

Isha Arora

Boston, MA 02120 | (617) 652 6728 | arora.isha4128@gmail.com

[Portfolio](#) | [LinkedIn](#) | [GitHub](#)

EXPERIENCE

Research Trainee | Northeastern University, Boston

Feb 2023 – Dec 2023

- Enhanced **state-of-the-art** deep learning models for grading **Prostate Cancer** using **10616 WSIs** (PANDA) and **192 WSIs** (DiagSet).
- Diagnosed with a new proposed configuration of EfficientNet-B1 model acquiring **0.66 QWK** and **0.81 weighted accuracy**.
- Engineered a use-case to extend to **Breast Cancer** with **9109 WSIs** (BreakHis) and RNA-sequencing data (TCGA) for **multi-modality**.

Research Student | Massachusetts General Hospital, Dekel Laboratory, Boston

Jan 2023 – Aug 2023

- Researched **traditional ML** and **deep learning** approaches using Electronic Medical Records to identify the risk factors for CB-PTSD.
- Inspected PTSD reporting metrics using **statistical modeling** assessed on **59 patients**.
- Orchestrated a data-driven approach to reporting metrics, resulting in an impressive **AUC-ROC of 0.94** and a **correlation value of 0.82**.

Associate Engineer – Technology | Virtusa Consulting Services Pvt. Ltd., India

Aug 2020 – Aug 2021

- Coordinated with **Wolters Kluwer USA** to develop a system hosting regulations in banking and insurance with an **Agile framework** using **PostgreSQL** for the database.
- Accelerated the lookup time for laws and the system aimed to help **improve client efficiency by at least 60% in 5 months**.
- Refined the data sent in by the client using **OpenRefine** and to assemble according to the US states and the relevant books.

Data Analytics Intern | Financial Software and Systems Pvt. Ltd, India

Dec 2019 – May 2020

- Conceptualized a project on spam detection on **4480 reviews** for a banking application.
- Implemented **Naive Bayes**, **Decision Tree**, and the **Apriori algorithm** for spam detection realizing an **accuracy of 61%**.
- Executed the **VADER algorithm** for sentiment analysis and **at least 50%** of the reviews marked as non-spam were **positive**.

PUBLICATIONS

Establishing the validity of a diagnostic questionnaire for childbirth-related post-traumatic stress disorder 2023

- Validate the use of the self-reporting PCL-5 checklist to assess CB-PTSD against the Clinician CAPS-5 for **59 patients**.
- Cutoff value of 28** maximized the **sensitivity (0.80)** and **specificity (0.93)**, and correctly **diagnosed 86%** of women.
- Youden J-index of 0.71** and an **overall diagnostic efficiency of 86%** helped identify the cutoff score.

PROJECTS

Exploring User Accessibility and Human-Machine Interaction Using EMG

[\[GitHub\]](#)

- Programmed a gesture recognition and user identification model using Electromyogram data to help disabled people with mobility issues.
- Formulated a neural network model for gesture recognition with an **accuracy of 0.91** and **F1-score of 0.9**.
- Yielded a LSTM user classification **accuracy of 0.94** and cross-day **rank-5 accuracy of 0.8**.

The Song Search

[\[GitHub\]](#)

- Formulated an **information retrieval system** for audio files by referencing the **MT3 model** from **TensorFlow Magenta**.
- Curated and crafted** a specific data set for audio data and proposed a model by finding efficient representation of songs.
- Attained an **accuracy of 74%** in the top 5 candidate set alongside a **MAP of 0.68**.

Deep Clustering for Unsupervised Learning of Visual Features - A Reproduction

[\[GitHub\]](#)

- Reproduced “**Deep Clustering for Unsupervised Learning of Visual Features**” by Facebook AI Research creating the DeepCluster network with **Power Iteration Clustering** and **AlexNet**.
- Clustered using a subset of **ImageNet dataset** of **64 classes** with **600 images each**, and an **external dataset** with **28000 images**.
- Analyzed the **NMI** between each new and previous cluster and attained an approximate value of **0.8**.

EDUCATION

Northeastern University, Boston, MA

Dec 2023

Master of Science in Data Science

GPA: 3.83

Vellore Institute of Technology, Vellore, India

Jul 2020

Bachelor of Technology in Computer Science and Engineering

GPA: 8.67/10

TECHNICAL KNOWLEDGE

Languages:

Python | R | RStudio | SQL | C++ | Java | MATLAB

Database:

MySQL | PostgreSQL | Oracle PL/SQL

Libraries and Frameworks:

AWS | Pandas | NumPy | Matplotlib | Scikit | seaborn | Keras | TensorFlow | PyTorch | OpenCV | ggplot | GitHub | NLTK | OpenRefine | PowerBI

Technical Applied Skills:

Statistics | Data Mining | Data Science | Machine Learning | Neural Networks | Computer Vision | NLP | Artificial Intelligence