Rohan Chandrashekhar

chandrashekhar.r@northeastern.edu | rohan-chandrashekhar-900b03212/ | Rohanc0412
Boston, MA 02135 | +1 (857) 919-2810

Education

Northeastern University, Boston, MA September 2024 – May 2026 (Expected)

Master of Science in Data Science

Courses: Algorithms, Introduction to Data Management and Processing, Large Language Model, Supervised Machine Learning, Natural Language Processing, DBMS, Data Mining in Engineering.

B.N.M Institute Of Technology (VTU), Bengaluru, India August 2018 – May 2022

Bachelor of Engineering in Electronics and Communication Engineering

Technical Skills

Language & Tools: Python, SQL, Java, FastAPI, REST APIs, Excel (Pivot Tables, Dashboards), Tableau, Jupyter, Git, CI/CD.

Machine Learning & Statistics: Scikit-learn, PyTorch, TensorFlow, GPT-4, LangChain, RAG, Conversational AI, NLP, Pandas, NumPy, Matplotlib, Seaborn, SHAP.

Data Engineering & Cloud: AWS (S3, EC2, RDS, Lambda), Apache Spark, Apache Kafka, MySQL, MongoDB, Docker, MLFlow, Model deployment and Monitoring, ETL Pipelines.

Work Experience

Northeastern University, Boston, MA May 2025 – Present

Teaching Assistant

- Supported 240+ students in courses: Programming with Data (DS 2000) and Intermediate Programming with Data (DS 2500), covering Python, OOP, API, data analysis, debugging, and applied machine learning.
- Led labs and office hours on Python, web scraping, APIs, visualization, regression, classification, and clustering.
- Curated hands-on coding exercises with real-world datasets, improving student project adoption and engagement.

Accenture Solutions Pvt Ltd, Bengaluru, India December 2022 – June 2024

Advanced App Engineering Associate

- Engineered high-performance, data-centric web applications using Java, Spring Framework, and Hibernate, optimizing backend operations through query refactoring that improved response times and reduced latency by 15%.
- Developed and maintained RESTful APIs for smooth microservice communication, integrated Apache Kafka for real-time data streaming, and managed MySQL and MongoDB databases with over 2M+ records to ensure reliability and scalability.
- Implemented 300+ unit and integration tests using JUnit and Mockito, integrated CI/CD pipelines achieving 95% code coverage, and decreased production defects by 40% through test-driven development.

Whitetail Technologies, Bengaluru, India January 2022 – November 2022

Data Scientist Intern

- Orchestrated and deployed machine learning models for classification, forecasting, and time series tasks that achieved over 88% accuracy and a 0.85 F1 score through feature engineering, hyperparameter tuning, and extensive backtesting.
- Automated ETL pipelines and applied MLOps practices including experiment tracking, Docker containerization, and CI/CD integration, improving data consistency, reducing preprocessing time by 60%, and shortening deployment cycles by 40%.
- Created interactive Tableau dashboards, reducing reporting time by 20% and improving data-driven decision-making.

Projects

Image-Guided Knowledge-Augmented Storytelling with RAG

- Constructed a multimodal RAG pipeline integrating CLIP/BLIP-2 captioning, Pinecone vector retrieval, and GPT-4 story generation to create fact-grounded narratives from images.
- Designed dual retrieval streams (vector search + real-time news via SerpAPI) and implemented a three-act narrative framework, reducing hallucinations and improving semantic alignment (summary–story similarity of 0.71).
- Applied skills in NLP, vector search, and LLM integration to enhance information retrieval and user experience.

AI-Powered Recommender for Enterprise Workflows

- Developed an AI-powered recommender system for enterprise workflows, suggesting task assignments, document references, and streamlined logistics routes.
- Hosted scalable ML models on Azure, achieving sub-200ms latency in testing and ensuring availability with cloud deployment.
- Boosted workflow efficiency with 18% better task matching, 35% faster document search, and 10% lower logistics costs.

Netflix IMDb Score Prediction and Tableau-Based Content Analytics

- Cleaned and analyzed 7,000+ Netflix title records to uncover relationships between runtime, genre, and audience engagement using Python (pandas, NumPy) and SQL, supporting data-driven content strategy.
- Built predictive IMDb rating models using scikit-learn and text analytics (TF-IDF + SVD), achieving R^2 = 0.89 and MAE = 0.23, and deployed inference endpoints via FastAPI for real-time score prediction.
- Architected Tableau dashboards and automated AWS pipelines (EC2, S3, RDS) to track predicted vs. actual IMDb scores and KPIs, enabling teams to identify high-ROI genres, improve targeting by 12%, and cut reporting time by 25%.