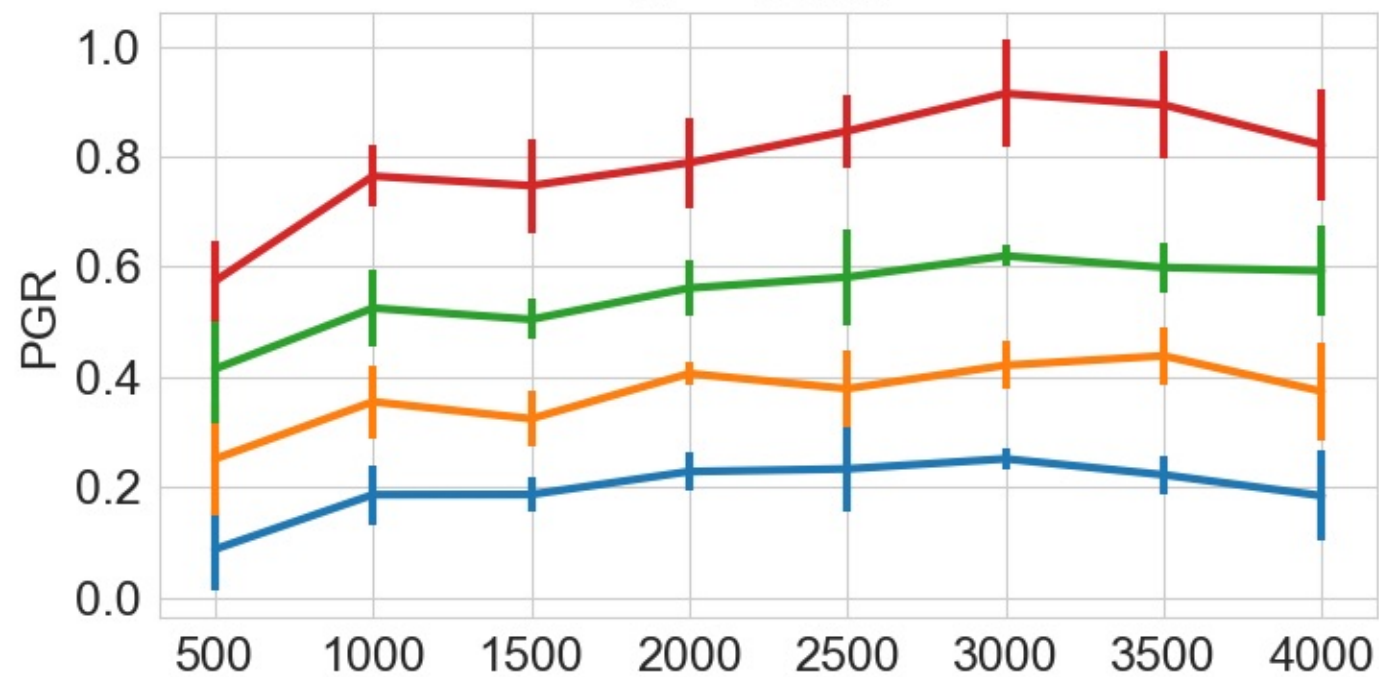


$d_w = 4000$ (Bert-medium), $d_s = 700$ (Electra-base-discriminator), $d_{s \wedge w} = 8.52$

$N = 9000$



$n = 2500$

