# Armen (Ahmadreza) Jeddi (Feb 2024)

Ahmadreza.Jeddi@gmail.com | GitHub | LinkedIn | Website | Google Scholar

## Summary

- Passionate problem-solver with 5+ of experience in predictive modeling and applying ML to solve high value problems. Proficient with designing, optimizing, and deploying ML models
- A Deep Learning / ML engineer with a Master's degree in Computer Science and an expert in data analysis with Python, NumPy and Pandas and ML frameworks such as **PyTorch**, **Tensorflow**, and OpenCV
- A researcher at heart, having publications in conferences such as **NeurIPS**, **CVPR**, and IEEE journals

## Experience

- **Zown** *July 2023- Jan. 2024*
  *Head of AI*
  - ❖ Implemented a computer vision pipeline that takes photos of houses from users, applies **LLM** models on the images to suggest multiple options for renovation for the user and then feeds the selected options from the user along with **segmentation masks** to a **ControlNet variant of Diffusion Models** to apply desired text guided renovations/staging on the images
  - ❖ Designed and fully implemented the most advanced real estate property valuation system in Canada with a median absolute error rate of %4, and then extending this to market forecasting using transformer-based tabular data regression backbones such as **TabNet**, and **Tab-Transformer**
  - ❖ Designed and fully implemented tree-based likelihood-to-sell models for optimized property pricing
  - ❖ Implemented an advanced data preprocessing and cleaning pipeline using tabular data imputation techniques
  - ❖ Implementing a GPT-4 chatbot for the users that interacts with users through text or online chat, performs API calls to our AI endpoints and collects the user behavior info in order to perform business analytics
  - ❖ All the implementations & research are done in GitHub and the Deployment is done with Azure and Google Cloud resources via docker and nginx served endpoints

- **Deeplite** *Jan. 2022- Aug. 2023*
  *Deep Learning Engineer*
  - ❖ Conducted extensive research on on low-bit **quantization** methods and their effect on the loss landscape and **sharpness** of neural networks (Paper submitted to NeurIPS 2023 - 1 weak accept & 2 borderline reviews)
  - ❖ Developed and deployed a SOTA hardware-aware **neural architecture search** framework, which creates hybrid neural networks with both transformer and CNN blocks, improving inference times up to %80 while maintaining performance levels
  - ❖ Applied data mining and semi-supervision on over 1.4 million data samples to achieve larger datasets for model pre-training, improving object detection model performances %2-9
  - ❖ Designed and developed passenger and pet detection models applying **pruning** and **quantization** improving processing FPS from 3 to 12 for BlackBerry's Ivy platform with a 2 mAP drop budget
  - ❖ Optimized person detection, **re-identification**, and **multi-object tracking** models for theft threat detection in smart cars, improving model inference time 4x by implementing **feature map sharing** and hardware specific optimizations – with 97% person detection accuracy and 92% object tracking accuracy

- **MARZ VFX**                                                                    *Oct. 2021- Jan. 2022*
  *Applied Machine Learning Research Scientist*
  - ❖ Designed and developed a pipeline for **facial re-enactment** of a target actor by using a source actor's audio applying face alignment, and **3D monocular face reconstruction** from 2D images
  - ❖ Implemented a **neural rendering** pipeline for generating the face of a target actor by applying UNET translation models that translate UV maps of a 3D face model into photorealistic 2D images

- **HMI lab, Huawei Technologies**                                                *Sep. 2020– Oct. 2021*
  *Associate Machine Learning Engineer*
  - ❖ Developed and optimized deep learning based gesture recognition models for smart TVs and cars by utilizing OpenCV, **PyTorch**, TensorFlow, **Albumentations**, and SOTA research & code bases
  - ❖ Applied data mining techniques such as Cleanlab to filter Out-Of-Distribution (OOD) data samples from a set of 700,000 hand gesture samples, improving classification model performances 2-5%
  - ❖ Developed and deployed a Nanodet based hand detection model on Huawei P40 smartphones, with an inference time of 4ms, hugely improving the inference time from prior Detectron2 based models
  - ❖ Implemented a very fast real-time LSTM based hand tracker with a tracking accuracy of over 94% which reduced our calls to the more costly hand detector model by 8 times

- **University of Waterloo**                                                       *Jan. 2019 – Aug. 2020*
  *Graduate Research Assistant*
  - ❖ Developed and published Adversarial Fine-tuning (**AFT**) in **CVPR 2021**, a SOTA approach to improve **adversarial robustness**. Models were trained in PyTorch.
  - ❖ Developed and published **Learn2Perturb**, a SOTA approach that improves the **adversarial robustness** of deep neural networks by introducing a **trainable noise injection** module.
  - ❖ Improved the **semantic segmentation** in the case of limited data. **Weak supervision, CRF-based loss regularization, transfer learning** and **DeepLab-V3** were used in this project.

## Skills Overview

- **ML Tools:** PyTorch [5 years, Proficient], TensorFlow [5 years, Proficient], OpenCV [5 years, Proficient], NumPy [6 years, Proficient], Keras [5 years, Intermediate]
- **Programming Languages**: Python [9 years, Advanced], Java [8 years, Proficient], Scala [6 years, Proficient], C [6 years, Intermediate], C++ [6 years, Beginner]
- **Big Data and Data Management Tools:** Hadoop [Proficient], Spark [Proficient], SQL [Proficient], Azure & Google Cloud
- **Other Tools**: Git, Visual Studio Code, Pycharm, Linux, CUDA toolkit

## Education

- **University of Waterloo**                                                       *Waterloo, Canada*
  *M.Sc. in Computer Science*                                                      *Jan. 2019 – Aug. 2020*
  **GPA:** 98 / 100
  **Thesis:** Learn2Perturb: Improving Adversarial Robustness on Deep Neural Networks through End-to-end Feature Perturbation Learning *with* **Professor Alexander Wong**
  **Courses**:   Pattern Recognition | Optimization in Data Science | Advanced Topics in Computer Vision (with **Yuri Boykov**)

- **Sharif University of Technology**                                              *Tehran, Iran*
  *B.Sc. in Software Engineering*                                                  *2014 – 2018*
  **GPA:** 17.86 / 20

**Notable Courses:**   Introduction to Artificial Intelligence | Stochastic Processes and Advanced Statistical Analysis | Data Mining | Modern Information Retrieval | Introduction to Big Data | Object Oriented Software Design

## Research Experience

**VIP (Vision and Image Processing) Lab** <span style="float:right">*University of Waterloo / Canada*</span>

*Graduate Research Assistant, Supervisor: Prof. Alexander Wong & MJ Shafiee* <span style="float:right">*Jan. 2019 – Aug. 2020*</span>

Deep computer vision models robust to adversarial examples, applying manifold learning and representation learning for the purpose of adversarial robustness

## Publications

- **A. Jeddi**, M.J. Shafiee, M. Karg, A. Wong, "Learn2Perturb: Improving Adversarial Robustness on Deep Neural Networks through End-to-end Feature Perturbation Learning", **CVPR 2020** main conference.
- M.J. Shafiee, **A. Jeddi**, A. Nazemi, P. Fieguth, A. Wong, "Deep neural networks and robust autonomous driving systems", **IEEE Signal Processing Magazine**, Special Issue on Autonomous Driving.
- *A. Jeddi, M. J. Shafiee, A. Wong, "A* Simple Fine-tuning Is All You Need: Towards Robust Deep Learning Via Adversarial Fine-tuning". **Best Paper Award at CVPR 2021** AML-CV Workshop.
- *A. Jeddi, "T*ackling the Problem of Limited Data in Semantic Segmentation"*. Available on Arxiv.*