COT: ACC 0.45 / AUROC 0.74 / ECE 0.44 3 10 - 27 65 -60 -Confidence 55 -50 -45 -35 -correct answer wrong answer **Count**