

# RAKSHAK KUNCHUM

8 Parkerhill Ave, Boston, MA 02120 | (857) 230-2483 | [krakshak7698@gmail.com](mailto:krakshak7698@gmail.com) | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

## EDUCATION

**Northeastern University**, Boston, MA

January 2023 – December 2024

Khoury College of Computer Science

Graduated – F1 OPT EAD (Exp. Date – 02/16/2026, STEM eligible)

Master of Science in Data Science, **GPA: 3.95/4.0**

Related Courses: Supervised Machine Learning, Data Mining, Data Management, Information Retrieval, Deep Learning, NLP.

**BMS College of Engineering**, Bangalore, India

September 2016 - August 2020

Bachelor of Engineering in Information Science Engineering, **GPA: 8.8/10**

Related Courses: Machine Learning, Statistics, Database Management Systems, Data Science, Algorithms, Python, Linear Algebra.

## PROFESSIONAL EXPERIENCE

**Capital One, McLean, Virginia**

June 2024 - August 2024

Data Science Intern – Anti-Money Laundering (AML) / Fraud / Risk team

- Saved 20 **FTE** investigators' effort and ~\$1M in operational costs for a Credit Card Fraud classification project by developing a machine learning pipeline with **data extraction, manipulation, feature selection, optimization and visualization** capabilities.
- Created transaction aggregation features based on customer activity on **Snowflake** and **Databricks**, amplifying model efficiency.
- Achieved a ROC-AUC score of 0.8 by optimizing Random Forest, XGBoost, LightGBM, and CatBoost models with advanced hyperparameter tuning techniques. Fine-tuned model thresholds to enhance precision, recall, and reduce false positives rates.
- Effectively communicated results to cross-functional leaders and product stakeholders from Fraud, AML and Investigation teams, translating **model outcomes** and **SHAP plots** into actionable insights and demonstrating clear business impact.

**Dataweave (Infoweave Analytics Pvt Ltd), Bangalore, India**

July 2020 - October 2022

Data Scientist

- Collaboratively automated 10+ agile end-to-end scalable data pipelines for **e-Commerce giants** on Pricing Intelligence, Brand Intelligence, Digital Shelf Analytics projects using **Airflow DAGs, Python, Scikit-learn, AWS, Kafka and Docker**.
- Devised complex **SQL** queries to handle large datasets using **AWS Athena, S3, Redshift** to create business value and KPIs data.
- Integrated JSON data sources, rest APIs, **machine learning algorithms, ETL frameworks** reducing manual effort by **30%**.
- Attained **50%** reduction in text/image **processing times** in the company's clustering algorithm by implementing **dask** library.
- Designed dashboards and customized interactive reports using **Tableau** with pricing recommendations for market competitiveness.

**Dataweave (Infoweave Analytics Pvt Ltd), Bangalore, India**

July 2019 - June 2020

Data Science Intern

- Developed and deployed **25+ Selenium** and **Python-based deep crawler bots** to optimize e-commerce data scraping processes for pricing intelligence and counterfeit product detection projects that unlocked business value for global retailers and brands.
- Led the successful delivery of structured data for 15+ data-as-a-service projects by employing elaborate data-wrangling methods.
- Assisted Classic Accessories brand to gain ~\$500K in additional sales by identifying and reporting counterfeit products.

## ACADEMIC PROJECTS

**FinAdvisor - Financial Advisory Tool for personal finance assistance** | Python/LLMs/Natural Language Processing/Vector DB

- Orchestrated a financial assistant bot by finetuning LLaMA, Mistral LLMs and integrated Pinecone DB for faster text retrieval.
- Deployed a Flask based UI with restful API, enhancing accessibility and user engagement with financial data advice and insights.

**Forecasting Credit Card Expenditure** | Python/Machine Learning/Boosting/Bagging

- Engineered and optimized regression based machine learning models to predict credit card spending using feature selection techniques, bagging, boosting, and grid search, enabling informed credit card limit decisions for the banking industry.
- Obtained an impressive R-squared of 0.86 and explained\_variance\_score of 0.8 showcasing high precision and model effectiveness.

## ACHIEVEMENTS

- 1st place at Harvard Data Science Initiative's Agri Data Hackathon 2024 in sponsorship with USDA NASS. ([Hackathon results](#))
- 3rd place at Northeastern University Data Science Hackathon Summer 2023.

## TECHNICAL SKILLS

- Programming/Databases:** Python, R programming, SQL, Java, MySQL (relational), BigQuery, Snowflake, Databricks.
- Toolkits/Software:** Azure, AWS (Sagemaker, Redshift, S3, Athena), Selenium, Elasticsearch, Apache Airflow, Docker, Kafka, Tableau, Power BI, Git, Bitbucket, Linux, Shell Scripting, CI/CD, MS Office, Agile development, MLOps, Rest APIs, Jupyter.
- ML/DL Libraries:** Pandas, NumPy, Scikit-learn, SciPy, Plotly, Tensorflow, Keras, Pytorch, Spark, Hive, OpenCV, NLTK.
- Data Science:** Regression, Classification, Clustering, Segmentation, Decision Trees, Random Forest, XGBoost, Transformers, Causal Inference, Statistical Modeling, Data Analysis, LLM, Optimization Techniques, Hypothesis Testing, A/B Testing.