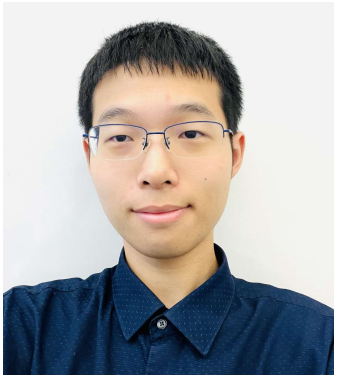
Qianyu (Kenneth) Zheng

Adress: Weizemkampstraße 133, 28199 Bremen
Telephone: +1 4049604914
Email: zhengqianyu0227@gmail.com
Linkedin: www.linkedin.com/in/zheng-qianyu-365b60250
Website: qzheng75.github.io



Motivation

I am a high-achieving master's student at Georgia Institute of Technology with a perfect GPA and extensive hands-on experience applying machine learning and data analytics to solve complex real-world problems across research institutions in the US and Europe. My expertise spans deep learning, cloud computing, data analysis and big data processing. I'm actively looking for opportunities to leverage my expertise in bridging the gap between computing and scientific discoveries.

Work Experiences

07/2025 – now

Data Analyst at Fraunhofer Institute for Wind Energy Systems in Bremen

- Implemented anomaly detection pipeline with Pandas/Scikit-learn to clean terabytes of turbine operational data.
- Built automated time-series processing pipelines for aggregating multiple data sources with complex filtering conditions.
- Scaled big data analytics using PySpark & SQL to process 6+ months of remote sensing data (radar, lidar, etc.).
- Performed geographical data visualization with Matplotlib, rasterio, and cartopy.

07/2024 – 08/2024

Dozent für Python/AI-„Scotch"-Bootcamp an der University of Maryland

- Durchführung einer 4-stündigen Deep-Learning-Vorlesung für über 200 AI-Anfänger zu PyTorch und den Grundlagen neuronaler Netze sowie Mitwirkung in einem Team von 20 Dozenten zur Verbesserung der Effektivität der AI-Ausbildung.

05/2024 – 08/2024

Computational Scientist at Leibniz Institute of Plant Biochemistry in Halle (Saale)

- Designed scalable data analysis algorithms for protein families containing 5+ million sequences using protein Large Language Models, clustering and network analysis algorithms.
- Participated in protein binding possibility estimation research. Implemented big data handling pipelines with duckdb and PySpark and deep learning solutions featuring transformer architecture with PyTorch.

05/2023 – now

Computational Scientist at Georgia Institute of Technology in Atlanta, USA

- Conducted research to improve reliability of Machine Learning Interatomic Potentials for molecular dynamics (MD) simulation with High Performance Computing, GNNs, PyTorch, deep learning, and research methodologies.

01/2023 – 05/2024

Teaching Assistant am Georgia Institute of Technology in Atlanta, USA

- TAed for Introduction to Object Oriented Programming course within a team of 42 TAs.
- Mentored 400+ students in OOP as Q&A forum lead among 42 TAs to provide programming assistance, achieving 4.9/5 effectiveness rating in end-of-semester student survey for four semesters.

Ausbildung

2022 – 2025 | GPA 4.0/4.0

Bachelor of Science in Informatik, Georgia Institute of Technology, Atlanta, USA

2025 – 2027 | GPA 4.0/4.0

Master of Science in Informatik, Georgia Institute of Technology, Atlanta, USA

Projekte

08/2024 – 01/2025

Natural Language Query für große Proteindatenbanken

- Entwicklung eines multimodalen Tools zur Textabfragen in der UniProt-Datenbank.
- Einsatz von LLM Llama 3.1 mit Langchain zur Generierung von Textabfragen als Trainingsdaten sowie PyTorch zum Trainieren eines CLIP-Modells (BERT + ESM) von Proteinsequenz- und Nutzerabfrage-Embeddings. Benchmark-Vergleich mit NER-basierter Methode mittels SpaCy zeigt eine doppelt so hohe Genauigkeit der entwickelten Methode.
- Entwicklung einer Flask-Anwendung (Flask, HTML/CSS), die mit Docker, AWS Fargate, Lambda, ECR, ECS und S3 bereitgestellt wurde. Die Anwendung ist jetzt online unter: www.nl2prot.org
- Erwerb von Erfahrung in Deep Learning, LLM-Finetuning, Cloud Computing und PyTorch.

08/2023 – 12/2023

Workout Of the Day (WOD) prediction with Data Science @ Georgia Tech

- Designed and implemented data cleaning and feature engineering pipelines for the downstream machine learning tasks.
- Leveraged modern optimization libraries to design an automated hyperparameter search pipeline for modeling, improving ML models' performance by 6% in WOD prediction.

Language

- English – equivalent to native level
- German – planning to take the B2 exam in December 2025

Certificates

Amazon Web Services: Certified Cloud Practitioner (CLF-C02)

- Foundational, high-level understanding of AWS Cloud, services, and terminology.

Amazon Web Services: Machine Learning Specialty (MLS-C01)

- Expertise in building and deploying machine learning solutions in the AWS Cloud.

Excel: Microsoft Excel Expert (MO-201)

- Expertise in Excel 2019, such as creating, managing, and distributing professional spreadsheets for a variety of specialized purposes and situations.