

qwen2.5: Accuracy (Mean  $\pm$  Std)

