<https://levelup.gitconnected.com/credit-card-fraud-detection-using-generative-ai-78824bf62717>

* Hope you are doing well. Thanks for all videos on Langchain and LLM, they are very helpful.
* I ahve a question on Constitutional AI , when you refer to create a new data set to retrain LLM in your video <https://www.youtube.com/watch?v=uoVqNFDwpX4&ab_channel=SamWitteveen>
* You mean with Propmt Template as below ?
* template = """
* Please use the following context to answer questions.
* Context: {context}
* ---
* Question: {question}
* Answer: Let's think step by step."""
* prompt = PromptTemplate(template=template, input\_variables=["context", "question"]).partial(context=context)
* Or use RL-AIF technique ?
* Do you have any links on using Langchain for RL-AIF?
* Thanks,
* Geeta
* THURSDAYSam Witteveen sent the following message at 6:16 PM [View Sam’s profile](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

[](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

* [Sam Witteveen](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw) 6:16 PM Hi Greeta,
* I am talking about using RLAIF to make a reward model and use that to make a whole new dataset. Both Anthropic and OpenAI are doing things like this for making instruct datasets and it is a very useful tool for making datasets for brands etc. So the data that you finetune a model on is "on brand".
* Hope that helps.
* Sam
* Geeta Desai sent the following message at 8:06 PM [View Geeta’s profile](https://www.linkedin.com/in/ACoAACG2vUsBBSOE9divGCVsz7bo6WD_4mHPoew)

[](https://www.linkedin.com/in/ACoAACG2vUsBBSOE9divGCVsz7bo6WD_4mHPoew)

* [Geeta Desai](https://www.linkedin.com/in/ACoAACG2vUsBBSOE9divGCVsz7bo6WD_4mHPoew) 8:06 PM Thanks for your repsonse. Data generation is the part I am confused. Generally we will train reward model with supervised data and then use PPO to train original model.
* reward\_model = RewardModel(model, tokenizer)
* comparison\_dataset = load\_dataset("allenai/scitldr")
* # Train the reward model
* reward\_trainer = trlX.RewardModelTrainer(
* reward\_model=reward\_model,
* train\_dataset=comparison\_dataset,
* train\_batch\_size=8,
* )
* reward\_trainer.train()
* How can data be generated with reward model?
* Thanks,
* Geeta
* TODAYSam Witteveen sent the following message at 5:29 AM [View Sam’s profile](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

[](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

* [Sam Witteveen](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw) 5:29 AM ok sorry reading it back. I am being confusing. You need to make a dataset to train your reward model this where the constitution prompts and rewriting prompts come in. You can also use the output of those prompting as a dataset for SFT. This can be (but not always ) used as what you are calling the "comparison\_dataset" in the trlX code above. This can be done to save on time and compute. A lot of this depends on how much data you have and how much compute you have access to. What are you trying to achieve overall?
* "Generally we will train reward model with supervised data" - are you getting this data from a CAI process?
* Geeta Desai sent the following message at 7:03 AM [View Geeta’s profile](https://www.linkedin.com/in/ACoAACG2vUsBBSOE9divGCVsz7bo6WD_4mHPoew)

[](https://www.linkedin.com/in/ACoAACG2vUsBBSOE9divGCVsz7bo6WD_4mHPoew)

* [Geeta Desai](https://www.linkedin.com/in/ACoAACG2vUsBBSOE9divGCVsz7bo6WD_4mHPoew) 7:03 AM Yes I am trying to get training data for reward model from CAI, but this is not the entire training data. E.g Text summarizatio on some documents is getting created ,I need to apply some rules on that .So I need to train reward model wth constitutional policy but data fro training base model is  those original documents.
* So how should I go about creating architecure for this?
* Sam Witteveen sent the following messages at 7:12 AM [View Sam’s profile](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

[](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

* [Sam Witteveen](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw) 7:12 AM Ok so there are multiple ways to do this. One way would be to use the CAI process to make a dataset of good outputs. (keep the rejected outputs for reward model training and if you are getting scores from the CAI model keep those) train a SFT on the good model outputs to have as a benchmark. Then make your reward model with all the outputs and then if needs be train with the RL and reward model. I would use the reward model to score the SFT trained model first to see how well that went.
* [View Sam’s profile](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

[](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

* [Sam Witteveen](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw) 7:14 AM "E.g Text summarization on some documents is getting created ,I need to apply some rules on that ." find a way to get these into your Reward model so it can score everything
* [View Sam’s profile](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

[](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw)

* [Sam Witteveen](https://www.linkedin.com/in/ACoAAABliokBds48A_IRkME2HtNPvFsL5Yb76kw) 7:18 AM What model are you using to do the evaluations for CAI?

I was not plannign to do scoring. Overall Prompt process generated

**Maximize compose field**

**Attach an image to your conversation with Sam Witteveen**

**Attach a file to your conversation with Sam Witteveen**