Ashwin Rachha

 $ashwin.rachha@gmail.com \mid linkedin.com/in/ashwinrachha \mid +1-540-824-8748 \mid github.com/ashwinrachhavt \mid Website \mid +1-540-824-8748 \mid github.com/ashwinrachhavt \mid Website \mid +1-540-824-8748 \mid github.com/ashwinrachhavt \mid +1-540-824-8748 \mid github.com/ashwinrachhavt \mid +1-540-824-8748 \mid +1-540-8248 \mid$

Education

Virginia Tech

M.S. (Thesis) in Computer Science; GPA: 4.0/4.0

Blacksburg, VA

Aug 2021 – Dec 2023

PICT

inci (incesse) sin comparen sesences, cilin 410/410

Pune, India

B.E. in Computer Science; GPA: 8.70/10.00

Jul 2016 - Nov 2020

Skills

Programming: Python, JavaScript, Go, SQL, MongoDB, Elasticsearch, Snowflake

ML/AI: PyTorch, Hugging Face, LangChain, LangGraph, Pinecone, FAISS, W&B

Web & Cloud: Django, React, NextJS, TypeScript, TailwindCSS, FastAPI, AWS, PostgreSQL

DevOps: Docker, Kubernetes, ArgoCD, Celery, Redis, Datadog, Sentry

Experience

Finally

Boca Raton, FL

Tech Lead & AI Product Engineer

Feb 2024 - Present

- Classify AI AI Bookkeeper: As founding engineer, built the flagship AI bookkeeper application from ground up using few-shot classification with Elasticsearch retrieval and Pinecone RAG database. Led a team of 3 engineers implementing LangChain workflows with Redis semantic caching and W&B observability. Processes 50,000+ transactions daily with 80% reduction in manual categorization time.
- Bank Connections Infrastructure: Architected core banking infrastructure using Plaid and Teller APIs that powers every
 application at Finally including Corporate Cards, Bookkeeping, and Bill Pay. Built OAuth flows, secure token management, and
 automated statement processing with OCR integration for PDF-to-CSV conversion.
- Enhanced Classification Pipeline: Integrated Plaid enriched transactions into Classify AI with merchant grouping optimizations and Celery task processing. Implemented advanced merchant categorization algorithms that significantly improved classification accuracy through additional transaction metadata and ML-driven pattern recognition.
- Cash-Based Underwriting System: Developed critical risk management system for Corporate Cards product that dynamically adjusts credit limits based on real-time cash balances and transaction patterns. This foundational underwriting method underwrote over \$3M in credit for 50+ companies in 3 months and became the standard approval basis for all Finally products.
- Real-time Reconciliation Engine: Built automated reconciliation system matching Plaid-synced bank transactions with accounting entries, reducing month-end close times by 50% and eliminating manual reconciliation errors for bookkeeping clients.
- Led cross-functional engineering team of 3 engineers building scalable infrastructure using Django, React, PostgreSQL, Kubernetes, and ArgoCD with comprehensive monitoring via Datadog and Sentry. Implemented CI/CD pipelines and established engineering best practices.

UNAR Labs

Portland, ME

- Machine Learning Engineer

 Jun 2023 Aug 2023

 Developed backend system enhancing multisensory information access for visually impaired individuals using computer vision and
- o Architected and optimized data pipelines for preprocessing, model training, and inference using OpenCV, PyTorch, Transformers, and FastAPI.
- Deployed scalable ML solutions on GCP with Docker containerization and HuggingFace model integration for accessibility applications.

Outreach

NLP technologies.

Seattle, WA

Machine Learning Engineering Intern (ML Platform Team)

May 2022 - Aug 2022

- Built scalable NLP inference system reducing BERT deployment time by 95% (3–4 days to 30 minutes) for topic detection and sentiment analysis on GKE.
- Developed PySpark + MLflow pipeline for text processing and deployed ONNX models on NVIDIA Triton.
- o Built Golang/Python microservices with Docker, unit tests, and CI/CD on CircleCI.

Mindbowser Inc

Pune, India

Software Engineer, ML

Dec 2019 - Feb 2021

 Developed facial expression recognition system for CRM meeting analysis using VGG-19 CNN transfer learning (73% accuracy) with MongoDB GridFS storage.

Projects

Gurukul LMS Platform — Django, NextJS, TypeScript, TailwindCSS, Prisma, Supabase, Docker, Vercel LIVE — CODE

- Master's Thesis project resulting in IEEE SouthEastCon publication LLM-enhanced learning environment with RAG and guardrails promoting Socratic thinking for DSA problem-solving.
- Built comprehensive coding platform with custom code editor, GPT-4 chatbot integration, problem tables, and intelligent tutoring system guiding students through structured learning paths.

Neuralflow - AI-Powered Productivity App — NextJS, TypeScript, Clerk Auth, Supabase

LIVE — CODE

- Built productivity application combining Pomodoro focus timer with Kanban board for task management, featuring Clerk SSO authentication and responsive design.
- Currently integrating AI categorization and workflow structuring using LLMs to automatically groom and organize user tickets from unstructured input data.

Publications

LLM-enhanced Learning Environments for CS: Exploring DSA with Gurukul — IEEE FIE PAPER — Nov 2024

 Published research at IEEE Frontiers in Education Conference (FIE24) exploring specialized LLMs with RAG and Guardrails for enhancing CS education. Research directly resulted in the development of Gurukul platform with adaptive learning capabilities.

Explainable AI in Education — IEEE SouthEastCon

PAPER — Apr 2023

• Published research on XAI approaches for education with algorithmic and design choice framework. Research directly informed the development of Gurukul LMS platform with RAG-enhanced tutoring system.

Community & Writing

Technical Writing & Community Contributions

- Kaggle Expert (Top 1% in Notebooks) contributing data science tutorials and competition solutions to the global ML community since November 2020. kaggle.com/ashwinrachha1
- Active technical blogger on Medium and personal blog, writing about machine learning, AI applications, and software engineering best practices with focus on practical implementation guides. medium.com/@ashwin_rachha