Divij Khaitan

9167434438 | divijkhaitan@gmail.com | divijkhaitan.github.io | linkedin.com/in/divijkhaitan

Education

Ashoka University Sonepat, Haryana

B.Sc. Hons Mathematics and Computer Science + Diploma in Advanced Studies and Research *Aug. 2021 – May 2025*

- CGPA: 3.82/4.0 (95%), Magna Cum Laude
- Teaching Assistantships: Trustworthy AI, Matrix Analysis, Discrete Mathematics, Data Structures
- Relevant Coursework: Advanced Algorithms, Theory of Probability, Reinforcement Learning, Matrix Analysis, Numerical Algorithms and Optimisation, Geometric Deep Learning, Differential Geometry
- Thesis: A New Framework to Asses the Robustness of Deep Neural Networks (Advisor: Subhashis Banerjee)

Technical Skills

Languages: Java, Python, C, SQL, JavaScript, HTML/CSS, R, LaTeX
Frameworks: React, Express, TensorFlow/Keras, PyTorch
Developer Tools: Git, Docker, VS Code, Shell scripting
Libraries: Pandas, NumPy/Scipy, Matplotlib, OpenCV, NetworkX, Scikit-Learn, xarray, jax

Experience

Research Assistant May. 2024 – Present

Centre for Data, Learning and Decision Sciences *Sonepat, Haryana*

- Enchanced the skill of **heatwave forecasting by 10%** using **CNNs, Transformers and GNNs** in torch
- Evaluated **anomaly detection** algorithms on **graph structured data** for **identifying financial fraud**
- Implemented and backtested **portfolio optimisation** methods on **20 years** of data and **1300+ stocks in jax**
- Constructed pipelines for **translating** and **fine-tuning foundation models** to improve forecasting skill

Undergraduate Research Assistant Dec. 2023 – Feb 2024

Koita Center for Digital Health - A *Sonepat, Haryana*

- Built and fine-tuned **deep learning models** including U-Nets for the **segmentation of 3D medical images**
- Trained ML models including SVMs and MLPs for **cancer classification**, improving skill by **12%**
- Containerized models into **Docker images** for **constrained testing** on a remote platform
- **Engineered features** by analysing **100X histopathology images**, **improving accuracy by 10%**

Undergraduate Research Assistant May 2023 – Aug. 2023

ML2CT, Ashoka University *Sonepat, Haryana*

- Benchmarked **HPC cluster hardware** by training a **110 million parameter** model on **6.7 million datapoints**
- Fine-tuned **encoder-only transformers** in **tensorflow** to **predict protein functions** with **70% accuracy**
- Created **BERT models** for **DNA promoter sequences** to predict gene expression

Projects

Robustness of Deep Neural Networks | *Python, Numpy, Torch, Scipy, OpenCV* Ongoing

- Investigating the phenomenon of **overfitting** in deep neural networks
- Analysing the **distribution of activations** in the hidden layers of **deep vision classifiers**
- Exploring applications of work to **OOD detection**, **adversarial robustness** and the **manifold hypothesis**

Applications of SVD | *Python, Numpy, OpenCV* Oct. 2023

- Implemented **efficient, numerically stable** algorithms for **Singular Value Decomposition**
- Curated a **private dataset of 200 images** and implemented **facial recognition** algorithms with **93%** accuracy
- Created a **library of 20 textbooks** and performed **information retrieval** using **latent semantic indexing**

RandNLA | *Rust, Git* Aug. 2024 - Dec. 2024

- Developed a **library** for randomised matrix algorithms including **low rank decompositions** and **least squares**
- Optimised memory management of **probabilistic sketching techniques** to obtain upto **20× faster** runtimes
- Designed and deployed a comprehensive suite of **automated tests** to a **CI/CD pipeline** for software reliability

Deep Learning for Single Cell Genomics | *Python, Torch, Tensorflow, Bash* Sep. 2023 – Dec. 2023

- Engineered graph-based models to predict single cell differential expression given a categorical variable
- Implemented **deep generative models** to impute **single cell perturbation** data in over **20000 dimensions**
- **Automated** differential expression analysis for almost **150 different groups** using **python, R and bash**

Consistent Hash | *C* Dec. 2022

- Built a **consistently-hashed binary search tree** of linked lists from scratch in C
- Implemented **dynamic addition**, **deletion** and **redistribution** for load balancing in distributed services
- Designed a **complete and exhaustive** suite of **test cases** to **ensure sound functionality** of the structure

Big Number Package | *Python* Apr. 2022

- Developed a package to provide **integer types of arbitrary size** in python using list operations
- Implemented the primitive arithmetic operations of **addition, subtraction, multiplication and division**
- Built algorithms such as the **extended GCD** algorithm, **fast modular exponentiation** and **Miller-Rabin**
- Created an **RSA encryption scheme** for the **secure transmission** of messages over an **unsecure channel**

Awards and Honours

Dean's List: Monsoon (2021, 2022, 2023, 2024), Spring 2023
Academic Excellence Award in Mathematics: Exceptional academic performance exceeding historical standards
Silver Award, HKIMO: Finished 97th in the world in the Hong Kong International Math Olympiad 2020