

# Rajvir Singh

(716) 573-7432 | [rajvirsinghfeb16@gmail.com](mailto:rajvirsinghfeb16@gmail.com) | [portfolio](#) | [linkedin](#) | [github](#) | [leetcode](#) | San Francisco Bay Area, CA, USA

## EDUCATION

**University at Buffalo, State University of New York**

*Master of Science in Computer Science*

Buffalo, NY, USA

August 2025 - Present

**Indian Institute of Information Technology Kota**

*Bachelor of Technology in Computer Science & Engineering*

Kota, India

August 2018 - June 2022

## TECHNICAL SKILLS

**Languages:** Java, Python, TypeScript, JavaScript, Clojure, ClojureScript, C++, C

**Backend & APIs:** Spring Boot, Dropwizard, FastAPI, REST, GraphQL (Apollo), gRPC

**Frontend:** React, Next.js

**Databases:** PostgreSQL, MongoDB, Redis, SQL (Query Optimization, Indexing)

**Distributed & Data:** Apache Spark, Hadoop (HDFS), Microservices, Distributed Systems

**Cloud & DevOps:** AWS, Docker, CI/CD (GitHub Actions, Jenkins), Linux

**Testing & Observability:** JUnit, Mockito, Grafana, Kibana

**Core CS:** Data Structures & Algorithms, System Design, Scalability, OOP

**LLM Integration:** LangChain, LangGraph

## EXPERIENCE

**Software Engineer**

**Fastr**

**Mar 2023 - Mar 2025**

- Engineered and optimized the **customer-facing Fastr editor** by extending Penpot with ADA compliance and responsive layout support using **Clojure + TypeScript**.
- Designed full-stack dashboards with **ClojureScript frontend + gRPC-SQL backend**, optimizing complex **SQL queries**, improving analytical speed by **35%** and reducing load times by **500 ms**.
- Resolved **30+ critical production issues** through root-cause analysis, structured logging, and validation improvements, reducing outages by **40%**.
- Automated CI/CD workflows using **GitHub Actions, Jenkins, and Docker**, improving deployment reliability and accelerating release cycles by **25%**.
- Established monitoring and alerting using **Grafana and Kibana**, reducing incident detection time by **20%** and lowering error rates by **15%**.
- Led peer code reviews and improved engineering standards, enhancing release stability and team delivery efficiency.

**Software Engineer**

**Cvent India**

**Jan 2022 - Feb 2023**

- Developed planner and guest interfaces using **React, TypeScript, and Next.js**, reducing navigation complaints by **30%**.
- Architected and maintained **4 backend services** in **Java + Dropwizard** serving **50k+ event users**.
- Integrated **Apollo GraphQL** with **REST APIs** to power real-time dashboards, achieving **20%** faster response time.
- Optimized backend performance via **caching and indexing strategies**, reducing query execution time by **40%**.
- Collaborated with cross-functional teams to deliver stable releases and improve sprint velocity by **15%**.

## PROJECTS

**Scalable Financial Analytics Pipeline** | *Hadoop (HDFS), Spark/PySpark, Docker, Spark ML*

**2025**

- Configured and deployed a **Docker-based multi-node Hadoop cluster** (NameNode/DataNodes) and ingested large-scale equities time-series data into **HDFS** for distributed processing.
- Orchestrated **PySpark** data-cleaning + feature engineering pipelines (deduplication, forward-fill for missing values, per-ticker normalization) to prepare OHLCV signals for modeling.
- Constructed and deployed **GMM-based anomaly detection** (score =  $1 - \max p(k|x)$  with 1%/5% thresholds), **Logistic Regression/Random Forest** for next-day direction, and **K-Means** segmentation (silhouette-tuned, best  $k = 2$ ).
- Trained a **Linear Regression** volatility predictor using log-volume features ( $RMSE = 0.0104$ ,  $R^2 = 0.346$ ).

**Symbolic Music Generation using Neural Networks** | *Python, Keras, Deep Learning*

**2022**

- Architected an end-to-end **sequence modeling pipeline** for piano melody generation from **18k MIDI files**, including pianoroll compression, temporal downsampling, and pitch transposition augmentation.
- Trained a 2-layer neural network (approx. 131K parameters) with autoregressive decoding and probabilistic sampling to reduce repetitive patterns.

**Land Use / Land Cover Classification** | *Google Earth Engine, Random Forest, SVM*

**2021**

- Formulated a **geo-spatial ML pipeline** in Google Earth Engine using Landsat 7 imagery to classify Built-Up vs Non Built-Up regions for Delhi (NCT).
- Derived NDVI/NDBI features and benchmarked CART, Random Forest, and SVM with 10-fold cross-validation (~66% accuracy).