Arsh Banerjee

New York, NY | P: 201-918-1312 | ab2736@cornell.edu | linkedin.com/in/arsh-banerjee/ | arshbanerjee.com

WORK EXPERIENCE

AI Scientist - Senior Engineer - Verizon Oct 2026 – Current

- **Velocity Assist:** Driving development of an internal agentic AI assistant that helps 4,000+ employees retrieve enterprise knowledge through natural language. Own a key workstream spanning agent orchestration, RAG pipelines, evaluation, and low-latency backend services, while partnering with stakeholders to shape the product roadmap.

AI Scientist - Engineer III - Verizon Jul 2025 – Oct 2026

- **Velocity Assist:** Led the development and launch of nearly 20 specialized AI skills (e.g., 10-K financial analysis, fraud prevention, and cellular connectivity prospecting) for an enterprise conversational interface designed to accelerate the sales cycle.
- **AI Quote:** Supported the development of an LLM-powered search interface that translates natural language customer prompts into automated cart builds; successfully scales to support 1M+ transactions annually.
- **Explainable AI (xAI) Framework:** Engineered a low-latency implementation of LLM-generated explanations for Verizon Business customers, providing real-time, profile-based justifications for product recommendations to increase buyer confidence.
- **Publications:** Banerjee et al. (2026). APLE: A Framework for Adaptable and Personalized Live Explainability for Recommender Systems. To appear in the Proceedings of the 18th ACM Conference on Web Science (WebSci'26).

Software Engineering Intern - Verizon Jun 2023 – Aug 2023

- Developed and deployed scalable automation solutions to monitor Verizon's core network devices, reducing time spent by 50%.
- Engineered robust data pipelines using SQL and RESTful APIs to integrate and process metrics from distributed databases.

Machine Learning Intern - Verizon Jun 2022 – Aug 2022

- Designed and deployed a cloud-native computer vision pipeline for robotic coordination, enabling real-time object tracking.
- Integrated AWS-hosted AI models for scalable inference, enhancing autonomous multi-robot navigation.
- Contributed to model lifecycle management, including data preprocessing, model deployment, and performance monitoring.

EDUCATION

CORNELL UNIVERSITY New York, NY

Master of Engineering in Computer Science; Cumulative GPA: 4.0/4.0 Aug 2024 – May 2025

Relevant Coursework: Deep Learning, Computer Vision, Productizing Machine Learning, Applied Machine Learning

PRINCETON UNIVERSITY Princeton, NJ

Bachelor of Science and Engineering, Magna cum laude Aug 2020 – May 2024

Major in Computer Science; Minor in Statistics and Machine Learning; Optimization; Cognitive Science

Cumulative GPA: 3.56/4.0; Served as a TA for COS 445 (Economics and Computing)

Relevant Coursework: Natural Language Processing, Computer Vision, Computer System Design, Optimization

Research:

- **LLM Evaluation Frameworks:** Built custom evaluation pipelines using OpenAI's GPT APIs to assess large language models on robustness, fairness, and prompt sensitivity.

CERTIFICATIONS & SKILLS

Technical Skills: Python, R, SQL, Java, Go, PyTorch, TensorFlow, Keras, HuggingFace Transformers, XGBoost, OpenCV, Scikit-learn, Pandas, NumPy, AWS, Docker, Git, Linux, OpenAI API; NLP, LLMs, Transformers

Certifications: AWS Certified Developer – Associate (DVA-C02), Obtained July 2024

RECENT PROJECTS

HoneyRoom – LLM RAG Application for Private Equity

- Built a web-based AI assistant that autocompletes the ILPA Due Diligence Questionnaire by indexing and embedding documents from a virtual data room (VDR) using a Retrieval-Augmented Generation (RAG) pipeline.
- Leveraged OpenAI's GPT-4, LangChain, and Neo4j to generate grounded answers with citations.
- Designed the front end with Streamlit and deployed backend inference and document handling on AWS Lambda, enabling investors to query and synthesize deal documents efficiently.

Text-to-SVG Pipeline with LoRA Fine-Tuning

- Developed a generative image pipeline using Stable Diffusion fine-tuned with LoRA to produce sketch visuals from prompts.