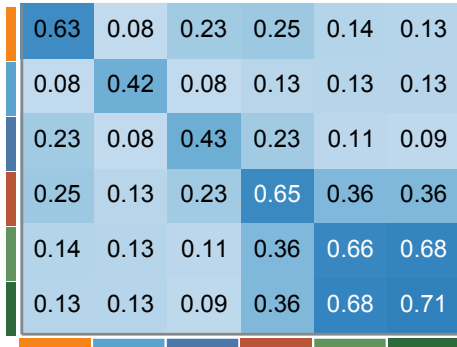


higher similarity →

GGR



VLAT



HOLF



LLaVA-1.5/Vicuna

BLIP-2/F-T5-XL

BLIP-2/F-T5-XXL

GPT-4V

Humans (less math)

Humans (more math)