

Varun Iyer

908-745-8400 | varun.iyer1305@gmail.com | [linkedin.com/in/varun.iyer](https://www.linkedin.com/in/varun.iyer) | github.com/varun.iyer |

Work Authorization: US Citizen

EDUCATION

Rutgers University—New Brunswick

Expected: May 2027

Bachelor of Science in *Computer Science*

New Brunswick, NJ

Minor: Data Science

Dean's List: Fall 2025

Relevant Coursework: Data Structures & Algorithms, Intro to Data Science, Intro to Linear Algebra, Principles of Information and Data Management, Statistical Inference for Data Science

SKILLS

Languages: Python, SQL, Java, C Programming

Data & ML Tools/Libraries: Pandas, NumPy, Scikit-Learn, PyTorch, Hugging Face Transformers, NLTK, Power BI, Tableau

Techniques: Machine Learning, Feature Engineering, NLP, Data Preprocessing & Cleaning, Model Evaluation, Statistical Analysis, Data Visualization

Platforms & Environments: AWS, Azure AI, Microsoft Foundry, SharePoint, Microsoft Copilot, Visual Studio Code, GitHub

Other Skills: Word, Excel, PowerPoint, Prompt Engineering

EXPERIENCE

Strategic Tech Solutions — Data Science Intern

Sep 2025 - Present

- Fine-tuned and feature-engineered a RoBERTa-based NLP classification model achieving 85% accuracy categorizing 30,000+ semi-structured consumer complaint records, using an 80/10/10 train/test/validation split (scikit-learn) with Python, PyTorch, Hugging Face Transformers, and AWS cloud infrastructure.
- Built an NLP data preprocessing pipeline (Python, NLTK, scikit-learn) using TF-IDF vectorization and scikit-learn evaluation metrics to clean, structure, and sentiment-tag unstructured text, standardizing inputs for downstream classification and analytics workflows.
- Partnered with a cross-functional intern team to build a Copilot-based chatbot on Azure AI and Microsoft Foundry, using vector embeddings and prompt engineering to retrieve and answer natural-language queries from various SharePoint-hosted documents.

PROJECTS

Gradio Interface — [Q&A RAG Chatbot](#)

Aug 2026

- Combined multiple RAG components → document loaders, text splitters, vector-embedding models, and a vector database into a single pipeline, constructing a fully functional Q&A chatbot under a frontend Gradio Interface.
- Leveraged LangChain and LLMs to retrieve and answer natural-language questions grounded in content extracted from more than 30 PDF documents.

Natural Language-to-SQL Generator — [LLM Query Pipeline](#)

Feb – May 2026

- Spearheaded backend development, implementing a prompt builder python file that performs basic text processing to create a prompt that writes and returns an SQL query based on real-world home mortgage disclosure act data (HMDA).
- Engineered a reusable SSH-tunnel and integrated it with the local LLM-Pipeline, enabling several end-to-end submissions of natural-language questions and execution of the appropriate SQL query against a remote Postgres database in the iLab virtual environment.

Multi-Client Interaction — [Interactive Chat Server](#)

Apr- May 2026

- Spearheaded development of a concurrent, multi-client chat server using TCP sockets and poll()-based I/O multiplexing, supporting 50+ simultaneous connections with minimal CPU overhead.
- Coordinated with one other team member to develop a lightweight custom protocol for real-time registration, status updates, and private messaging, with structured error handling achieving 100% uptime under multi-client stress testing.

Rising Tides — [Regional Risk Analysis](#)

Jun – Jul 2024

- Developed a Java program using a custom 3-layer hierarchical linked list (state → county → community) to structure and analyze regional environmental risk data.
- Synthesized cross-domain datasets (demographics, pollution, flood risk) to identify correlation factors across hundreds of data points.
- Designed multi-level traversal and insertion algorithms to reduce search complexity for regional data queries.

CERTIFICATIONS

- **IBM Data Science Professional Certifications (Coursera)** — Data Science Methodology, What is Data Science, Tools for Data Science, Python for Data Science, Databases & SQL for Data Science with Python, Data Visualization with Python, Data Analysis with Python, Generative AI Applications using RAG & LangChain.

LEADERSHIP & ACTIVITIES

Boy Scouts of America: Senior Patrol Leader

Sep 2021 – Mar 2022

- Served as the troop's senior student leader, running weekly in-person and virtual meetings for up to 25 scouts and taking ownership of consistent life-skills instruction across the group.
- Held the Patrol Leadership Council accountable to its own goals, running monthly progress check-ins tied back to the term's Annual Planning Conference so no patrol lost track of its commitments.
- Owned end-to-end planning and execution of a multi-day canoeing and backpacking expedition for a 15-scout roster along the Mullica River, coordinating directly with adult troop leaders and patrol leaders to deliver the trip safely from start to finish.

Kumon North America : Instructor

May 2020 – June 2023

- Directed individualized and small-group instruction for 30-35 students at a time, adapting teaching approach per student to build measurable gains in math and reading proficiency.
- Maintained a tracking system (100+ attendance, assignment, and participation records in Excel) that instructors and parents relied on for an accurate read on each student's (ages 6-14) progress.

Rutgers Design Lab Club

Sep 2023 – Dec 2023

- Worked in a 20–25-person cohort applying structured design-thinking frameworks to open-ended problem statements, from ideation through proposed solution.
- Partnered with peers to surface high-potential problem areas and shape solutions aimed at measurable social impact.