Soham Kundu

skundu6@wisc.edu • sohamkundu.dev • linkedin.com/in/sohamkundu27 • (763) 406-1195 • US Citizen

Education

University of Wisconsin-Madison GPA: 3.7/4.0

B.S. Computer Engineering, B.S. Computer Science *Expected May 2027*

- Honors/Awards- MadData25 **Hackathon 1st Place** Winner (40+ Teams, 200+ Participants), Top 3 at University Madness Startup Pitch Comp (9 Universities, 25+ Teams), Eagan Foundation Scholarship - 5% Acceptance
- Coursework- Data Structures & Algorithms, Object Oriented Programming, Linear Algebra, Discrete Math, Multivariable Calculus, Digital System Fundamentals, Signals and Computation, Physics E+M & Mechanics

Experience

Blue Cross Blue Shield - *Software Engineering Intern* June 2025 – August 2025

- Developed an in-call support app for customer service reps that transcribes live calls, retrieves answers from internal databases, and delivers AI-generated responses in under **5 seconds**, serving **3M+** Blue Cross Blue Shield members
- Engineered a **Python**-based RAG pipeline with LangChain (custom agents, RetrievalQA) using hybrid keyword + semantic search, Azure real-time transcription, and **Azure** OpenAI LLMs, cutting CSR lookup time by **95%+**
- Integrated the pipeline into a desktop app with automated post-call summaries, enabling seamless real-time call assistance

FiPet - *Lead SWE/Co-founder (Part-Time) | fipet.dev* May 2025 – August 2025

- Architected a cross-platform mobile app with AI-powered, gamified financial education modules tailored for Gen Z audiences
- Scaled pre-launch to a **200+** user waitlist, using **ML-driven** personalization for adaptive learning
- Led a **15**-person engineering team to build the MVP using React Native, Firebase (Firestore, Auth, Cloud Functions), and **AI/LLM pipelines** for automated bonus quest generation and progress tracking

Computational Optics Group - *Machine Learning Researcher* January 2025 – May 2025

- Developed **ML** pipelines with optical sensor data, enabling non-line-of-sight object imaging and reconstruction
- Achieved **20%** reduction in neural network inference overhead, improving edge deployment feasibility on embedded systems
- Engineered embedded Python software for lasers and high-speed cameras for real-time photon capture, and trained/fine-tuned convolutional neural networks in Python/MATLAB on **3** large-scale datasets with varied scenes, noise, and sensor setups

Wisconsin Autonomous - *Software Engineer* September 2024 – May 2025

- Architected a perception system for autonomous vehicles enabling real-time lane, cone, and boundary detection
- Achieved **sub-100 ms** inference latency for safe high-speed navigation by optimizing ML algorithms with OpenCV
- Integrated the perception module into a software pipeline coordinating sensor fusion, control systems, and vehicle decision-making, collaborating with a **30+** person team on Git-based workflows and **CI/CD** pipelines

Projects

ClaimReady | *useclaimready.ai* | *github.com/sohamkundu27/MadData25*

- Co-founded a **5x** award-winning AI web application that generates complete home inventories and reduces insurance claim valuation time from **20+ hours** to under **2 mins**, accelerating post-disaster recovery for homeowners and adjusters
- Scaled to **350+** users by deploying an image valuation pipeline, via Docker AWS EC2, using YOLO11 for detection, Gemini API for brand/price identification, and SupabaseDB; processing **1,500+** images and valuing **$800K+** in items
- Recognized as "investor-ready" by judges; acknowledged by YC partners and UW-Madison CS Dept for technical excellence

Hyperacing | *hyperacing.us* | *github.com/Hype-Racing*

- Created and scaled a Formula 1 analytics platform to **1000+ users**, delivering AI insights and live telemetry in real-time
- Integrated Firebase Auth for secure login and session management, and used F1 WebSockets to stream real-time race data
- Engineered a Next.js frontend with dynamic dashboards, a live F1 news feed, and live analytics

Voice-Controlled Robot Arm | *Demo* | *github.com/sohamkundu27/UCBerkeleyAIHackathon*

- Architected a **full-stack** voice interface to control a robot arm by transcribing real-time speech into robot commands
- Validated the entire pipeline with async audio input and physics-based simulation built in Python, achieving sub-200 ms end-to-end response time and **90%** task completion accuracy for natural, hands-free operation
- Engineered a Claude **LLM** pipeline to convert prompts like "clear the table" into JSON robot actions executed via MCP

Greenhouse IoT Monitoring System | *github.com/sohamkundu27/GIoTMS*

- Innovated an IoT system tracking temp, humidity, and brightness on Raspberry Pi in Python, optimizing GPIO control
- **10K+** data points transmitted daily in real time by configuring Azure IoT Hub and **REST APIs**
- Deployed a Dockerized Django-React app on Azure App Services, enabling remote monitoring by parents in India

Technical Skills

Languages - Python, Java, JavaScript, C/C++, SQL
Frameworks/Libraries - Django, Flask, Matplotlib, NumPy, OpenCV, Pandas, PyTorch, Scikit-learn, TensorFlow, .NET
Developer Tools - AWS, Azure, CI/CD, Docker, Git/GitHub, Google Cloud, Linux, REST APIs, Test-Driven Development