

XINGYU (ALFRED) LIU

1420 Centre Ave. APT 1603, Pittsburgh, PA, 15219 | 412-482-2106 | xingyu.alfred.liu@gmail.com | xingyuliu.me

EDUCATION

Carnegie Mellon University

Pittsburgh, PA

- Doctor of Philosophy in Materials Science & Engineering - GPA: 3.87/4.0 Dec. 2021 (expected)
 - Master of Science in Materials Science & Engineering - GPA: 3.81/4.0 Dec. 2017
 - Secondary Master of Science in Machine Learning - GPA: 3.67/4.0 Dec. 2021 (expected)
- Related Courses: 10707-Topics in Deep Learning, 10805-ML with Large Datasets, 10725-Convex Optimization, 10716-Advanced ML Theory & Method, 36705-Intermediate Stats, 10701-Intro to ML (PhD), 10703-Deep Reinforcement Learning & Control

Wuhan University of Technology

Wuhan, Hubei, P.R.China

- Bachelor of Science in Materials Science & Engineering - GPA: 91.11/100 Jun. 2016

TECHNICAL SKILLS

- Programming Language: Python (proficient), C++, Bash, SQL
- Python Packages: PyTorch, Tensorflow, Keras, Scikit-Learn, GPyTorch, Pandas, PyMC3
- Large Dataset Management & Platforms: Spark, MongoDB, AWS, Git, Balsam, Slurm

RESEARCH EXPERIENCE

Research Assistant, Advisor: Dr. Noa Marom

Carnegie Mellon University

- **Looking for Singlet Fission Materials with MatML Workflow (Python, MongoDB)** Sep. 2018 - Present
 - Designed and implemented Materials Science and Machine Learning (MatML) workflow for automated customizable database screening, model training and inference with **MongoDB** and **Balsam**.
 - Project recognized as 1 of 10 Early Science Project (ESP) being conducted at Argonne National Laboratory.
 - Resulted in 2 first-author publications and 2 second-author publications that received 24 total citations (i.e., *Journal of Physics: Condensed Matter*, *The Journal of Physical Chemistry C*; impact factor: 4.2 - top 10%).
- **Band Gap Prediction for Organic Molecular Crystals (Python, GPyTorch)** Spring 2019 - Present
 - Trained **Gaussian process regression** model for molecular crystals band gap prediction with structural features MBTR, achieved RMSE 0.45 eV, lower than the top-performed NN model with the same objective.

WORK EXPERIENCE

Machine Learning Engineer

Invictus BCI Inc.

- **Hand Motion Prediction with Electromyography Data (Python, PyTorch, AWS)** Dec. 2020 - Jan. 2021
 - Explored and applied feature extraction to electromyography (time series electrical activity) data for hand motion prediction with fine-tuned LSTM, improved the accuracy by 5%.

ACADEMIC PROJECTS

- **Disaster Tweets Classification with BERT and LSTM (Python, Tensorflow)** Dec. 2020 - Jan. 2021
 - Applying BERT and LSTM to disaster tweets classification task with 10,000 labeled observations.
 - Fine-tuned pre-trained BERT model with multiple model sizes and hyperparameters. Achieved 80% accuracy. Implemented and fine-tuned LSTM model and achieved 76% accuracy with the same dataset.
- **Million Song Dataset Hotness Prediction with Machine Learning (PySpark, AWS)** Nov. 2020 - Dec. 2020
 - Designed song hotness prediction pipeline for 581,965 instances parsed from Million Song Dataset. Perform data preprocessing using functions such as **OneHotEncoder** implemented in **MLlib**.
 - Linear regression with different regularization techniques, **Random Forest**, and **Gradient Boosted Trees** were implemented and compared. Total training cost 2 hours on 4 AWS m5.xlarge nodes.
 - Gradient Boosted Trees yielded lowest RMSE as 0.1321, improved almost 20% compared to the baseline model. Concluded that the song hotness is highly correlated with artist hotness.
- **Atomistic NN for Accelerated Crystal Structure Prediction (Python, PyTorch)** February 2019 - May 2019
 - Implemented an **Atomistic Neural Network** with **Gaussian symmetry functions** for thermodynamic energy prediction. MSE of 1.50 kJ/mol was achieved.
 - Realized transfer learning to predict the energy for 8 molecules per cell materials using 4 molecules per cell structures as the training set.