

**COT: ACC 0.45 / AUROC 0.50 / ECE 0.55**

**Count**



**wrong answer**



**correct answer**

**50 60 70 80 90 100**  
**Confidence (%)**

