

MIAN WU

✉ nothern42@sjtu.edu  [nothern](#)  [Mian Wu](#)  [Nothern](#)

INTERESTS

My research evolved around constructing sophisticated models to delve into the intricacies of deep learning mechanisms across various computational and data environments. My interest broadly lies in image generation, generative artificial intelligence and natural language processing.

EDUCATION

University of California, Berkeley

Sep. 2023 – May 2024

Visiting Student

Shanghai Jiao Tong University

Sep. 2021 – Flexible

B.S.E Electrical and Computer Engineering

RESEARCH EXPERIENCE

Dissecting the Impact of Dataset Scale on Model Training

Jan. 2024 - Present

Advisor: Zhuang Liu

University of California, Berkeley

- **Objectives:** Explain the reasoning behind effects of training sample sizes and dataset compositions on the manifestation of biases within machine learning models, specifically focusing on classification tasks.
- **Contributions:** Conducted parallel training of ConvNeXt Tiny models, focusing on sample extraction and classification across different datasets. Analyzed the relationship between training sample size and the model's performance regarding dataset bias in downstream tasks. Assessed the experimental outcomes by adjusting the dataset and model architecture, thereby gaining a deeper understanding of how dataset scale impacts model performance and the emergence of bias.

Deep Learning in Phononic Crystal Design

Oct. 2023 - Present

Advisor: Lei Shao

Shanghai Jiao Tong University

- **Objectives:** Leverage deep learning techniques to design phononic crystals that exhibit larger relative bandgaps. The goal is to systematically optimizing phononic crystal configurations, merging deep learning with materials science to enhance physical feasibility.
- **Contributions:** Collected images of phononic crystals with varying bandgap widths to create a dataset. Explored various neural network architectures, including Convolutional Neural Networks (CNNs) and Vision Transformers (ViT), to assess their effectiveness in generating desired phononic crystal configurations. Ultimately, selected and implemented a Conditional Wasserstein Generative Adversarial Network (cWGAN) and a network for calculating generated images' bandgap widths. Implemented a strategy similar to reinforcement learning, iteratively generating phononic crystal designs with progressively larger bandgaps, and studied the underlying physical principles and learned generation patterns.

Investigation of the Transformer Optimization

Feb. 2023 - Jun. 2023

Advisor: An Zou

Shanghai Jiao Tong University

- **Objectives:** Improve GPT-2 performance using early-exit strategies.
- **Contributions:** Implemented early-exit layers within the Transformer model and conducted ablation studies. Applied various early-exit strategies in the GPT-2 architecture, quantified the overall performance, and assessed the advantages of using early-exit strategies.

[4]

SKILLS

Toolchain

Colab, Linux, Docker, RLLib

Programming Languages

C/C++, Python, Matlab, Mathematica, L^AT_EX, Java

Frameworks & Libraries

Transformers, Pytorch, Scikit-Learn, OpenAI

HANDS-ON PROJECTS

Bert CC – *Enhancing Text Classification with Fine-tuned Bert*

Jan. 2024 - Mar. 2024

Fine-tune BERT for text classification tasks. Extracted dataset from 12345 case reports and performed cleaning and encoding. Modified model functions to accommodate tokenization, and added a linear classification layer for training. Conducted training on Colab, integrated into a Python script for practical application, demonstrating strong performance in classifying report content.

SVRG – *Optimizing Deep Learning Training with SVRG idea*

Nov. 2023 - Dec. 2023

Implement the Stochastic Variance Reduced Gradient (SVRG) algorithm in PyTorch. Inspired by the approaches discussed in seminal papers (*arXiv:1812.04529, NeurIPS 2013*), explore SVRG as a general method for gradient modification. Develop a multi-layer MLP model trained on the CIFAR-10 dataset. Conduct ablation studies on SVRG with standard optimizers (SGD and Adam). pursue optimization by adjusting snapshot frequencies and refining hyperparameters.

RPredictor – *2022 X-Game Shanghai ICV&EV Bigdata Competition*

Sep. 2022 - Nov. 2023

Develop a clustering approach to segment electric vehicle operational data into multiple subsets for customized model training. Implement Gaussian Process and Support Vector Machine models on a cloud platform to accurately predict risk labels for electric vehicles, securing a 6th-place.

HONORS & AWARDS

Sep. 2023	Berkeley Global Access Scholarship	<i>University of California, Berkeley</i>
Sep. 2022	Undergraduate Excellent Scholarship	<i>Shanghai Jiao Tong University</i>
Nov. 2022	Silver Medal	<i>The University Physics Contest (UPC)</i>