

Yijing (Carol) Zhang

5252 S Cornell Ave, Chicago IL | +1(858) 203-8697 | carolzhang18@outlook.com | [LinkedIn](#)

EDUCATION

The University of Chicago

Chicago, IL

Master of Science in Financial Mathematics (GPA: 3.8/4.0)

Dec 2024 (Exp)

- Courses: Portfolio Theory & Risk Management, Data Science in Finance, Option Pricing, Probability & Stochastic Processes, Numerical Methods, Fixed Income Derivatives

University of California, San Diego

San Diego, CA

B.S. in Probability and Statistics; B.A. in Economics; Minor in Data Science (GPA: 3.7/4.0)

Dec 2022

- Courses: Calculus, Linear Algebra, Mathematical Real Analysis, Statistical Modeling, Stochastic Process, Hypothesis Testing, Econometrics, Programming & Data Structures, Data Science Algorithms, Machine Learning

SKILLS

Computing: C++, C, Python, Jupyter, R, MATLAB, Java, STATA, SPSS, LaTeX, MS Office

Knowledge: Financial Markets, Machine Learning, Statistical Modeling, Data Analytics, Econometrics

Trading Products: Futures, Fixed Income, Equities, Options

EXPERIENCE

Bodhi Research Group

Chicago, IL

Quantitative Research Intern

June 2024 – Present

- Developed and tested the Emerging Manager Portfolio on GitHub, implementing the pop to drop algorithm, gain to pay algorithm, omega function, and utilizing Conditional Value at Risk (CVaR) for robust risk management.
- Analyzed portfolio turnover rates, average holding periods, and trading patterns, evaluating stability and profitability of repetitive security transactions, contributing to enhanced portfolio management strategies
- Developed and tested a portfolio strategy using monthly purchase and holding assumptions, conducted PCA on returns, and performed performance attribution analysis, providing insights into factor loadings and currency-adjusted returns.

Irrational Capital

West Conshohocken, PA

Quantitative Research Intern

Jan 2024 – Mar 2024

- Initiated advanced NLP modeling for Glassdoor text analysis, employing Langdetect for language filtering, and streamlined text cleaning with lowercasing, tokenization, and stopword removal. Enhanced processing efficiency through parallel computing and refined data with lemmatization.
- Implemented LDA model using tf-idf for word vector adjustment, emphasizing rare terms, and count vectorizer for frequency vectors, providing a robust framework for text analysis. The approach minimized common word impact while ensuring comprehensive term representation.
- Rigorously evaluated model performance with precise metrics, ensuring the generation of accurate relevance and sentiment scores for predefined topics, showcasing the system's effectiveness in deriving meaningful insights from text data.

Google

Sunnyvale, CA

Part-time Research Assistant

Mar 2022 – June 2022

- Conducted an in-depth analysis of the bank's credit card customer datasets to proactively predict churn rates and reveal the underlying factors; Executed thorough exploratory data analysis and engaged in meticulous feature engineering on a comprehensive dataset comprising 10,000 customers, systematically curating relevant features to facilitate subsequent model training.
- Employed machine learning classification models like Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting, to accurately forecast customer churn trends, ensuring robust model selection and evaluation.
- Employed rigorous evaluation metrics, such as the confusion matrix, classification accuracy, AUC, and F1 score, to meticulously assess and compare prediction results; Delivered comprehensive reports summarizing the findings, methodology, and results, contributing valuable insights to inform strategic decision-making within the organization.

RESEARCH

Adversarial Robustness and Explaining Neurons in Deep Representations

May 2022 – May 2023

- Employed a procedure to automatically explain logical in Python and perceptual abstractions encoded by individual neurons in deep networks trained on image classification
- Utilized the procedure to generate features on images in Jupyter, which could increase the accuracy of the testing results; tested the image recognition procedure on one of the university's club website's log-in verification sites