Sixian Jia

540-682-0838 sixian@umich.edu -linkedin-Sixian

EDUCATION

Mechanical Engineering (Ph.D.) GPA: 4.0 — **University of Michigan - Ann Arbor**
Research Focus: ***Machine Learning***, ***Statistics***, *Data Science* — *Aug. 2020 - May. 2026*

Computer Science (MS) GPA: 4.0 — **University of Illinois at Urbana-Champaign**
Coursework: Machine Learning, Natural Language Processing — *Aug. 2022 - May 2023*

Mechanical Engineering (BS) GPA: 3.81 — **Virginia Tech**
Coursework: Linear Algebra, Algorithms and Data Structures — *Aug. 2016 - May 2020*

QUANT ACTIVITIES

Events: Jane Street Women+Summer Social, IMC Invited Visit, Jane Street Game Night at Michigan, Citadel Engineering Tech Talk, DRW Trading Simulation Game, AmplifyME simulation with Morgan Stanley
Quant Courses: Akuna Capital Options 101, JP Morgan Chase & Co. **Quantitative Research** Job Simulation
Competitions: SIG Coding Challenge, UMich Quant Convention, **Akuna Capital's Quant Trading Challenge**

WORK EXPERIENCE

Machine Learning Science Intern (NLP), Wayfair — *June. 2024 – Aug. 2024*

- Implemented **2200+** lines to the components of **GenAI** framework that tested with **image** and **text-only** tagging
- Developed a **28-class** subject tag prediction model for product images (**Recall: 0.78**) and created a pipeline for **320-class** material tag predictions from text descriptions (**Recall: 0.72**).
- Significantly enhanced the performance of a flash model by implementing **batch prediction** and **parallelized learning** which led to faster model deployment and more efficient resource utilization, achieving a **4x** reduction in processing time
- **Delivered a machine learning pipeline** that was successfully deployed **online**, automating the product tagging process, reducing **78% manual labor** and saved **$1.2M annually**

Quant Researcher, Quant at Illinois — *Feb. 2023 – Jan. 2024*

- Trained machine learning models for **time series forecasting** based on the last 30 days of google stock data
- Built an **LSTM-based Seq2Seq Variational Autoencoder** with a **91.98% accuracy** in forecasting stock prices.
- Achieved a **90.7% accuracy** in stock price prediction by combining **CNNs with sequential data processing**.

SELECTED PUBLICATIONS

- **Sixian Jia** et. al. Hybrid physics-guided data-driven modeling for generalizable geometric accuracy prediction and improvement in two-photon lithography. (2024) *Journal of Manufacturing Processes* **Impact Factor: 10.2**
- Zhiqiao Dong, **Sixian Jia** et. al. Filtered Kriging: Improve Kriging Interpolation for Periodic Manufacturing Surfaces with a Pre-filter. (2024) *Journal of Manufacturing Processes* **Impact Factor: 10.2**

SELECTED PROJECTS

Career Preparation for Quant Research || Python, Machine Learning, ***Time Series*** — *Aug. 2023 - Now*

- Implemented a **multi-layer RNN** using TensorFlow, leveraging BasicRNN cells to model temporal sequences of stock prices, allowing for robust prediction of financial data trends
- Experimented with various RNN cell types, including **Basic RNN**, **LSTM**, and **GRU** cells, to optimize the model's performance based on different financial time-series datasets
- Employed the **Adam optimizer** to minimize the Mean Squared Error loss function, leading to accurate predictions of stock prices on training, validation, and test datasets
- Conducted regular evaluation of model performance across epochs using **MSE metrics** on both training (**0.000059**) and validation (**0.000284**) datasets, fine-tuning the model to achieve optimal prediction accuracy (**0.000215**)

Data-driven Machine Learning in Manufacturing Process || Python, PyTorch, Git — *Feb. 2021 - Aug. 2022*

- Lead **data cleaning** initiatives e.g., addressing inconsistencies, handling missing values, and **feature extraction** from manufacturing datasets to drive insightful **decision-making**, ensuring data integrity for accurate analyses
- Built physics-informed **regression** models, including non-linear regression, linear regression, **Gaussian Process** to predict and increase the accuracy of the 3D printing process, the geometric deviation is reduced by **85.5 %**
- Achieved an R^2 score greater than **0.97** for the regression task and the **MAPEs** are about **5%** for non-linear regression, can overcome extremely data-scarce situations, reducing the number of required experiments by **94.4%**

SKILLS

Programming: **Python**, SQL, C/C++, R, MATLAB
Packages: Scikit-learn, Statsmodels, Pandas, NumPy, Scipy, Matplotlib, Seaborn, **Pytorch**, **TensorFlow**
Specialties: **Machine Learning**, Optimization, **Statistical Learning**, Natural Language Processing, **Probability**