Neha Joshi

Boston, United States| (+1) 6179924170
joshineha1905@gmail.com | https://www.linkedin.com/in/nehajoshi404/
Available from May 2025 for Full Time roles

EDUCATION

Northeastern University, Boston, United States
Master's of Science in Computer Science 01/2023 - 05/2025
Courses: Algorithms, Programming Design Paradigms, Data Mining Techniques, Foundations of Cloud Computing, NLP
University of Pune, Pune, India
Bachelor of Technology, Electronics Engineering 08/2016 - 10/2020
Courses: Robotics, Object Oriented Programming, Data Structures and Algorithms, Machine Learning, Computer Vision

TECHNICAL SKILLS

Languages	Java, Python, JavaScript, Typescript, C++, Shell Scripting, SQL
Web dev & Databases	Nginx**,** Kornshell**,** Tableau, SQL(PostgreSQL, MySQL), NoSQL(MongoDB, Cassandra)
Tools & Technologies	Git**,** AWS (Lambda, RDS, S3, EC2, OpenSearch, Quicksight, ALB), Apache Flink, Airflow

WORK EXPERIENCE

Staples, Boston, United States
Machine Learning Co-op 07/2024 - 12/2024

- **Designed and deployed scalable Python** FastAPI-based **RESTful APIs** to automate inference of LSTM & BERT-based pricing models on AWS, **reducing latency by 40%** through Kubernetes & Docker.
- **Optimized API performance** using **caching (Redis) and multithreading**, reducing response times by **30%** and improving system throughput.
- **Enhanced price forecasting pipeline on** Databricks by optimizing **PySpark & Hive** operations, **cutting data processing time by 80%,** and improving scalability.

Thoughtworks, Pune, India
Senior Software Engineer 04/2022 - 12/2022

- Engineered high-performance, **large-scale distributed ETL pipeline** using Apache Spark (Scala), Hadoop, and Python to build proprietary pricing recommendation algorithms, processing billions of records daily.
- **Reduced API response times by 20%** via **Redis caching**, eliminating redundant microservice calls and enhancing system efficiency.
- Optimized complex SQL queries in PostgreSQL**, improving query performance** by 50% through indexing, query restructuring, and effective use of EXPLAIN for analyzing execution plans.

Software Engineer 04/2021 - 03/2022

- Constructed a cross-system automation framework written in Java, based upon Java Messaging Service(Pub Sub), Java Spring, Junit 5, and Selenium which reduces the time to validate testing workflow manually by 80 %.
- Designed and built **backend APIs** for an internal Feature Toggle System using **Django,** reducing feature rollout time by 30% and deploying it on AWS EC2 for scalable performance.
- Developed a **Python**-based analytical tool for real-time streaming of user log data via **Kafka,** enabling in-depth platform usage analysis and decommissioning 2 on-premise servers, reducing infrastructure costs.

Sportsseam, Pune, India 06//2020 - 03/2021
Software Engineer

- Developed and implemented an end-to-end **ad detection pipeline** using Python and Rockset, leveraging audio fingerprinting and Query Lambdas to identify and verify ad occurrences within a specified channel stream.
- Achieved 93% accuracy by developing an Image Segmentation-based algorithm using **PyTorch** for human hair color prediction, with a customized training and testing pipeline in **AWS**.
- Eliminated manual data collection by designing and implementing robust JSON-based data ingestion pipelines. Automated video dataset generation using Python and ffmpeg.

PROJECTS

- **Multiplayer Kill Dr Lucky Board Game:** Developed a multiplayer **Java** game **using MVC architecture**, implementing design patterns (Builder, Factory, Singleton) and Java Swing UI, with robust testing using Junit 5.
- **Stack Overflow Replica:** Built a full-stack replica using **Node.js, MongoDB, and Angular.** Scaled **MongoDB** models **for 10,000 concurrent requests/min and cut data latency by 25% with Axios.**