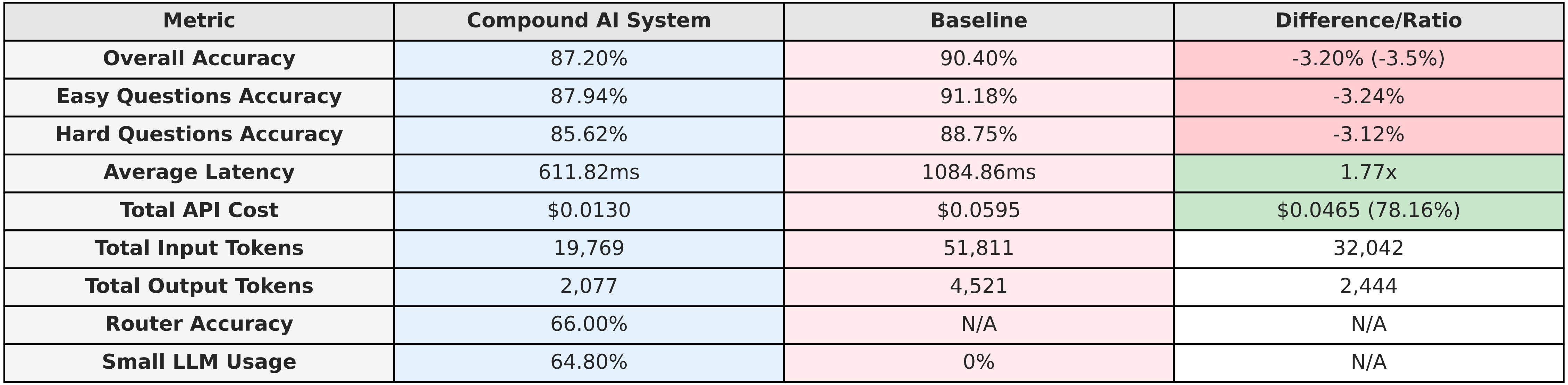
Key Performance Metrics



Metric	Compound AI System	Baseline	Difference/Ratio
Overall Accuracy	87.20%	90.40%	-3.20% (-3.5%)
Easy Questions Accuracy	87.94%	91.18%	-3.24%
Hard Questions Accuracy	85.62%	88.75%	-3.12%
Average Latency	611.82ms	1084.86ms	1.77x
Total API Cost	$0.0130	$0.0595	$0.0465 (78.16%)
Total Input Tokens	19,769	51,811	32,042
Total Output Tokens	2,077	4,521	2,444
Router Accuracy	66.00%	N/A	N/A
Small LLM Usage	64.80%	0%	N/A

Positive values in the difference column indicate improvements by the Compound AI system.