

# Varshith Pilli

✉ varshith6946@gmail.com

☎ +91 8978930678

📍 Chennai, India

🌐 varshith-pilli

👤 varshithpilli

## SUMMARY

Aspiring Generative AI and Backend Engineer with hands-on experience in building LLM-powered applications, including a cloud-based hybrid legal assistant for constitutional law retrieval. Skilled in developing robust backend systems using FastAPI, Flask, and PostgreSQL with SQLAlchemy ORM. Experienced in creating RESTful APIs, JWT-based authentication, and deploying scalable solutions with a focus on modular design and efficient data handling. Passionate about exploring advanced LLM architectures, fine-tuning techniques, cloud deployment strategies, and building AI-driven tools that combine natural language understanding with practical, real-world applications.

## SKILLS

- Programming: Python, C/C++, Java, Go, JS
- Technologies/Tools: Docker, PyTorch, Linux, Git, FastAPI, RESTful APIs Flask, PostgreSQL, Neo4j
- GenAI: HuggingFace, LangChain, OpenRouter, PineConeDB

## EXPERIENCE

### VIT Chennai

Research Intern (AI for Education)

June 2025 - Present

#### Knowledge Distillation – Team Lead & Coordinator

- Led a 13-member cross-functional team alongside two co-leads to develop an AI solution within a university research initiative, successfully addressing challenges in data scraping and model fine-tuning.
- Served as the primary bridge between faculty advisors and team members, facilitating clear communication, timely progress updates, and alignment with research objectives.
- Oversaw the design and implementation of the knowledge pipeline structured around entity-relationship triplets, enabling efficient representation and utilization of information in the AI model.

## PROJECTS

### Text Media – Text-Based Socializing App

July 2025

- Developed and developed backend APIs using FastAPI, implementing user authentication with JWT and data persistence via PostgreSQL with SQLAlchemy.
- Integrated basic CRUD operations for users to create, read, update, and delete posts, focusing on secure and efficient data management.
- Currently progressing on backend development, with frontend implementation planned to deliver a full-featured socializing experience.

### Constitutioner – Generative AI

July 2025

- Designed and developed a system that processes and chunks the Indian Constitution, embedding these chunks and indexing them in Pinecone to enable fast, similarity-based retrieval of relevant legal sections based on user queries, resulting in a 30% reduction in query time.
- Developed both GUI (with markdown formatting) and CLI interfaces, integrating with OpenRouter API to stream real-time, context-aware responses from an LLM, improving accessibility and user engagement.
- Applied prompt engineering techniques to reduce hallucinations by instructing the model to explicitly state when no relevant constitutional context exists.

### Linear Algebra Toolkit – WebDev

March 2025 - April 2025

- Developed and deployed a lightweight Flask-based backend supporting 10+ core linear algebra operations, reducing manual computation time by ~40% and enabling easy API integration for students and developers.
- Implemented and tested RESTful APIs for core mathematical functions, gaining hands-on experience in backend development and API design.
- Ensured robustness and usability by creating modular, well-documented endpoints for seamless client interaction, providing step-by-step solutions.

## EDUCATION

### Vellore Institute of Technology, Chennai

Bachelor of Technology (CSE) - 8.97

August 2023 - July 2027

### Bhashyam College Of Education

Secondary Education - 95.1

June 2021 - May 2023