

# CHIRAG MALHOTRA

Boston, MA | +1 857-395-1120 | malhotra.chdelhi@gmail.com  
[linkedin.com/in/chirag-malhotra-delhi/](https://www.linkedin.com/in/chirag-malhotra-delhi/) | [github.com/chiragml](https://github.com/chiragml) | [chiragml.github.io/me/](https://chiragml.github.io/me/)

## Education

### Northeastern University

Boston, MA

*Masters of Science in Computer Science, GPA: 3.92*

*{Expected} May 2025*

- Relevant Courses: Machine Learning, Database Management Systems, Object Oriented Design, Data Structures and Algorithms, Foundations of software Engineering, Large Scale Parallel Data Processing

## Technical Skills

**Languages:** Java, Python, C, C++, SQL, JavaScript, HTML/CSS, DLang, Scala, ROS1

**Frameworks/Architectures:** Pytorch, Tensorflow, DWSA, GLoVe, ResNet, Unet, YOLO, AWS, Hadoop, Apache Spark

**Concepts:** PPO, prototype learning, feature extraction, Q-learning, LLMs, fine tuning, model evaluations **Libraries:** pandas, NumPy, Matplotlib, seaborn, spacy, stablebaseline 3, ML agents, QLora, Lora

## Experience

### CivicAI Lab

Boston, MA

*Research Assistant | Python, Poetry, React, Node js, Pandas, Numpy*

*January 2025 – present*

- Working towards a ICCV paper **banchmarking** VLMs instruction following abilities.
- Utilizing **GCP's Vertex AI and Lambda labs** to generate responses on **Nvidia A100** in parallel using **multi-threading**.

### International Panel on the Information Environment (IPIE)

Remote

*Gen AI engineer | Python, Poetry, React, Redux, JS, Django, kmeans, Dashboard*

*January 2025 – present*

- Building **GenAI** based **dashboard** to **cluster and analyze** affiliate data at IPIE to help guide talent acquisition.
- Developing and deploying the **containerized** dashboard with **React and Redux** on frontend and **Django** for the backend.

### Northeastern University

Boston, MA

*Graduate Teaching Assistant | Reinforcement Learning & OOD | Python, Q-learning, Java, OOPs*

*Jan 2023 – present*

- Improved student performance in Object-Oriented Design and reinforcement learning by organizing tutoring sessions, grading assignments, leading to a combined **12% improvement** in course outcomes.

*Graduate Research Assistant | Python, Docker, Spacy, DWSA, GLoVe, ResNet*

*June 2023 – Nov 2023*

- Applied temporal for video data and contextual embeddings for text to achieve **over 5 % improvement** to model accuracy.
- Utilized DWSA architecture with soft-ordered prototype learning and Spacy to map video and text features effectively.

### myRide

Boston, MA

*Machine Learning Engineer Co-op — Advanced RAG, ROS1, AWS, Python, C++, TensorFlow* *August 2024 – December 2024*

- Developed perception and path planning systems for a self-driving car using **ROS1, roscco and OpenCV**.
- Implemented digital control of vehicle systems, leveraging **OSCC (Open Source Car Control)** for core functionality
- Designed a **cloud-based language model** on **AWS** to serve as the central intelligence system for autonomous decision-making

### LightBird AI

Boston, MA

*GenAI Engineer Co-op | python, LLMs, spacy, NLP, prompt engineering*

*May 2024 - August 2024*

- Fine-Tuned BERTopic Model to improve **accuracy by 20 %** for domains as well as the clients.
- Implemented intent engine powered by domain expert model allowed **reduce on-boarding time by 17%**.
- Conducted quantitative and qualitative evaluations, ensuring models met enterprise safety and reliability standards.

## Projects

### 2D Unet for semantic segmentation | [Ongoing](#)

- Building a **2D Unet from scratch**, implementing all the important components and comparing it with **segmentation\_models\_2D's Unet** architecture for **microscopy** application.

### Diffusion Gif Generator | <https://github.com/chiragml/Diffusion-Gif-Generator>

- Built a diffusion-based image-to-video generator, cutting GIF generation time from **40 to 9 minutes** with **CPU offloading**, memory and attention optimization.
- Achieved **77%** runtime reduction using **adaptive pruning (30%)** with sensitivity analysis, maintaining high-quality output.

### 6 Degrees of Separation | <https://github.com/chiragml/6DegreesOfSeperation>

- Designed and built a **Spark** application to parse and traverse a **900 Billion** edges graph from twitter indicating social network.
- Applied Dijkstra's algorithm to the graph in parallel and found out the optimal cluster configuration on **AWS EMR and S3**.
- Used **RDDs** to handle and distribute **2gb csv(stored as .txt)** data over **8, 10 and 12 nodes** for comparative analysis.