

Teja Swaroop Sayya

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

Education

University of North Carolina

May 2025

Masters of Science in Computer Science - AI/ML

Certifications

- MongoDB Certified Associate Developer — AWS Certified Machine Learning Engineer

Experience

Open Source Developer

June 2023 – Present

- Collaborating on backend improvements for Rocket.Chat and FitCheck App, optimizing database queries and streamlining API integrations.
- Contributed to the TensorFlow Addons repository, implementing a custom loss function and fixing compatibility issues with TensorFlow.

Software Engineer - Amazon(Client) - Infosys

Oct 2022 – May 2023

- Engineered scalable backend services using Spring Boot and microservices architecture for enhanced modularity in enterprise applications
- Contributed to the development of an ML-based recommendation system, focusing on fine-tuning model parameters, optimizing data pipelines, and ensuring seamless data ingestion for accurate predictions.
- Built advanced dashboards using ReactJS, visualizing key metrics and insights from the ML recommendation engine.
- Assisted in integrating the recommendation system with RESTful APIs, enabling real-time personalized suggestions for users.
- Refined data preprocessing and pipeline efficiency with data science and engineering teams to boost system performance.
- Collaborated with teams to refine data preprocessing and optimize pipeline efficiency for improved system performance.

Data Engineer - Loreal(Client) - Infosys

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.
- Built a data quality framework with Apache Spark for schema validation on 1TB+ datasets, ensuring data accuracy.
- Engineered a robust data quality framework using Apache Spark to validate schemas and profile large datasets for integrity.
- Used Spark to process 1 billion+ unstructured records with text mining algorithms for user table insights.
- Managed data pipelines in GCP with PySpark and BigQuery, streamlining flows and reducing risks in cloud infrastructure.
- Extensively worked with Hadoop, PySpark, AWS EMR, and the broader Hadoop ecosystem (Hive, HDFS, Spark) to architect and manage big data solutions, ensuring scalability and performance optimization.

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, LLM, RAG

Big Data/Cloud Hadoop, PySpark, Hive, Sqoop, AWS, GCP, MySQL, MongoDB

Developer Tools Docker, Git, 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.