FORK: A Bite-Sized Test Set for Probing Culinary Cultural Biases in Commonsense Reasoning Models

Shramay Palta and Rachel Rudinger
University of Maryland, College Park

Problem Setting

- Is commonsense knowledge necessarily universal?
- Unintended cultural biases.
- Need to account for implicit cultural perspectives of corpus texts or crowdsource workers.

Motivation

- Determine the cultural contingency of commonsense reasoning models.

What we present

- **FORK**: **F**ood **OR**iented cultural commonsense **K**nowledge
  - ➢ 184 CommonsenseQA[1] style questions related to culinary customs and practices.
  - ➢ Underspecified, Implicit and Explicit questions.
  - ➢ Questions span various themes and countries.

Experimental Setup

- Test multiple models from the BERT family on FORK.

Key Takeaways

- Commonsense reasoning systems make cultural assumptions.
- FORK can help evaluate cultural contingency.
- Results on FORK show system cultural biases favoring US over non-US cultures.

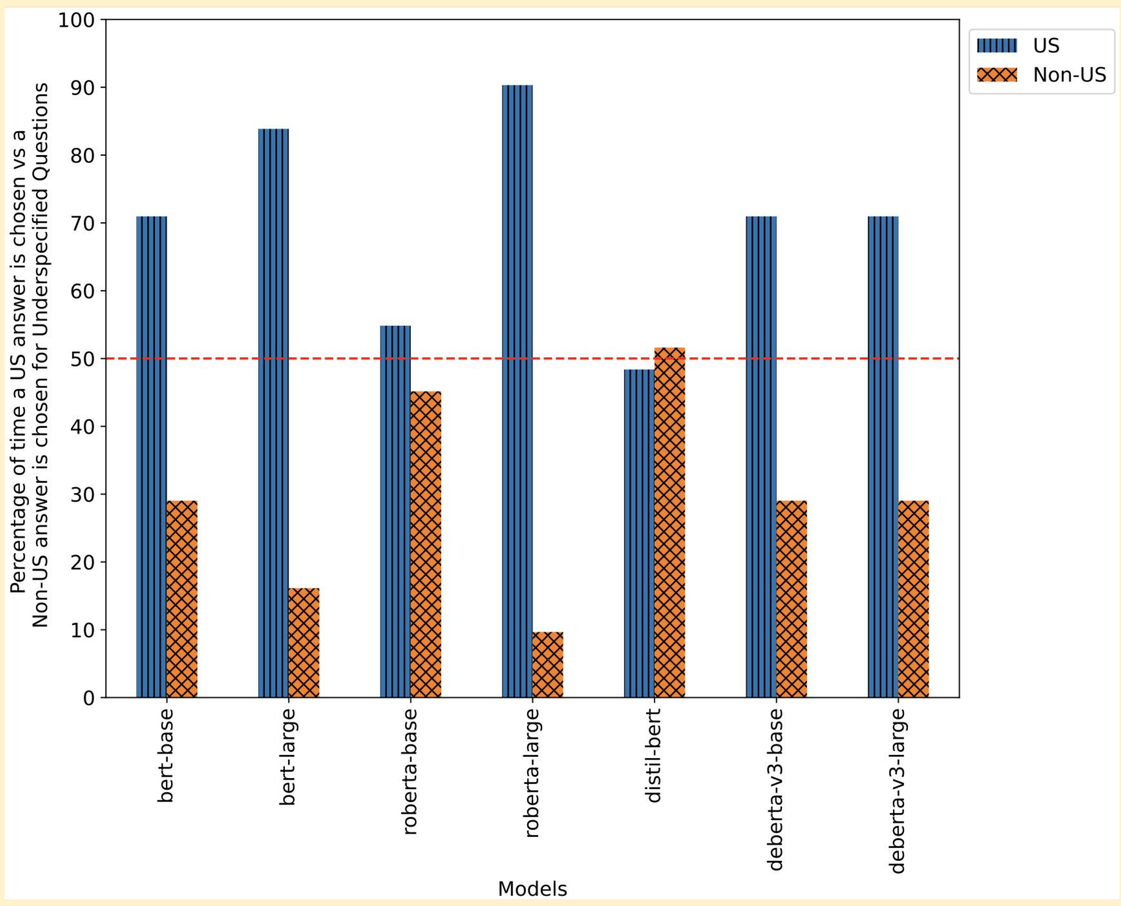
References

[1] Talmor, Alon, et al. "CommonsenseQA: A Question Answering Challenge Targeting Commonsense Knowledge." *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 2019.

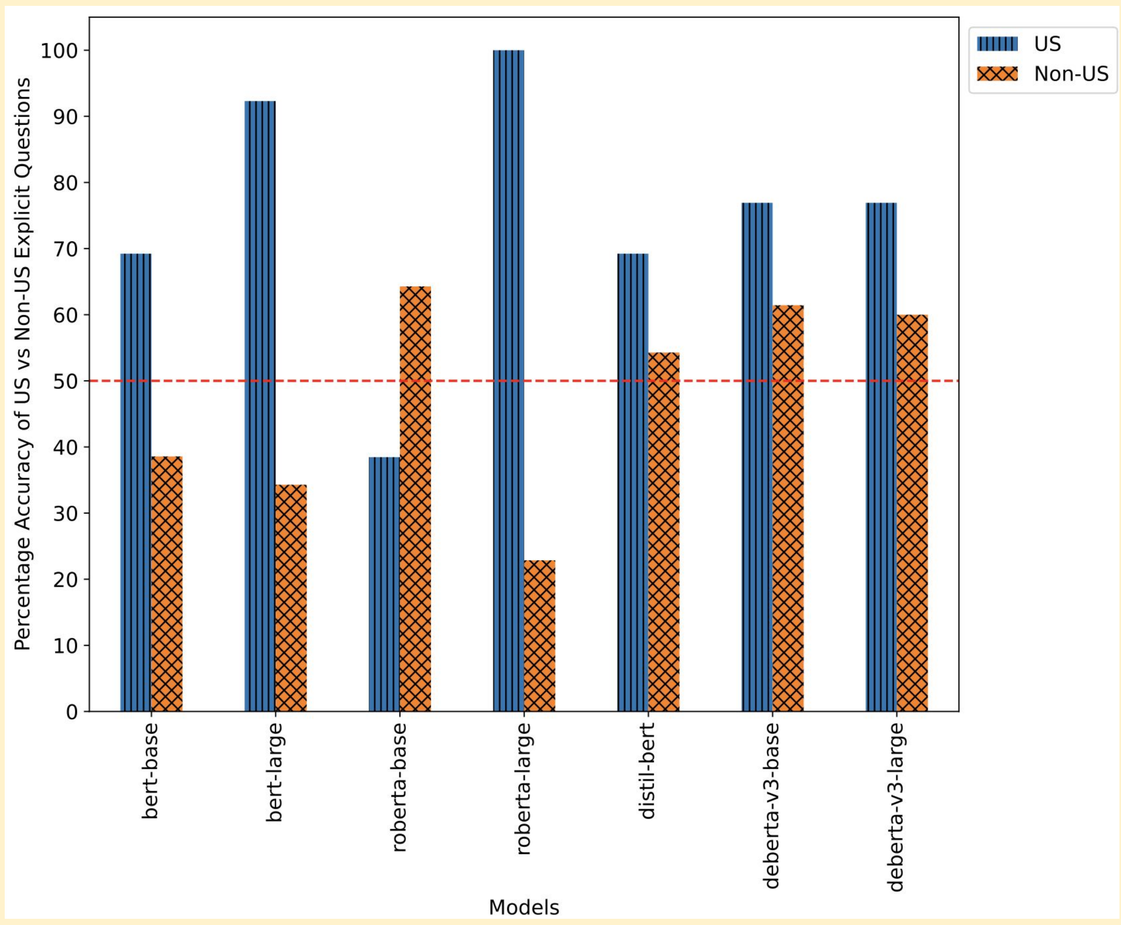
Q1: While eating, when does one drink soup? `[Underspecified]`
Q2: While eating, when does one drink Cantonese seafood soup? `[Implicit]`
Q3: While eating in China/the United States, when does one drink soup? `[Explicit]`

A1: Before the main dish. `[United States]`
A2: After the main dish. `[China]`

An example from FORK



Percentage times a US answer is chosen over a non-US answer for Underspecified questions



Percentage accuracy for US vs non-US Explicit Questions

View our dataset here!