

Lokesh Kanna Rajaram

lrajaram@buffalo.edu | [Portfolio](#) | [LinkedIn: Lokesh kanna](#) | +1(716) 520 7064

EDUCATION

University at Buffalo, The State University of New York

Buffalo, NY, USA

Master of Science

12/2025

- **Concentrations:** Engineering Science - Data Science
- **Coursework:** Numerical Mathematics, Introduction to Probability, Statistical Data Mining, Database Fundamentals, Data Intensive Computing, Data Model Query Languages, Introduction to Machine Learning, Computer Vision.

Anna University

Chennai, India

Bachelor Of Engineering

Graduated, 04/2023

- **Concentrations:** Geoinformatics, Minor: Computer Science
- **Coursework:** Data Structures & Algorithms, Satellite Weather Forecasting and Modelling, Computer Organization & Programming, Satellite Image Processing, Object-Oriented Programming, Decision Support System, Database Management System.

PROJECTS & EXPERIENCE

AMAZON BOOK REVIEW USING BIG DATA PIPELINE

05/2025

- Designed and implemented a scalable big data pipeline to process Amazon book reviews using Hadoop and Apache Spark on Docker clusters, enabling distributed ingestion, ED, and ML-ready transformation of over 1 million records.
- Automated data ingestion and storage into HDFS from raw CSV sources using CLI tools and batch upload strategies, ensuring robust fault-tolerant storage for parallel data access across nodes.
- Optimized text feature pipelines with Tokenizer, StopWordsRemover, HashingTF, and IDF in Spark ML library, balancing dimensionality, which accelerated downstream model training by 40% and achieved up to 90.4% accuracy.

OPTIMIZED BULK STOCK SELLING STRATEGIES WITH MACHINE LEARNING

12/2024

- Boosted stock price and volume prediction accuracy by 15% (to 78%) by designing and deploying end-to-end machine learning pipelines with Random Forest, Gradient Boosting, and LSTM on a 4-year NVIDIA dataset.
- Built efficient and scalable algorithmic trading strategies (VWAP and TWAP) to reduce market impact and enhance profitability for bulk stock transactions, showcasing applied data engineering and modeling skills.
- Enhanced feature quality and model performance by integrating advanced technical indicators (RSI, Bollinger Bands) and performing data-driven exploratory analysis and visualizations to inform strategy.

TEXT SUMMARIZER USING DEEP LEARNING

07/2024

- Improved system scalability and performance for text summarization models by deploying secure, high-availability infrastructure using AWS EC2 and IAM, and conducting extensive performance tuning.
- Accelerated deployment cycles and reduced manual overhead by building an automated CI/CD pipeline with GitHub Actions and containerizing the application using Docker for reproducible builds.
- Delivered a production-ready, user-friendly summarization tool by managing cloud-based deployment on EC2, integrating workflow automation, and collaborating with teammates to ensure seamless continuous delivery.

TIME SERIES ANALYSIS OF GROUNDWATER CHANGE USING GRAVITY RECOVERY AND CLIMATE EXPERIMENT

04/2024

- Spearheaded geospatial data processing and transformation by integrating multi-format datasets (GLDAS, NetCDF, GeoTIFF) using Python, and visualized temporal-spatial patterns using ArcMap and QGIS for climate-impact analysis.
- Uncovered a long-term groundwater decline of 5.877 cm/year across the Cauvery River basin by analyzing GRACE satellite time-series data from 2003 to 2022, supporting climate-informed policy insight.
- Predicted an average annual groundwater thickness variation of -26.179 cm by building data models to assess significant changes over time, applying statistical analysis to large-scale geospatial data.

SKILLS

Programming Languages & Databases: Python, R, MySQL, MATLAB, Hadoop, Spark, Map Reduce, Hive

Tools & Platforms: AWS, Docker, Kubernetes, GitHub, Git, Jenkins, GCP, PowerBI, Tableau.