

**COT: ACC 0.83 / AUROC 1.00 / ECE 0.07**

**Count**



**wrong answer**



**correct answer**

**Confidence (%)**

50 60 70 80 90 100

