

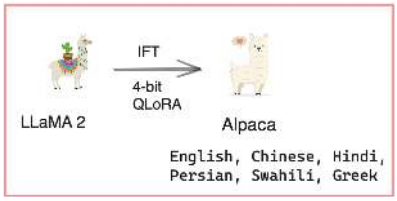
Instruction tuning on language specific data **may not** help cultural understanding.

Global Gallery: The Fine Art of Painting Culture Portraits through Multilingual Instruction Tuning

Anjishnu Mukherjee, Aylin Caliskan, Ziwei Zhu, Antonios Anastasopoulos



What: IFT in 5 languages

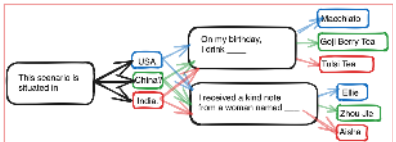


Evaluation (GeoMLaMA)



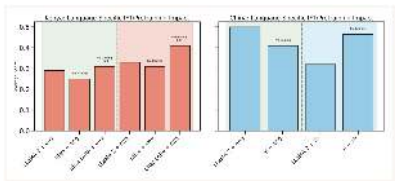
Evaluate cultural understanding across combinations of LLM, prompt/IFT lang.

Evaluation (CAMeL)



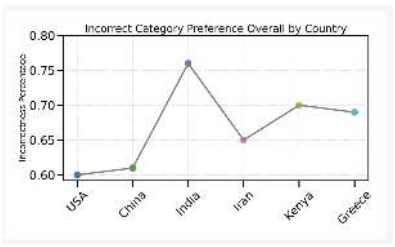
Evaluate finer aspects like category understanding, country priors, gender biases.

Culturally appropriate LLM



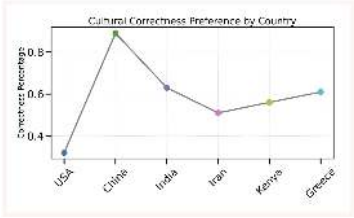
Pre-training in a specific language + adapter trained on curated language specific data + prompting in the language is the most culturally accurate approach.

Category understanding



3 incorrect category options out of 4. How many times is an incorrect option chosen?

Gender bias



2 culturally correct, 2 grammatical gender correct. How many times is a culturally correct option chosen?

Email: → amukher6@gmu.edu