

SRIBHUVAN REDDY YELLU

+1(470) 628-4804 ♦ Atlanta, GA

sribhuvan.y@gmail.com ♦ [linkedin.com/in/sribhuvanyellu](https://www.linkedin.com/in/sribhuvanyellu) ♦ github.com/Sribhuvan-25 ♦ sribhuvanyellu.com

EDUCATION

Master of Science in Computer Science, Georgia State University

Jan 2024 - May 2025

Bachelor of Science in Computer Science, Georgia State University

Aug 2020 - Dec 2023

• Graduated Summa Cum Laude

SKILLS

Languages Java, Python, JavaScript, TypeScript, Swift, Kotlin, HTML5, CSS

Frameworks Next.js, React, React Query, Angular, Node.js, Spring Boot, Flutter, TensorFlow, PyTorch

Database SQL, Firebase, MongoDB, MySQL, PostgreSQL, SQLite

Tools Azure DevOps, AWS, Docker, GraphQL, Git, Kubernetes

PROFESSIONAL EXPERIENCE

ML/AI Engineer Intern

May 2025 - Present

Pearson VUE, Bloomington, MN

- Engineered real-time AI proctoring with Gemini-2.5-Flash at 30 FPS 92% detection confidence across 6 violation categories.
- Boosted API throughput 4x, added thread-safe alerts and 95% frame-quality capture for 12–22s real-time violation display.
- Developed Python facial-recognition module with Gemini-2.0-Flash API to verify student IDs against live webcam feeds.
- Built GCP Vertex AI pipeline to fine-tune Gemini face-comparison model, achieving 92% accuracy on test size of 1,680.

ML/AI Engineer Intern

Nov 2024 - April 2025

Pearson VUE, Bloomington, MN

- Built a Copilot agent in Teams Toolkit that surfaces exact doc answers in seconds, replacing minutes-long manual searches.
- Prototyped CrewAI agents to auto-summarize Teams meetings and email minutes, eliminating manual note-taking.
- Spearheaded the implementation of DiskANN indexing PostgreSQL in RAG service to cut down indexing times by 35%.
- Built an MCP server on Bookings-API with Claude + Perplexity for one-chat scheduling and related recommendations.

Software Engineer Intern

May 2024 - Nov2024

Pearson VUE, Bloomington, MN

- Developed System Test Receipt Dashboard in Angular and Spring Boot, reducing test-launch troubleshooting time by 40%.
- Engineered a scalable backend service in Java Spring Boot enhance internal operations, increasing user efficiency by 20%.
- Optimized CI/CD pipelines with Azure DevOps, TeamCity, and Helm, cutting deployment time by 30%.
- Updated SQL scripts for key data updates across all clients, ensuring 100% data integrity during critical operations.

Software Engineer Intern

May 2023 - Aug 2023

NCR Corporation, Atlanta, GA

- Developed frontend features in Next.js, enhancing acquisition & retention, impacting 95% of customer transactions.
- Executed and optimized microservices architecture using GraphQL, and RESTful APIs, improving backend scalability.
- Created and integrated end-to-end testing suites using Jest and React Testing Library, ensuring high code quality.

PROJECTS

FitFusion-AI [*Streamlining Personalized Health and Fitness Guidance with AI*]

- Managed a Retrieval-Augmented Generation(RAG)-based system integrating Pinecone vector database &OpenAI GPT-4.
- Achieved 87% relevance accuracy by enabling real-time, context-aware query responses using a RAG pipeline.
- Reduced query latency to 9 seconds by optimizing embedding generation, document retrieval, and GPT-4 response synthesis.
- Designed iterative, context-aware query processing, achieving 90% user satisfaction from real-time feedback.

Drive-Thru Vision [*Revolutionizing Drive Thru Systems*]

- Secured a top 5 position and received an honourable mention at the NCR Global Hackathon.
- Engineered an ML-based CV model with OpenCV & Tesseract OCR, achieving 90% license plate accuracy.
- Applied the model to an NCR product, using a Flask backend and MongoDB for data administration.
- Built an interactive frontend using React.js, ensuring a seamless user experience for drive-thru operators and customers.

AWARDS & RESEARCH

- Co-authored a research paper on Dynamic Contextual Attention Network for Endoscopic Polyp Diagnosis.
- Secured a top 5 position and received an honorable mention at the NCR Global Hackathon.