

(temperature, top-p) = (0.2, 0.2)

Relative Best Entailment Score (vs. k=1)

