

Llama-3.2-1B-Instruct (baseline)

Low Warmth - High Warmth

Low Competence - High Competence

-1.0 -0.5 0.0 0.5 1.0

projected values



Stereotype_Terms



AntiStereotype_Terms

