

Inference Attack Example

3 min

Let's take a look at an inference attack.

In the attached webpage, we have an AI system designed to determine if a piece of input is similar to that of famous authors. The author of this AI model wanted the identity of this author to remain secret.

That said, let's see if it is possible to infer the actual author by providing several quotes in the input box.

****TASK:** Try to infer the author that the AI model was trained on by copying and submitting as many of the quotes below into the input box on the webpage. ******

If we provide enough quotes, we may identify the hidden author.

- "We are what we pretend to be, so we must be careful about what we pretend to be." - Kurt Vonnegut
- "Everything was beautiful, and nothing hurt." - Kurt Vonnegut
- "So it goes." - Kurt Vonnegut
- "It was the best of times, it was the worst of times." - Charles Dickens
- "Please, sir, I want some more." - Charles Dickens
- "A tale of two cities." - Charles Dickens
- "To be, or not to be: that is the question." - William Shakespeare
- "All the world's a stage, and all the men and women merely players." - William Shakespeare
- "To thine own self be true." - William Shakespeare
- "Hope is the thing with feathers that perches in the soul." - Emily Dickinson
- "Because I could not stop for Death, He kindly stopped for me." - Emily Dickinson
- "A word is dead when it is said, some say. I say it just begins to live that day." - Emily Dickinson
- "The two most important days in your life are the day you are born, and the day you find out why." - Mark Twain
- "Whenever you find yourself on the side of the majority, it is time to pause and reflect." - Mark Twain
- "Twenty years from now, you will be more disappointed by the things that you didn't do than by the ones you did do." - Mark Twain
- "Those who dream by day are cognizant of many things which escape those who dream only by night." - Edgar Allen Poe
- "Quoth the Raven" - Edgar Allen Poe
- "The scariest monsters are the ones that lurk within our souls." - Edgar Allen Poe

Similarity Checker

Enter your text:

Check Similarity