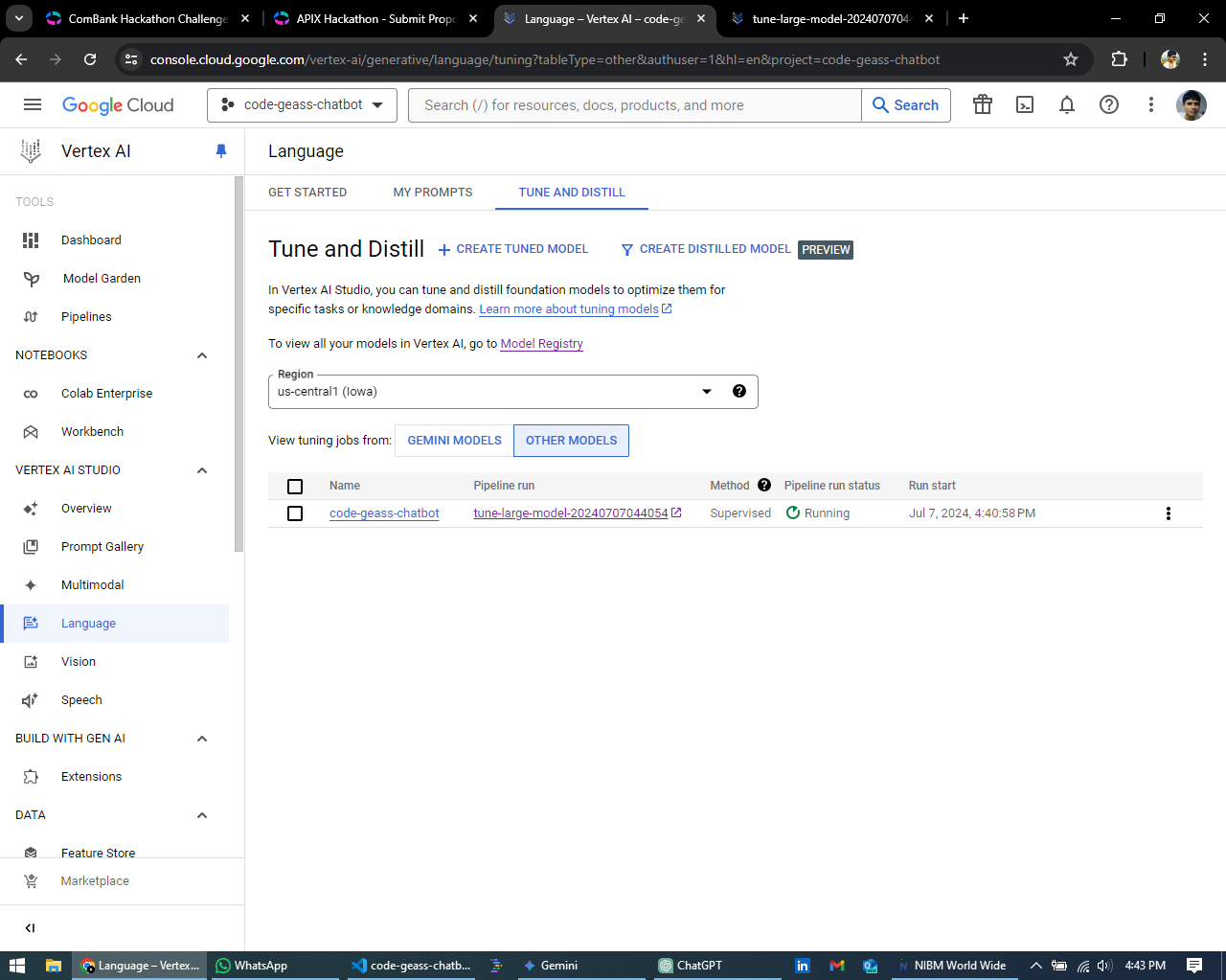
**Code Geass Chatbot Report**

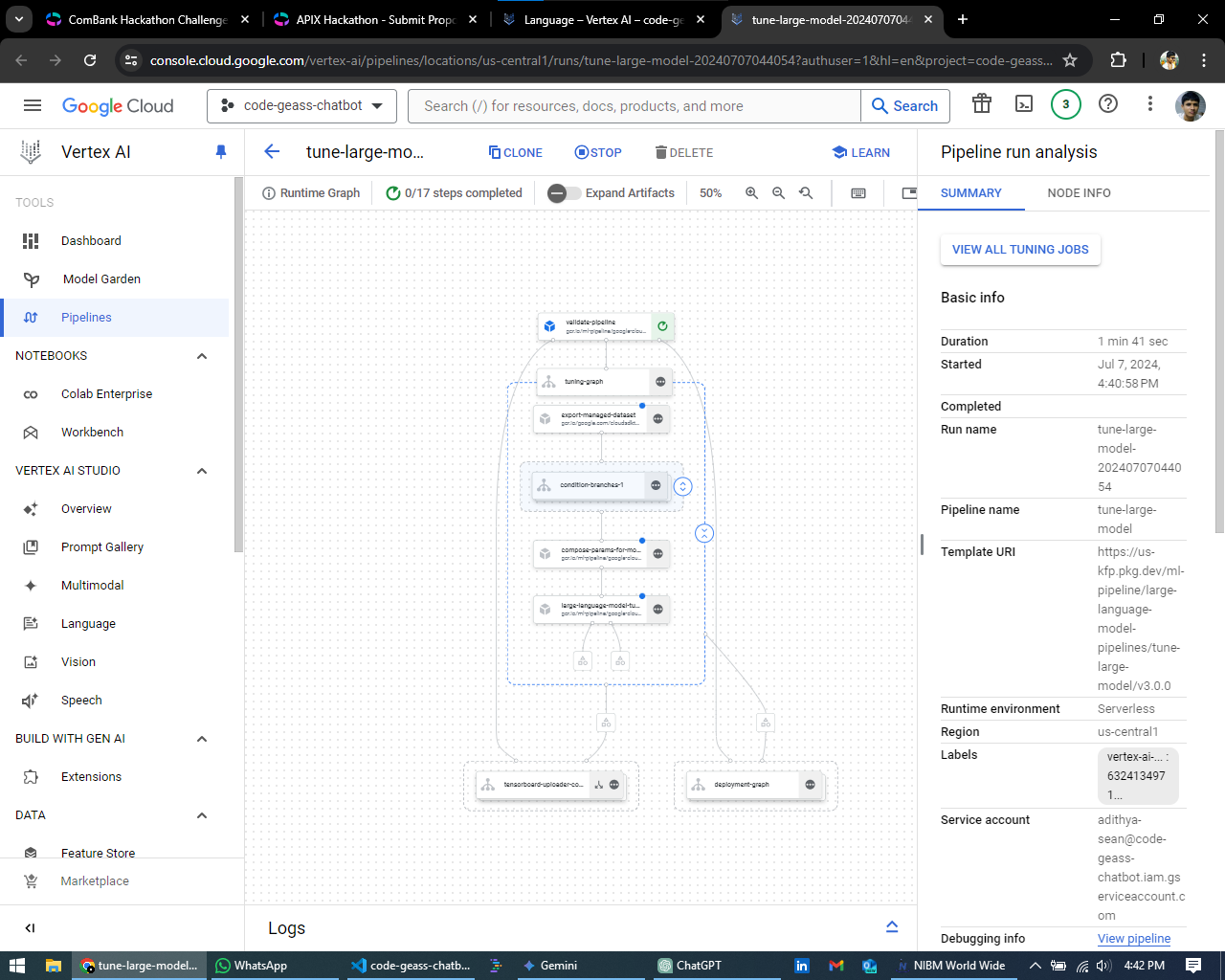
**Project Goal:**

**Develop a Singlish-language chatbot capable of understanding and generating natural, contextually relevant responses to user queries. The chatbot is designed to cater specifically to the communication style of Sri Lankan users, who frequently use Singlish in online interactions.**

**Technical Approach:**

* **Model: Palm2 Text Bison fine-tuned on Google's Vertex AI platform.**
* **Data Collection:**
  + **Synthetic Singlish datasets generated from existing language models proficient in basic Sinhala communication.**
  + **Transliteration of Sinhala text from various sources into Singlish.**
* **Fine-tuning: Iterative training of the model on the curated Singlish dataset, with ongoing optimization for performance and accuracy.**

****

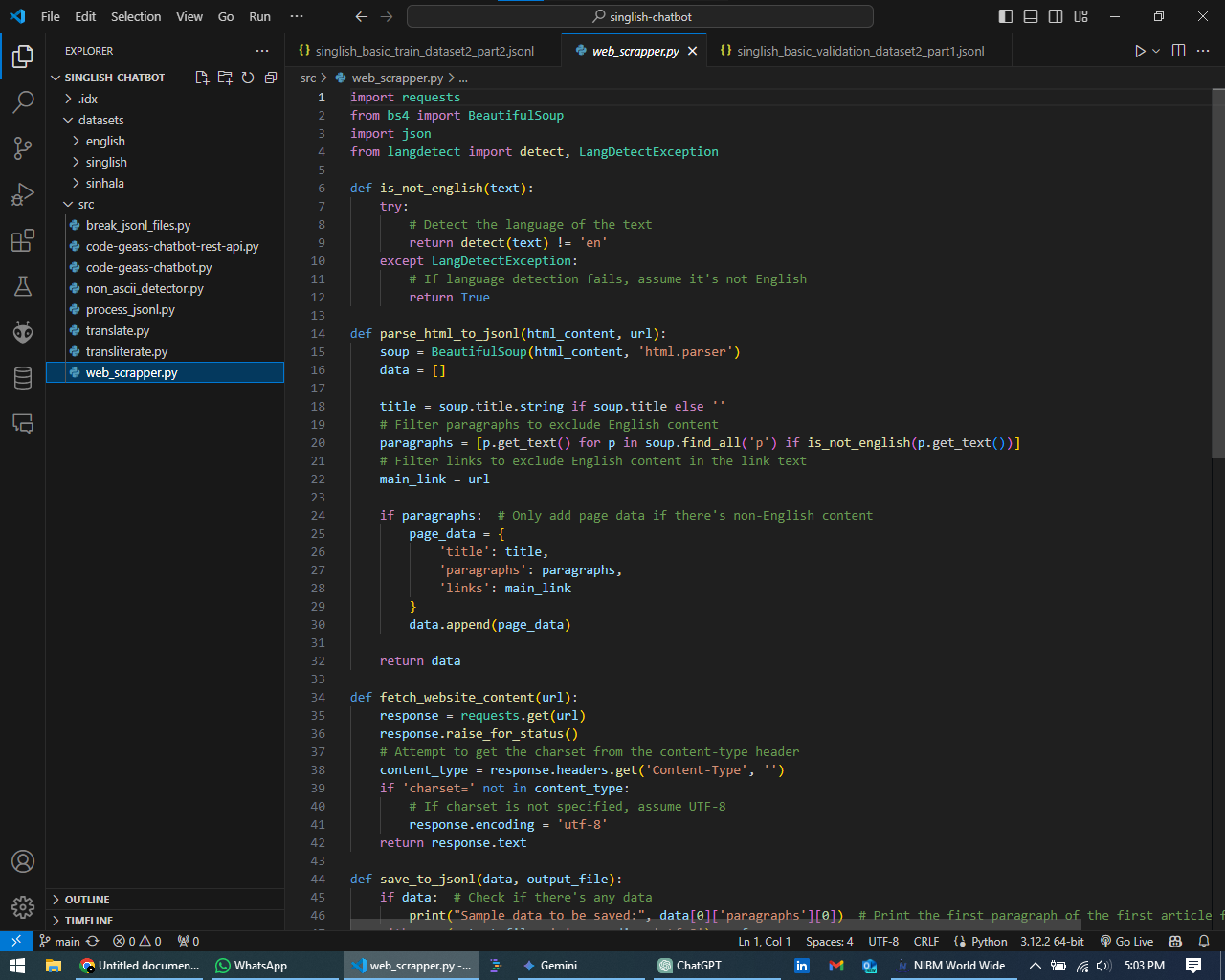
****

**Current Progress:**

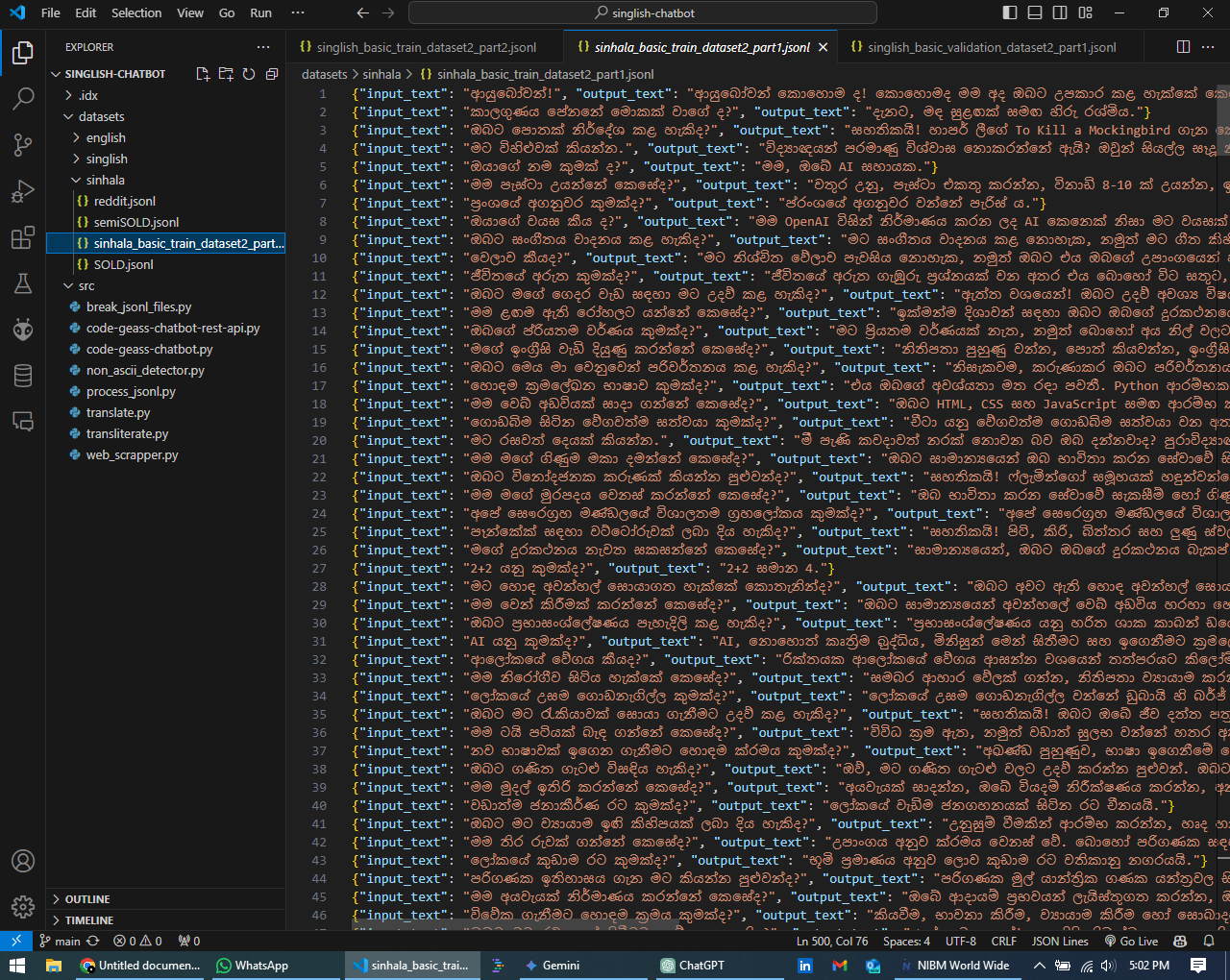
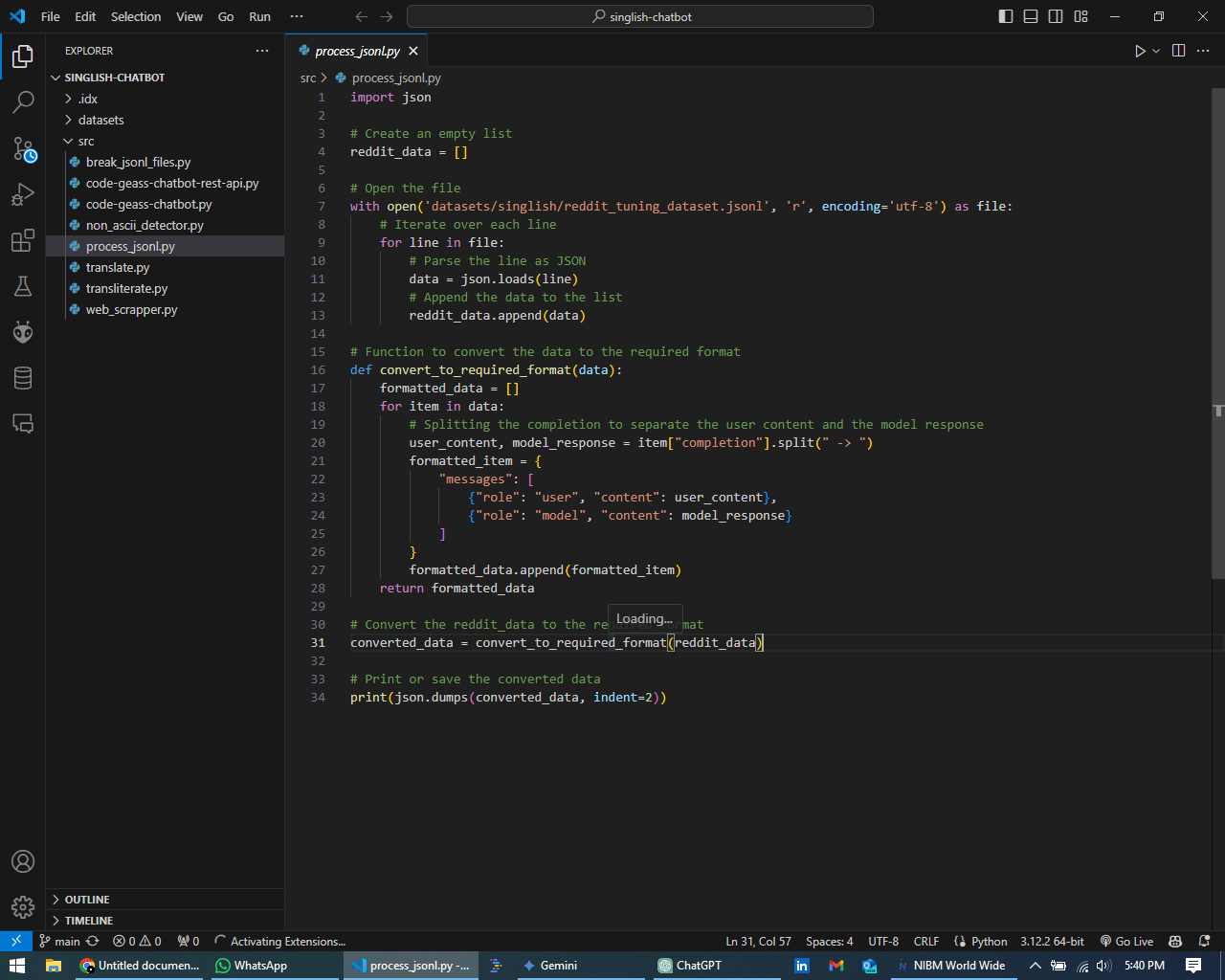
* **The prototype chatbot demonstrates basic conversational capabilities in Singlish.**
* **It can understand and respond to simple queries and greetings.**
* **While limited by initial resource constraints, the model shows promise for generating meaningful Singlish text in familiar contexts.**
* **Demonstration Video -** [**https://photos.app.goo.gl/pLv72q9z3MuCBWBX6**](https://photos.app.goo.gl/pLv72q9z3MuCBWBX6)
* **REST API -** [**curl -X POST http://112.134.172.20:5000/predict -H "Content-Type: application/json" -d "{\"text\":\"lankawe usama kanda mokadda?\"}"**](http://localhost:5000/predict)

**Challenges and Solutions:**

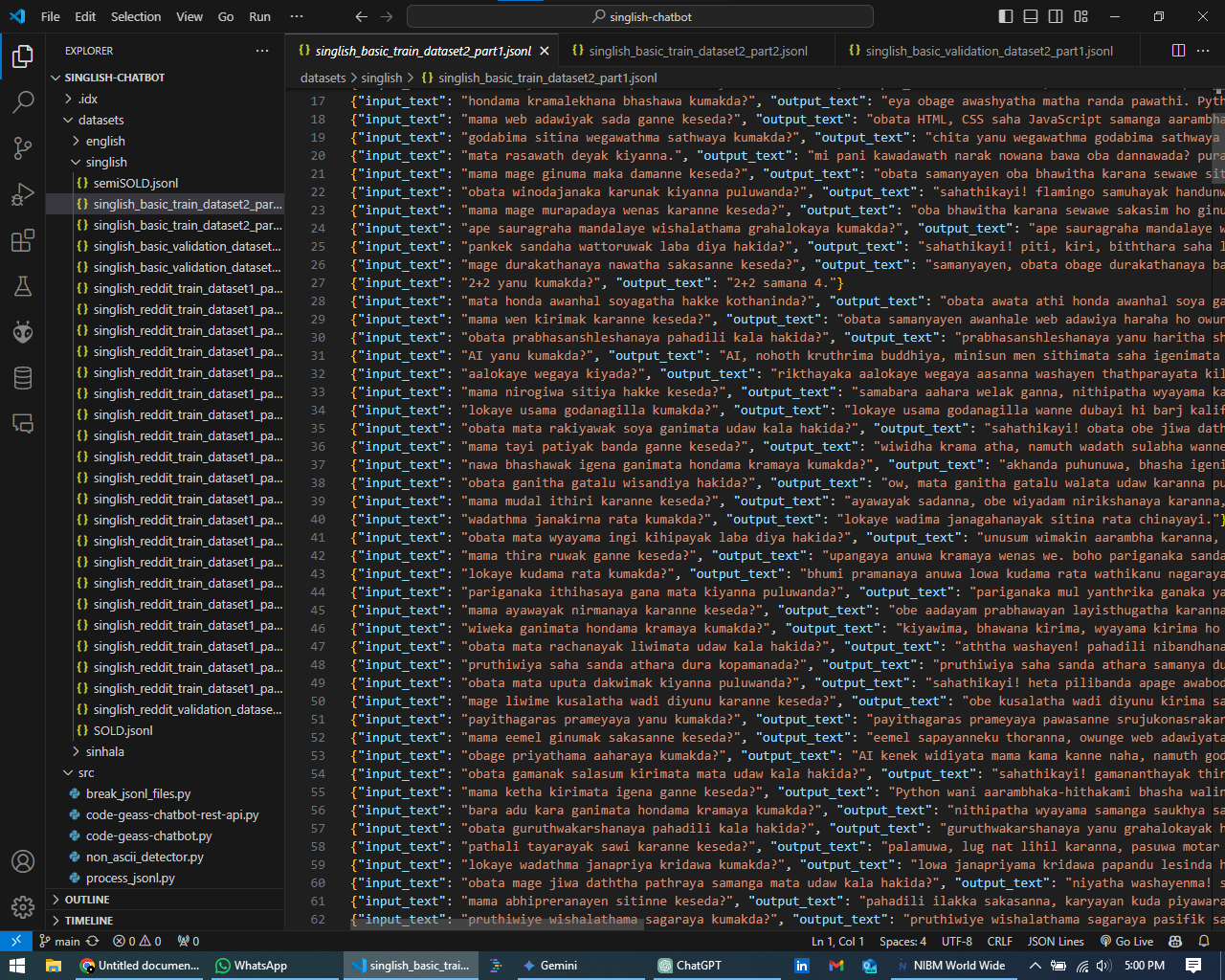
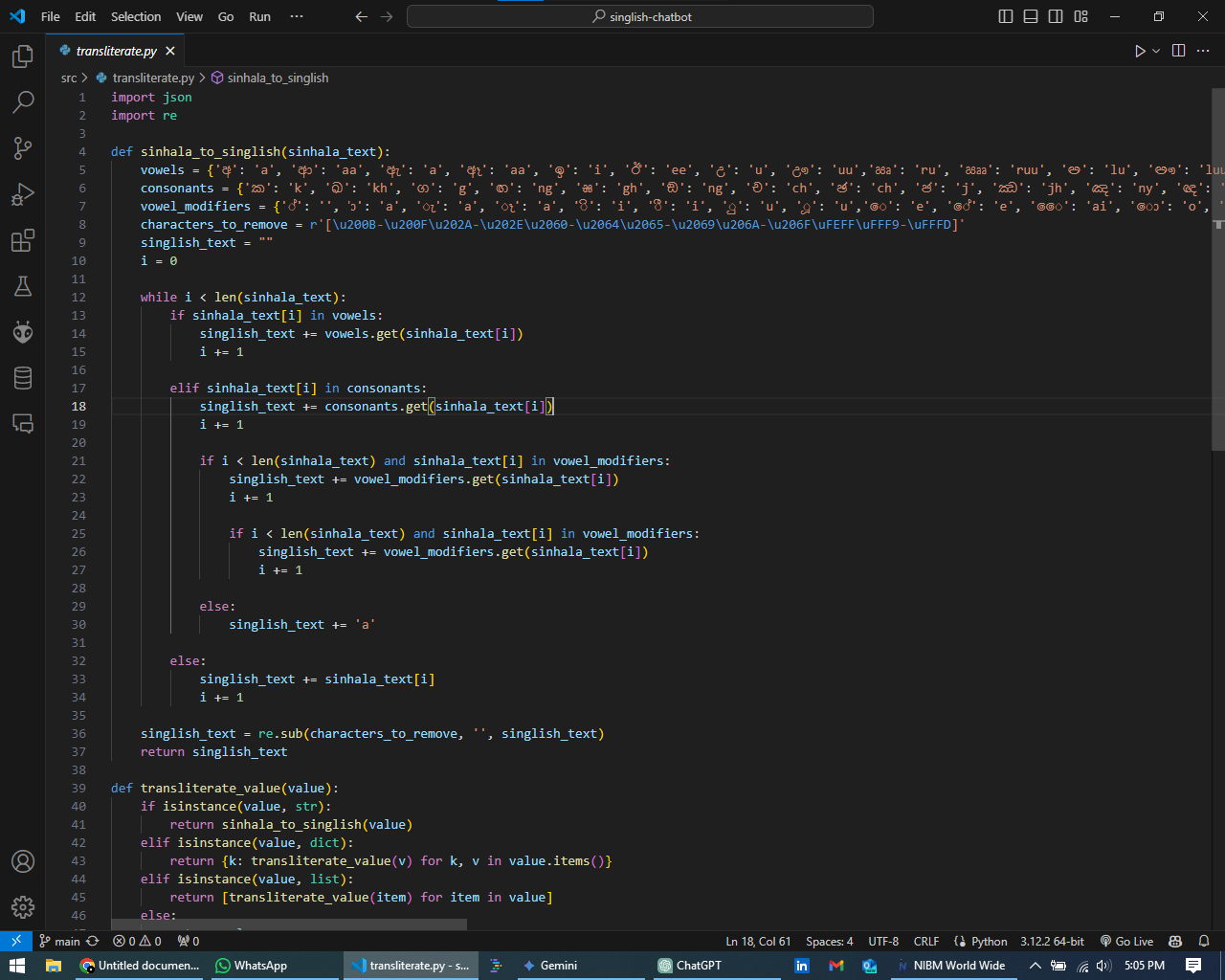
* **Limited Singlish Resources: Addressed by generating synthetic data and utilizing existing language models for transliteration.**
* **Technical Expertise: Leveraged Google's Vertex AI platform for ease of use and sought guidance from online resources and experienced peers.**
* **Resource Constraints: Optimized model architecture and training strategies, utilized free tiers and credits, and explored potential partnerships for additional resources.**

**Webpage Text Scraping Script Which Can be Automated to Run on a Server**

**This Script converts web pages text into JSON Lines format compatible with language models for training purposes.**

****

**After That We Can Use Character mapping to convert sinhala text into singlish**

****

**This whole process can be automated in the cloud.**

**Future Enhancements:**

* **Further Fine-tuning: Expanding the training dataset with more diverse Singlish text to improve the model's understanding and generation capabilities.**
* **Advanced Features: Exploring the integration of information retrieval capabilities to enable the chatbot to answer complex questions and provide relevant information.**
* **User Testing and Feedback: Gathering feedback from Singlish speakers to continuously refine the chatbot's performance and ensure it meets the needs of its target audience.**

**Conclusion:**

**The Singlish chatbot project is a promising initiative to bridge the language gap in Sri Lanka's digital communication landscape. By leveraging advanced AI technology and a data-driven approach, the project aims to create a valuable tool that enhances communication, fosters cultural understanding, and promotes digital inclusion for Singlish speakers in Sri Lanka.**