Diya Ilinani

✉ di2078@nyu.edu   ⚲ https://github.com/i-diya/   ☐ 646-371-4454

EDUCATION

New York University (Courant Institute of Mathematical Sciences)   New York, NY
Master of Science in Computer Science   *Sep.2021 – May. 2023*

Indian Institute of Technology   Dharwad, India
Bachelor of Technology in Computer Science and Engineering   *Aug. 2017 – July. 2021*

PROGRAMMING SKILLS / COURSEWORK

- **Languages and Tools**: Python, C, C++, CUDA, OpenMP, PostgreSQL, Java, JavaScript, Scala, R, MATLAB. HTML, CSS, PHP, MySQL, Bash, Verilog, Scheme, Radare2, AutoCAD.

EXPERIENCE

Citigroup Inc.
Software Developer   *Summer 2022*

- Developed a sandbox testing platform, enabling untested trading bots to migrate to a **Docker** container autonomously. Enhanced the performance testing efficiency using **InfluxDB.** and RabbitMQ
- Authored Python scripts to generate reports of malicious activities by parsing trading bots for unauthorized **OS accesses** and calls exceeding 80 different types, improving the pre-testing security measures.

Candle Research Lab.
Research Assistant to Dr.Sparsh Mittal, IIT Roorkee   *Fall 2020*

- Crafted Python scripts to assess the efficiency of Deep Neural Networks accelerator architectures using power, energy and resource utilization metrics.
- Conducted analysis to **optimize data reuse (by taking data-types into cognizance)** and **Processing Element Utilization** in DNN accelerators, using timeloop to emulate a DNN workload on accelerators.

PROJECTS

Multi-modal Learning for Early Detection of Alzheimer's Disease.
Python (pandas, scipy, nibabel, scikit-learn), XGBoost, 3D CNN, Auto-encoder   *Fall 2022*

- Constructed late fusion multi-modal models, utilizing clinical data and **3D MRI images** for early Alzheimer's detection by **predicting the stage of cognitive impairment in an year**.
- Developed joint models deploying **3D CNNs** and **Auto-encoders** that improve the quality of features for clinical data, enhancing the AUROC from **54%** for the MRI images model to **86%** for the joint **VGG** model. Achieved an accuracy of **92%** for the clinical data model.

DF-GAN+: Improved Learned Embeddings for DF-GAN.
PyTorch, Generative Adversarial Networks   *Fall 2022*

- Implemented contextual word embeddings from **Large Language Models (LLMs)** like XLM, RoBERTa, GPT-2, BERT, combined with image-level features using Deep Fusion Blocks to generate better expressions of color descriptions.
- Validated hypothesis that extensive training corpora of implemented LLMs yield superior quality embeddings and lead to higher quality images . XLM demonstrated best performance, improving the FID metric (used to determine realism in GANs) to **17.81** over the baseline of **19.20**.

Malware Detection: Static and Dynamic analysis of windows executables.
Python, JSON, Radare2   *Fall 2020, Spring 2021*

- Devised static and dynamic analysis techniques for malicious Windows executables, leveraging a dataset of 2,054 malicious and 323 benign files.
- Achieved a detection accuracy of **98.18%** for static analysis using **SVC** and **88.16%** for dynamic analysis. Conceived a novel method using the frequency of occurrence of **API calls** that determine vulnerability, and basic block **control flow graphs** to predict the malignant nature of a file for static analysis.