Guanwen Qiu

Email : guanwenq@seas.upenn.edu
Webpage: https://nayutaqiu.github.io/

Education

- **University of Pennsylvania** — Philadelphia, PA
  Master of Science in Engineering in Computer and Information Science, GPA: 3.7 — *May 2023 - Dec 2024 (Expected)*
- **University of Pennsylvania** — Philadelphia, PA
  Master of Science in Engineering in Data Science, GPA: 4.0 — *Jan 2022 - May 2023*
- **University of California, Irvine** — Irvine, CA
  Bachelor of Computer Science (Cum Laude), Major GPA: 3.9 — *Sep 2016 - Dec 2020*

Publication

- **Complexity Matters: Dynamics of Feature Learning in the Presence of Spurious Correlations, ICML 2024, GuanWen Qiu**, Da Kuang, Surbhi Goel
- **Deep Tree Reconstruction: A supervised representational approach to phylogenetic tree reconstruction, Preprint, GuanWen Qiu**, Da Kuang, Junhyong Kim

Research Experience

- **Understanding representation learning through Boolean function** — UPenn, PA
  Independent Research with Prof. Surbhi Goel — *May 2023 - Current*
  - Propose a novel theoretical framework with accompanied dataset that encapsulate feature as Boolean functions with different complexity and varying correlation to study the learning dynamics of neural network in presence of spurious correlation
  - Conducting carefully defined experiments on neural network under the framework to elucidate critical aspects of deep learning, such as simplicity bias and lazy learning.
  - Study the influence of spurious feature on the representation learned by neural network under SGD both empirically and theoritically.
  - Point out critical weakness of popular spurious dataset and debasing algorithms.
- **Phylogenetic Tree Reconstruction Based on Metric Learning** — UPenn, PA
  Independent Research with Prof. Junhyong Kim — *May 2024 - Current*
  - Developed a novel embedding framework incorporating topological constraints of tree metrics and transformer to enhance phylogenetic tree reconstruction.
  - Conducted extensive experiments on diverse datasets to validate the framework's robustness and adaptability.
  - Explored and applied advanced embedding techniques, leveraging theoretical and practical insights to refine algorithmic performance.
  - Performed rigorous theoretical analyses to demonstrate the feasibility and soundness of the proposed algorithm.
- **Music Generation through Deep Learning** — UPenn, PA
  Independent Research under Prof Pratik Chaudhari — *Jan 2022 - June 2022*
  - Compared the performance of different deep learning architectures in generating music MIDIs including Transformer, LSTM, CNN and Variational Auto Encoder (VAE) using Torch and LaKh MIDI Dataset.
  - Designed a highly flexible music midi generation architecture combining LSTM and VAE CNN to produce 5 tracks (Drums, Strings, Piano, Guitar and Bass) MIDI with long term musical development. The model also can be used by musicians to do style transfer and melody generation.
  - Used the pre-trained novel generative model as a downstream model to classify MIDIs into 10 different genres with higher than 80% accuracy.

Teaching experience

- TA CIS522: Deep Learning in Data Science
- Head TA CIS555: Internet and Web Systems

Skills Summary

- **Languages**: Java, Python, C++, SQL, Unix scripting
- **Tools**: Pandas, PyTorch, Numpy, CVXPy, sk-learn
- **Graduate ML Course Taken**: Big Data Algorithm, Modern Convex Optimization, Modern Topics in Uncertainty Estimation, Principles of Deep Learning, Computational Learning Theory, Computational Complexity, Probability Theory, Information Theory