

NEEL VORA

+1(682) 374-6877 ◊ San Francisco, CA, USA

neelvora27@gmail.com ◊ [linkedin/neelvora27](https://www.linkedin.com/in/neelvora27) ◊ wsslab.org/neelvora ◊ [github/freaksie](https://github.com/freaksie) ◊ [scholar/neelvora](https://scholar.google.com/citations?user=neelvora)

Results-oriented Computer Science graduate with expertise in machine learning and software development. Dedicated to drive innovation and efficiency in machine learning model development and deployment. Seeking roles where I can leverage my expertise in ML algorithms that contribute towards impactful projects.

EDUCATION

The University of Texas at Arlington, TX, USA

Aug 2022 - May 2024

Master of Science in Computer Science (Thesis)

(GPA: 4/4)

Dharmsinh Desai University, India

Aug 2019 - May 2022

Bachelor in Information Technology

(GPA: 8.21/10)

EXPERIENCE

Lawrence Berkeley National Lab, Machine Learning Engineer Intern, (CA, USA)

Aug 2023 - Present

As part of the Advance Quantum Testbed (AQT) team, I worked on the development of a machine-learning model for rapid quantum information processing and deploying these models on FPGAs and the cloud infrastructure for real-time and remote inference (*Mentors: Yihun Xu, Gang Huang*)

- Implemented non-Markovian noise modeling techniques to accurately characterize noise in quantum signals, resulting in a drastic reduction of the qubit readout operation time to **1 μ s**.
- Utilized LSTM-based recurrent neural networks to classify states across 8 qubits, achieving a **98%** fidelity.
- Enhanced the Digital Local Oscillator (DLO) by optimizing it through a data-driven Hierarchical Navigable Small World (HNSW) weighting technique. This led to a reduction in readout time and further improvement in fidelity.
- Enhanced memory allocation and data representation when deploying the LSTM model on FPGA at LBNL's quantum facilities, resulting in an inference overhead of just **30ns**.
- This achievement enables scientists to receive real-time feedback from the ML model regarding quantum states.

WSSLab, Machine Learning Research Intern, (MA, USA)

May 2023 - Aug 2023

In a Neuroscience-focused ML team, I played a key role in exploring compression algorithms tailored for seizure signals. Additionally, I focused on deploying these models onto edge devices for practical implementation. (*Mentor: VP Nguyen*)

- Developed Light Variational Auto-Encoder (LVAE) architecture to compress physiological signals, generating diverse latent spaces while minimizing information loss using KL divergence and reconstruction loss.
- Attained a compression ratio of **1:293**, outperforming state-of-the-art compression algorithms like DCT, JPEG2000.
- Trained an XGBoost model on the generated latent space produced by the encoder, achieving a high accuracy of **91%** in seizure detection, thereby validating the effectiveness of the latent space.
- Developed signal collection and compression pipeline on ARM cortex V8 and Nvidia Jetson Nano, enabling **real-time, on-chip seizure monitoring** while reducing space required to store seizure signals by 40 times.

The University of Texas at Arlington, Research Assistant, (TX, USA)

Aug 2022 - May 2023

Engaged in projects within the Sensor System Lab and Data Science Lab, with a focus on object tracking and natural language processing, respectively. (*Mentors: VP Nguyen, Jacob Lubner*)

Data Science Lab

- Finetuned LLM (**Llama 13B**) on medical data for discrete information retrieval and implemented Retrieval-Augmented Generation (**RAG**) to enhance efficiency in information retrieval tasks.
- Utilized **Llama Index** for Personally Identifiable Information (PII) masking atop RAG before LLM inference, ensuring privacy protection of sensitive data during information retrieval processes.
- Developed a Kubernetes-based pipeline to manage workflow, leveraging efficient data management to reduce latency by **15%**, facilitating seamless integration of the fine-tuned LLM model

Sensor System Lab, (NSF funded project)

- Constructed a novel **multi-modal** system based on convolutional recurrent neural networks (**CRNN**), integrating both acoustic and vision data for continuous UAV tracking.
- Leveraged the complementary strengths of both modalities to enhance tracking precision and reliability by **26%** in low light and blockage conditions, surpassing the state-of-the-art vision-based models
- Developed an audio and video sensing pipeline in Python to collect and process hours of UAV monitoring videos and stored it in an Amazon S3 bucket using **Boto3** for subsequent utilization in training and fine-tuning models.

The Tann Mann Gaadi, *Machine Learning Intern, (India)*

Sept 2021 - Jun 2022

Worked within the Applied ML team focused on Recommendation Systems, specifically developing a deep learning-based ranking algorithm to enhance ads to content relevancy for OTT platforms. (*Mentors: Rahul Nathan, Chiranjiv Roy*)

- Developed a Key-Frame detection algorithm to enhance information gain by filtering redundant frames in video data.
- Fine-tuned a transformer-based image captioning model on a 20-hour of processed custom dataset, utilizing forward feature selection for improved performance.
- Devised interpolation techniques for creating a similarity matrix between ad and content captions, and created a platform for client audience engagement analytics using **Amazon QuickSight**.
- Developed a React-based web app integrating the model with NodeJS, deployed on AWS, and stored data on an Amazon S3 bucket, resulting in a significant **28%** boost in advertisement revenue.

PUBLICATIONS

FPGA-based Machine Learning for In-situ Qubit State Discrimination on QubiC

American Physical Society, (APS'2024) [aps/mar24/N50.10](https://arxiv.org/abs/2403.12587)

Real-Time Diagnostic Integrity Meets Efficiency: A Novel Platform-Agnostic Architecture for Physiological Signal Compression (*Pre-Print ArXiv 2023 [arXiv:2312.12587v2](https://arxiv.org/abs/2312.12587v2)*)

DroneChase: A Mobile and Automated Cross-Modality System for Continuous Drone Tracking

ACM, Micro Aerial Vehicle Networks, Systems, and Applications (DroNet '23) doi.org/10.1145/3597060.3597237.

An Unobtrusive and Lightweight Ear-worn System for Continuous Epileptic Seizure Detection

(*Pre-Print ArXiv 2024 [arXiv:2401.05245](https://arxiv.org/abs/2401.05245)*)

PROJECT

Stockopedia [[Code](#)]

- Developed a JavaScript and Flask-based web platform for real-time stock analysis and data visualization using Chart.js.
- Integrated advanced machine learning techniques, including a Long Short-Term Memory (LSTM) model, for accurate stock price prediction.
- Achieved an 89% accuracy in stock price prediction by training the LSTM model on two decades' worth of historical data from the Yahoo Finance API.

SKILLS

Programming Languages: Python, C++, Java, MATLAB, JavaScript, R

Databases: MySQL, MongoDB, VectorDB, GraphQL.

Frameworks & Libraries: Pytorch, TensorFlow, Keras, Pandas, Librosa, OpenCV, Scikit-Learn, React, Flask, NodeJS, Django, Spring Boot, RestAPI, Chart.JS, Hadoop.

Tools & Services: Git, AWS, GCP, Azure ML, CUDA, PowerBi, Postman, Docker.

EXTRA-CURRICULAR ACTIVITIES

- Visiting research scholar at The University of Massachusetts Amherst under supervision of [Dr.VP Nguyen](#).
- Reviewer at The 30th Annual International Conference On Mobile Computing And Networking ([MobiCom '24](#))
- Mentored group of 7 juniors as ML Team Head at Google Developer Student Clubs (GDSC).
- Ranked among top 3% participants in competitive programming competition 'Gujarat State Hackathon'.