

vlat jaccard_similarity between Agents

agent_type_2

llava-hf/llava-1.5-7b-hf	0.91	0.58	0.55	0.40	0.45	0.44
Salesforce/blip2-flan-t5-xl	0.58	0.95	0.58	0.48	0.37	0.37
Salesforce/blip2-flan-t5-xxl	0.55	0.58	0.96	0.46	0.37	0.37
GPT-4V	0.40	0.48	0.46	0.96	0.58	0.59
Human/Math-2-1	0.45	0.37	0.37	0.58	0.67	0.68
Human/Math-3	0.44	0.37	0.37	0.59	0.68	0.69

llava-hf/llava-1.5-7b-hf

Salesforce/blip2-flan-t5-xl

Salesforce/blip2-flan-t5-xxl

GPT-4V

Human/Math-2-1

Human/Math-3

agent_type_1