

Claire, Ian, Em

Peer review

Intro:

- Consider adding a description of your simple linguistic features.
 - This note depends on your intended audience: are you targeting people who already have a linguistic background? If not, maybe consider:
 - Examples of hedging words
 - Examples of certainty markers
 - What is the relation of pronoun usage to certainty?
 - What is the relation of sentence length to certainty?

Methods:

- Add an example or two of a question and response, so the reader can understand the dataset better
 - This is again to the point of considering your audience, though this would be helpful regardless
 - Maybe clarify what determines which questions fall into which category for overlapping categories? For instance, there is a lot of conceptual overlap between “superstition” and “paranormal”.
- “This complete dataset preservation ensured maximum statistical power while maintaining data integrity.” feels like an unnecessary sentence, although perhaps I am just confused by the phrase “maximum statistical power”?
- Maybe format your feature definitions as a table? (just so it isn’t a huge block of text, but this is definitely a preference thing)
- Use some of your pretty graphs in the paper (you worked hard on them!)
 - Add alt text

Results:

- Clarify the religion sentence** (religion, as far as I am aware, tends to not have “agreed upon” facts), perhaps define some of these question categories

Discussion:

- Consider talking about the implications of using a dataset on hugging face

Overall:

- Precise & Consistent style throughout the piece!
- Methodology seems appropriate & well-thought out
- Formatting is consistent, could potentially be cleaner?
 - consider using varying header sizes to demonstrate section breaks, rather than just bold/bold+underline. This is purely stylistic/personal choice however!
 - The bolded sub-headers within each section are great organizationally, and keep the piece very clear
- Something to consider:
The register used throughout this paper is a great example of something that communicates certainty/truthful authority. Your analysis not only has implications for LLM design, but also is well situated in a broader discussion regarding academia and

misinformation (specifically, demonstrating in line with previous literature that surface-level cues play a major role in how humans determine credibility). This is not to say that this type of discussion necessarily belongs in your paper – the strict & precise scope/focus is a strength – but it may be relevant (or, I hope, at least interesting!) when writing to think about how the work speaks to your representation of it.