

# Zhengguang Wang

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

## Summary

- **Languages:** Bilingual in English and Chinese, Intermediate French
- **Programming Languages:** Python, Java, C, R, x86 assembly language; HTML/CSS/Javascript
- **Relevant Packages:** PyTorch, Transformers, Pandas, Scikit-learn, Django, Flask
- **Research Interest:** Natural Language Processing, Explainable Deep Learning and Sensitivity Analysis
- **Personal Website:** <https://zhengguangw.github.io/>

## Education

University of Virginia | Major in *Computer Science and Statistics* Graduation Date: Expected May 2024  
Echols Scholar at the University of Virginia (top 5% at UVA), GPA 3.895/4.0, GRE 329/340

## Research Experience

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

### **Research: Consistency of Large Language Models' Political Leaning and Reliable Text Summarization**

- Applied Selenium and BeautifulSoup to scrape news headlines from Allsides, Breitbart, and the DailyBeast; queried OpenAI API with headlines and designed prompts with both GPT 3.5-turbo and GPT 4
- Used Pytorch and Transformers package to fine-tune a DistillBERT to classify the GPT outputs on UVA Rivanna HPC GPU, then applied Spearman rank test to calculate consistency of GPT outputs across prompts
- Applied LoRA method to fine-tune a Llama-7b for text summarization; plan to integrate Statistical NLP ideas such as POS tagging to reduce hallucination
- Involved in the ILP group effort to build a LLM Parameter Efficient Fine Tuning (PEFT) codebase

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

June 2023-

### **Research: Consistency and Performances of Perturbation-based Sensitivity Analysis Methods**

- Prepared the presentation of *Global Pervasive Computational Epidemiology (GPCE) Interpreting County-level Covid Infections using Deep Learning for Time Series* for IEEE ICDH'23; won the third prize
- Led a team to build an interactive project website with an embedded U.S. county map with Javascript D3 in which user could click on counties to compare ground truth and model predictions
- Submitted a Workshop paper "*Interpreting Time Series Transformer Models in Multi-Horizon COVID-19 Infection Forecasting and Age Sensitivity Analysis*" as a coauthor; conducted sensitivity analysis using methods like feature ablation and integrated gradients and trained time-series deep learning models for benchmarking
- Submitted a research statement "*Evaluation of Interpretability Methods for Time-Series Deep Learning with Sensitivity Analysis*" 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 with statistical concepts (A/B Testing, sampling, unbiased estimator, type I error, etc.); assisted in lectures twice per week by answering in-class questions
- Taught data wrangling, control-flow, functions, and objects in Python

## 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 customers' addresses
- Utilized ArcGIS and geocoding to generate four interactive maps of LEAP sites in Virginia; conducted geo-spatial analysis and presented the finding to the Executive Director

**Economic Analyst Intern** | *China Minmetals Corporation* | Beijing, China

Jun 2021-Aug 2021

### **Identifying Merger & Acquisition (M&A) pattern by conducting Time-Series and Hypothesis testing**

- Accessed the S&P Database to retrieve M&A metadata such as case time, monetary amounts, minerals, locations; employed Python Pandas and Excel V-Lookup to clean data; made visualizations and numerical summaries to examine the spatial-temporal pattern; analyzed different countries' preferences in M&A
- 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

# Zhengguang Wang

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

## **Publications**

---

- Evaluation of Interpretability Methods for Time-Series Deep Learning with Sensitivity Analysis  
**Zhengguang Wang**  
*AAAI'24 Undergraduate Consortium, Feb.2024*
- Interpreting Time Series Transformer Models and Sensitivity Analysis of Population Age Groups to COVID-19 Infections  
Md Khairul Islam, Tyler Valentine, Timothy Joowon Sue, Ayush Karmacharya, Luke Neil Benham,  
**Zhengguang Wang**, Kingsley Kim, Judy Fox  
*AAAI'24 Workshop AI4TS:AI For Time Series Analysis, Feb.2024*



COURSE NUMBER	COURSE TITLE	GRADE	CREDITS	COURSE NUMBER	COURSE TITLE	GRADE	CREDITS
---------------	--------------	-------	---------	---------------	--------------	-------	---------

**Issued / Mailed To:**

ZHENG GUANG WANG

CS	1110	Introduction to Programming	A	3.0
EGMT	1530	Engaging Difference	A	2.0
Course Topic:		Faith in Democracy?		
EGMT	1540	Ethical Engagement	A	2.0
Course Topic:		Faith in Democracy?		
ENWR	1510	Writing and Critical Inquiry	B+	3.0
Course Topic:		Writing about the Arts		
LPPL	3480	Leadership: US Policy to China	A	3.0
MATH	1320	Calculus II	A	4.0
ZFOR	3512	International Study	N	0.0
Course Topic:		Exchange		
Curr Credits	17.0	Grd Pts	65.900	GPA 3.876
Cuml Credits	25.0	Grd Pts	97.900	GPA 3.916

**2021 Fall**

National Id: \*\*\*\*\*XXXX  
Birthdate: 10/16/XX

School:	College & Graduate Arts & Sci			
Major:	Arts & Sciences Undeclared			
ECON	2010	Principles of Econ: Microecon	A-	3.0
PLCP	3500	Topics in Comp Politics	A	3.0
Course Topic:		Japanese Politics		
STAT	3080	From Data to Knowledge	A	3.0
STAT	3110	Foundations of Statistics	A+	3.0
STAT	3220	Intro to Regression Analysis	A	3.0
Curr Credits	15.0	Grd Pts	59.100	GPA 3.940
Cuml Credits	40.0	Grd Pts	157.000	GPA 3.925
Honor:	Dean's List			

**Test Credits**

Test Credits Applied Toward Arts & Sciences Undergraduate

Transferred to Term 2020 Fall as  
COMM 1800 Foundations of Commerce 4.00

Other Test Credit 3.00  
**Test Credit Total: 7.00**

**Transfer Credits**

Transfer Credit from Cornell University  
Applied Toward Arts & Sciences Undergraduate  
**Transfer Credit Total: 4.00**

Transfer Credit from Fudan University  
Applied Toward Arts & Sciences Undergraduate  
**Transfer Credit Total: 11.00**

**Beginning of Undergraduate Record**

**2020 Fall**

Academic disruption due to global pandemic. Default A-F grading option for all undergraduate classes. Students able to select CR=C or higher/GC=C- to D-/NC=no credit grading option prior to knowing final grade. Dean's List suspended for term.

School:	College & Graduate Arts & Sci			
Major:	Arts & Sciences Undeclared			
EGMT	1510	Engaging Aesthetics	A	2.0
Course Topic:		Sounds of Resistance		
EGMT	1520	Empirical Engagement	A	2.0
Course Topic:		How Do You Measure a Rainbow?		
STAT	2120	Intro to Statistical Analysis	A	4.0
ZFOR	3512	International Study	N	0.0
Course Topic:		Exchange		
Curr Credits	8.0	Grd Pts	32.000	GPA 4.000
Cuml Credits	8.0	Grd Pts	32.000	GPA 4.000

**2021 Spring**

Academic disruption due to global pandemic. Default A-F grading option for all undergraduate classes. Students able to select CR=C or higher/GC=C- to D-/NC=no credit grading option prior to knowing final grade. Dean's List suspended for term.

School: College & Graduate Arts & Sci  
Major: Arts & Sciences Undeclared

**2022 Spring**

School:	College & Graduate Arts & Sci			
Major:	Applied Statistics			
Concentration:	Data Science			
CS	2100	Data Structures and Algo 1	A+	4.0
CS	2120	Discrete Math and Theory 1	A-	3.0
ECON	2020	Principles of Econ: Macroecon	A	3.0
ECON	3010	Intermediate Microeconomics	B+	4.0
FREN	1020	Elementary French II	A-	4.0
Curr Credits	18.0	Grd Pts	67.100	GPA 3.728
Cuml Credits	58.0	Grd Pts	224.100	GPA 3.864
Honor:	Dean's List			

**2022 Fall**

University-wide disruption. Student had option to select CR/GC/NC grading rubric. CR meets all degree requirements; GC meets limited requirements; NC is no credit (F). CR/GC/NC not included in GPA calculation.

School:	College & Graduate Arts & Sci			
Major:	Applied Statistics			
Concentration:	Data Science			
CS	3100	Data Structures and Algo 2	A-	3.0
FREN	2010	Intermediate French I	A	3.0
MATH	3350	Applied Linear Algebra	A	3.0
STAT	3120	Intro Mathematical Statistics	A	3.0
STAT	3250	Data Analysis with Python	A	3.0
UNST	3510	Internships, Self & Organizations	CR	1.0
Course Topic:		Self & Organizations		
Curr Credits	16.0	Grd Pts	59.100	GPA 3.940
Cuml Credits	74.0	Grd Pts	283.200	GPA 3.879
Honor:	Dean's List			

**2023 Spring**

School: College & Graduate Arts & Sci



THE INFORMATION CONTAINED IN THIS TRANSCRIPT IS CONFIDENTIAL AND SHOULD NOT BE RELEASED WITHOUT THE STUDENT'S WRITTEN CONSENT



COURSE NUMBER	COURSE TITLE	GRADE	CREDITS	COURSE NUMBER	COURSE TITLE	GRADE	CREDITS
---------------	--------------	-------	---------	---------------	--------------	-------	---------

Major: Applied Statistics  
 Concentration: Data Science  
 Major: Computer Science

CS	3140	Software Dev Essentials	A	3.0
CS	4774	Machine Learning	A+	3.0
SARC	5400	Data Visualization	A-	3.0
STAT	4120	Applied Linear Models	A	3.0
STAT	5170	Applied Time Series	A	3.0
Curr Credits	15.0	Grd Pts	59.100	GPA 3.940
Cuml Credits	89.0	Grd Pts	342.300	GPA 3.890

Honor: Dean's List

**2023 Fall**

School: College & Graduate Arts & Sci  
 Major: Applied Statistics  
 Concentration: Data Science  
 Major: Computer Science

CS	2130	Computer Systems and Org 1	A-	4.0
CS	3240	Advanced Software Development	A	3.0
CS	4993	Independent Study	A+	3.0
STAT	4630	Statistical Machine Learning	A+	3.0
STAT	4996	Capstone	A	3.0
Curr Credits	16.0	Grd Pts	62.800	GPA 3.925
Cuml Credits	105.0	Grd Pts	405.100	GPA 3.895

**2024 Spring**

School: College & Graduate Arts & Sci  
 Major: Applied Statistics  
 Concentration: Data Science  
 Major: Computer Science

CS	3120	Discrete Math and Theory 2		3.0
CS	3130	Computer Systems and Org 2		4.0
CS	4501	Spec Top: Computer Science		3.0
Course Topic:		Cryptocurrency		
STAT	3280	Data Visual and Management		3.0

End of Undergraduate Record



THE INFORMATION CONTAINED IN THIS TRANSCRIPT IS CONFIDENTIAL AND SHOULD NOT BE RELEASED WITHOUT THE STUDENT'S WRITTEN CONSENT