

IMRAN AHMED

(773) 983-5863 | imranahmed.csai@gmail.com | [linkedin.com/in/imran12234](https://www.linkedin.com/in/imran12234) | github.com/imran12234 | [imrans.live \(portfolio\)](https://imrans.live) | USA CITIZEN

EDUCATION

DePaul University	Sep 2021 – Jun 2026
Master of Science in Artificial Intelligence	Chicago, IL
Bachelor of Science in Computer Science, GPA: 3.7 , Dean's List for 6 Quarters	
Relevant Coursework: Algorithms, Data Structures, Distributed Systems, Database Systems, Machine Learning, Neural Networks	
Certification: AWS Certified Cloud Practitioner (CLF-C02), Sep 2026	

RELEVANT PROJECTS

Vautyn: AI Agent Forensics and Cloud Attribution vautyn.com	Aug 2026 – Present
Python, FastAPI, AWS Lambda, CloudTrail, SQS, EventBridge, DynamoDB, S3, Terraform, React	
- Built a forensic correlation engine linking AI agent tool calls to AWS CloudTrail events using STS SourceIdentity session binding and action, resource, policy, and timestamp matching - Designed attribution logic that distinguishes unique matches, ambiguous candidates, and insufficient evidence, preserving uncertainty when a specific tool call cannot be established - Provisioned an event-driven AWS pipeline with Terraform, supporting delayed-telemetry retries and idempotent processing of replayed events - Validated attribution behavior across 9 local scenarios, including competing calls, delayed telemetry, and missing evidence; passed 371 automated local acceptance tests	
SplitAI: AI-Powered Expense Splitting Platform Live Demo: splitai.dev	Mar 2026 – Present
Java, Spring Boot, PostgreSQL, React, TypeScript, Claude API, Docker, GitHub Actions	
- Achieved 100% field accuracy on amount, participants, and split type across 60 labeled expense descriptions, by building a Claude API pipeline that parses natural language into validated JSON for equal, exact, and percentage splits - Cut settlement transactions by 45% on average and 83% for 20-member groups across 1,373 randomized test groups, by implementing a debt simplification algorithm guaranteed never to exceed naive pairwise settlement - Developed a Spring Boot and PostgreSQL backend with RESTful APIs and JWT authentication to support expense-splitting workflows - Deployed the application on DigitalOcean using Docker and GitHub Actions CI/CD, supporting 15 active users and validating functionality with a 45-test suite	
ChiGo: AI-Powered Trip Planning Web App Live Demo: chi-go-v2.onrender.com	Mar 2025 – Present
Python, Django, PostgreSQL, OpenAI API, Google Places API, Google Routes API, Docker	
- Won Best Use of AI and Best Design among DePaul senior capstone projects - Built a web app that turns a 10-question travel survey (budget, cuisine, activity level, neighborhood) into a personalized multi-day Chicago itinerary, with mapped stops and one-click activity swaps - Generated itineraries with GPT-4o-mini at 100% first-try structured JSON output across 20 tests, enriching 316 of 316 activities with Google Places coordinates and addresses; deployed on Render with Docker	
Cross-Family LLM Distillation with Multi-Task LoRA	Mar 2026 – Jun 2026
Python, PyTorch, Hugging Face Transformers, PEFT/LoRA, bitsandbytes, Qwen 2.5, Claude API	
- Distilled Claude Sonnet into Qwen 2.5 1.5B and 3B models for 5 NLP tasks at roughly 750x lower inference cost; generated 12,500 Claude-labeled synthetic examples for about \$9 - Fine-tuned with 4-bit multi-task LoRA; the 3B student closed 86 to 99% of the teacher performance gap on BANKING77, AG News, and IMDb, including IMDb as an unseen task - Built an evaluation harness and promote/block release gate that blocked CLINC150 and DBpedia-14 where quality regressed, versioning teacher model, prompt, and sampling config for reproducible runs	
WORK EXPERIENCE	
Black Fox Security Solutions	Jun 2024 – Aug 2024
Software Development Intern	Aurora, IL
- Developed Python scripts to automate file-validation checks and flag invalid inputs, supporting repeatable file-upload testing - Performed functional testing of file-upload features to evaluate input handling, identify defects, and verify expected application behavior - Authored technical documentation covering internal development workflows and testing procedures, providing a shared team reference	
TECHNICAL SKILLS	
Languages: Python, Java, TypeScript, C++, SQL	
Backend: FastAPI, Spring Boot, Django, Node.js, Express	
Frontend: React, Next.js, HTML, CSS	
Cloud: AWS Lambda, EventBridge, SQS, S3, CloudTrail, STS	
Databases: PostgreSQL, DynamoDB, MySQL, MongoDB	
Infrastructure & Tools: Terraform, Docker, GitHub Actions, CI/CD, Git, Linux	
Machine Learning: PyTorch, Hugging Face Transformers, PEFT/LoRA, scikit-learn, NumPy, pandas, Claude API, OpenAI API	