

Arnav Ahuja

aa5790@columbia.edu — linkedin.com/in/arnav-ahuja/ — arnavahuja.github.io/ — (716) 994-1303

Education

- **Columbia University** New York, US
Master of Science in Data Science Aug 2025 - Dec 2026
Relevant Coursework: Financial Analysis, Algorithms, Data Analysis, Visualization
- **Birla Institute Of Technology and Science** Pilani, IN
Bachelor of Technology Computer Science Aug 2018 - Jul 2023
Master of Science in Mathematics
Relevant Coursework: Probability, Statistics, Optimization, Operations Research, Linear Algebra, Graphs

Work Experience

Barclays Chennai, IN
Graduate Developer — Model Services Center of Excellence Aug 2023 - Apr 2025

- Provisioned credit risk datasets powering 50+ quantitative models used by 10+ teams across Barclays
- Deployed an AWS data pipeline to automate dataset delivery, saving 80% of manual preparation time
- Built infrastructure to integrate and run 100+ quantitative models, decreasing setup time from 1 week to 1 day
- Mentored an intern on project and presentations, resulting in appreciation by stakeholders
- Secured 1st place among 2,000+ participants in India in a global Barclays Generative AI Hackathon

Western Australia Department of Health Remote
Associate Researcher Jan 2023 - May 2023

- Analyzed ~19,000 attributes across 373 suburbs to identify key factors influencing community health in Australia
- Implemented hierarchical and PCA-based clustering, reducing health attribute dimensionality by 95%
- Identified 3 representative suburbs from 373, driving data-backed health policy insights for 3 million residents
- Authored a thesis presenting 5 key findings supported by in-depth analyses, resulting in actionable policy insights

Amazon Bengaluru, IN
Applied Scientist Intern — Selection Monitoring Jun 2022 - Dec 2022

- Built reinforcement learning baseline models to predict 4 user actions using 5 million web pages with 30% accuracy
- Designed and implemented a webpage segmentation approach achieving 84.2% accuracy in identifying user actions
- Leveraged a graph based approach for exhaustive product selection on 100k+ competitor e-commerce websites

Publications

- **Use of spatio-temporal features for earthquake forecasting of imbalanced data**
(IEEE) *International Conference on Intelligent Innovations in Engineering and Technology (ICIET)*
Arnav Ahuja, Aaditya Sharma, Sumanta Pasari
- **Disease Identification in Tomato Leaf using pre-trained ResNet and Deformable Inception**
(Springer) *5th International Conference on Computational Intelligence in Data Science*
Arnav Ahuja, Jennifer Ranjani, Aditya Tulsyan

Academic Experience

Neural Networks and Fuzzy Logic: Teaching Assistant Aug 2022 - Dec 2022

- Guided a class of 100 students, created assignments and ensured improved learning outcomes
- Supervised 3 student teams of 3–4 members each, providing guidance and support throughout the course project

Computer Vision: Crop Disease Identification Jan 2020 - Dec 2020

- Devised a novel Inception-ResNet model to identify 15 diseases in plant leaves with 98.16% accuracy on 50k images
- Generated a dataset of real images via data augmentation, boosting model accuracy by 15
- Co-authored a white paper

Deep Learning: Earthquake Forecasting Aug 2020 - Dec 2020

- Analyzed earthquake time series data from 5 regions to extract seismicity indicators for forecasting models
- Developed a neural network and achieved 90.4% accuracy in forecasting earthquake probability in the Himalayas

Technical Skills

- **Programming Languages** C, C++, Java, Python, MATLAB, SQL, Shell Scripting
- **Data Science Libraries** Boto3, PyTorch, TensorFlow, Keras, Pandas, Numpy, OpenCV
- **Deep Learning** LLMs, CNNs, RNNs, GRUs, LSTMs, Reinforcement Learning, Generative AI
- **Platforms/Tools** AWS (all services), Google Colab, Jupyter Notebook, MATLAB