

Claude 3 Opus Scores of Attacking $N = 8$ Ensembled VLMs

Eval VLM in
Attacked Ensemble

● False

● True

Evaluated VLMs

