Suyash Nagumalli

Urbana, IL ● usasuyash@gmail.com ● linkedin.com/in/suyashnagumalli/ ● github.com/faddock ● suyashnagumalli.com

EDUCATION

University of Illinois at Urbana-Champaign | Grainger College of Engineering
Master of Computer Science in Machine learning   Expected Graduation: May 2025
Bachelor of Science in Computer Science   May 2024
Relevant Coursework: Data Structures, Algorithms, Machine Learning, Deep Learning, Software Engineering, LLM post pre-training, Artificial Intelligence, UI Design, Database Systems, Internet of Things, Compilers, (CUDA Parallel Programming, Computer Security)

SKILLS

Languages: Proficient: Python(7yrs), SQL(4yrs), C#(2yrs) Intermediate: C++(2yrs), JavaScript(2yrs), Java(1yr) Beginner: Dart, Swift, Go
Software: AWS, GCP, Docker, Git, GitHub Action CI/CD, Jenkins, Postman, Agile, Jira, Langchain, HuggingFace, TensorFlow, PyTorch, NumPy, Pandas, Matplotlib, CV2, MySQL, Neo4j, ELK, ElasticsearchDB, FastAPI, Streamlit, React, Angular, Flutter, iOS, Unity, VR, Figma

EXPERIENCE

AbbVie | Pharmaceutical Company   Champaign, Illinois
Software Engineer (GenAI)   **Aug 2023 - Present**
- Built a prompt library with 15 REST APIs using FastAPI, architected the backend in NoSQL, developed the frontend in Streamlit, secured executive approval leading to a $2 million contract with Accenture, and advised early-stage vision and development
- Chunked and vectorized gigabytes of clinical study reports and investigative brochures, expert in prompt engineering, developed a rag pipeline to benchmark four embedding models using open-source ragas toolkit, all-MiniLM-L6-v2 fastest with 90% faithfulness
- Developed an integrated dual-LLM chat interface that saves 30 seconds per query, boosting daily productivity by 50 minutes
- Led a cross-functional design initiative with UI/UX and business teams to plan a V2 prompt management software
- Pioneered Langchain JavaScript and Python compatibility for internal LLM deployments
- (part-time during schoolyears and breaks, full-time during summer 2024)

VR Software Engineer   **Jan 2023 – Jul 2023**
- Built a POC VR scenario for doctors to converse with eight underrepresented patient avatars during informed consent
- Optimized LLM response generation pipeline, reducing latency from 7 seconds down to 800ms
- (Part-time during schoolyears and breaks, full-time during summer 2023)

Healthcare Engineering Systems Center | UIUC   Champaign, Illinois
VR Software Research Intern – Extra-Corporeal Membrane Oxygenation (ECMO) Project   **Jun 2022 – Jul 2022**
- Built a VR learning module in Unity to train doctors on the ECMO machine for providing life support to patients
- Got positive feedback from doctors at OSF Healthcare for developing 30 minutes of interactive content and 2 quizzes

Showingly | Showing Management Platform   Denver, Colorado
Software Engineer Intern – Frontend Web Development   **Jun 2021 – Aug 2021**
- Designed and built a CRM using ReactJS, MongoDB, and Agile methodology, for managing 1000+ agent records
- Learned React-Redux, Git, Scrum, and coding best practices and submitted 12 Pull Requests as a freshman intern

PROJECTS

Improving RAG Retrieval methods | LLM post pre-training CS 598 Project (~70 hours) link   **Oct - Dec 2024**
- Comprehensive literature review, explored limitations, proposed hybrid retrieval approach combining CDE and graph RAG
- Preliminary results show graph RAG (2024) outperforms CDE (2024) and RAG (2020) by 30% on Rouge-L F1 score

Audit of Twitter | Social Spaces CS 567 Paper (~70 hours)   **Oct - Dec 2024**
- Ran audit of X's "For You" feed, analyzing 201k tweets across nine countries, working with two PhD students
- Observed 73.1% political right-leaning bias, 29.9% US-centric bias & amplification of users @alexjones, @elonmusk

Considerthis.ai | UC Berkeley AI Hackathon (~30 hours) link   **Jul 2024**
- An emotionally intelligent multi-agent conversation tool used by educators to empower students to stay curious
- Successfully collaborated with UCBerkeley, UWashington and CUBoulder students to pitch to six judges

Spndmate | Cozad New Venture Challenge 2024 (~60 hours) link   **Feb – May 2024**
- A fresh perspective to group-bill splitting and personal budgeting for young adults, powered by Splitwise
- An Android/iOS app built using Flutter, Android emulator and Material UI, with Google Sheets backend API

Spotify LED Controller | Internet of Things CS 437 Project (~30 hours) link   **May 2023**
- Program to sync LED lights to song playing on Spotify, enhancing user immersion; successfully tested with 30+ songs
- Written in Go on Raspberry Pi; uses Spotify API, Bluetooth, and a custom album-art based coloring algorithm

ACTIVITIES

Emergent Abilities of LLMs papers / Led 90-minute presentation and discussion of two papers with 10 LLM researchers   **2024**
Python Private Tutor / taught Python basics to three sixth graders; Positive feedback from parents: "Great Teacher"   **2024**
Engineering Open House Volunteer / strong leadership skills in engaging 50+ middle-schoolers and parents   **2024**
iOS App Development / "Timely" / a personal productivity tracker for monitoring time spent on daily activities   **2024**
First place Hackathon Winner / PowerBI & Data Analytics project with AbbVie / ~50 competitors   **2024**
ElasticON Chicago Conference Attendee / Insight into ELK latest features, professional networking   **2024**
Music / Released two original songs on Spotify, Ableton production, ~2000 streams, 19 monthly listeners   **2024**