Siddharth Arya

647-400-4151 | sid19arya@gmail.com | linkedin.com/in/sid19arya | github.com/sid19arya

Education

University of Toronto Toronto, ON

B.S. in Computer Science (Minor in Statistics and Math) GPA: 3.89 *Expected May 2025*

- Recipient of University of Toronto International Scholar Award - Scholarship
- Coursework: Data Structures, Discrete Maths, Algorithms, Digital Circuit Design, Machine Structure and Assembly-language Programming, Linear Algebra, Software Design (Clean Architecture, Design Patterns), Probability, Statistics, Intro to Data Science, Intro to Machine Learning, Operating Systems, Computer Vision, Natural Language Processing (NLP), Deep Learning

Technical Skills

Languages: Python, C, SQL (Postgres),Java, JavaScript, Typescript, HTML/CSS, R
Frameworks: React, Node.js, ExpressJS, JUnit
Developer Tools: Git, Docker, Microsoft Azure, AWS, Linux(Shell scripting), VS Code, Slurm Workload Manager
Libraries: PyTorch, Pandas, NumPy, Matplotlib, Scikit-Learn, Pytest, Hugging Face (transformers)

Experience

Research Intern: Machine Learning May 2024 - August 2024

Data Science Institute, University of Toronto *Toronto, ON*

- Engineered a novel Machine Learning based method to **monitor and evaluate performance of deployed Deep-Neural-Networks**, achieving a **93% True Positive Rate** in foreseeing model failure, ensuring **proactive model reliability** and **performance** - in collaboration with peers at the Vector Institute of Technology
- Organized and cleaned data for over $\sim$ **200,000 patients** into **900 features** (lab results, vitals, demographics) using **SQL** and **Numpy**, and trained neural networks to achieve $\sim$**95%** accuracy in predicting 14-day mortality
- Led a comprehensive benchmark study evaluating the performance of various shift detection methods, implementing solutions in **PyTorch** and **Scikit-learn** for both real-world medical and semi-synthetic data shifts, funded by the **Data Science Institute at the University of Toronto**
- Presented research findings at **Showcase Day** among a select group of grant recipients, highlighting the efficacy of shift detection methods and the importance of this research towards **Reliable AI**

Research Projects

Low-light, High-speed Imaging | *Computational Imaging* September 2024 - *Current*

- Supervised by **Prof. Kyros Kutulakos** at **Toronto Computational Imaging Group**
- Conducted experiments to characterise key behaviour in SPAD-512 (single photon avalanche diode) camera
- Developed on **Fourier Probing Regime** on photon timestamp data for **higher SNR** ratio of probed frequencies

Image Captioning for Selective Prediction | *Trustworthy AI* September 2024 - *Current*

- Supervised by **Prof. Marsha Chechik** at **Department of Computer Science** at the University of Toronto
- Explored auxiliary **Image Captioning** model to tag inputs at test-time in order to construct **Functional Requirements** for **Machine Vision Component** in **Facial Recognition Software**

Explaining Distribution Shift using SHAP | *Explainable AI* January 2024- May 2024

- Supervised by **Prof. Rahul G. Krishnan** at **Vector Institute of Technology**
- Documented degradation of **predictive model** performance when evaluated on patients from different hospitals
- Characterised the shift in **distribution-shift** using **SHAP values** to capture functional relationship of variables

GRA-KNN: Feature Imporance in KNN | *Supervised Learning* September 2023- December 2023

- Implemented and Benchmarked several techniques - **KNN**, **Principle Component Analysis** and **Decision Trees** - on their efficacy for classification on the MNIST Dataset using **Scikit-Learn**
- Explored a Novel Modification on the KNN, which **reduced input features by 87%** while ultimately achieving **93.09% accuracy** on classifying inputs on the MNIST dataset