Arnav Ahuja

✉ arnavahuja21@gmail.com — in linkedin.com/in/arnav-ahuja/ — ⓘ arnavahuja.github.io/

Education

- **Columbia University, New York** — *2025-2026*
  Master of Science in Data Science

- **Birla Institute Of Technology and Science, Pilani, Rajasthan** — *2018-2023*
  Bachelor of Technology Computer Science
  Master of Science in Mathematics (Dual Degree)
  Overall CGPA: 8.29/10

Work Experience

Barclays — *Aug'23 - April'25*
Position: Graduate Developer — Team: Model Implementation Team

- Provisioned credit risk based **datasets to run quantitative models** for multiple teams within Barclays
- Spearheaded the development of a unified messaging service on AWS **to automate dataset delivery**
- Built and managed infrastructure to **integrate and run over 100 quantitative models for the bank**

Western Australia Department of Health — *Jan-May'23*
Position: Research Intern

- Investigated ~**19000 attributes for 373 suburbs** in the Australian continent for improving community health
- Implemented **hierarchical clustering and PCA based clustering** for attribute correlation
- Obtained a specific suburb from the data for in-depth analysis and evaluation of policy effectiveness

Amazon — *June-Dec'22*
Position:Applied Scientist — Team:Selection Monitoring

- Developed **reinforcement learning based baseline models** with **30% accuracy** for identifying user actions
- Constructed **WebPage Segmentation** based approach with **84.2% accuracy** for finding user actions on web-pages
- Utilized a **graph based approach for exhaustive product selection** on 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* (*ICIIET*) LINK
  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* LINK
  Arnav Ahuja, Jennifer Ranjani, Aditya Tulsyan

- **Forecasting Earthquakes Using Neural Network Models**.
  (*Springer Nature*) *Disaster Management in Complex Himalayan Terrains - Natural Hazard Management, Methodologies and Policy Implications.* LINK
  Arnav Ahuja, Sumanta Pasari

Academic Projects

Classical ML: Detecting Diabetic Retinopathy using ML — *Jan-May'22*

- Analyzed **single nucleotide polymorphism** data for identifying the susceptibility to **diabetic retinopathy**.
- Implemented **Lasso Regression and Random Forest algorithm for feature selection** in SNPs.
- Used **machine learning algorithms like kNN, SVM, Gradient Boosted DT** for predicting the susceptibility.

Computer Vision: Crop Disease Identification — *Jan-Dec'20*

- Devised a **new Inception Resnet deep learning architecture** to identify diseases in the leaf of tomato plant
- Achieved an **accuracy of 98.16%** which is higher than the traditional resnet model (97.5%)
- Created a new dataset of real images using **data augmentation** which significantly increased the accuracy

Deep Learning: Earthquake Forecasting *Aug-Dec'20*
- o Interpreted **time series data of earthquakes** in five different regions to extract the seismicity information
- o Implemented a neural network **model which forecasts earthquakes** using seismicity indicators in the regions
- o Achieved an **accuracy of 90.4%** for forecasting the probability of an upcoming earthquake in the Himalayas

Computer Vision: Facial Recognition Based Attendance System *May-July'20*
- o Designed a **facial-recognition based attendance system** to help curb the spread of COVID-19
- o Used **openCV library (Haar Cascade Algorithm)** for facial recognition

Statistical Analysis: Epidemiological Analysis of COVID-19 *Jan-May'20*
- o Analyzed COVID – 19 data with respect to the **SIR epidemic model of disease spread**
- o Estimated the defining characteristic parameters of the model by **minimizing squared error loss**
- o Calculated the **reproductive number to be close to 1.2**

Deep Learning: English to Hindi Language Transliteration *May-June'20*
- o Trained an **Encoder-Decoder model** which transliterated English alphabets to Hindi language font
- o Deployed **Gated Recurrent Units with attention mechanism** to enhance the performance of the model

Computer Vision: Occlusion Analysis & Filter Visualization *May-June'20*
- o Analyzed the filter in a CNN for detecting the important parts of an image
- o Performed **occlusion sensitivity analysis** on various images

Mentorship Experience

Teaching Assistant *BITS F312 : Neural Networks & Fuzzy Logic*
Prof. Surekha Bhanot, BITS Pilani *Aug-Dec'22*
- o Guided a class of approximately 100 students and was responsible for their assignments
- o Supervised multiple groups of 3-4 students in their projects

Technical Skills

Programming Languages	- C, C++, Java, Python, MATLAB, SQL, Shell Script
Data Science Libraries	- Boto3, PyTorch, TensorFlow, Keras, Pandas, Numpy, openCV
Machine Learning	- CNNs, RNNs, GRUs, LSTMs, Transformers, Reinforcement Learning
Platforms/Tools	- AWS (all services), Google Colab, Jupyter Notebook, MATLAB

Extracurricular Activities

- o **Member of Student's Academic Council** - Organized various talks and activities to augment the academic and research culture of the campus *[2020-2022]*
- o **Volunteer at National Service Scheme (NSS)** - Mentored underprivileged students and tutored them on their curriculum subjects *[2018-2020]*
- o **Game Developer at Coding Club** - Developed several games on the unity platform as well as designed the characters on the blender platform *[2018-2019]*