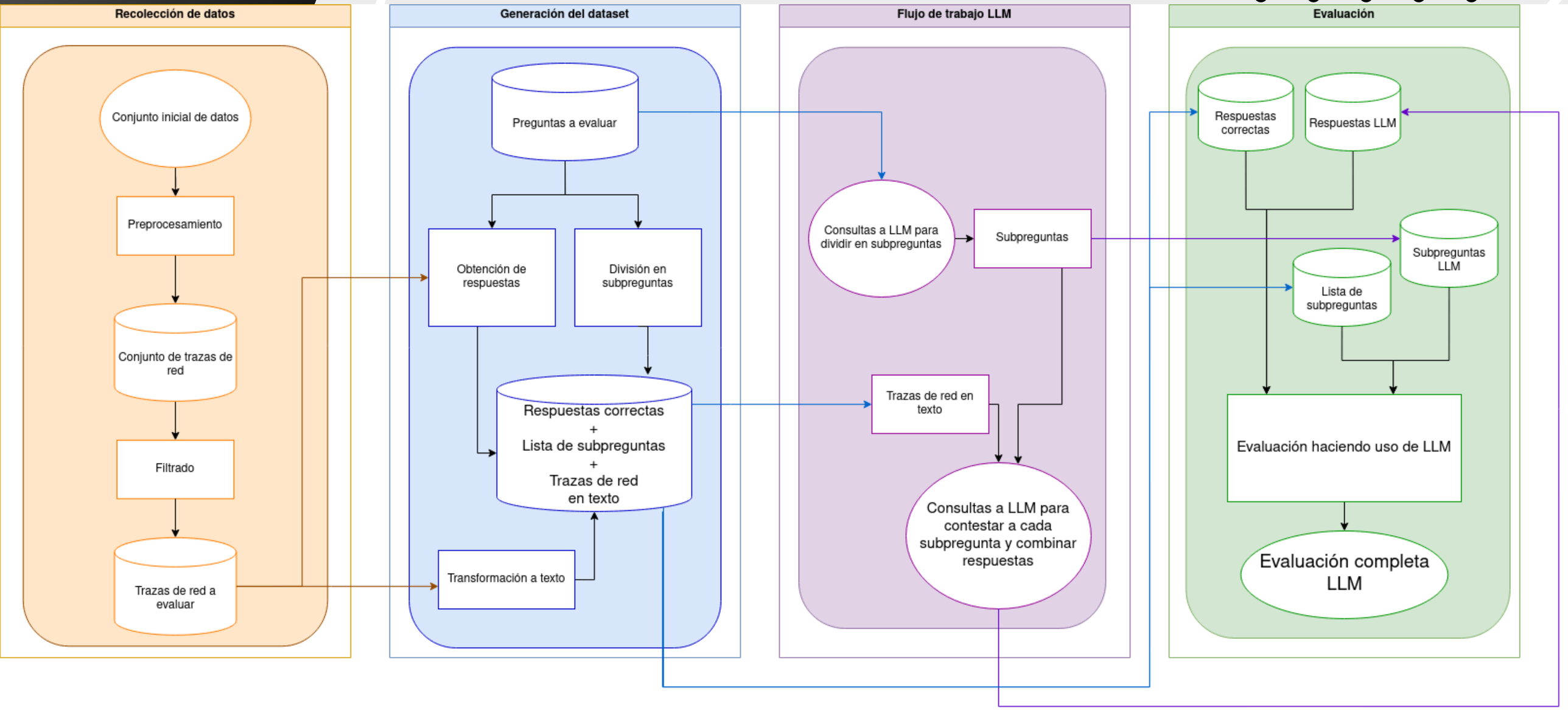
DISEÑO E IMPLEMENTACIÓN

1. ¿Cuál es el número total de paquetes en la traza?
2. ¿Cuántos comunicadores únicos hay en la traza?
3. ¿Cuál es la dirección IP que participa en la mayor cantidad de comunicaciones?
4. ¿Cuál es el tamaño total de bytes transmitidos?
5. ¿Cuál es el tamaño promedio de los paquetes en bytes?
6. ¿Qué protocolo predomina en la captura: ICMP, TCP, o UDP?
7. ¿Cuánto dura la comunicación en segundos?
8. ¿Cuál es el promedio de paquetes enviados por segundo?
9. ¿Cuál es el promedio de bytes enviados por segundo en la comunicación?

DISEÑO E IMPLEMENTACIÓN

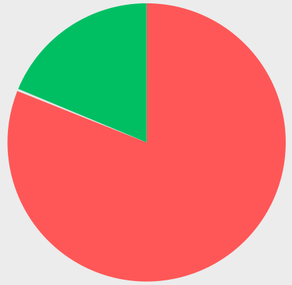


DISEÑO E IMPLEMENTACIÓN



LLAMA2 – 7B

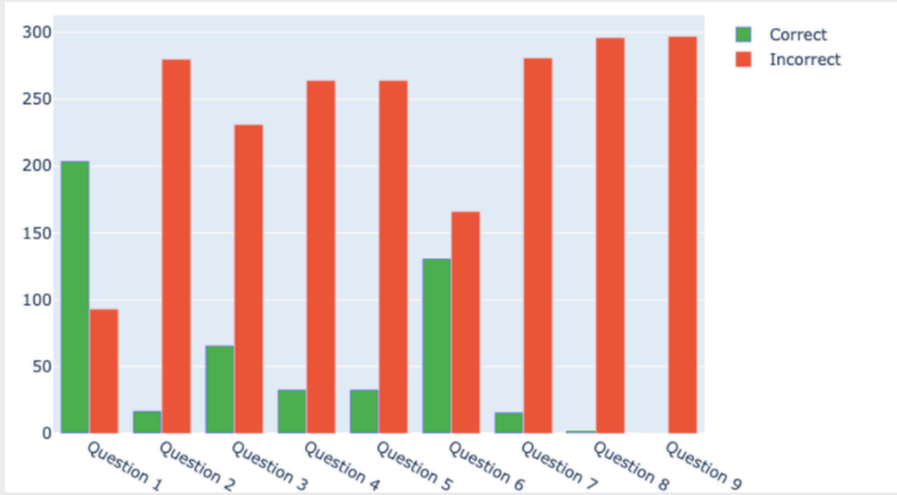
Original



18.7%
Correcto

81%
Incorrecto

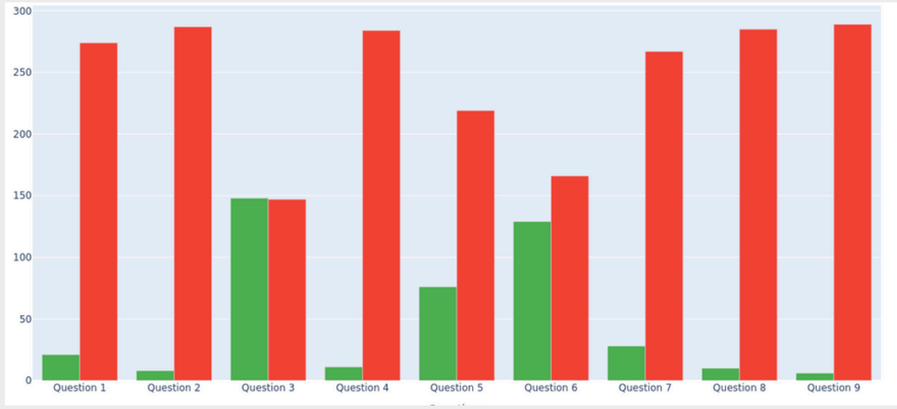
0.262%
Problemático



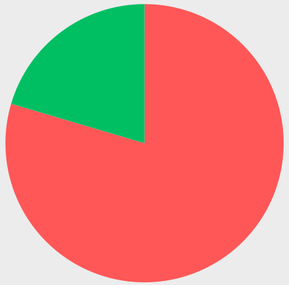
División
(sin calculadora)

16.5%
Correcto

83.5%
Incorrecto

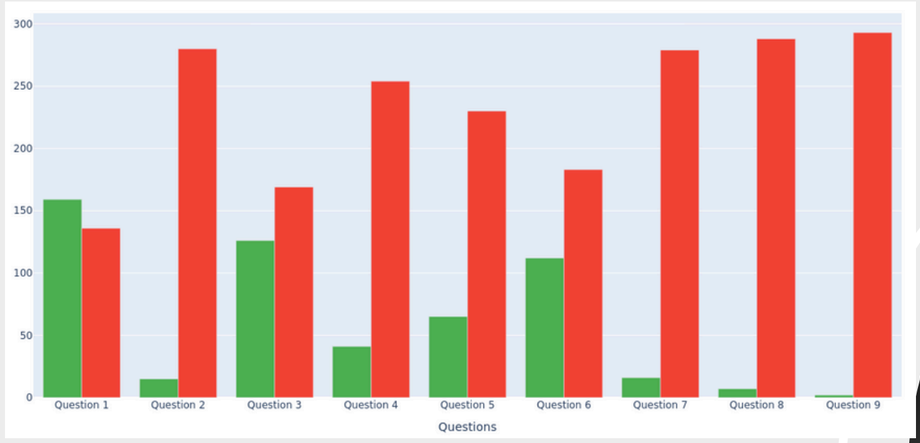


Mixto
(sin calculadora)



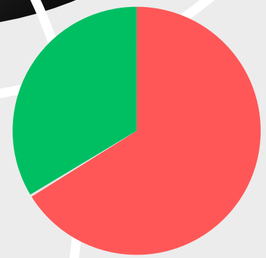
20.4%
Correcto

79.6%
Incorrecto



MISTRAL – 7B

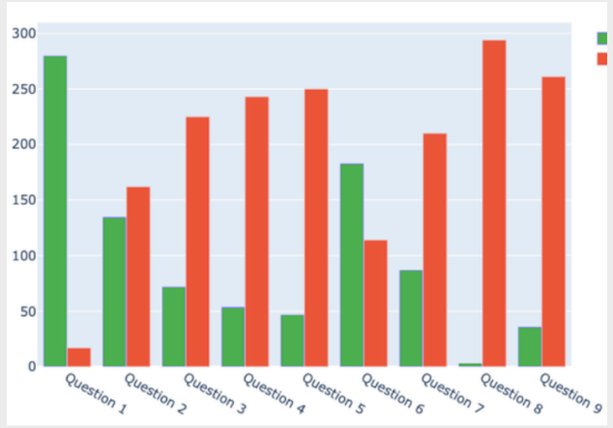
Original



33.6%
Correcto

66.1%
Incorrecto

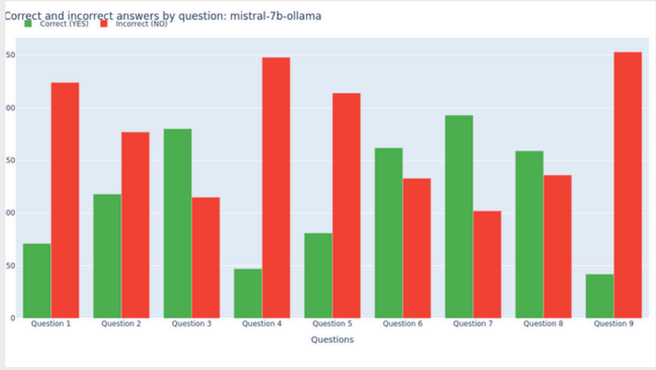
0.3%
Problemático



División (sin calculadora)

39.7%
Correcto

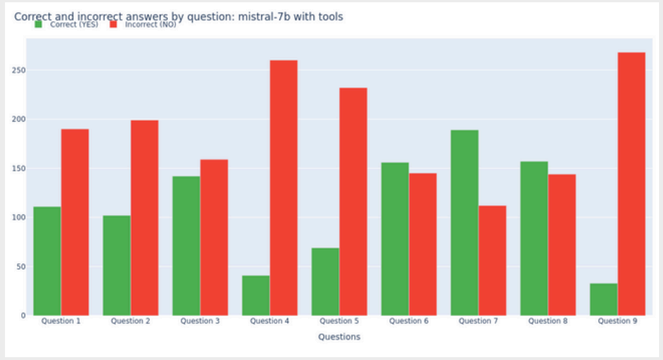
60.3%
Incorrecto



División (con calculadora)

36.9%
Correcto

63.1%
Incorrecto



Mixto (sin calculadora)

47.7%
Correcto

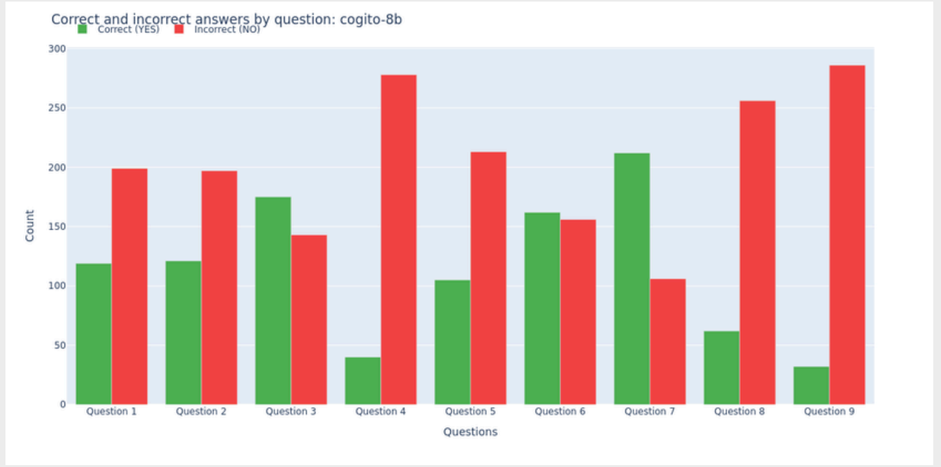
52.3%
Incorrecto

COGITO – 8B

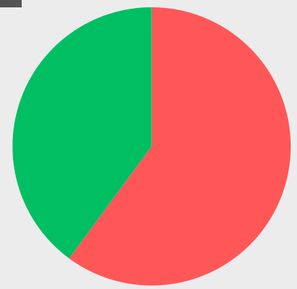
División (sin calculadora)

35.9% Correcto

64.1% Incorrecto

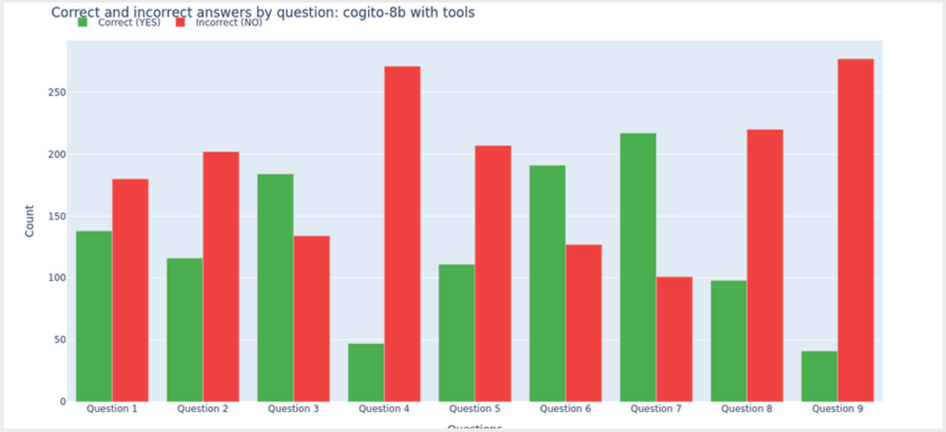


División (con calculadora)



39.9% Correcto

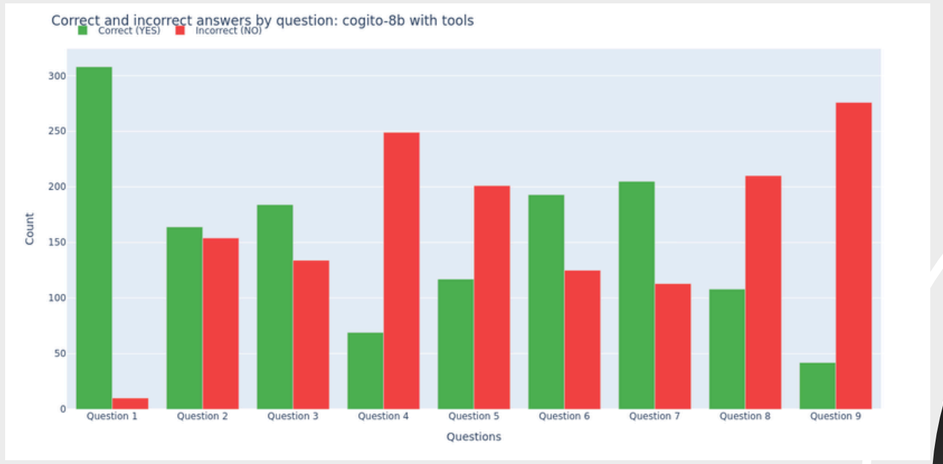
60.1% Incorrecto



Mixto (con calculadora)

48.6% Correcto

51.4% Incorrecto



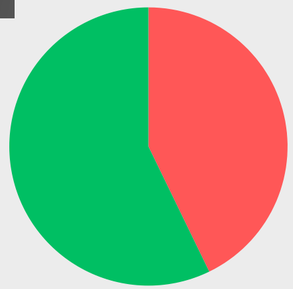
QWEN2.5 - 32B

División (sin calculadora)

53.1% Correcto

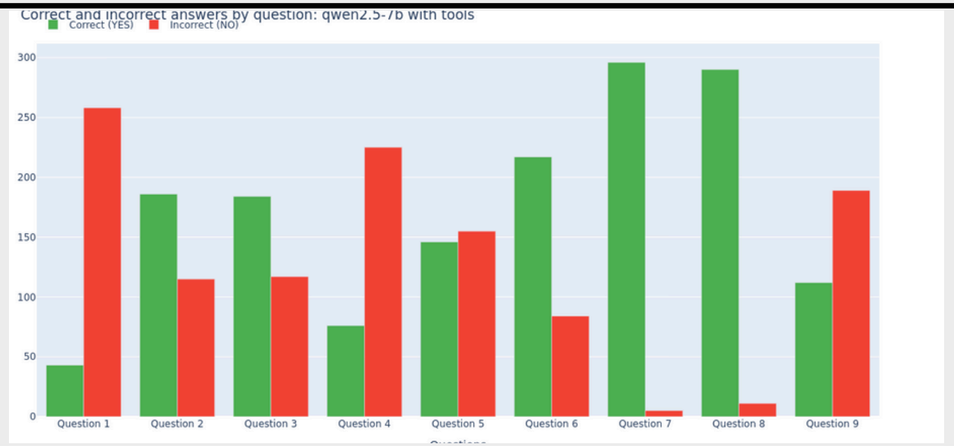
46.9% Incorrecto

División (con calculadora)



57.2% Correcto

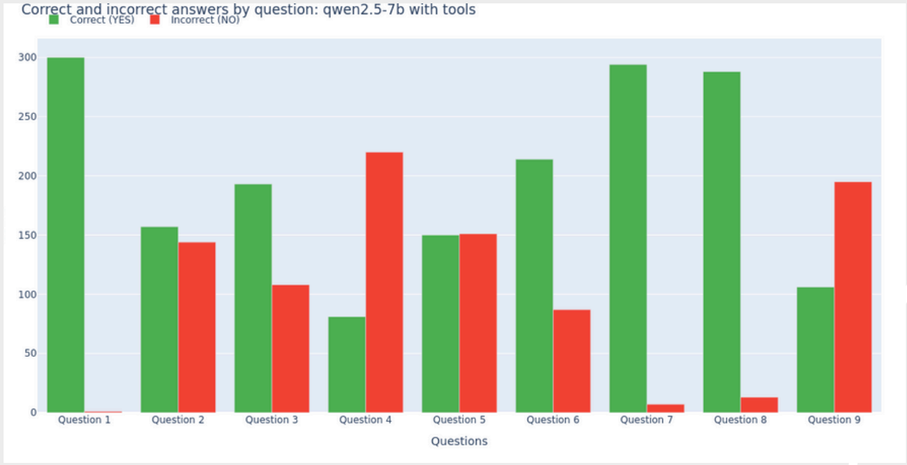
42.8% Incorrecto



Mixto (con calculadora)

65.8% Correcto

34.2% Incorrecto

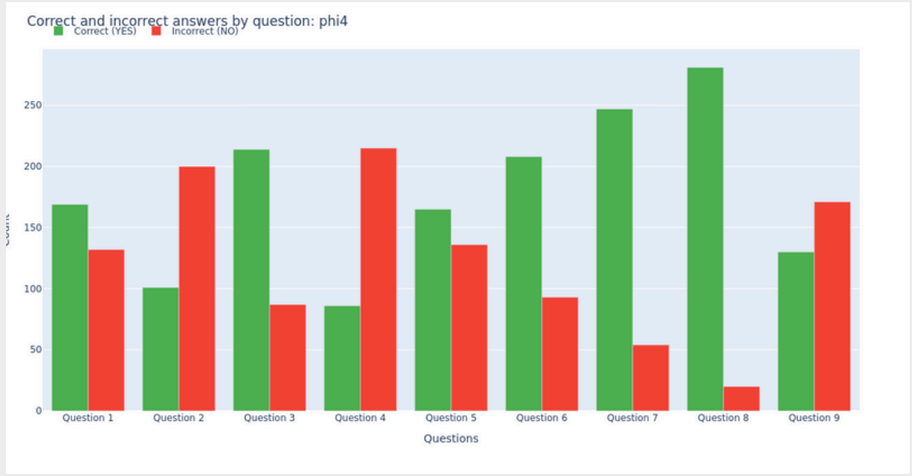


PHI4 – 14B

División (sin calculadora)

59.1%
Correcto

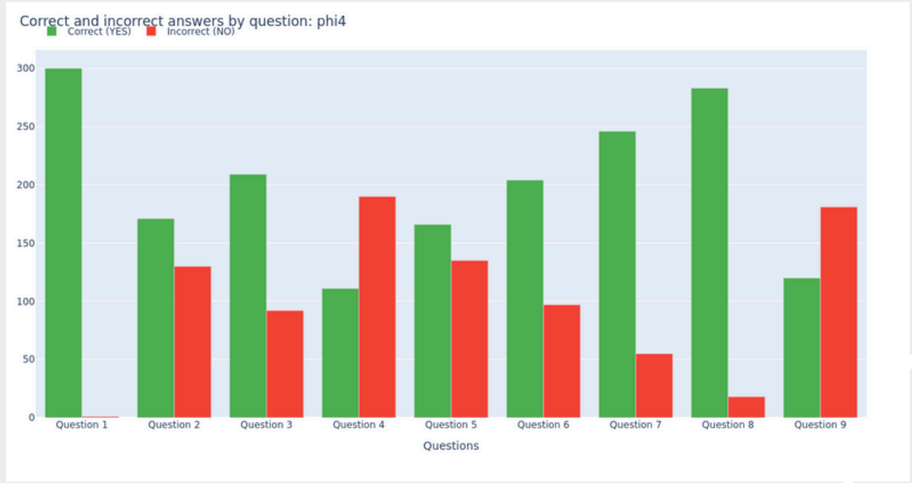
40.9%
Incorrecto



Mixto (sin calculadora)

66.8%
Correcto

33.2%
Incorrecto



RESULTADOS FINALES

Modelo	Resultados originales	División sin herramientas	División con herramientas	Enfoque mixto sin herramientas	Enfoque mixto con herramientas
Llama2	18.74 %	16.46 %	—	20.45 %	—
Mistral 7B	33.56 %	39.66 %	36.91 %	47.73 %	—
Cogito 8B	—	35.92 %	39.94 %	—	48.57 %
Qwen2.5 32B	—	53.08 %	57.22 %	—	**65.82 %**
Phi4	—	59.10 %	—	**66.81 %**	—

CONCLUSIONES

- Descomposición de tareas
  - Poca utilidad para modelos débiles.
    - Mala división.
  - Contraproducente para preguntas simples.
  - Útil para preguntas más complejas.
- Uso de herramienta matemática externa
  - Si el modelo no sabe utilizarla, resulta ser contraproducente.
  - A pesar de que hay mejora, no es muy notable.
  - Problemas especificados en estado del arte.
- Enfoque mixto
  - El más prometedor:
    - Sin división para las preguntas simples.
    - Con división para las preguntas complejas.
  - Los mejores resultados.
- Uso de LLMs
  - Buenos resultados
  - Campo de estudio bastante prometedor
  - Potencial para superar desafíos de los SOC

LÍNEAS DE TRABAJO FUTURO

- Mejorar desempeño de los modelos a la hora de usar herramientas.
- Pruebas en entornos más realistas.
- Evaluación de modelos más grandes, actuales, o menos cuantizados.
- Prueba de nuevos enfoques:
  - Fine-tuning.
  - Retrieval-Augmented Generation (RAG).
  - Modelos con razonamiento.
  - etc.

PyData Granada

GRACIAS

¿Alguna duda?

LinkedIn