

# ZIYUE WANG

+1 917 436 9176 | ziyue.wang0125@gmail.com | ziyuewang.com | github.com/ziyuewang25

## INDUSTRY EXPERIENCE

### GOOGLE

Mountain View, USA

*Software Engineer, Labs*

*May 22 – Now*

- Curating high quality code training data for large language models
- Built end-to-end models to rank important tests given a code change and reduced the rollback rate by 20+%
- Built and integrated ranking service into Google-wide testing framework
- Built a new Google-wide standard Python code indexing tool to provide code cross-references based on AST

### BNP PARIBAS CIB

New York City, USA

*Associate, Algorithmic Quant Researcher, Electronic Equity Group*

*Mar 21 – Apr 22*

- Processed trillions of order book data and built level 2 limit order book data using C++ (STL, Boost)
- Built short-term alpha signal with CNN, attention and LSTM models using Tensorflow, Pandas and Numpy and improved current signal performance by 20% with 50% less computational time
- Processed billions of records of order book and trade data in multi-thread using kdb+/q with functional-oriented design and fed them into deep learning model in streaming version

### AIGEN INVESTMENT MANAGEMENT

New York City, USA

*Quantitative Research Intern, Quantitative Research Division*

*Jun 20 – Aug 20*

- Parsed five types of portfolio financing emails and cleaned ten years of security lending data from five different brokers using Python (smtplib, email, re) with detailed documentation for alpha research

## OTHER EXPERIENCE

### AI Safety Research and Study

*Arp 23 – Now*

- Self-studied and finished [Deep Learning Curriculum](#) and [AI Safety Fundamentals Alignment Course](#)
- [Eval Hackathon result paper: Can Large Language Models Solve Security Challenge?](#)
- [Red Teaming Language Models with Language Models](#)
- [Finetune LLM with RLHF to generate positive tone message from Shakespeare Corpus](#)
- [Implementing PPO from scratch in Progen environment](#)
- [Replicating Scaling Laws by using MNIST data](#)
- [Mechanistic Interpretability: Find the induction heads in GPT-2](#)

### Kaggle Data Science Competition Master, Top 0.2% Worldwide

*Jan 21 – May 22*

- [G2Net Gravitational Wave Detection](#), 3/1219 (**0.2%**). Applied signal and image processing technique and built novel 1D CNN models to detect whether GW exists in the time series data; [Code](#)
- [Tensorflow Underwater Starfish Detection](#), 11/2025 (**0.5%**). Applied YOLOv5 model to detect underwater small object and enhanced model performance with pseudo labeling, ensemble learning and tracking [Post](#)
- [Jane Street Market Prediction](#), 84/4245 (**1.9%**). Deciphered 130 anonymous features and applied DenseNet, XGBoost and Autoencoder to predict stock short-term future return
- [CommonLit Readability Prize](#) 200/3663 (**5.5%**). Conducted extensive LLM model structure optimization, training, ensembling and evaluation for literature rating

## EDUCATION

### BARUCH COLLEGE, CITY UNIVERSITY OF NEW YORK

New York, USA

*MSc in Financial Engineering, GPA 3.79*

*Aug 19 – Dec 20*

Winner, 9<sup>th</sup> annual International Association for Quantitative Finance Competition, [Paper](#)

### RENMIN UNIVERSITY OF CHINA

Beijing, CHN

*BSc in Applied Mathematics; BEcon in Financial Engineering, GPA 3.78 (3/52)*

*Sep 15 – Jun 19*

## PROGRAMMING & ACTIVITIES & INTERESTS

Language: Python, Go, C++, Kdb/q+, SQL, Unix Shell, Git, L<sup>A</sup>T<sub>E</sub>X

DL Frameworks: Pytorch, JAX/Flax, TensorFlow, Keras

Leadership: President, Business and Finance Career Development Club, RUC

Interests: Rock Climbing, Hiking and Reading