Yash Thakkar

ythakkar@cmu.edu • yash-thakkar.com • linkedin.com/in/yashthakkar21/ • San Francisco Bay Area

EDUCATION

Carnegie Mellon University, Pittsburgh, PA   Expected Graduation: December 2025
Master of Information Systems Management, **CGPA - 4.04**
Coursework: Machine Learning, Deep Learning, Distributed Systems, Data Science, Generative AI, Statistics, Agile

Dwarkadas J. Sanghvi College of Engineering, Mumbai, MH, India   Graduated: May 2022
Bachelor of Engineering in Mechanical Engineering, **CGPA - 3.92**

EXPERIENCE

Software Engineer Intern   May 2025 - August 2025
Newfront Insurance, San Mateo, CA
Applied LLM pipelines to automate benefit plan documents processing at scale

- ○ **AI-Powered Document Processing -** Reduced benefits plan data extraction and review turnaround by >**90%** and saved **$175k/year** by developing AI-driven pipelines using Next.js, FastAPI, GraphQL, PostgreSQL, and OpenAI APIs
- ○ **LLM Pipeline Optimization** - Improved **recall** by **7%** and **true positive accuracy** by **5%** by chunking extractions and building specialized OCR and extraction agents orchestrated in parallel using asyncio and semaphores
- ○ **Monitoring & Analytics -** Increased operational reliability by implementing **20+ Datadog alerts** and an analytics dashboard to track adoption, usage, and cost savings, enabling data-driven product improvements

Software Engineer   September 2024 - December 2024
Carnegie Mellon University, Pittsburgh, PA
Built and optimized an academic web platform for reusability, accessibility, and scale

- ○ **Frontend Development -** Boosted **user engagement** by **23%** by developing core frontend features using React, Redux, TypeScript, and Tailwind CSS
- ○ **Accessibility Enhancement -** Expanded potential **user base** by **2.5x** by integrating i18n for multilingual support

Head of Software Engineering   January 2022 - August 2024
Advect Systems Pvt. Ltd, Remote, India
Led product development and automation pipelines for a 3D printing & drone e-commerce startup

- ○ **Software Development Life Cycle** - Led the **end-to-end development** of Nanofactory, a 3D printer management app with a decentralized architecture using Svelte (JavaScript library), Flask (Python Backend), and WebRTC
- ○ **CI/CD & Automation -** Reduced **time-to-market** by **40%** and improved **team productivity** by **34%** by implementing a CI/CD pipeline (Docker, GitHub Actions, Cloudflare) and Python API integrations
- ○ **Test Coverage & Quality Assurance -** Reduced **bugs** by **13%** by expanding test coverage with Playwright and Pytest

Business Analyst   August 2022 - November 2022
Quantiphi Analytics Solution Pvt. Ltd, Mumbai, IN
Data Insights & Product Strategy in Analytics Consulting

- ○ **Requirements Gathering -** Reduced **dev costs** by **35%** by translating business needs into technical specifications

PROJECTS

Speech Recognition using Transformers - GitHub   November 2024 - December 2024
- ○ **Transformer Model Development -** Achieved a **10.9% character error rate** (CER) by building an attention-based end-to-end speech transcription model using transformer architectures.

AI-Powered Design Generation - GitHub   September 2024 - December 2024
- ○ **Generative AI Pipeline -** Reduced **Amazon listing time** by **39%** by developing a platform that converts sketches into production-ready images, using BLIP, Llama, BART Large-CNN, and Stable Diffusion-XL

SKILLS & CERTIFICATIONS

Programming Languages: Python, Java, C++, TypeScript, JavaScript, SQL, Bash
Tools: Docker, Amazon Web Services, Google Cloud Platform, GraphQL, Linux, Datadog, Hadoop, Spark, FastAPI
AI & ML: TensorFlow, PyTorch, Scikit-learn, Neural Networks, Matplotlib, Pandas, Numpy.
Certifications: Data Structures and Algorithms, Database Management Systems.