

RAKSHAK KUNCHUM

Boston, MA | (857) 230-2483 | krakshak7698@gmail.com | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

EDUCATION

Northeastern University, Boston, MA

January 2023 – Present

Khoury College of Computer Science

Expected Graduation: December 2024

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

Related Courses: Supervised Machine Learning, Data Mining, Data Processing and Management, Algorithms, 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 Foundations, Python Programming.

TECHNICAL SKILLS

- **Programming Languages:** Python, R programming, SQL, Java.
- **Databases:** MySQL (Server and Workbench), MongoDB.
- **Cloud Computing Platforms:** AWS (EC2, S3, Athena, Redshift).
- **Toolkits/Software:** Selenium, Airflow, Kafka, Docker, Tableau, Git, Excel, Shell Scripting (Unix/Linux), MS Office Suite.
- **Data Science Libraries:** Pandas, NumPy, Scikit-learn, Matplotlib, Plotly, Seaborn, Tensorflow, Pytorch, PySpark, OpenCV.
- **Data Science Techniques:** Regression, Decision Trees, Random Forest, Bagging, Boosting, Ensemble Models, Clustering, PCA, Neural Networks (CNN, RNN, LSTM, GAN), Deep Neural Nets, Transformers, Statistical Models, Hypothesis Testing.
- **Soft Skills:** Effective Communication, Analytical Skills, Problem-Solving, Leadership-Oriented, Adaptive and Curious.

PROFESSIONAL EXPERIENCE

Capital One Financial Services, McLean, Virginia

June 2024 - August 2024

Data Science Intern

- Developed and tested automated **feature creation, selection, and algorithm optimization** pipeline for a Supplemental Fraud SAR risk sloping project, potentially saving 20 FTE investigators' effort and costs.
- Experimented with various **machine learning** algorithms, applying advanced hyperparameter tuning and **optimization** techniques with a best model ROC-AUC score of 0.79 on validation and out-of-time test samples.
- Achieved high precision and recall with a low false positive rate, demonstrating the model's reliability and effectiveness.
- Created transaction aggregation features based on customer authorization activity, significantly enhancing model efficiency.

Dataweave (Infoweave Analytics Pvt Ltd), Bangalore, India

July 2020 - October 2022

Data Engineer

- Collaboratively automated more than 10 end-to-end data science pipelines using **Airflow, Python, Kafka and AWS** to integrate web crawlers, data sources, APIs and internal ETL frameworks resulting in enhanced project outcomes.
- Programmed complex **SQL** queries to handle large datasets using **AWS Athena & S3** to create impactful business data for clients.
- Attained 50% reduction in text/image **batch processing** in the company's clustering algorithm by implementing **dask** library.
- Developed and deployed **15+ Selenium** and **200+ Python crawlers** to mine e-commerce data, enhancing product offerings.

ACADEMIC PROJECTS

Multimodal Sentiment Analysis with Vision-Language Models | Python/LLMs/Deep Learning/Web Scraping/NLP

- Utilized few-shot learning and prompt engineering techniques to enhance Bakllava model for predicting Amazon customer ratings.
- Integrated review text and images for comprehensive model training and improved prediction accuracy.

US Air Pollution Time Series Analysis | Python/Deep Learning/Time Series

- Applied advanced time-series analysis and forecasting approaches (SARIMAX, LSTM) to predict future US air pollution trends.
- Attained a low Root Mean Squared Error (RMSE) of 0.87 and Mean Absolute Error (MAE) of 1.45 for LSTM predictions.

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

- Engineered and optimized a regression-based machine learning model to predict credit card spending using feature selection techniques, bagging, boosting, and grid search, enabling data-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 its effectiveness.

Deep Learning Approaches to Detect Pneumonia | Python/Deep Learning/Computer Vision/Transfer Learning

- Ran a comparative analysis of three machine learning models (CNN, U-Net, Mask-RCNN) for pneumonia detection.
- Achieved accurate identification and localization of infected regions by leveraging the Mask-RCNN model in Chest Radiographs.

EXTRACURRICULARS

Data Science Hub, Northeastern Graduate Student Government, Hiking, Volleyball, Video Gaming and E-Sports.