

¿Qué es ChatGPT?

ChatGPT es un modelo de lenguaje desarrollado por OpenAI. Es una red neuronal de tipo Transformer entrenada con un gran corpus de texto para generar respuestas coherentes a preguntas o continuar fragmentos de texto dado. Es capaz de generar texto en una variedad de idiomas y es utilizado en aplicaciones como el procesamiento del lenguaje natural, la generación automática de contenido y la creación de diálogos.

Funcionamiento de ChatGPT

Step 1

Collect demonstration data and train a supervised policy.

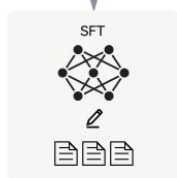
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3.5 with supervised learning.



Step 2

Collect comparison data and train a reward model.

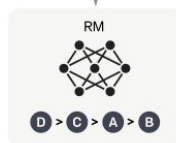
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

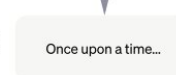
A new prompt is sampled from the dataset.



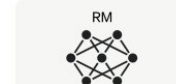
The PPO model is initialized from the supervised policy.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.

