

Shimanto Bhowmik

shimantobhowmik.github.io | (585) 360-8638 | sb6778@g.rit.edu | linkedin/in/shimanto-bhowmik | github/ShimantoBhowmik

EDUCATION

Rochester Institute of Technology

Bachelor of Science in Computer Science. GPA: 4.0/4.0

Rochester, N.Y.

Aug. 2021 – Dec. 2025

WORK EXPERIENCE

Software Engineering Intern

May 2025 – Aug. 2025

Amazon AGI

Seattle, WA.

- Spearheaded parallel experimentation for **Alexa+**, implementing traffic splitting and consistent hashing (**Kotlin/Guava**) to boost ML experimentation velocity by **96% (days to minutes)**.
- Modernized **Alexa+** experimentation platform from single-lane to multi-tenant, boosting experimentation accuracy by **60%** by implementing deterministic identity routing and isolation guardrails.

AI Researcher

Sept. 2024 – May 2025

Stanford Artificial Intelligence Laboratory

Remote

- Expanded Bengali NLP evaluation resources by translating **20 major LLM benchmark datasets** into Bengali, enabling systematic evaluation of models for an underrepresented language spoken by over **250 million speakers**.
- Identified primary failure modes limiting Bengali LLM quality via **10-model, 8-benchmark evaluation**; tokenizer analysis surfaced materially higher tokens/word versus English, tracing issues to Unicode normalization and vocab coverage.

Software Engineering Intern

May 2024 – Sept. 2024

Amazon AGI

Seattle, WA.

- Drove platform modernization with a model release pipeline that enforced semantic versioning and automated checks, **compressing time-to-release by 60%** and streamlining the path from training to production.
- Operationalized automated testing end-to-end (pipeline regressions and post-deploy checks), **achieving a 90% reduction in update time** and solidifying model reliability across versions.

Machine Learning Engineer

Sept. 2023 – May 2024

Sage AI

Remote

- Amplified customer engagement by **42%** through **end-to-end ownership** of a full-stack CRM (**React, Rust, DynamoDB, Cognito**), delivering secure onboarding, real-time analytics and zero-downtime releases.
- Engineered a production-grade SEC filings ingestion pipeline in Python coupled with LLM-driven financial analysis, increasing **analysis speed/precision by 23%** and raising **bank decision accuracy by 36%**.

Research Assistant

Aug. 2023 – May 2024

RIT Computer Science

Rochester, N.Y.

- Increased student engagement by **25%** through an **NSF(National Science Foundation)-funded** rollout of an R sandbox and R interactive sessions alongside Python, enabling side-by-side comparison of data-science learning outcomes.
- Enhanced platform interactivity by **15%** by integrating **Thebe** and **Binder** with version-locked **R/Python** kernels for reliable in-browser execution and reproducible labs.

Software Engineering Intern

May 2023 – Aug. 2023

Amazon AGI

Seattle, WA.

- Led design and rollout of a rubric-driven feedback pipeline and evaluation dashboards for Alexa's next-gen Q&A LLM, **enhancing model performance by 9%** under standardized evaluation criteria.
- Decreased end-to-end model testing time by **12%** through a parallel, tabbed multi-LLM interface enabling concurrent prompts and rapid comparison as verified by Applied Scientists.

PROJECTS

Vision Voice ([Link](#)) | *HuggingFace, BLIP, TTS, Gemini, FastAPI, Angular*

- Implemented end-to-end pipelines for image-to-text, text summarization and text-to-speech by integrating **BLIP** and **Gemini** on **HuggingFace** to generate accessible image descriptions.
- Deployed SpeechT5-based **TTS** to convert captions to audio, delivering **up to 7.5 hours/day** in potential time savings for audio-description workflows.

Math Mania ([Link](#)) | *Machine Learning, TensorFlow, TensorBoard, JavaScript, HTML/CSS*

- Built an interactive platform for my niece, turning drills into game-like challenges with progress tracking to increase engagement.
- Designed and trained from scratch a **TensorFlow** model on **MNIST** dataset for digit classification using **TensorBoard** to refine architecture and hyperparameters.

TECHNICAL SKILLS

Programming Languages: C, C++, Java, Python, JavaScript, TypeScript, SQL, Rust, C#, .NET, Go, PHP, Swift, Kotlin

Frameworks & Libraries: Spring Boot, Node.js, Express.js, React, Next.js, Fast API, Django, GraphQL, Jest, JUnit

Databases & Data Tools: PostgreSQL, MySQL, MongoDB, Redis, Elasticsearch, Firebase, Pandas, NumPy, Matplotlib

Cloud, DevOps & Testing: Docker, Kubernetes, AWS, GCP, Azure, CI/CD, Git, GitHub

Machine Learning & Other: RLHF, NLP, LLM, Generative AI, TensorFlow, PyTorch, Keras, Scikit-learn, Transformers, LangChain, OpenCV, OpenAI API, OAuth, Apache Spark

AWARDS AND PUBLICATIONS

- Bhowmik, S. et al. "Evaluating LLMs' Multilingual Capabilities for Bengali: Benchmark Creation and Performance Analysis"
- NSF-funded Research Assistant - National Science Foundation
- Outstanding Undergraduate Scholar - RIT
- Asia Pacific ICT Award (APICTA) - China