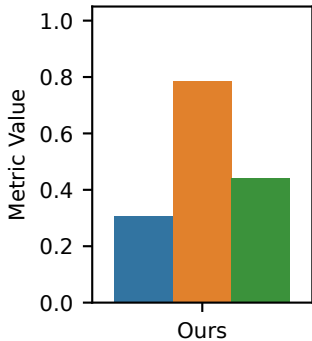
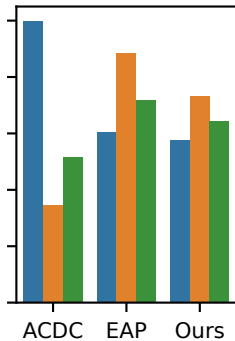


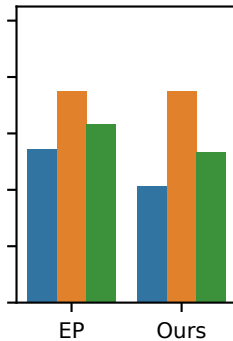
pythia-160m, IOI



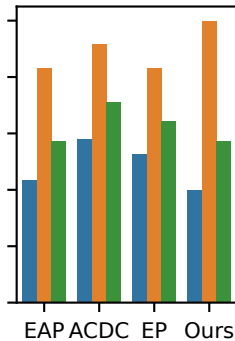
gpt2-small, IOI



gpt2-small, GP



gpt2-small, GT



Precision



Recall



F1-Score