Ritvik Gupta

+1-617-938-7292 Gritvik8@gmail.com LinkedIn GitHub - Portfolio

Data Scientist and AI Research Fellow with 2+ years of experience in machine learning, scientific modeling, and large-scale data analysis. Specialized in designing scalable ML systems, AI-driven solutions, and cloud-integrated analytics using Python, PySpark, and AWS. Strong research foundation in optimization, statistical modeling, and healthcare/public sector AI applications.

Education

Boston University | *Masters, Applied Data Analytics* **2023 - 2025**
SRM Institute of Science and Technology **2019 - 2023**
Bachelor of Technology (B.Tech.), Computer Science and Engineering

Skills

- **Programming:** Python, SQL, PySpark, Bash
- **ML/AI:** scikit-learn, XGBoost, TensorFlow, MLlib, FAISS
- **Data Engineering:** Apache Spark, Kafka, Hive, Hadoop, Snowflake
- **Cloud & MLOps:** AWS (S3, EMR), MLflow, Docker, CI/CD, Databricks
- **Analytics & Tools:** Power BI, Git, Jupyter, VSCode

Work Experience

The Chicago Education Advocacy Cooperative | *Data Science Fellow* **June 2025 - Present**
- Built automated dashboards and predictive forecasting models using Python and statistical techniques to support cross-departmental planning and improve strategic decision-making.
- Delivered actionable insights through comprehensive analysis of multi-source datasets, directly influencing operational efficiency, budget allocation, and policy development.

BU Faculty of Computing & Data Sciences (CDS) | *Graduate Teaching Assistant* **Aug 2024 - May 2025**
- Led workshops and mentoring sessions for over 80 students, focusing on Python-based data engineering and machine learning applications relevant to business and analytics.
- Contributed to curriculum design and course optimization by incorporating real-world datasets and performance-based coding assessments.

ENGAGEAthon | *Intern* **May 2024 - Aug 2024**
- Automated and standardized sales and marketing data pipelines from 9 sources using Python and Django, reducing manual reporting time by over 60% and improving data reliability.
- Collaborated with business stakeholders to enhance visibility into KPIs and implemented tool improvements that streamlined performance tracking across teams.

Boston University | *Teaching Assistant* **Jan 2024 - December 2024**
- Mentored students in machine learning, web extraction, and code optimization, focusing on Python-based data pipelines, model evaluation, and iterative data analysis for real-world problem-solving.
- Supported students in overcoming technical challenges, fostering academic success, and enhancing their practical skills and expertise in data science and software development.

Enhanced Communications & Technologies PVT. LTD. | *Python Developer Intern* **Mar 2021 - May 2021**
- Developed a cloud-integrated AI chatbot and a message automation tool connected to a live database, enabling real-time updates and efficient WhatsApp-based communication.
- Improved internal communication workflows by automating message scheduling and reducing manual intervention through hands-free scripting solutions.

Technical Projects

Scalable Recommendation Engine with PySpark & AWS
- Implemented ALS, FAISS, and KMeans in a hybrid recommendation system to deliver personalized article suggestions, enhancing user engagement and content discovery.
- Utilized PySpark, AWS S3, and EMR to create a scalable data pipeline for efficient processing of large-scale datasets, improving recommendation accuracy and system performance.

Predictive Stock Market Algorithm
- Developed a stock prediction model leveraging daily Yahoo Finance data, incorporating time-series forecasting techniques and feature engineering; achieved 85% accuracy in directional movement prediction.
- Utilized statistical modeling (ARIMA, linear regression) and data visualization to uncover trends, evaluate volatility, and refine algorithmic trading signals for strategy optimization

Publications

- *Https://www.ijser.in/archives/v10i12/SE221206135738.pdf*
  Applied deep learning to public health safety, highlighting relevance to healthcare analytics.
- *Https://www.ijser.in/archives/v11i8/SE23822002659.pdf*
  Designed secure data systems, underscoring compliance and process integrity.