

ADITYARAM KOMARANENI

New York, NY | (646) 427-7052 | ak5480@columbia.edu | adityaramk.com | linkedin.com/in/adityaramk | github.com/adityaram-2003

Education

Columbia University

Master of Science in Data Science

Coursework: ML Theory, Applied Machine Learning, NLP, Exploratory Data Analysis, Generative AI

Expected Dec 2026

New York, NY

Gokaraju Rangaraju Institute of Engineering & Technology

Bachelor of Technology in Computer Science and Engineering; GPA: 8.9/10 (WES equiv. 3.9/4.0)

Ranked 1st of 66 in academics in Summer 2022 and Spring 2023

Nov 2020 – Aug 2024

Hyderabad, India

Work Experience

Columbia University – Data Science Institute

Data Science Research Scholar

Jan 2026 – Present

New York, NY

- Developing an end-to-end computer vision pipeline for automated phytoplankton identification from ocean imagery
- Implemented preprocessing data and quality controls (label normalization, train/val/test splits, augmentation) and trained image classification models using Python and deep learning frameworks to improve robustness and generalization
- Prototyped baseline image classification models and evaluation workflows, documented results and insights for research stakeholders to guide next-step model and data iteration

Tata Consultancy Services

System Engineer

Aug 2024 – Jul 2025

Hyderabad, India

- Optimized loan approval prediction model by identifying and implementing data preprocessing techniques: PCA, LDA
- Collaborated within team on comprehensive testing and optimization of ML models: Logistic Regression, Naive Bayes Classifier, KNN, K-means, and SVM, and presented key findings on overall accuracy increase by 11%
- Led a team of three, organized team efforts to communicate analytical findings on loan database, resulted in a final 89% accuracy and a 9% runtime reduction

Projects

Automated Phytoplankton Classification (CAPCS) | Lamont-Doherty Earth Observatory

Dec 2025 – Present

- Structured 12,315 microscopy images across 10 phytoplankton classes into a consistent dataset + label metadata format; merged 2 source roots and established annotation rules/class mapping to reduce label ambiguity across sources
- Organized clustering outputs by creating 71 cluster folders and saving 1,626 clustered images to support downstream error discovery, dataset refinement, and class separability analysis
- Performed diagnostics and interpretability reviews (confusion matrix analysis, per-class precision, recall, F1) across 10 classes; documented priorities to enhance dataset coverage and reduce recurring failures

Loan Approval Prediction | Tata Consultancy Services

Aug 2024 – Jul 2025

- Conducted research and development of a loan approval model, directing team's focus on data analysis, and finance, leading to an 89% accuracy based on metrics: credit score, income, and dependents
- Designed a comprehensive data cleaning workflow leveraging Z-Score Normalization and IQR outlier elimination, improving quality of 40,000+ entry loan database and resulting in 8% increased accuracy using Gradient Boosting

Research Publications

Uncertainty-Aware Molecular Property Prediction Using Heterogeneous Molecular GNNs

DOI: 10.1007/978-981-97-9132-3-16

2025

Springer Nature

Detection of Bone Fractures from Medical Imaging Data Using Machine Learning

Manuscript under review

2024

IEEE

Performance Comparison of Depth Limited Search and A* Algorithm: A Case Study

DOI: 10.1051/e3sconf/202339101140

2023

E3S Web of Conferences

Skills

Programming: Python, SQL, R, Java, JavaScript, HTML, CSS

Frameworks: Pandas, NumPy, Matplotlib, Seaborn, Plotly, scikit-learn, Flask, Streamlit, TensorFlow, Keras, PyTorch

Tools: Git, GitHub, Google Colab, Jupyter, VS Code, PyCharm, RStudio, Tableau, OpenML, Quarto, Excel

Certifications: Harvard CS50, Stanford Online: SQL0001 Databases, IBM: ML with Python