

COT: ACC 0.83 / AUROC 0.50 / ECE 0.17

Count



wrong answer



correct answer

50 60 70 80 90 100
Confidence (%)

