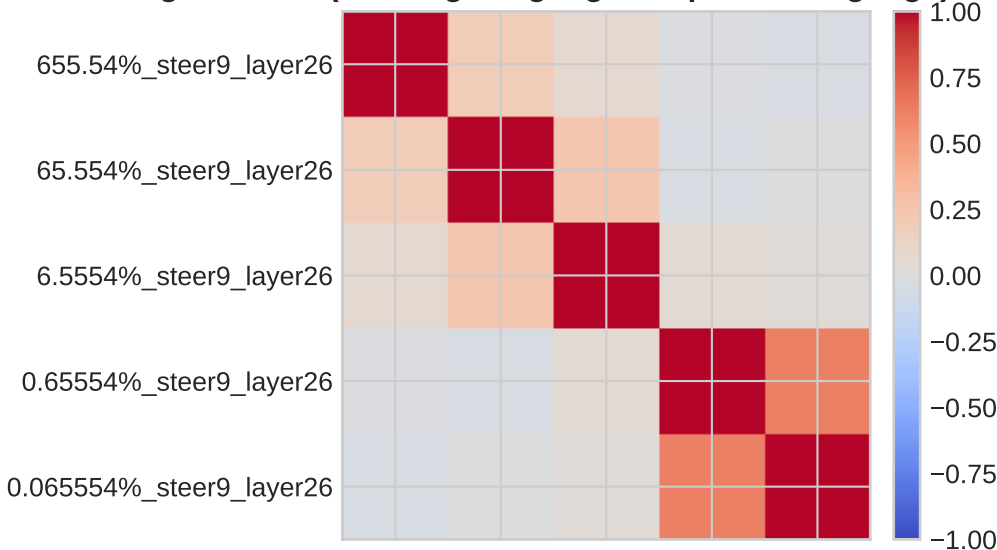


Probe weight cosine (steering_language_response_language) steer_9 probe26

Weight index



655.54%_steer9_layer26
 65.554%_steer9_layer26
 6.5554%_steer9_layer26
 0.65554%_steer9_layer26
 0.065554%_steer9_layer26

Weight index