



Zero-Shot



Aligner

Llama3-8B -

28%



15%

Llama3-70B -

9%



1%

Mistral-7B -

34%



37%