Internship weekly report

Mathis Embit

MILES, LAMSADE

June 17, 2024

Reading

[Yad+24] tries to disantangle aleatoric (multiple correct answers) and epistemic (when the model lacks of knowledge) uncertainty. They suppose that when the answer is highly dependent on the context, it means that the model is not sure of its knoledge and therefore that the epistemic uncertainty is high. They do this formally using mutual information.

Idea

Unrelated idea:
Rather than projecting like

$$\underset{e \in W_E}{\arg\min}\, d(x, e)$$

we may do

$$\underset{e \in W_E}{\arg\min}\, D(f(x), f(e))$$

where f is the output of the LLM, for example the probabilities of generating the target, and D is a divergence.

[Yad+24]   Yasin Abbasi Yadkori et al. *To Believe or Not to Believe Your LLM*. 2024. arXiv: 2406.02543.