

RAKSHAK KUNCHUM

Boston, Massachusetts | (857) 230-2483
kunchum.r@northeastern.edu | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

EDUCATION

Northeastern University, Boston, MA

January 2023 - Present

Khoury College of Computer Sciences

Expected Graduation: December 2024

Candidate for Master of Science in Data Science, GPA: 4.0/4.0

Related Courses: Supervised Machine Learning, Data Mining, Data Processing and Management, Deep Learning, Algorithms.

BMS College of Engineering, Bangalore, India

September 2016 - August 2020

Bachelor of Engineering in Information Science Engineering, CGPA: 8.8/10

Related Courses: Machine Learning, Statistics, Database Management Systems, Linear Algebra, Analysis and Design of Algorithms, Data Structures, Data Mining.

TECHNICAL SKILLS

Programming Languages: Python (Pandas, NumPy, Tensorflow, Pytorch, Keras, Scikit-learn, NLTK, Dask, OpenCV, Matplotlib, Seaborn, Plotly), R (dplyr, tidyr, ggplot2, modelr), MySQL, NoSQL (MongoDB), Java.

Toolkits/Frameworks: Selenium, Airflow, Kafka, AWS (EC2, S3, Athena), Tableau, Git, Excel.

Data Science/Analysis Techniques: Regression, Decision Trees, Random Forest, Boosting, SVM, Ensemble Learning, Clustering, Dimensionality Reduction, Topic Models, Neural Networks (CNN, RNN, LSTM, GAN), ARIMA, SARIMAX, Hypothesis testing.

Data Engineering: Shell Scripting, Web Scraping, Data Wrangling, Workflow Automation.

PROFESSIONAL EXPERIENCE

Dataweave (Infoweave Analytics Pvt Ltd), Bangalore, India

July 2020 - October 2022

Data Engineer

- Automated data engineering projects, applying project-specific automation techniques to streamline processes, enhance efficiency and reduce manual intervention, improving team productivity and project outcomes.
- Pioneered the implementation of the Dask framework to optimize batch processing of image data in the company's clustering algorithm, achieving a remarkable 50% reduction in processing times and demonstrating innovative problem-solving skills and effective teamwork.
- Spearheaded the setup of end-to-end data pipelines using Apache Airflow, enabling seamless data workflows.
- Piloted the comprehensive workflow for counterfeit detection showcasing leadership and collaboration skills. Achieving over 95% accuracy in identifying and taking down counterfeit products.

Dataweave (Infoweave Analytics Pvt Ltd), Bangalore, India

January 2020 - June 2020

Data Engineering Intern

- Built diverse crawler plugins and designed efficient solutions for scheduling data crawls, enabling effective data scraping from multiple e-commerce websites and ensuring optimal data extraction and flow.
- Led the delivery of structured data by employing data-wrangling methods to clients under data-as-a-service project.

ACADEMIC PROJECTS

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

- Applied time-series analysis techniques to the US pollution dataset, by focusing on significant pollutants.
- Leveraged advanced time-series forecasting approaches (SARIMAX, LSTM) to predict future air pollution trends in California, providing valuable data-driven insights for environmental monitoring and strategic decision-making.

Trending News Prediction using NLTK and Web Scraping | Python/Machine Learning/NLP

- Developed a web scraper to gather news articles from a specific website and processed the data using NLTK.
- Employed various classification machine learning models (logistic regression, k-nearest neighbors, decision tree, random forest, xgboost) to determine the optimal model for predicting and analyzing news trends.

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

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

EXTRACURRICULARS

Data Science Hub, Northeastern Graduate Student Government, Badminton, Hiking