Knowledge Grounding in Retrieval-Augmented LMs Quick Update

Martin Fixman et. al

City, University of London

30 August 2024

New questions

More questions were added, including:

- What was the duration of {`historical_event`}?
- What are the dimensions of {`painting`}?
- What is the melting point of {`element`}?
- What's the main nationality of {`person`}?

New Questions

Category	Questions		Objects		Total
book	5	+	45	=	225
city	7	+	60	=	420
element	5	+	35	=	175
historical_event	3	+	56	=	168
painting	7	+	39	=	273
person	7	+	47	=	329
principle	5	+	30	=	150
Total	39		312		2040

Better calculation of probabilities

Now we properly use[1] perplexity rather than average-of-token-probabilities.

$$\text{CE} = -\frac{1}{N}\sum \log P(x_i)$$
$$\text{Perplexity} = e^{\text{CE}}$$

[1] Due to a last-minute problem with precision on some large answers, today's plots will still have average-of-token-probabilities. Sorry!

Better calculation of probabilities

Now we properly use[1] perplexity rather than average-of-token-probabilities.

$$\text{CE} = -\frac{1}{N}\sum \log P(x_i)$$
$$\text{Perplexity} = e^{\text{CE}}$$

.

[1] Due to a last-minute problem with precision on some large answers, today's plots will still have average-of-token-probabilities. Sorry!

Different flavour of Llama

Part of the hypothesis[2].

Will a larger model tend to prefer parametric answers?

I'm not using two models of Llama:

`meta-llama/Meta-Llama-3.1-8B-Instruct`

`meta-llama/Meta-Llama-3.1-70B-Instruct`

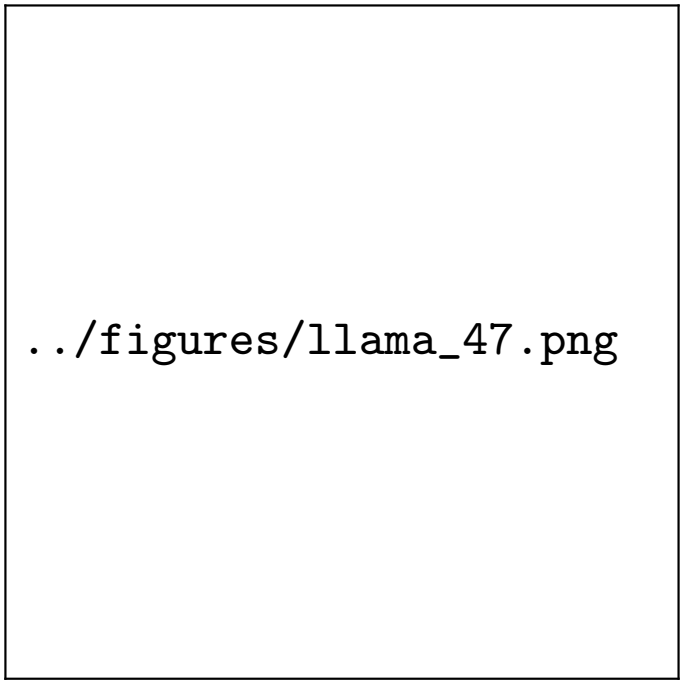
More models will come soon!

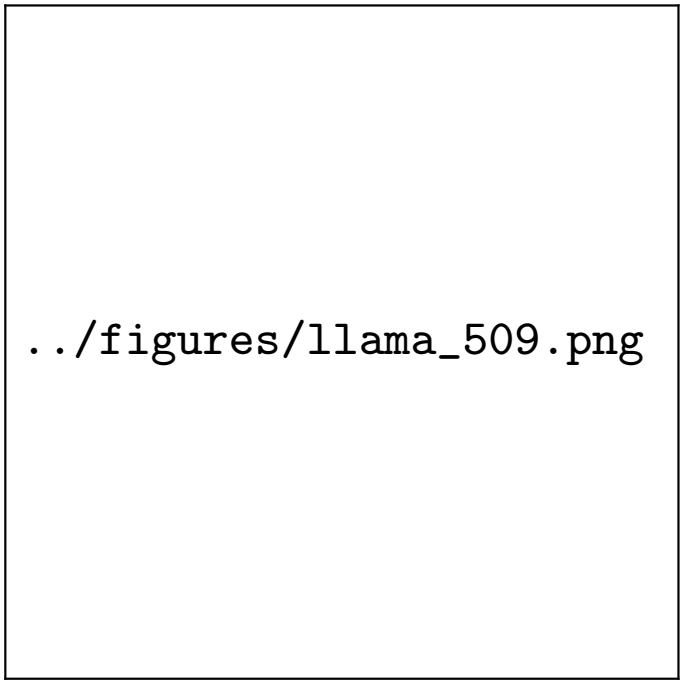
[2] The full hypothesis will come later.

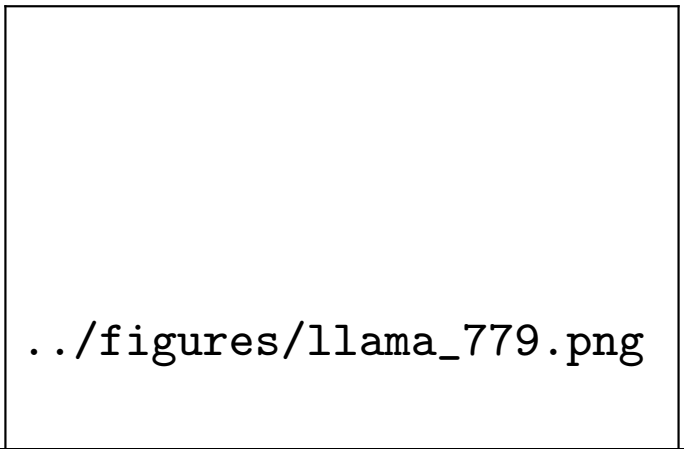
Calculating probabilities for both parametric and counterfactual answer, regardless of which one gets answered.

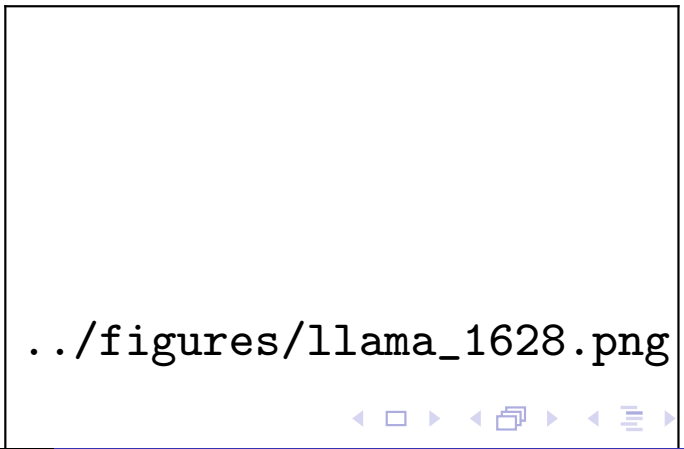
Some interesting results

Larger models tend to have a higher chance of choosing a parametric answer when given counterfactual information.









Some interesting results

Certain counterfactual answers have low probability, even when being added as part of the context.

Next Steps

Very important: **Define a hypothesis!**
A lot more!