Animish Jain

732-668-0300 | aj655@cornell.edu | animishjain.com | linkedin.com/in/animishjain | github.com/AniJ15

Education

Cornell University, College of Engineering Ithaca, NY

Bachelor of Science in Computer Science Aug 2024 – May 2028

Master of Engineering in Computer Science (3+1 Program)

Relevant Coursework: Object Oriented Programming and Data Structures, Introduction to Machine Learning, Computer Vision, Discrete Structures, Engineering Data Science, Data Structures and Functional Programming

Experience

Machine Learning Engineering Intern, GEICO Jun 2026 – Aug 2026

- Building a retrieval-augmented memory pipeline for LLM agents in Python using Hugging Face sentence transformers (BGE, MiniLM), dense vector embeddings, and cosine-similarity search to retrieve semantically similar historical execution traces during inference, enabling adaptive context injection for improved agent reasoning
- Developing a modular case library framework in Python (curator, store, retriever, HTTP agent wrapper) integrated with LangSmith tracing and offline evaluation pipelines for scalable agent experimentation and rapid iteration
- Implemented embedding pre-computation, JSON-based case schemas, and sub-100 ms semantic retrieval supporting 5,000+ historical cases while enabling automated A/B and regression testing with $< 2\%$ performance degradation

Machine Learning Engineering Intern, Anote May 2025 – Aug 2025

- Benchmarked zero-shot object-detection pipelines using YOLOv8, DINO, and Faster R-CNN, implementing workflows in Python with PyTorch and Ultralytics to compare mAP, IoU, and other ML metrics across models
- Designed and deployed a smart orchestrator reasoning system using Python, Docker, and Ollama; engineered multi-agent workflows with LangChain and custom routing logic, reducing LLM inference latency by ∼25%
- Integrated MCP into LLM pipeline using Python, FastAPI, JSON Schema, and LangChain, expanding agent tool interoperability by 40% and enabling tool-augmented reasoning across internal and external system endpoints

Software Engineering Intern, Pfizer Inc. Sep 2023 – Feb 2024

- Queried over 50,000 internal license-request records using Splunk and SQL for downstream model training
- Developed a multimodal language sentiment analysis algorithm in Python using TensorFlow and scikit-learn with 5 individual machine learning models to evaluate the justification behind license and permission requests
- Applied optimization techniques including L1/L2 penalties, dropout regularization, and learning rate scheduling to improve model performance, achieving 97.6% accuracy for automated request approvals and denials

Projects

Equity Options Volatility & Risk Engine | *Python, NumPy, Pandas, SciPy* Apr 2026

- Built a vectorized SPY options pricing engine for cross-model comparative analysis implementing Black–Scholes, Binomial Tree, and Heston models, evaluated on 6.6M real contracts across 757 trading sessions from 2019–2021
- Conducted a large-scale out-of-sample implied volatility surface study with liquidity filtering on 548K+ SPY options; Found that binomial tree achieved best model with 44.8% of predictions within ±10% of observed prices
- Analyzed model robustness under calm vs. high volatility regimes (March 2020), finding Black–Scholes and binomial accuracy improved by ∼15–17% under high volatility while Heston declined ∼2%

GreenTab | *Python, JavaScript, Firebase, Chrome APIs, Supabase* Nov 2025

- Built a full-stack sustainability analytics Chrome extension using JavaScript, Manifest V3, and Chrome APIs to monitor real-time browsing activity and estimate the environmental impact of spending time on the internet
- Implemented user authentication using Google OAuth and persistent data storage with Supabase (PostgreSQL + SQL queries) to manage user profiles and browsing-impact telemetry to create personalized dashboards for users

Skin Cancer Classification App | *Python, TensorFlow, Scikit-Learn, OpenCV, SwiftUI* Sep 2024 – Jan 2025

- Developed a multiplicative-weights update ensemble algorithm to optimally combine predictions from heterogeneous classifiers, improving overall accuracy, precision, and robustness against model-specific bias by over 10%
- Built and deployed an iOS application using SwiftUI, integrating the trained ensemble model pipeline to allow real-time inference on mobile devices with a clean and accessible UI supporting seamless on-device classification

Technical Skills / Hobbies

Technical Skills: Python, Java, Keras, Pandas, Numpy, Tensorflow, Scikit-Learn, LangSmith, LangChain, SQL OpenCV, MCP, JavaScript, Docker, FastAPI, Ollama, Git, Splunk, Firebase, Supabase, Kubernetes, Linux, OCaml

Interests: New York Giants, Poker, DJ Mixing, Electric and Acoustic Guitar, Bhangra Dance, Bollywood Music