

BINGSEN (BALE) CHEN

☎ (86) 189-3005-0896 ✉ bale.chen@nyu.edu [LinkedIn](#) [GitHub](#) [Home Page](#)

Education

New York University Shanghai, Shanghai, China 09/2020 – 05/2024

New York University, New York, USA 01/2023 – 05/2023

*B.S. in Data Science (concentration in Artificial Intelligence) Minor: Computer Science** Cumulative GPA: 3.96/4.00

Honors: Global Elite Scholarship (Top 30 incoming freshmen); Recognition Award 2020-2023; Dean's List 2020-2023.

Selected Courses: Natural Language Understanding and Computational Semantics (Graduate Level), Modern Topics in Statistical Learning Theory (Graduate Level), Mathematical Statistics, Reinforcement Learning, Data Structures

* I will take Discrete Math, Computer Architecture, and Software Engineering in Spring 2024 to finish the minor.

Research Experience

Dialogue State Tracking with Large Language Models 09/2023 – Present

Ongoing Data Science Capstone Research | Supervisor: Prof. Yik-Cheung (Wilson) Tam [GitHub](#)

- Currently experimenting with Llama-2 and Falcon to tackle dialogue state tracking on the MultiWOZ benchmark with QLoRA and mixed-precision training.

Finetuning Large Language Models with RLHF and Guided Exploration 06/2023 – 10/2023

Research Assistant at the DAIL Lab | Supervisor: Prof. Yik-Cheung (Wilson) Tam [GitHub](#)

- Conducted individual research on aligning Large Language Models (LLMs) with Reinforcement Learning from Human Feedback using guided exploration for faster and stabler convergence.
- Trained Llama-7B model on summarization and instruction following task using LoRA, quantization, and DDP.
- Summarized and presented 20+ latest papers in RLHF and LLM alignment to Prof. Tam during weekly meetings.

Trash or Treasure: How to Utilize Emojis in Social Media Sentiment Classification 05/2022 – 09/2022

Dean's Undergraduate Research Fund Summer Research | Advisor: Prof. Mathieu Laurière [Blog](#) | [Report](#) | [GitHub](#)

- Evaluated the emoji-compatibility of pretrained BERT-based encoders and how emojis can be incorporated into sentiment analysis to improve accuracy.
- Proposed 5 data preprocessing methods for emojis that robustly boost sentiment analysis accuracy by 1.2%.
- Published a blog post on Towards Data Science medium account and won the Best Project in STEM and Media in the Fall 2022 Undergraduate Research Symposium.

Shanghai Key Lab of Urban Design and Urban Science, Shanghai, CN 05/2022 – 12/2022

Research Associate (Part-time) | PI: Prof. Chenghe Guan

- Led a research project on spatial equity analysis on public EV charging stations in Shanghai supervised by Dr. Jialin Liu
- Provided Deep Learning assistance to urban researchers, including CNN-based water quality prediction with remote sensing data, and applying RoBERTa to Twitter sentiment analysis and data filtration.

Selected Projects

Beyond DetectGPT: Towards AI-generated Text Detection in a Black-box Setting Spring 2023

Course Project in Natural Language Understanding (Graduate-level) | Instructor: Dr. Sophie Hao [Paper](#) | [GitHub](#)

- Coordinated with a group of 4 Master's and undergraduate students to conduct an independent research project.
- Proposed a fine-tuned OPT method that shows the potential of generalizing DetectGPT framework to a black-box setting and adopted accuracy under false positive control to tackle the false positive rate issue with AUROC.

Plant Seedling Classification Competition: A Deep Learning Approach Spring 2023

Course Project in Modern Topics in Statistical Learning Theory (Graduate-level) | Instructor: Prof. Qi Lei [Report](#) | [GitHub](#)

- Won 6th place among 22 groups of graduate students in this Kaggle-style competition with an F1-score of 0.983.
- Experimented with transfer learning, data augmentation, ensemble methods, model scaling, and RayTune.

Qualitative and Quantitative Analysis of Soft Actor-Critic with REDQ and Reset Mechanism Fall 2022

Course Project in CSCI 375 Reinforcement Learning | Instructor: Prof. Keith Ross [GitHub](#)

- Analyzed the configurations of Randomized Ensembled Double-Q (REDQ) Learning, Reset Mechanism, and hyperparameter tuning in the Soft-Actor Critics algorithm in deep reinforcement learning.
- Presented the paper "The Primacy Bias in Deep Reinforcement Learning" by Nikishin et al. and its Reset Mechanism.

Optimizing Shanghai's Lockdown Policy: Economy, Mental Health, and Severe Cases Considered Spring 2022
Course Project in DATS 240 Intro to Optimization and Math Programming | Instructor: Prof. Zhibin Chen [GitHub](#)

- Proposed a mathematical model that builds upon the SEIR model to simulate the COVID outbreak in Shanghai 2022, with economics, public mental health, and severe cases considered.
- Optimized Shanghai 2022 lockdown policy using a Genetic Algorithm and conducted sensitivity analysis to understand optimization results.

Professional Experiences

IAV Automotive Engineering, Shanghai, CN 08/2023 – Present
AI Innovation Intern

- Prototyped a real-time style transfer and object detection application using OpenCV, OpenVINO, and YOLOv5, winning the company-wide AI innovation campaign, currently developing a Proof-of-Concept on a Kubernetes cluster
- Researched and presented in-vehicle virtual assistants in China and the recent developments of LLMs bi-weekly.
- Initiated a LangChain-based Context-aware music recommender idea independently, currently in the prototyping phase.

Pingan Bank, Shanghai, CN 06/2021 – 08/2021
Data Analyst Intern

- Retrieved and processed data daily with HiveQL to support 6 business decisions on a plugin display system.
- Developed 16 auto-updating visualized reports on a Tableau-like platform for mid-year reports and daily monitoring.
- Maintained 84 tables in a database by monitoring daily data updates, altering fields based on business requirements, and managing storage usage.

Teaching Experiences

NYU Shanghai, Shanghai, CN 08/2023 – Present
Learning Assistant for CSCI 360 Machine Learning | Instructor: Prof. Shengjie Wang

- Hosted weekly office hours and biweekly review sessions to tutor students on course materials, homework, and project.
- Conveyed students' feedback and commonly asked questions from office hours to faculty on a weekly basis.

NYU Shanghai, Shanghai, CN 08/2023 – 10/2023
Course Assistant for CSCI 375 Reinforcement Learning | Instructor: Prof. Mathieu Laurière [GitHub](#)

- Developed 7 Python notebooks of tabular and deep RL algorithms for homework templates and in-class demonstrations.
- Communicated with Prof. Laurière weekly to offer suggestions on course design from a previous student's perspective.

Leadership / Extracurricular Experiences

AI Interest Group @ NYU Shanghai 05/2023 – Present
Co-founder

- Co-founded a school-wide student interest group and organized 11 AI-related workshops and reading seminars.
- Held a beginner workshop on prompt engineering and an advanced seminar on InstructGPT and RLHF individually.

NYU Shanghai Residential Life 09/2021 – 12/2022
Resident Assistant

- Managed 2 floors of 40 ethnically diverse residents by designing bi-weekly events, enforcing residential policies, and troubleshooting health emergencies as well as interpersonal conflicts.
- Cooperated with a team of 6 supervisors and 30 resident assistants to conduct special projects including new recruits training, school-wide quiz bowl, and recruitment.

NYU Shanghai Center for Career Development 09/2020 – 05/2022
Student Worker Leader

- Managed and coordinated a team of 17 to facilitate events and develop career-related resources.
- Hosted the opening of the annual conference with 4 distinguished guests and moderated an alumni dialogue session.
- Executed a video project with a team of 3, producing a short video series for service introduction and promotion.

Skills & Interests

Artificial Intelligence: High proficiency in PyTorch and Hugging Face framework

Technical Skills: Python (NumPy, Scikit-learn, Pandas, Matplotlib), SQL, Git, Linux, Slurm, L^AT_EX, Kubernetes

Language: Mandarin (Native), English (Full Professional Proficiency)

Interests: A Cappella Choir, Trekking, Soccer, Indoor Rock Climbing