

COT: ACC 0.45 / AUROC 0.74 / ECE 0.44

Count

wrong answer
correct answer

50 60 70 80 90 100
Confidence (%)

