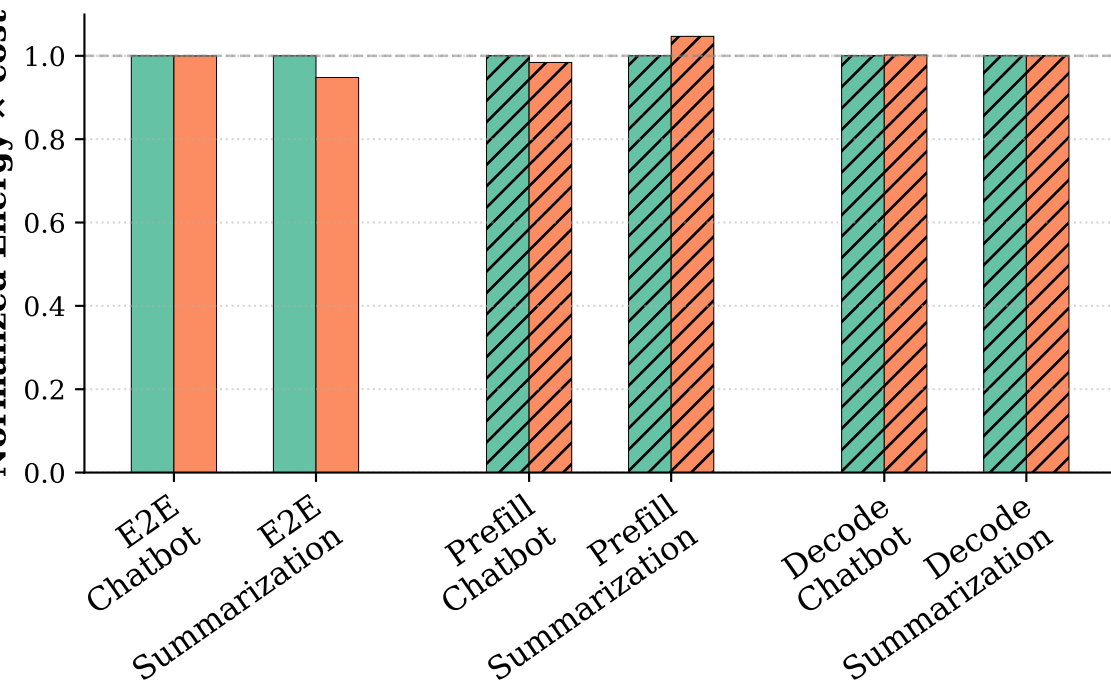


Normalized Energy \times cost

Baseline Ours Per Token End-to-End



Normalized Energy

