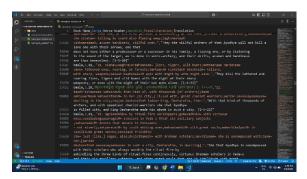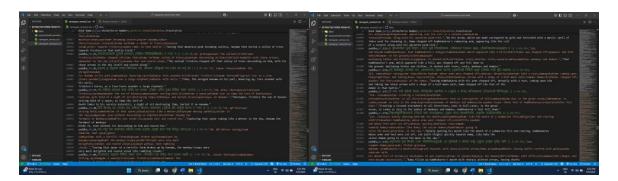# Stage 1 Progress Details

Targeted Web Scraping

Harvested the full text of Valmiki Rāmāyaṇa Kandas 1–7 from https://valmikiramayan.net/; recorded the URL and timestamp for traceability and version control.
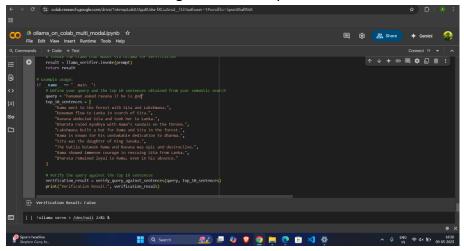

Initial Model Testing

Applied the selected finetuned BERT AND LLAMA3.1 + semantic search architecture with successful results achieving 100% accuracy in kanda 1

Ready for Scale-Up

Confirmed that all components are fully modular and prepared for rollout across the remaining Kandas.

Plan for Next Stage

Model Deployment Across Kandas 2–7

Run the validated BERT + LLaMA 3.1 pipeline on all remaining shlokas.

Documentation & Presentation

Prepare comprehensive documentation of methodology, tests, and results.

Create a concise presentation summarizing progress and next steps.