Harshita Kukreja

Email: hk3203@nyu.edu | Phone: +1 (917) 325-8134 | **in** harshitakukreja8 | 🐦 Harshita Kukreja | NY, USA

EDUCATION

New York University, Courant Institute of Mathematical Sciences   **New York, NY**
MS Computer Science [GPA: 3.81/4.0]   Sep 2021 – May 2023

Indira Gandhi Delhi Technical University for Women   **Delhi, India**
B.Tech Computer Science and Engineering [Aggregate: 81.3%]   Aug 2016 – July 2020

EXPERIENCE

Machine Learning - Research Associate   **New York, NY**
NYU Langone Health   Jul 2023 - Present
Contrast-Enhanced MRI using Non-Contrast 3D Images
- Spearheading the **Machine Learning** efforts for generating agentless contrast-enhanced MRIs.
- Leveraging **PyTorch Lightning** along with **nibabel** and **MONAI** to build, train and evaluate **3D UNet** model on **NYUMets** database managing development with **git**.
- Resultant MRIs aim to reduce image processing times leading to less patient discomfort and more efficient medical data analysis.

Machine Learning Graduate Researcher   **New York, NY**
NYU Grossman School of Medicine   Aug 2022 - May 2023
Adapting GLOM for Videos
- Investigated improvement in **video**-based tasks for neuroscience research through representation learning using part-whole hierarchies in the **GLOM** model.
- Adapted the concept **Deep Learning** model's implementation for images to video using **PyTorch Lightning**, **TorchVision**, and **TorchVideo**.

Less Common Mutation Prediction in Lung Cancer
- Performed **Exploratory Data Analysis** on Lung Adenocarcinoma cell scans to predict less common EGFR mutation.
- Preprocessed the high resolution whole slide image files (**5GB** per image) with **OpenSlide**.
- Tailored the registration process for pixel-to-pixel correspondence using **Python**, **OpenCV**, and **scikit-image**.

Software Engineer Intern   **Boston, MA**
Tech For Good Inc.   Jun 2022 - Aug 2022
- Engineered network congestion reduction with **LeCAR** Machine Learning algorithm to **cache** data.
- Reduced network latency, bridging the digital divide of internet accessibility for underrepresented communities.
- Achieved a **10%** increase in cache-hit ratio, optimizing network efficiency.

Deep Learning Research Associate   **Delhi, India**
Netaji Subhas University of Technology   Jun 2019 - Aug 2021
Denoising and Classification of Biometric Images   [Paper]
- Diminished Non-Local Means filter performance time by **75%**, improving biometric image denoising on CASIA-Iris-M1 and Tongjo Palmprint datasets.
- Achieved higher PSNR values of upto **40.86** in **4.4 seconds** and classification accuracy of **98.39%** for denoised image classification on the pretrained ResNet50 model in PyTorch.

Privacy Protection of Biometric Templates   [Paper 1] [Paper 2]
- Engineered DeepCrypt, combining **deep CNN** with **cryptographic hash** function enabling cloud storage of biometric templates.
- Achieved matching authentication performance of **99.56%** on the LFW and Casia databases, alleviating the risk of identity theft and enhancing security of the system.

Multimodal Biometric System for User Identification   [Paper]
- Generated a reliable multimodal biometric authentication system by fusing iris (IITD IRIS dataset) and face (CASIA Face dataset) modalities at the feature level using **TensorFlow** and **Keras**.
- Improved image recognition rate by attaining an accuracy of **99.8%** along with mitigating security limitations.

Software Developer Intern   **Delhi, India**
Leiothrix Technologies Pvt. Ltd   Jan 2018 - May 2018
- Designed a full-stack web-based solution to connect users with mental health specialists using **HTML** and **CSS** for frontend and **PHP** for backend, improving accessibility and support for users by utilizing **Skype API** to set up meetings and send invites.
- Streamlined the dynamic **RSS** feed to provide relevant articles on mental health based on user preferences and behavior, resulting in a **20%** increase in user engagement.

PROJECTS

Social Media Analysis of GLP-1RA   [POSTER]
Digital Health Lab, Johns Hopkins Medicine International   [NLP, BERTopic, Tweepy]
- Constructed a web scraping pipeline for collecting and preprocessing tweets using **Twitter API v2**, **Python**, and **Tweepy**.
- Applied topic modeling using **BERTopic**, **dimensionality reduction**, and **TF-IDF** scores to identify key themes.
- Performed sentiment analysis with **RoBERTa** to quantitatively assess the sentiments expressed in tweets.

Promotions Resource in an E-Commerce Website [CODE]

Course: DevOps & Agile Methodologies [Docker, PostgreSQL, Kubernetes]

- Designed the back-end for the Promotions team of an eCommerce website as a collection of **REST API** services for a client.
- Worked on **Docker** containers with **Kubernetes CI/CD** pipelines to integrate code with **nosetests** before deploying it, simulating an Agile development with biweekly sprints.

Autism Spectrum Disorder Screening using Predictive Analytics [CODE]

Course: Data Analytics & Visualization in Healthcare [Tableau, R, Python, Scikit-Learn]

- Performed **EDA** using **Tableau**, **R**, **Pandas**, and **Matplotlib** on the Autism Spectrum Disorder (ASD) dataset to screen children, adolescents, and adults likely to have ASD.
- Showcased exemplary **supervised machine learning** approaches including **Logistic Regression**, **Random Forest** and **Decision Trees** to predict ASD, with Random Forest achieving the best **F1-score** of **99%**.

Parallelizing Non Convex Optimization [CODE]

Course: GPU Architecture & Programming [GPU, CUDA, Optimization, C++]

- Achieved a **980x** speedup in non-convex problem optimization by optimizing Genetic Algorithm and Gradient Descent, handling a dataset of **65,000** data points.

Stackable CNN and Transformers

New York University [PTorch Lightning, Vision Transformer, CNN]

- Developed stackable **CNN** - **Transformer** blocks to build deeper Vision Transformer networks.
- Evaluated the framework on **COCO** dataset for object detection tasks.

Multimodal Siamese Neural Network [CODE]

Course: Introduction to Deep Learning Systems [Tensorflow, Keras, Pillow, Jupyter]

- Built a multimodal biometric security model in **TensorFlow**, utilizing Siamese Convolutional Neural Networks (CNNs) for **One-Shot Learning**, to provide similarity scores for face-sign image pairs.
- Employed advanced deep learning techniques to enhance biometric security measures, contributing to robust and efficient authentication systems capable of handling **multimodal** data inputs.

SKILLS

Languages	Python, Java, C++, C, R, SQL, HTML, CSS, JavaScript, Embedded C, LaTeX
Libraries/Tech	PyTorch, PyTorchLightning, TensorFlow, Keras, Jupyter, NumPy, Pandas, scikit-learn, Matplotlib, Seaborn, Tableau, Linux, Git, Agile, CI/CD, Docker, Kubernetes, SQLAlchemy, A/B Testing, IBM Cloud
Coursework	Data Visualization & Analytics, Deep Learning Systems, Computer Vision, Machine Learning, DevOps, Database, Algorithms, Data Structures, Compilers, Object Oriented Software Engineering, Networking

ACADEMIC SERVICE

Teaching Assistant Sep 2022 - May 2023

Handled grading of assignments and tutoring for students, created and hosted hands-on exercises, and provided feedback based on students' answers for Machine Learning, Computer Vision, Artificial Intelligence, and Algorithms courses.