Nitesh Thota

Fairfax, VA | (703) 656-3582 | nthota2@gmu.edu | LinkedIn | GitHub | Website

Education

M.S. in Computer Science **Aug 2023 - May 2025(Expected)**

George Mason University, Fairfax, VA GPA: 3.8

Relevant Coursework: Software Engineering, Component-Based Software Development, DevOps Practices, Database Systems, Natural Language Processing, Data Mining, Analysis of Algorithms

B.Tech. in Electronics and Communication Engineering **Aug 2017 – Apr 2021**

SRM Institute of Science and Technology, Tamil Nadu, India GPA: 3.5

Relevant Coursework: Algorithms, Data Structures, Database Management Systems, Computer Networks, Web Development

Technical Skills

Programming Languages: Java, Python, C, C++, JavaScript, TypeScript
Web technologies & Frameworks: HTML, CSS, SpringBoot, Spring MVC, Microservices, Node.js, React, Angular
Machine Learning: TensorFlow, PyTorch, OpenCV, Scikit-Learn
Data Processing and Analysis: Apache Spark, Apache Hive, Apache Hadoop, Apache Kafka, NumPy, pandas, SciPy
Cloud Computing and DevOps: Amazon Web Services (AWS), Google Cloud Services (GCS), Kubernetes, Docker, Jenkins, Ansible, Terraform
Tools, Software and OS: Android Studio, Git, Tableau, Postman, VS Code, IntelliJ, Eclipse, Linux, Windows
Databases: MySQL, Oracle, Postgres, SSMS, MongoDB

Professional Experience

Modak Analytics | Hyderabad, India **July 2021 - July 2023**

Role *Software Development Engineer*
Client *Humana Inc*

- Engineered, designed, and implemented robust **automated ETL jobs** using **Python** and **Java** to streamline **data ingestion, transformation, and hydration** across diverse data sources including **SSMS, MySQL, Big Query, GCS, and Oracle**. Delivered seamless **data delivery** to destinations like **HDFS, ADLS, and Google Cloud Platform (GCP)**, reducing **pipeline latency by 20%**.
- **Optimized complex data pipelines**, improving **performance and scalability** to manage **50% larger data volumes**. Utilized **industry-standard file formats** like **Parquet, Avro, and CSV** for efficient storage and retrieval in **big data environments**.
- Contributed significantly to developing **advanced data analytics dashboards**, utilizing tools like **Tableau** and **Python** to enable **dynamic reporting and actionable insights**.
- Enhanced existing codebase to comply with **Greenlight API (GLAPI)** benchmarks, significantly improving **CI/CD pipeline efficiency by 15%** and reducing deployment times using **Azure DevOps**.
- **Automated Python scripts** for unzipping and securely transferring data to **Google Cloud Storage (GCS)**, ensuring reliable daily execution through scheduled Cron jobs.
- Developed a scalable **Python-based Google Cloud Function**, triggered by **Pub/Sub messages**, capable of processing millions of CSV entries from fragmented files. Achieved **40% faster file processing** and consolidated data into comprehensive files uploaded to **GCS buckets**.
- Designed and implemented **RESTful APIs** with **Spring Boot** for the **Nabu product**, ensuring seamless integration with Humana's systems and **improving backend performance by 25%**.
- Built a responsive web application with Spring Boot, integrating **JPA/Hibernate** for database interactions and deploying the solution on **Kubernetes** for scalability.

Academic Projects

Survey Management System – Full Stack Web Application **Sep 2024 - Dec 2024**

- Built a scalable survey management application using **Spring Boot** for the backend and frontend frameworks (**Vue.js and Angular**) to create interactive and responsive **user interfaces**.
- Implemented **RESTful APIs** and integrated **JPA/Hibernate** with **MySQL** for efficient storage, ensuring seamless **CRUD operations**.
- Unified deployment by integrating the front and backend into a single Spring Boot JAR, containerized with **Docker** and hosted locally for verification.
- Optimized a **CI/CD pipeline** using **Jenkins** for automated testing and enhanced scalability with **Kafka** for data streaming and **Terraform** for infrastructure management.

PatriotPilot: NLP-Based University Chatbot **Sep 2024 - Dec 2024**

- Developed and integrated a chatbot for George Mason University's online resources using a **Retrieval-Augmented Generation (RAG)** framework, deployed as a pop-up on a prototype GMU CS website for seamless user interaction.
- Scraped, structured, and preprocessed data from university websites into JSON pairs, optimizing preprocessing to reduce noise and retain semantic relevancy.
- Embedded text using **E5-Large-v2**, stored embeddings in a **FAISS index**, and fine-tuned **Qwen 2.5 14B Instruct LLM** with instruction-response pairs.
- Achieved recall and **precision scores >70%**, ensuring accurate retrieval and response generation.