

# Teja Swaroop Sayya

+19802304200 — tejasayya@gmail.com — linkedin.com/in/teja-sayya/ — github.com/tejasayya

## Education

### University of North Carolina

Masters of Science in Computer Science - AI/ML

May 2025

### Certifications

- MongoDB Certified Associate Developer — AWS Certified Machine Learning Engineer

## Experience

### Open Source Developer - GSoC

June 2023 – Present

- Optimized Rocket.Chat queries, reducing execution time by 40% and API latency by 30%.
- Streamlined FitCheck App API integrations, cutting latency by 25% and boosting engagement by 15%.
- Enhanced TensorFlow Addons with a custom loss function, improving accuracy by 12% and fixing 5+ compatibility issues.

### Backend Software Engineer- Amazon

Oct 2022 – May 2023

- Built scalable backend services with Spring Boot, handling 10,000+ users at 99.9% uptime.
- Improved ML recommendation accuracy by 18% and reduced training time by 25%.
- Optimized data pipelines, cutting ingestion latency by 40% for 1M+ users.
- Created ReactJS dashboards for 5+ business units, improving decision-making by 20%.
- Reduced pipeline runtime by 30% and improved data quality by 15%.

### Big Data Engineer - CVS Health

June 2021 – Oct 2022

- Automated ETL workflows with Tableau Prep, reducing manual effort by 95%, saving 35+ hours monthly.
- Optimized 50+ Sqoop jobs, managing 5TB of data in HDFS and Hive, ensuring low-latency queries.
- Configured and monitored network security measures, including firewalls and intrusion detection systems.
- Reduced query times by 60% on 10TB datasets using Spark RDD transformations and actions.
- Built a data quality framework with Apache Spark for schema validation & profiling 1TB+ datasets, ensuring data integrity.
- Managed data pipelines in GCP with PySpark and BigQuery, streamlining flows and reducing risks in cloud infrastructure.
- Processed 10TB+ daily with 90.5% reliability using Hadoop and AWS EMR.

## Skills

**Programming** Data Structures and Algorithms (DSA), Java, Python, JavaScript

**Web Frameworks** Spring Boot, FastAPI, Next.js, REST APIs, Microservices

**AI/ML Libraries** TensorFlow/Keras, PyTorch, Hugging Face, LangChain, RAG (Retrieval-Augmented Generation)

**Big Data/Cloud** Hadoop, PySpark, Hive, HDFS, AWS EMR, Sqoop, Apache Airflow, AWS, GCP, MySQL, MongoDB

**Developer Tools** Docker, Git, Jenkins, Google Colab, Anaconda, CUDA

## Projects & Research

### Chat with Teja Sayya — RAG(Retrieval Augmented Generation): [Link](#)

- Built a RAG-based AI agent with LLM functionality and custom data, enabling scalable and context-aware conversations.

### MakeItTalk — Deep Learning models: [Link](#)

- Improved facial animation accuracy by 15% using CNNs, RNNs, and GANs.

### Connect4 with AI — Alpha-Beta Pruning Search Algorithm: [Link](#)

- Developed an AI-powered Connect Four game with a 95% win rate.

### Alzheimer's Disease Analysis — Machine Learning: [Link](#)

- Achieved 85% prediction accuracy using XGBoost and Random Forest.

### Real-Time Chat Application — Spring Boot, Web Sockets + React: [Link](#), [Demo](#)

- Built a chat app supporting 500+ users with <100ms latency.
- Developed a real-time chat application with Spring Boot, WebSockets, and React, enabling seamless communication in chat rooms with MongoDB persistence and Docker-based microservices.

### Photo App — ReactJs: [Link](#)

- Developed a scalable React-based photo sharing app using MongoDB and Material UI, supporting 10,000+ daily users with <100ms image load times and 30% faster API responses.