

SIRIUS (YUNHAN) ZHANG

(858)-260-9968

yz1048@georgetown.edu

EDUCATION

GEORGETOWN UNIVERSITY

Master of Science in Data Science and analytics

Georgetown, DC

GPA 4.0/4.0

05/24

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Bachelor of Science in Data Sciences

La Jolla, CA

Upper Major GPA 3.68/4.00

03/22

- Additional major in Joint Major of Mathematics & Economics, covering core courses of both mathematics and economics
- Provost's Honors

EXPERIENCE

MCDONOUGH SCHOOL OF BUSINESS, GEORGETOWN UNIVERSITY

Research Assistant, Data Analyst

Georgetown, DC

10/22-Present

- **Clustering & Labeling:** Assessed the DualCl model's accuracy on training data; Translated the original dataset and labeled the dataset with modified DualCl model; Re-structured dataset for later A/B testing on market reformation analysis.
- **Data Processing:** Developed an automated script to filter out unwanted data from over 100G files; Designed a new file management logistic; Cleaned filtered xml files and formulated for later regression analysis on companies' cash flow quality.

GLOBAL POLICY & STRATEGY, UC SAN DIEGO

Research Assistant, Data Analyst

La Jolla, CA

08/20-06/22

- **Automation:** Developed an automated script utilizing logging and exception handling to efficiently scrape electricity trading data from multiple sources. Streamlined the data transformation pipeline, converting raw data into a clean and well-organized dataset within a structured folder system. Successfully accelerated the scraping process by 80%.
- **Programming:** Developed a new algorithm to match and merge different datasets by comparing Chinese characters and English words; Enabled Dr. Davidson's lab to combine spreadsheets from different languages and facilitated research progress.
- **Data Analysis:** Incorporated 12-year data to analyze the correlation of electricity trade price and coal prices via Linear Regression model; Found recent abnormality with this model and proposed a new research topic.

ZHONGTIAN FINANCE GROUP INC.

Data Analyst Intern

Shenzhen, CN

07/19-08/19

- **Research:** Investigated China's A-share re-entry procedure of red chip companies with VIE structure and IPOs of Energy firms across the country to analyze major success reasons; examined eligibility and mapped client re-entry to the A-market.
- **Analytical:** Facilitated an energy firm going public by utilizing previous research; Utilized 5+ years of retrieved financial data to validate the data and supported the team for such process.

RESEARCH/PROJECTS

BERT MODEL ANALYSIS ON FINANCIAL NEWS CLASSIFICATION AND STOCK PREDICTION

09/21-03/22

- **Leadership:** Designed the structure, pipeline, data, and model usage of the project; Directed weekly basis meeting and composed report collaboratively to solve encountered problems quickly and keep the project within the schedule.
- **Natural Language Processing:** Fetched and cleaned stock price data and reports from 5+ institutions; Compared NLP models, including AutoPhrase, LSTM, and BERT using NLTK; Fine tuned hyper-parameter to predict the next movement of a certain stock.

ADAPTIVE LAYER-WISE LEARNING RATE DECAY FOR BERT FINE-TUNING?

09/21-12/21

- **Deep Learning:** Implemented a tuned BERT model in tensorflow by comparing different learning rates and layer settings on the IMDB dataset and achieved a 94% accuracy on if the user liked a certain movie based on historical comments.

SCALABLE LOAN DEFAULT DETECTION

03/21-06/21

- **Big Data:** Generated a pipeline from the labeling process to a classification model that is scalable to a 20G loan dataset by using scalable tools, including pySpark and Dask;
- **Cloud Computing:** Deployed such model to AWS EC2 and EMR cloud computing service, and achieved an overall 85% accuracy on classifying if the loan delinquency status is default by parallel computing and feature engineering

COURSEWORK/SKILLS

- **Mathematics & Statistics:** Probabilities, Statistics, Differential Equations, Linear Algebra, Dynamic Programming, Time Series
- **Economics & Finances:** Econometrics, Operation Research, Financial Markets. Intermediate Economics, Financial Accounting
- **Data Sciences:** Data Management, Scalable Analytics, Recommender Systems, Web Mining, Machine Learning, Deep Learning, Natural Language Processing, Data Visualization, Data Structures, Algorithms, Artificial Intelligence, Data Analysis & Presentation
- **Programming Languages:** Python (SciPy, matplotlib, sklearn, webDriver, tensorflow, pyTorch), R, Java, MATLAB, SQL, STATA