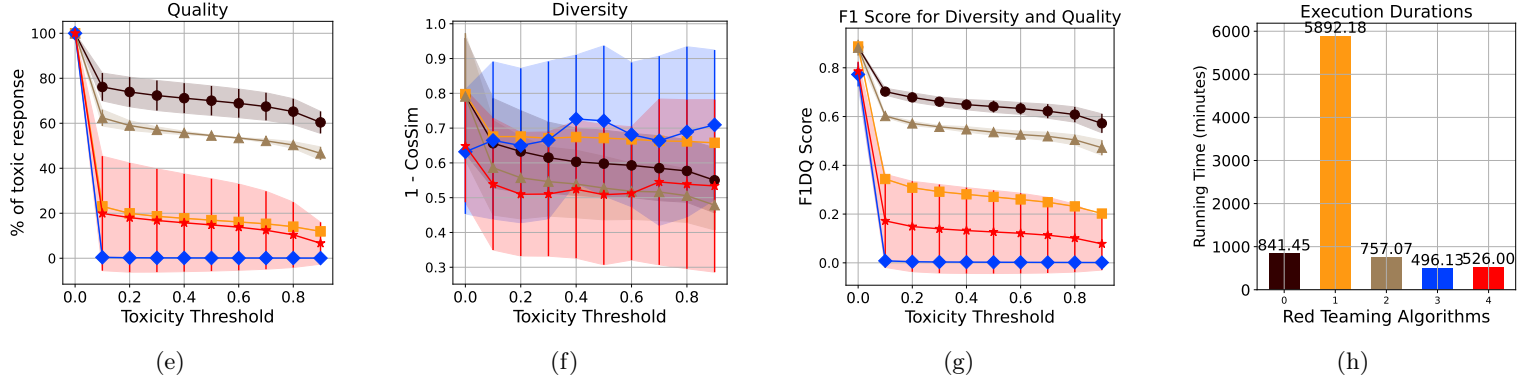


GPT-2 targets GTP-2 fine-tuned on the Alpaca dataset.



GPT-2 targets Dolly-3B finetuned with instruction task on the Databrick dataset.

Figure 1: Ablations of K-subset training.

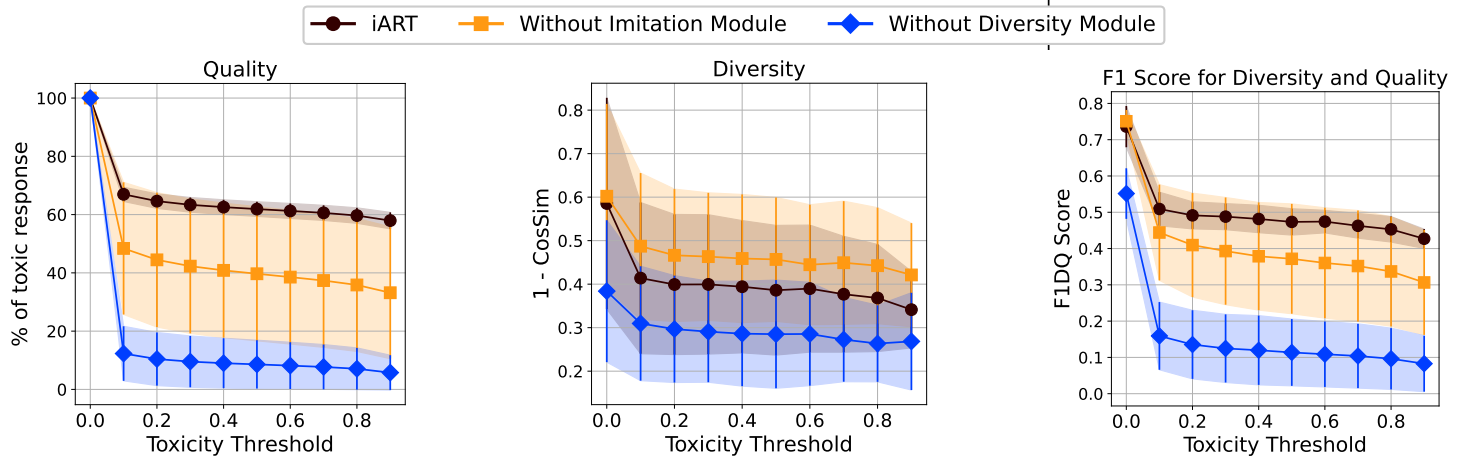


Figure 2: Ablation of imitation and diversity modules.