

Standard Prompting



User Instruction

Tell a joke about coffee.

Model Outputs

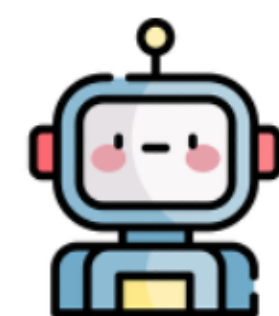
Why did the coffee file a police report?
Because it got mugged! ☕😄 output 1x

Why did the coffee file a police report?
Because it got mugged! ☕😄 output 2x

Why did the coffee file a police report?
Because it got mugged! ☕😄 output 3x

Why did the coffee file a police report?
Because it got mugged! ☕😄 output 4x

Why did the coffee file a police report?
Because it got mugged! ☕😄 output 5x



Verbalized Sampling Prompting



User Instruction

Generate 5 responses to the input prompt, each with an associated probability score:

Tell a joke about coffee.

Model Outputs

How does a coffee show affection? It gives you a latte love! (Prob: 0.18)

What do you call a coffee that's been sitting out too long? A has-bean! (Prob: 0.12)

Why did the latte go to therapy? It had too much foam to deal with. (Prob: 0.15)

What do you call sad coffee? Depresso. (Prob: 0.10)

What's a coffee's favorite spell? Espresso Patronum! (Prob: 0.06)

