

---

Q 1. What do you think about this weeks topic?

For me the most important point that was brought up is that SOTA models might be topping all the leaderboards but be barely usable in the real world. Even though I don't think every leaderboard should be based on energy consumption, latency , size and so on, those are still very important features of a product for any end-user or even researcher trying to replicate results.

Q 2. What is still not clear?

Why don't people just introduce new leaderboards that are similar to the established ones but either only accept models that are smaller than a certain size or factor in other attributes of the model (like size, energy consumption, training time, etc.)

Q 3. Why is this topic relevant to the NLP community?

I am not sure if that is the case but I imagine the issue is similar in other machine learning fields with Performance being the deciding factor while other aspects are being neglected. This is more of a general problem with there being too little focus on accessablity, fairness and ecological impact.

Q 4. What have you learned that you can apply on your current or future works?

The issues mentioned in the paper are mostly refering to huge pre-trained models which I will probably not be building/training in the near future. If it does come to that I will be sure to also mention other attributes of the model beyond its accuracy on some evaluation set.