Aviral Yadav

Phone: 7007934595 | aviy2402@gmail.com | LinkedIn | Github

Education

Vellore Institute of Technology, Bhopal – ***B Tech in Computer Science and Engineering (Cyber Security and Digital Forensics) | CGPA – 7.84*** August 2022 – June 2026

Skills

Programming Languages: C++, Java, Python

Web Development: HTML, CSS, JavaScript, React, Node.js, Flask

Tools & Platforms: Git, AWS, Excel

Framework/Libraries: Pandas, NumPy, Matplotlib

Experience

QuantAI, Auckland, New Zealand - *Data Scientist Intern* May 2025 – July 2025

- Designed and developed backend services and automated data pipelines using Python and Flask, handling large-scale structured and unstructured datasets.
- Built and deployed web scraping solutions to collect and preprocess data from multiple sources, improving data availability for analytics.
- Collaborated with team using Git for version control and Agile sprints to deliver production-ready code.

Projects

Scrapeazon - *[Flask| Python | BeautifulSoup| Requests | REST API]* May 2025

- Developed a **Flask-based REST API** to automate **Amazon product data extraction**, solving the challenge of manual data collection for **price monitoring, market research, and competitor analysis**.
- Achieved **>95% parsing accuracy** by extracting key attributes — product name, price, rating, review count, availability, discount, and Prime status.
- Implemented **scalable scraping logic** capable of navigating **20+ pages** and retrieving **100+ unique listings per query** without duplication.
- Enhanced system reliability through **5-level retry logic, randomized user-agents, and exponential backoff**, reducing request failure rate by **~80%** compared to naive scraping.

Car Finder *[Python | Requests | BeautifulSoup | Flask | Flask-CORS | JSON]* June 2025

- Developed a Python scraper to aggregate data from **2** car dealership websites **(Andrew Simms and NZ Cheap Cars)**, delivering up to **100** combined results per search.
- Aggregated a comprehensive dataset of up to **100** vehicles per query by parsing **8** key attributes—including price and odometer reading—from as many as **6** pages of search results.
- Ensured high reliability by implementing a **3-attempt** retry mechanism with exponential backoff and rotating through **3** unique user-agents to prevent blocking.

Academic Research Assistant – *[Python | Flask | AI & Machine Learning]* July 2025

- Developed a user-friendly, responsive web application using **Flask** that allows users to upload PDF files and engage in a real-time conversational chat about the documents.
- Constructed a RAG chain using the **LangChain** framework. When a user asks a question, the system first retrieves the most relevant document chunks from the **FAISS** index. This context is then fed to a Large Language Model to generate a precise and well-informed answer.
- Integrated the **Llama3-8B** model running on the high-speed **Groq LPU™ Inference Engine**. This ensures that answers are not only accurate but also generated with extremely low latency, providing a seamless user experience.

Certifications

- **Bits and Bytes of Computer Networking –** Coursera
- **Cyber Physical System –** NPTEL
- **VLSI Design –** Maven Silicon