

Feedback:

I would say that in my opinion your aim is more clear the way it's written in the "Background and problem statement" section rather than the "official" version, but for the rest the objectives seem to be outlining the path you have to follow in order to reach your goal.

I just cannot understand practically how are you going to find those models that will help you to identify toxic comments, so I'm not able to evaluate effectively how feasible is your purpose, but maybe this is not the point of this first presentation and you can explain this later.

=> how to find models?

Since you mentioned deep learning and conversation AI, it would have been good to present more references regarding artificial intelligence even if they are just potential references.

=> more references

First things first, you've cited a lot of material from pretty decent sources. However, instead of citing 4 sources on the impact of toxicity, you could have maybe focused more on the computer science part which is the subject, ultimately. You could have sold the importance of your topic with even one citation.

=> maybe less citations in intro?

Second point is also a detail, and maybe up for debate, but I think you could have swapped one of your objectives with the aim. Namely the ambitious but rather non-actionable objective, "achieve state-of-the-art on Conversation AI's public dataset". Since you can aim for the stars, and actually "investigating text classification techniques" would be a reasonable objective. Put it this way: If you were a project manager, you could ask someone to investigate text classification techniques but you can't really tell them to merely achieve success. Success is the dream end state here, and it's all fine, but it's not a task itself. So, in my opinion, those two could swap places.

=> drop state of the art?

The last and perhaps most substantial point is about feasibility and focus. So, are your goals and methods feasible and focused enough? Perhaps you know it is, in which case you could elaborate on how. For instance, perhaps detecting certain covert bullying would be extremely hard programmatically, so you might constrain the problem space to e.g. overtly offensive language. Furthermore, coming back to the aims and objectives, I don't think it's necessary in our context to "achieve state-of-the-art on Conversation AI's public dataset" since that might not even be feasible given how many massive platforms with a lot more resources deal with this same issue, I think it would be fine to just find the best option out of what you currently have available.

=> focus on feasibility

I am quite confused about the aim, whether it will only detect the toxic texts or prevent it also. To avoid cyber bullying it's a step or a complete solution to the problem. The problem works with the language, do you consider about different types of language? or you are focusing on a specific language.

=> add language. Explain detection not prevention

## Page Break

"Explore various text-classification techniques on the task of identifying toxic comments online." Does this mean that you will be testing various methods and comparing their fitness to others or that you will be researching what techniques other people have been using on their research and comparing other research to other and trying to find some conclusion on in.

=> clarify questions/question verbs

Info needed:

Recent advances in text-classification, data-set, literature on how to perform this task (how to choose model etc).