[[A screenshot of a computer

Description automatically generated](https://substackcdn.com/image/fetch/f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2F2e6c2f94-a539-4851-be8c-5b525c1e227c_2568x3462.png)](https://substackcdn.com/image/fetch/f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2F2e6c2f94-a539-4851-be8c-5b525c1e227c_2568x3462.png" \t "_blank)

1. Training. To train a ChatGPT model, there are two stages:

- Pre-training: In this stage, we train a GPT model (decoder-only transformer) on a large chunk of internet data. The objective is to train a model that can predict future words given a sentence in a way that is grammatically correct and semantically meaningful similar to the internet data. After the pre-training stage, the model can complete given sentences, but it is not capable of responding to questions.  
  
- Fine-tuning: This stage is a 3-step process that turns the pre-trained model into a question-answering ChatGPT model:  
  
1). Collect training data (questions and answers), and fine-tune the pre-trained model on this data. The model takes a question as input and learns to generate an answer similar to the training data.  
2). Collect more data (question, several answers) and train a reward model to rank these answers from most relevant to least relevant.  
3). Use reinforcement learning (PPO optimization) to fine-tune the model so the model's answers are more accurate.

1. Answer a prompt

* Step 1: The user enters the full question, “Explain how a classification algorithm works”.
* Step 2: The question is sent to a content moderation component. This component ensures that the question does not violate safety guidelines and filters inappropriate questions.
* Steps 3-4: If the input passes content moderation, it is sent to the chatGPT model. If the input doesn’t pass content moderation, it goes straight to template response generation.
* Step 5-6: Once the model generates the response, it is sent to a content moderation component again. This ensures the generated response is safe, harmless, unbiased, etc.
* Step 7: If the input passes content moderation, it is shown to the user. If the input doesn’t pass content moderation, it goes to template response generation and shows a template answer to the user.