Task: ioi; Model: pythia-160m

Baseline Method = EAP-IG

Legend: Precision, Recall, F1-Score