

Evaluation Rubric

Each summary was evaluated based on the following 3 criteria.

Each criterion is scored from 0 (poor) to 2 (excellent).

Criterion	Score	Meaning
Accuracy	0	Summary is incorrect or misleading
	1	Partially correct, vague or missing key details
	2	Fully accurate and reflects the main idea of the review
Tone	0	Robotic, unclear, or unprofessional tone
	1	Somewhat readable, lacks polish
	2	Polite, fluent, and human-like
Task Success	0	Did not follow instructions (too long, incomplete, or off-topic)
	1	Somewhat followed the prompt (minor format/tone issues)
	2	Fully met the task (polite, 1-2 lines, clear and complete summary)

Total Score per prompt = Accuracy + Tone + Task Success (Max: 6)