

COT: ACC 0.40 / AUROC 0.58 / ECE 0.55

Count

wrong answer
correct answer

Confidence (%)

