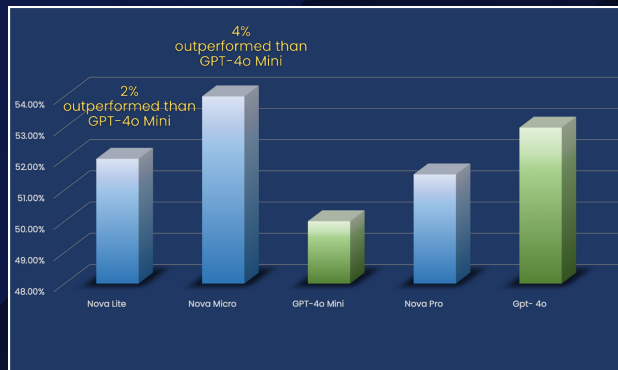


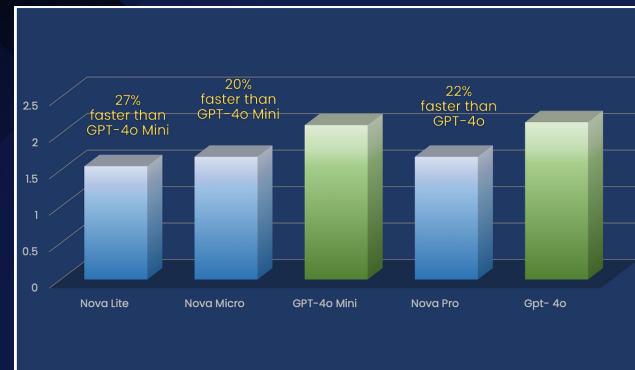
BENCHMARKING AMAZON NOVA AND GPT-4O MODELS WITH FLOTORCH

FloTorch compared Amazon Nova models against GPT-4o using the CRAG dataset, evaluating accuracy, latency, and cost effectiveness.

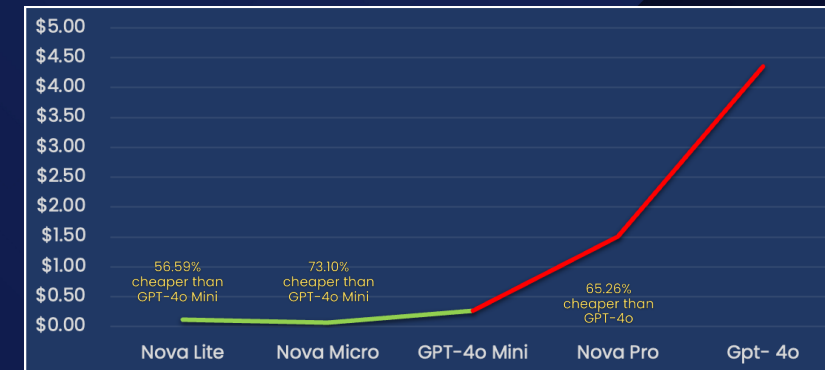
% Accuracy on Subset of CRAG Dataset



Average Latency (in Seconds)



Cost for Running Million Queries



Key Findings

Amazon Nova Pro vs. GPT-4o

- 65.26% cost savings with Amazon Nova Pro.
- 21.97% faster than GPT-4o.
- 1.5% lower accuracy, with GPT-4o leading.

Amazon Nova Micro & Lite vs. GPT-4o Mini

- 56.59% (Lite) and 73.10% (Micro) cost savings compared to GPT-4o Mini.
- 20% faster than GPT-4o Mini.
- 2 and 4 points higher accuracy with Lite and Micro, respectively, over GPT-4o Mini.

Check out the complete Benchmark Analysis: [Amazon Nova vs GPT-4o](#)
To Learn More, visit the [Amazon Nova Product Page](#)

Start Experimenting with FloTorch, available on: