

Discussion

Study 1 (Automated scoring of communication skills in physician-patient interaction)

- The study uses automated linguistic annotation to evaluate physician–patient communication. **What other real-world domains could benefit from automatically annotated communication or interaction features?**
- If an automated scoring system were used for high-stakes assessment (e.g., medical licensing, teaching evaluation), **what ethical or fairness concerns should we consider?**
- Automated annotation greatly increases scalability. **What trade-offs might exist between scalability, interpretability, and validity of the scores?**

Discussion

Study 2 (Restaurant reviews)

- The paper shows that naive word-count methods cannot distinguish contextually different meanings (e.g., *a stinky restaurant* vs. *the stinky tofu*). **Can you think of other examples where the same word appears in two different contexts but should NOT be interpreted the same way?**
 - What does this tell us about the limitations of word-count-based analysis?
 - What advantages does LLMs bring for real-world applications (e.g., sentiment analysis, communication scoring, language assessment)?