Arushi Gupta

732-207-0530 arushi.gupta1@rutgers.edu linkedin.com/in/arushi-gupta11/ github.com/ArushiG11 arushigupta11.netlify.app

Professional Summary

Software Engineer with 2+ years of experience in data analysis and software development. Pursuing an M.S. in Computer Science with focus on AI applications. Skilled in Python, Django, React, and Power BI with a strong foundation in data pipelines and visualization. Adept at building scalable, user-focused solutions and currently exploring AI-driven applications.

Education

Graduate Exchange Program - Computer Science — Princeton University **Spring 2025**

Coursework: Systems & Machine Learning

M.S. in Computer Science — Rutgers University - New Brunswick, NJ **Jan 2024 – Present**

Coursework: Intro to AI, Database Systems for Data Science, Natural Language, Advanced Robotics

Technical Skills

Programming Languages: Python, C++, Java, JavaScript, TypeScript, SQL
Web & API Development: HTML, CSS, Node.js, React.js, FastAPI, Django, REST APIs, A2A (JSON-RPC), OAuth 2.0
Machine Learning & AI: TensorFlow, PyTorch, MindsDB, Hugging Face, BERT, NumPy, Pandas, Seaborn, Matplotlib
Data Engineering & Visualization: Power BI, Apache Kafka, MySQL, PostgreSQL, MongoDB, Firestore
Cloud & DevOps: AWS, Azure, Azure DevOps, Docker, Kubernetes, CI/CD, Jenkins, Git, GitHub
Software Engineering & Practices: Data Structures & Algorithms (DSA), SDLC, Agile, Scrum, Jira, DevOps

Experience

Robot Toolworx - Software Engineer Intern **May 2025 – Aug 2025**

- Designed and implemented a **RAG** pipeline connecting **Ingester, Datastore, and Compiler** services to enable semantic search and context-aware AI responses.
- Developed a data embedding layer using **OpenAI** and **pgvector**, enabling efficient vector storage and similarity search across 3D model metadata and extracted text.
- Set up **Langfuse** and **OpenTelemetry** for monitoring, tracking performance and data flow across the pipeline.

MAQ Software - Software Engineer **Jun 2022 – Dec 2023**

- Developed an **analytics dashboard** using **Power BI** for the client's sales team to visualize partner performance, revenue streams, and customer engagement, attracting over **2,000+** monthly visits.
- Refactored and indexed SQL queries to cut notebook execution time **from 45 to approximately 20 minutes**.
- Designed and automated **ETL pipelines** using Azure Data Factory, reducing refresh and execution time by **58–60%**.
- Migrated report data sources to Azure Data Lake Storage (ADLS), **reducing data refresh time by 60%**.
- Executed **validation test cases**, ensuring data integrity through **BVTs and staging checks.**

Projects

CodeStory – AI-Powered Code Quality Analyzer — Python, Scikit-learn, CodeBERT, FastAPI, Docker **Sep 2025**

- Built ML pipeline to analyze GitHub repos and predict **bug-prone modules** using code metrics and repository history.
- Integrated **CodeBERT and Gemini API** to summarize code modules and suggest refactoring improvements.
- Visualized model predictions and explanations through **Streamlit dashboard**, estimated to reduce manual review time by roughly 40%.

AutoJobTrack – Smart Job Application Tracker — Gmail API, Gemini, Cloud Run, Firestore, Cloudflare **Oct 2025**

- Created a tool that connects to a user's **Gmail** and automatically tracks job-related emails, extracting details like company name, role, and interview dates.
- Used **Gemini** to read and structure email content, storing results in **Firestore** and syncing them to **Google Sheets**
- Built the backend on **Cloud Run** with real-time updates through Gmail **Pub/Sub**, and hosted an interactive dashboard on **Cloudflare Pages** for viewing applications.

Achievements

Winner – RAISE-25 Hackathon **Apr 2025**

Edward J. Bloustein School of Planning and Public Policy, Rutgers University

- Analyzed **15,000+ multilingual news articles** using a fine-tuned **BERT model**, achieving **87% sentiment classification accuracy** to study global perceptions of AI under the theme *"Utopian or Dystopian"*.
- Developed an **interactive dashboard** integrating sentiment trends with socio-economic indicators, providing actionable insights on **AI's educational and policy impact across regions**.