Graded Assignment

Week 3 Quiz

Review Learning Objectives

coach Beta

Ready to review what you've learned before starting the assignment? I'm here to help.

Help me practice Let's chat

Assignment details

Due Jul 21, 11:59 PM +07

Attempts Unlimited

Try again

Submitted Jun 29, 8:33 PM +07

Your grade

To pass you need at least 80%. We keep your highest score.

90%

View submission See feedback

Like Dislike Report an issue

1. Which of the following are true in regards to Constitutional AI? Select all that apply. **1 point**

- [x] In Constitutional AI, we train a model to choose between different responses.
- [x] To obtain revised answers for possible harmful prompts, we need to go through a Critique and Revision process.
- [ ] For constitutional AI, it is necessary to provide human feedback to guide the revisions.
- [x] Red Teaming is the process of eliciting undesirable responses by interacting with a model.

2. What does the "Proximal" in Proximal Policy Optimization refer to? **1 point**

- ( ) The algorithm's ability to handle proximal policies.
- ( ) The algorithm's proximity to the optimal policy
- (•) The constraint that limits the distance between the new and old policy
- ( ) The use of a proximal gradient descent algorithm

3. "You can use an algorithm other than Proximal Policy Optimization to update the model weights during RLHF." **1 point**

Is this true or false?

- (•) True
- ( ) False

4. In reinforcement learning, particularly with the Proximal Policy Optimization (PPO) algorithm, what is the role of KL-Divergence? Select all that apply. **1 point**

- [ ] KL divergence is used to train the reward model by scoring the difference of the new completions from the original human-labeled ones.
- [ ] KL divergence encourages large updates to the LLM weights to increase differences from the original model.
- [x] KL divergence measures the difference between two probability distributions.
- [x] KL divergence is used to enforce a constraint that limits the extent of LLM weight updates.

5. Fill in the blanks: When fine-tuning a large language model with human feedback, the action that the agent (in this case the LLM) carries out is _____ and the action space is the _____. **1 point**

- ( ) Processing the prompt, context window.
- ( ) Generating the next token, the context window
- (•) Generating the next token, vocabulary of all tokens.
- ( ) Calculating the probability distribution, the LLM model weights.

6. How does Retrieval Augmented Generation (RAG) enhance generation-based models? **1 point**

- ( ) By optimizing model architecture to generate factual completions.
- ( ) By applying reinforcement learning techniques to augment completions.
- ( ) By increasing the training data size.
- (•) By making external knowledge available to the model

7. How can incorporating information retrieval techniques improve your LLM application? Select all that apply. **1 point**

- [x] Overcome Knowledge Cut-offs
- [ ] Faster training speed when compared to traditional models
- [ ] Reduced memory footprint for the model
- [x] Improve relevance and accuracy of responses

8. What are correct definitions of Program-aided Language (PAL) models? Select all that apply. **1 point**

- [x] Models that offload computational tasks to other programs.
- [ ] Models that enable automatic translation of programming languages to human languages.
- [x] Models that assist programmers in writing code through natural language interfaces.
- [ ] Models that integrate language translation and coding functionalities.

9. Which of the following best describes the primary focus of ReAct? **1 point**

coach

( ) Enhancing language understanding and decision making in LLMs