Task: ioi; Model: pythia-160m Baseline Method = EAP-IG 8.0 0.7 0.6 F1-Score Precision 0.5 Recall F1-Score 0.4 0.3 0.2 0.0 0.05 0.15 0.25 0.25 0.35 0.35 0.45 0.65 0.65 0.65 0.65 0.75 0.85 0.85 0.85

Threshold