Rajat Kumar Thakur

Delhi, India

☎ 8810544717 ✉ rajatlovescloud@gmail.com linkedin.com/in/rajat-kumar-thakur github.com/rajat-kumar-thakur

Education

Indian Institute of Information Technology Vadodara **November 2022 – June 2026**
Bachelor of Technology in Computer Science and Engineering (CGPA: 9.46) *Gujarat, India*

Experience

Software Engineer Intern **September 2025 – Present**
Cableteque Corporation *California, USA*

- Engineered a scalable search agent for advanced querying using MCP servers to optimize and achieve 97.6% accuracy.
- Collaborated to developing, and deploying Agentic models in production, and collaborating with a 5-member team.

Summer Intern **May 2025 – July 2025**
IIT Gandhinagar *Gujarat, India*

- Fine-Tuned language models for code generation and execution on resource-constrained devices, achieving inference by StarCoder model in under 3 seconds.
- Optimized small language models for resource-constrained devices, reducing inference latency by 46.3% for TinyLlama.

Projects

LLM Semantic Query Engine [Link] | *FastAPI, Gemini API, Pinecone, SQLite* **August 2025**

- Architected and deployed a full-stack PDF Q&A application using FastAPI. Improved performance by implementing efficient retrieval strategies that reduced response time to 2.1 seconds and supported the project's full life cycle.
- Integrated parsing, semantic chunking (1000+ chunks per document), and APIs to deliver structured answers with citations, reasoning, and confidence metrics.

Collaborative Whiteboard App [Link] | *React, WebSockets, TypeScript* **June 2025**

- Engineered a real-time collaborative whiteboard using React and WebSockets; architected the backend to handle high concurrency and support over 50 simultaneous users per session, ensuring a seamless and interactive experience.
- Implemented session persistence and export features, enabling users to save and share over 200 drawings.

Restaurant Management System in Assembly [Link] | *x86 Assembly Language* **May 2025**

- Improved order processing time in a restaurant management system implemented with Microsoft Macro Assembler for the 32-bit architecture.
- Orchestrated Microsoft Macro Assembler to manage billing history, retrieve thousands of records, and automate bill generation for restaurants.

Pose Estimation for Time-Critical Applications [Link] | *Computer Vision* **March 2025**

- Devised a pipeline that is 4x faster for pose estimation in point clouds using surface variation, Harris corner detection, and pose estimation using corner points.
- Achieved a 70% reduction in computation time (from 26.8 seconds to 7.9 seconds) compared to state-of-the-art methods.

Relevant Coursework

- Data Structures
- Software Engineering
- Image Processing
- Database Management
- System Software
- Data Analytics
- Algorithms
- Machine Learning
- Cryptography
- Computer Networks
- Artificial Intelligence
- Operating Systems
- Computer Architecture
- OOPS
- 5G Communication
- Parrallel Computing

Technical Skills

Languages: C, C++, Python, R, JavaScript
Databases: SQL, MongoDB, PostgreSQL
Developer Tools: VS Code, Google Cloud Platform, ArchiMate, StarUML, Jupyter Notebook, TensorBoard
Technologies/Frameworks: Next.js, React, Node, Linux, Git, Postman, Express, Selenium, Tailwind, TensorFlow
Cloud Platforms: Google Cloud Platform (GCP), AWS
Libraries & Tools: NumPy, Pandas, PyTorch, GitHub Actions

Extracurricular

- Mentored over 190 students as a **Teaching Assitant** in developing skills in C and designing scalable software solutions
- Secured **2nd place** in the Battle Bots and Sentience events at a Fest for designing a robotic hand and a combat robot.
- Led an 8-member team as **General Secretary of the Academics Committee** to plan and execute academic events.
- Attained mastery of complex algorithms through completion of 500+ challenging **LeetCode** problems.LeetCode
- Earned an NVIDIA Certificate of Competency in evaluation and light customization of LLMs.Certificate
- Completed a comprehensive 12-hour online training in deep learning offered by NVIDIA.Certificate