

RAHUL SHARMA

College Park, Maryland | 240 610 7281 | rahul11s@umd.edu | [@LinkedIn](#) | [@Git](#)

Summary

Data Scientist with proven expertise in transforming complex datasets into actionable insights that drive real business growth. Skilled in leveraging advanced analytics, statistical methods, and ML/DL techniques to deliver innovative data-driven solutions.

Education

University of Maryland - College Park, US

Dec 2025

Master of Science in Data Science - (GPA: 4.0/4.0)

- Relevant Courses: Probability & Statistics, Principles of Data Science, Principles of Machine Learning, Big Data Systems, Algorithms for Data Science, Data Representation and Modelling

Indian Institute of Technology (BHU) Varanasi, India

July 2020

Bachelor of Technology in Engineering

- Relevant Courses: Eng. Mathematics, Computer Programming, Numerical Techniques, Mathematical Methods

Work Experiences

Jet2 Travel Technology, Pune, India

June 2022 - Aug 2024

Data Scientist

- Processed and summarized raw call transcript data, **fine-tuned** a classification model, and utilized **few-shot learning** techniques to classify transcripts, achieving **88% Jaccard recall** and significantly reducing manual effort.
- Engineered topic modeling for analysis and utilized **isolation forest** with a grid search approach to detect disaster events in real-time using LivePerson text data, generating statistical features and achieving a **92% Specificity**.
- Developed risk scoring system by forecasting risk scores using multiple attributes; employed **LSTM/GRU** models with a walking forward approach. Implemented ensemble modeling with **meta-learner**, resulting in improved decision-making.

Fibe (Formerly EarlySalary), Pune, India

Sept 2021 - June 2022

Data Scientist

- Optimized Behavior Scorecard with 1,000+ bureau variables, achieving **0.94 ROC-AUC** through optimized logistic model.
- **Automated** Behavior Scorecard workflows using the Knime platform by integrating feature engineering and model deployment pipelines; developed performance monitoring dashboards with key metrics, reducing processing time by **15x**

Advanced Technology Consulting Service, Jaipur, India

Nov 2020 - Sept 2021

Associate Data Scientist

- Developed collection reports utilizing MSSQL, achieving **10-fold acceleration** through full automation deployment.
- Trained a car damage detection model using **Mask R-CNN** with transfer learning on customized datasets, achieving **85% mAP**, significantly enhancing claim verification accuracy for an insurance firm.

Axio (Formerly CapitalFloat), Bangalore, India

May 2020 - Nov 2020

Data Science Intern

- Engineered comprehensive address dataset from JSON using NLP techniques and lifestyle features, applying statistical analysis and feature selection to model customer default trends with a **95% F1 score**.
- Architected dynamic, data-driven dashboards using **SQL** and Power BI, transforming complex datasets into actionable insights through advanced metric formulation, enabling strategic decision-making.

Projects

- **Multimodal E-commerce Recommendation System, Fall'24, UMCP:** Developed Fine-tuned CLIP's vision and text encoders with BLIP-generated captions and segmentation masks to enhance text-image alignment. Integrated adaptive query reweighting and FAISS-based nearest-neighbor search, achieving 88% Mean Opinion Score in retrieval relevance.
- **Fake News Detection from Social Media, Summer Intern'19, IIT(BHU) Varanasi:** Engineered and labeled a Twitter dataset through meticulous fact-checking. Applied NLP techniques for feature extraction, including Word2Vec and GloVe, and implemented Gradient Boosting, Random Forest, and LSTM models. Achieved **94% recall** with an Adaptive Multilevel Ensemble Model for comprehensive content classification.

Skills

- **Programming:** Python (Adv.), SQL (Adv.), PySpark (Int.), C/C++ (Int.), MATLAB (Int.)
- **Frameworks:** Pandas, Numpy, NLTK, Scikit-Learn, TensorFlow, PyTorch, Transformers, OpenCV, Statsmodels.
- **Databases:** MS SQL Server, Snowflake, MySQL, Hadoop, Spark.
- **Tools:** AWS(SageMaker, EC2, Redshift), Dataiku, GCP, Knime, Confluence, Git, Azure, Databricks, Docker, Tableau.
- **Specializations:** LLMs, GenAI, Deep Learning, Statistics, NLP, ML, Predictive Modeling, Big Data Technologies.

Extracurricular Achievements

- **Kashi Utkarsh Initiative:** IIT-Varanasi (2016-19)- Volunteered teaching primary education to underprivileged children.
- **Inter-IIT Sports Tournament:** IIT-Madras (2017)- Secured bronze in 800m race; participated in 4*400 relay race event.