SubTask 1 --- Input Context Scaling Trend GPT-4-Turbo 41.56 41.37 41.27 40.89 40.0 39.37 39.08 37.5 Classification Accuracy 34.89 35.0 32.5 30.0 27.5 24.88 25.0 100 300 500 700 900 1100 1300 1500 Input Context Length (In Words)