Zhengguang Wang

Email: zw4re@virginia.edu Telephone: (434)422-2399

Summary

- **Languages:** Bilingual in English and Chinese, Intermediate French
- **Programming Languages:** Python, Java, C, x86 assembly language; HTML/CSS/Javascript
- **Relevant Packages:** Pytorch, Transformers, Pandas, scikit-learn
- **Research Interest:** Natural Language Processing, Interpretable Time Series and Sensitivity Analysis
- **Personal Website:** https://zhengguangw.github.io/

Education

University of Virginia | Major in *Statistics* and *Computer Science* Graduation Date: Expected May 2024
Echols Scholar at the University of Virginia (top 5% at UVA), GPA 3.89/4.0, GRE 329/340

Research Experience

Information and Language Processing Lab| *CS Department, UVA* June 2023-

- Conducted evaluation into the *Consistency of Large Language Model's Political Leaning*, in which my research pipeline involves using Selenium and Beautifulsoup to web-scarpe headlines from news media such as Allisdes and Breitbart, querying OpenAI API with both GPT 3.5, and fine-tuning a Distillbert to analyze the model outputs; aiming for a paper submission to the 2023 NAACL

UVA-MLSys | *CS Department, UVA* June 2023-

- Participated in the *Global Pervasive Computational Epidemiology (GPCE) Interpreting County-level Covid Infections using Deep Learning for Time Series* project; won the IEEE ICDH 3rd prize
- Led a team to build an interactive project website with an embedded U.S county map, in which user could hover over counties and click counties to see ground truth and our model performance in a temporal manner
- Submitted a poster "*Time Series Sensitivity Analysis of Population Age Groups in Multi-Horizon COVID-19 Infection Forecasting"* to AAAI'24 in group; ran experiments on Sensitivity Analysis method such as Morris method and feature ablation and train time-series deep learning model on remote HPC using GPU
- Submitted a research statement "*Consistencies and Performances of Perturbation-based Sensitivity Analysis Methods*" to AAAI'24 Undergraduate Consortium

Teaching Experience

Data Science Teaching Assistant | *Statistics Department, UVA* | Charlottesville, Virginia Aug 2022-Dec 2022
Teaching Machine Learning Algo, KNN classification, A/B Testing, Regression, and simulation in Python

- Held Office Hour twice per week to help students understand statistical concepts (sampling, unbiased estimator, type I error, etc); performed simulations to demonstrate these concepts
- Taught data wrangling and control-flow in Python; implemented KNN classifier and A/B Testing for example usage; wrote sample scripts for Bootstrapping and Monte-Carlo Simulation

Professional Experience

Analyst Intern | *Local Energy Alliance Program* | Charlottesville, Virginia Jun 2022-Aug 2022
Web-Scraping with Python Selenium and Mapping with ArcGIS; Cleaning 7-year Solarize Program Data

- Used Python Pandas package to merge and tidy the 7-year Solarize Program Data; wrote Selenium scripts to retrieve county names and geographical coordinates from customer addresses using google map API
- Utilized ArcGIS and geocoding to generate four interactive maps of LEAP sites with features of amounts in Virginia; conducted geo-spatial analysis and present the finding to the Executive Director

Analyst Intern | *China Minmetals Corporation* | Beijing, China Jun 2021-Aug 2021
Identifying Merger & Acquisition (M&A) pattern by conducting Time-Series and Hypothesis testing

- Examined China's oversea mineral M&A pattern based on Big Data in the post-Financial Crisis era
- Accessed the S&P Data Base to retrieve M&A metadata such as case time, monetary amounts, minerals, locations; employed Python Pandas and Excel V-Lookup to clean data
- Conducted Time-Series Analysis on each year's M&A monetary amount using methods like differencing and detrending to remove time trend from data and employed Auto Regressive Moving Average model
- Applied statistical analysis techniques such as two-sample proportion t-test and Chi-Square test to compare the preferences of China's M&A with the rest of world
- Illustrated the global comparisons on a world map using Python Pyechart then compiled the findings and published a 16-page report on the Monthly publication of Economic Research Institute