

Sabbir Hossain

hossain.sabbir17@gmail.com | (647) 545-8842 | <https://linkedin.com/in/itsabbir> | <https://github.com/itsSabbir>
Toronto, ON | Open to relocation (US/Can) | TN Visa Auth.

Summary

Versatile Software Engineer and Data Scientist experienced across the full SDLC in DevOps/SRE practices. Proven ability applying software engineering principles to design, build, and deploy scalable systems, distributed applications, web services, and data solutions in cloud (AWS) environments. Adept at cross-functional collaboration and remote teamwork. Seeking entry-level Software Engineering, DevOps, SRE, Data Engineering, or Data Science roles utilizing Python, Java, C/C++, R, SQL, Docker, Kubernetes, CI/CD, Terraform, and artificial intelligence and machine learning.

Skills

Programming Languages: C, Python, R, Java, C++, JavaScript, TypeScript, Bash/Shell, HTML, CSS, SQL, MATLAB

Technologies & Tools: Linux, AWS, Docker, Kubernetes, JSON, Serverless (AWS Lambda), Terraform, Git, PostgreSQL, MongoDB, Redis

Software Development: Algorithms, Data Structures, OOP (SOLID), CI/CD, Agile, Cloud Computing, TDD, Microservices, ETL, Automated Testing

Frameworks & API Development: Node.js, React, MERN Stack, Flask, Django, Shiny, PyTorch, REST APIs, SOAP, GraphQL, WebSockets

Databases, Testing & DevOps Concepts: CRUD Operations, Data Modeling, SQL Optimization, Unit/Integration Testing, IaC, Monitoring Basics

Professional Expertise: Technical Documentation, Cross-functional Collaboration, Bioinformatics Analysis, Problem-Solving, Leadership

Other Skills: Microsoft Office (Excel, Word, PowerPoint), LaTeX, SDLC, DVCS, JIRA, Confluence, Technical Writing, Data Visualization

Experience

Junior Full Stack Software Engineer — Johns Hopkins University, Baltimore, MD (Remote)

Sept 2022 – Present

- Architected & maintained open-source full-stack platform (Python, R, JS, C) using microservices (Docker/K8s) & SOLID; reduced analysis load 83% via caching (Redis).
- Engineered scalable ETL pipelines processing 750+ TB multi-omics data on HPC using Python, R, SQL & ML (SVM, RF); accelerated biomarker discovery 40% & cut analysis time 40%.
- Implemented automated data quality/anomaly detection (K-Means, DBSCAN via TensorFlow) in CI/CD; improved data integrity 30% & validated biomarker analysis software.
- Optimized GraphQL/REST APIs (API Gateway, caching) reducing latency 35%; built interactive portals (React, Next.js/TS, D3) enhancing user engagement 56%.
- Deployed fault-tolerant microservices on AWS (EC2, S3, Lambda) using IaC (Terraform); established CI/CD (Jenkins, Git/Actions) with 95% TDD coverage achieving zero-downtime deployments.

Backend Software Developer Intern — Outlier (AI Startup), Oakland, CA (Remote)

Mar 2024 – Nov 2024

- Built & refactored GraphQL/REST APIs (FastAPI/Python, Spring Boot/Java), executing monolith-to-microservices migration; designed serverless functions (AWS Lambda) reducing infra costs 30%.
- Developed validation scripts (Python/Java) for AI-generated code applying optimizations; improved model efficiency 5% & reduced code errors 10%.
- Assisted senior engineers debugging production bottlenecks (**troubleshooting**); contributed to 20% faster critical endpoint response times & implemented integration tests for GenAI models.

Software Development Research Assistant — University of Toronto, Toronto, ON (Hybrid)

Sept 2019 – Apr 2024

- Engineered full-stack bioinformatics tools (Python, R, C++, Java) applying OOP; automated workflows saving 30+ hrs/week across 7 teams.
- Implemented microservices architecture (GraphQL/REST APIs over PostgreSQL); applied **data modeling** & optimized queries reducing data retrieval latency 25%.
- Established DevOps practices (Docker/Kubernetes) for HPC deployments cutting environment setup time 50%; led Agile adoption improving team throughput.

Projects

Image Processing Pipeline Server

- Architected multi-threaded C server (POSIX threads, sockets) handling 100+ clients (<100ms latency); implemented TDD & CI/CD.

Key Tech: C (pthreads), Python, Linux Sys Progm, Multithreading, Sockets, TDD, CI/CD.

Bioinformatics Pipeline for Gene Expression Analysis

- Built containerized (Docker/K8s) bioinformatics pipeline using Nextflow for reproducible RNA-seq analysis, reducing manual effort 40%.

Key Tech: Nextflow, R (Bioconductor), Python, Docker, K8s, Workflow Automation.

Stock Market Prediction Pipeline

- Architected real-time stock prediction system integrating Kafka streaming data & ML models (RF, XGBoost); deployed via Dockerized Flask API.

Key Tech: Python (Scikit-learn, Flask), Kafka, Docker, ML, REST API.

Programming Language Learning Hub (Open Source)

- Created & maintained open-source platform (React, Node.js); managed 40+ contributors, reaching 1000+ developers globally.

Key Tech: React, Node.js, Git/GitHub, Community Mgt, Technical Writing.

Awards & Achievements

- Plenary Speaker, NCRC Harvard 2024 (*1 of 12 selected from 5,000+ applicants*)
- Best Detailed Oral Presentation, ABRCMS 2023 (*Top in division; 80/3,500+ attendees*)
- Best Poster Presentation, ABRCMS 2024 (*vs. 150+ graduate presenters*)

Education

Bachelor of Science, (Hons) Computer Science, Bioinformatics & Computational Biology — University of Toronto

June 2024

- GPA: 3.96 / 4.0
- Relevant Coursework: Data Structures & Algorithms, Software Design & Engineering Principles, Systems Programming, Algorithm Design & Analysis, Theory of Computation, Operating Systems, Database Systems, Machine Learning, Distributed Systems, Cloud Computing, Computer Networks, Applied Bioinformatics, Systems Biology, Statistics & Probability, Calculus, Programming Languages (Python, C, R, Java), Web Technologies (HTML/CSS), Microsoft Office Suite (Excel, Word, PowerPoint).