

Ankit Badatala

703-989-0542 | ankitbadatala2019@gmail.com | www.linkedin.com/in/ankit-badatala

EXPERIENCE

Strategy	May 2025 – Present
<i>Senior AI Engineer, Personalization and Insights</i>	
<ul style="list-style-type: none">• Co-engineered ReAct and ReWOO LLM frameworks, leveraging 20+ specialized agents and MCP tool calling, allowing users to extract business insights from gigabytes of structured data.• Led a global team of 10 to build and deploy 3 enterprise LLM chatbots serving 5,000+ DAU across 23 countries.• Ensured LLM framework is model agnostic so agents can dynamically switch between GPT-4, LLaMA, and GROQ—cutting Time to First Token by 45%, reducing API costs by 25%, and boosting response accuracy by 15%• Optimized RAG pipelines with ClickHouse Database vector search, reducing hallucinations by 50% and decreasing Time to First Token by 20%• Led a team of 5 to productionalize a computer vision pipeline to convert hand-drawn sketches of graphs into a presentation ready dashboard.<ul style="list-style-type: none">* RF-DETR Object Detection model, mAP@0.50 score: 0.84* Feature Extraction + RBF-SVM Image classification model, optimized the model to shrink size to 6kb while maintaining 100% accuracy, ROC AUC of 1.0, and F1 Score of 0.9841• Drove adoption of AI products from zero to 20 global companies within three quarters, partnering directly with Fortune 100 companies such as Cisco and Pfizer to ensure seamless production deployments.• Helped build voice feature for Strategy App (Company mobile app with 29,000 DAU)<ul style="list-style-type: none">* speech to Text, Text to Text (Finetuned Mistral 7B LLM), Text to Speech Pipeline* Hosted Voice and LLM on AWS SageMaker. Optimized inference using vLLM infra	
JPMorgan Chase	Aug 2023 – May 2025
<i>Software Engineer, Personalization and Insights team</i>	
<ul style="list-style-type: none">• Developed a full-stack web app to perform data validation across hundreds of data pipelines. This tool is used daily by 40 ML engineers to improve data quality monitoring.<ul style="list-style-type: none">* Followed MVCS Java architecture and deployed to production using Kubernetes with 90% JUnit code coverage• Architected and deployed a solution to enforce a large scale data retention policy, impacting over 100,000 Chase customers, by using Terraform, Python, AWS, and Spinnaker• Built a python application to optimize AWS EMR cluster runtimes and pinpoint over-fitted instance types, reducing data processing costs by \$120,000 a year• Built a python application to identify and clean over 400 terabytes of noncurrent version data in AWS S3 buckets, saving \$72,000 a year in cloud storage costs• Orchestrated over 20 ETL pipelines in Control M for processing and storing petabytes of data	
JPMorgan Chase	Jun 2022 – Aug 2022
<i>Software Engineer Intern, Trading Desk Technology team</i>	
<ul style="list-style-type: none">• Developed a full-stack web app that consolidates key market data from 15 trading desks and integrates into internal risk analysis tools. Adopted by more than 70% of traders on the floor in the first month.<ul style="list-style-type: none">* Used React, Java, Springboot and deployed to production using Kubernetes with 85% JUnit code coverage	
Strategy	May 2021 – Jan 2022
<i>Software Engineer Intern, Information Security team</i>	
<ul style="list-style-type: none">• Developed a Python application that uses sentiment analysis and facial recognition to flag over 1,000 social media accounts impersonating the company/CEO for cryptocurrency scams.	

PROJECTS

AWS DeepRacer Competition | Python, Reward Functions, Reinforcement Learning

- Developed an AWS powered reinforcement learning model for an autonomous racing competition.
- Secured a top 1% ranking out of 200,000 developers globally

TECHNICAL SKILLS

Languages: Python (6 yrs), Java (4 yrs), SQL, Bash, HTML, JavaScript, CSS

AI/ML: LLMs (GPT-4, LLaMA, GROQ), RAG architectures, Multi-agent frameworks, Prompt engineering, Vector databases (ClickHouse), Computer Vision (YOLO, RF-DETR), Voice Models (Sesame, Whisper), Scikit-learn, ONNX, TensorFlow, PyTorch

ML Ops: AWS SageMaker, vLLM, GPU infra, benchmarking

Cloud/Infra: AWS, Terraform, Kubernetes, Spark, MongoDB

Frameworks/Tools: React, Spring Boot, JUnit, Control-M, Spinnaker, Splunk, Postman API, Android, Jules

Certifications: AWS Cloud Practitioner

EDUCATION

Georgia Institute of Technology

Master of Science in Computer Science, Machine Learning

Expected Grad: Dec 2026

- Online and Part-Time. Courses: Machine Learning for Trading

Ohio State University

Bachelor of Science in Computer Science and Engineering

- National Buckeye Merit Scholarship Recipient and Dean's List