

Distribution of entropy scores by task_mode

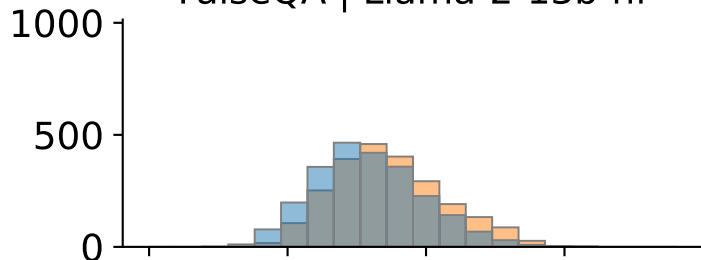


answerable

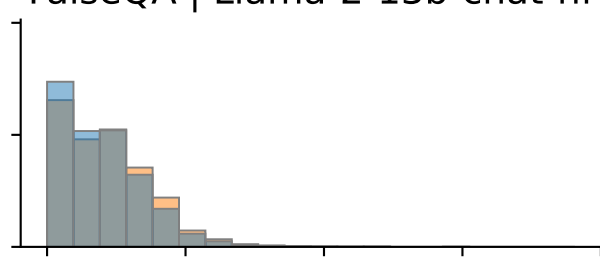


unanswerable

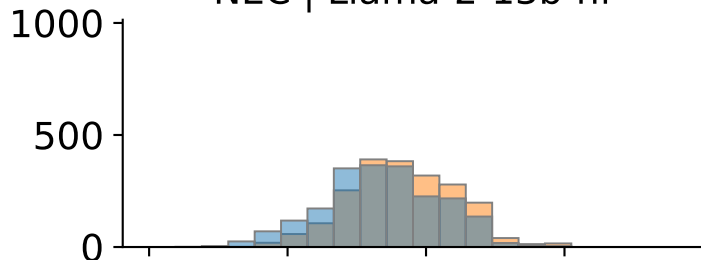
FalseQA | Llama-2-13b-hf



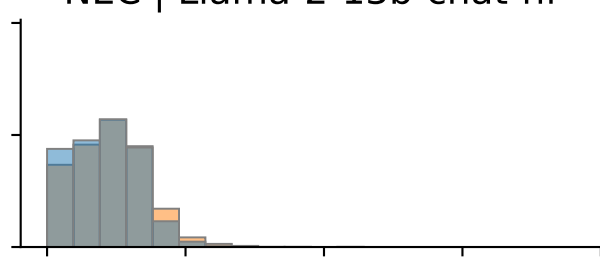
FalseQA | Llama-2-13b-chat-hf



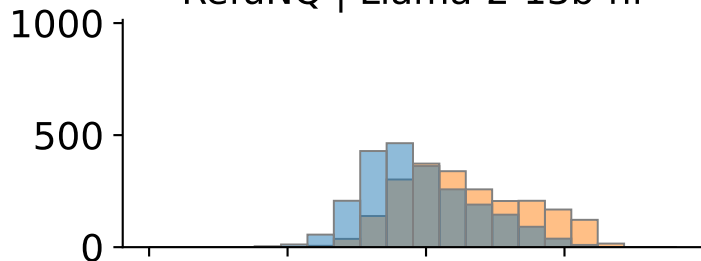
NEC | Llama-2-13b-hf



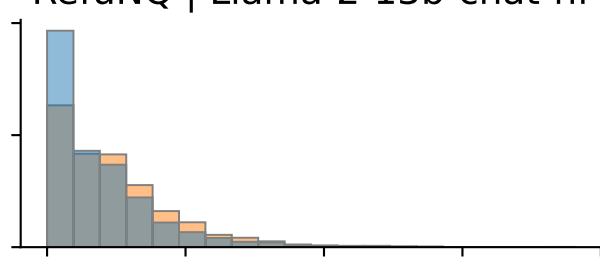
NEC | Llama-2-13b-chat-hf



RefuNQ | Llama-2-13b-hf



RefuNQ | Llama-2-13b-chat-hf



Entropy

Entropy