Sophia Sun

1 (617) 777-0906 | sophiasun@eng.ucsd.edu | huiwenn.github.io

EDUCATION

University of California San Diego, San Diego, CA — Ph.D.   2018 - present
Artificial Intelligence Group, Computer Science and Engineering
Advisor: Prof. Rose Yu

Wellesley College, Wellesley, MA — B.A.   2016 – 2018
Double Major in Mathematics and Computer Science.

Massachusetts Institute of Technology, Cambridge, MA — *Cross Registration*   2016 – 2018
Coursework: 6.006 Introduction to Algorithms, 6.034 Artificial Intelligence, 21M.380 Music and Technology,
6.S094 Deep Learning for Self-Driving Cars (winter session), 6.867 Machine Learning (graduate level)

RESEARCH EXPERIENCE

Graduate Student Researcher, UC San Diego AI Lab   San Diego, CA   2018.9 – present
PI: Prof. Rose Yu
* Machine learning for spatiotemporal data.
* Uncertainty quantification for machine learning models; decision making under uncertainty.

Research Assistant, Johns Hopkins University   Baltimore, MD   2017.6 – 2018.6
PI: Prof. Suchi Saria; NSF Research Experience for Undergraduate program
* Poster: *Predicting in-Hospital Cardiac Deterioration from Heterogeneous Medical Time-Serie*s.
* Data pipelining and machine learning for of multi-modal and multi-resolution medical time-series data.

Research Assistant, music21 (MIT) http://web.mit.edu/music21/   Cambridge, MA   2016.10 – 2018.2
PI: Prof. Michael Cuthbert
* Investigated the expression of sentiments in songs based on lyrics, melodic, harmonic, and other musical data
  extracted from sheet music through computational methods. (Paper in Empirical Musicology Review)

INDUSTRY EXPERIENCE

Applied Scientist Intern, AWS AI   2023.6 – 2023.9
(time-series forecasting, uncertainty quantification, anomaly detection)
* Designed unsupervised anomaly detection algorithm for AWS database usage through conformal prediction and
quantile tracking methods; provided theoretical bounds for decision-making regret.

Research Scientist Intern, Amazon Robotics RAD   2020.6 – 2020.9
(reinforcement learning, multi-agent planning, optimization)
* Machine learning for multi-robot trajectory planning in amazon's fulfillment centers.
* Help built simulation for multi-robot trajectory planning and improved overall throughput by 3% in simulation.

Amazon Alexa Prize Challenge, team UC San Diego   2019.9 – 2020.5
(reinforcement learning, natural language processing)
* Principal member of one of the 10 finalist teams. Incorporation of RL for more natural LLM conversations.

Software Engineer Intern, Google DeepMind   London, U.K.   2018.6 – 2018.9
(Golang, backend engineering)
* Secure Electronic Health Records data pipelining, analysis, and verification.
* Improved interpretability and verifiability of analytical algorithms deployed in Deepmind's Streams App.

Data Science Intern, Musical.ly (now TikTok)   Shanghai, China   2016.6 – 2016.8
(Java, Hadoop on Spark, SQL, JavaScript/JQuery)
* Social media data analytics and monitoring pipelines using Hadoop on Spark and AWS's Elastic MapReduce.

PUBLICATIONS

Sophia Sun, Abishek Sankararaman, Rose Yu. "Online Unsupervised Anomaly Thresholding for Data Under Distribution Shift" In Submission, 2023

Sophia Sun, Rose Yu. "Copula Conformal Prediction for Multi-horizon Time Series." In Submission, 2023

Sander Tonkens*, **Sophia Sun***, Rose Yu, Sylvia Herbert. "Scalable Safe Long-Horizon Planning in Dynamic Environments Leveraging Conformal Prediction and Temporal Correlations*." IEEE International Conference on Robotics and Automation workshop on Long-term Human Motion Prediction (ICRA),* 2023

Sophia Sun, Robin Walters, Jinxi Li, Rose Yu. "Probabilistic Symmetry for Multi-Agent Dynamics." *Learning for Dynamics and Control Conference (L4DC),* 2023

Sofy Yuditskaya*, **Sophia Sun***, and Margaret Schedel*. "Synthetic Erudition Assist Lattice." *New Interfaces for Musical Expression (NIME), 2021*

Sofy Yudistkaya*, Derek Kwan*, **Sophia Sun***. "Karaoke of Dreams: A multi-modal neural-network generated music experience." *Proceedings of the Joint Conference on AI Music Creativity (CIMC+MUME), 2020*

Bodhisattwa Prasad Majumder, Shuyang Li, Jianmo Ni, Huanru Henry Mao, **Sophia Sun**, Julian McAuley. "Bernard: A Stateful Neural Open-domain Socialbot." *Alexa Prize Proceedings*, 2020

Sophia Sun, Michael Scott Cuthbert. "Emotion painting: lyric, affect, and musical relationships in a large lead-sheet corpus." *Empirical Musicology Review 12.3-4*. 2018

REVIEWER

ICML 2022 - 2023, NeurIPS 2022 – 2023, ICLR 2023

TALKS & WORKSHOPS

Conformal Methods for Quantifying Uncertainty for Time-series Data Talk at AWS AI	Sunnyvale, CA	2023
Conformal Methods for Quantifying Uncertainty in Spatiotemporal Data Talk at UCSD CSE	San Diego, CA	2022
Probabilistic Symmetry for Improved Trajectory Forecasting Talk at UCSD CSE Research Open House AI spotlight	San Diego, CA	2022
Creative Coding for Robotics and Art Outreach keynote at Sweetwater High School	San Diego, CA	2022
Reinforcement Learning 101 Three-part workshop series for artists at Arts Counsel Korea (ARKO)	online	2020
Natural Language Processing for Creatives Workshop at mars.college tech/art residency	Bombay Beach, CA	2020

TEACHING

Lecturer & head TA , CSE151b Deep Learning, UC San Diego	Spring 2023
Head TA, CSE 291 Deep Generative Models, UC San Diego	Fall 2022
Teaching Assistant, CSE101 Design & Analysis of Algorithms, UC San Diego	Summer 2022
Teaching Assistant, CSE20 Discrete Mathematics, UC San Diego	Summer 2022
Teaching Assistant, CSE251C Machine Learning Theory, UC San Diego	Spring 2021
Teaching Assistant, CSE291 Topics in Search and Optimization, UC San Diego	Winter 2020
Tutor, CS230 Data Structures / CS231 Algorithms, Wellesley College	2016 - 2018

LEADERSHIP

UCSD ACM-W Student Chapter, UC San Diego	Chair, 2020 – 2022
Graduate Women in Computing (GradWIC), UC San Diego	Vice President, 2019 – 2022
The Center for Ethics in Science & Technology, UC San Diego	Board Member, 2018 – 2020

SKILLS

Machine Learning: Algorithms for **time series & dynamics modeling**, probabilistic forecasting, uncertainty quantification, deep generative models; ML code frameworks **TensorFlow** and **Pytorch**.
Data Science: SQL, PostgreSQL; distributed system **Hadoop on Spark**
Software Engineering: Competent in **Java** and **Python**, experience in **C, C++, R, Go, JavaScript/JQuery**
Robotics: Experience with **ROS** and physics simulation engine **Mujoco**.