Sofian Zalouk

saz43@cornell.edu | szalouk.github.io

Research Interests

Uncertainty quantification, fairness, and generative modeling in machine learning, with an emphasis on creating trustworthy AI systems for critical domains, such as healthcare.

Education

Cornell University 2024 – Present
Ph.D. in Computer Science *Ithaca, NY*
Advisors: Kilian Q. Weinberger, Chris De Sa

Stanford University 2024
M.S. in Computer Science, Concentration: Artificial Intelligence *Stanford, CA*
Advisors: Stefano Ermon, Andrew Ng
GPA: 4.14/4.00

University of Toronto 2020
B.A.Sc. in Electrical Engineering *Toronto, Canada*
Ranked 1st out of over 100 students in department.
GPA: 3.96/4.00

Academic Experience

Graduate Research Assistant Sep 2020 – Present
Stanford Artificial Intelligence Laboratory *Stanford, CA*

- Uncertainty quantification and diffusion models with Stefano Ermon
- Machine Learning for healthcare with Andrew Ng and Sharon Zhou

Research Assistant Summer 2022
Ford-Stanford University Collaboration *Stanford, CA*

- Active learning and image segmentation with Andrew Ng and Ram Rajagopal

Undergraduate Researcher Sep 2019 – May 2020
University of Toronto *Toronto, Canada*

- Monocular depth estimation with Andreas Moshovos

Industry Experience

Software Engineer (Intern) Sep 2018 – Sep 2019
Intel *Toronto, Canada*

- Worked on LLVM compiler to analyze and accelerate deep learning tasks on FPGAs
- Improved Intel's OpenCL memory management, reducing runtime and memory overhead by an order of complexity

ASIC Design Engineer (Intern) Summer 2018
NVIDIA *Santa Clara, CA*

- Design and verification of processors for deep learning.

Teaching Experience

Stanford CS 228 (Probabilistic Graphical Models) TA Winter 2024
Stanford CS 229 (Machine Learning) TA Spring 2021
Stanford CS 236G (Generative Adversarial Networks) TA Winter 2021

Publications

Calibration by Distribution Matching: Trainable Kernel Calibration Metrics 2023
Charles Marx*, **Sofian Zalouk***, Stefano Ermon
NeurIPS – Conference on Neural Information Processing Systems, 2023
*Equal contribution

A System for Automated Vehicle Damage Localization and Severity Estimation Using Deep Learning 2023
Yuntao Ma, Hiva Ghanbari, Tianyuan Huang, Jeremy Irvin, Oliver Brady, **Sofian Zalouk**, Hao Sheng
Andrew Ng, Ram Rajagopal, Mayur Narsude
IEEE Transactions on Intelligent Transportation Systems (Impact Factor: 9.551)

Projects

Measuring and Reducing Bias in LLMs introduced by RLHF | GitHub, Poster, Report Spring 2023

- Awarded "Best Project" in Stanford's CS 224R for identifying and mitigating bias in LLMs due to RLHF
- Conducted comprehensive analysis to assess language polarity, stereotype bias, and pronoun-based bias
- Identified and mitigated increased bias in larger models using a post-hoc self-debiasing method

Data Augmentation for Speech Recognition | GitHub, Report Winter 2021

- Implemented MaskCycleGAN-VC, the state-of-the-art method for many-to-many voice conversion
- Established the project as the leading implementation of MaskCycleGAN-VC on GitHub, evidenced by being the **most starred repository in its category**
- Developed data augmentation pipeline for Automated Speech Recognition
- Generated African American Vernacular English utterances from generic American English to address data scarcity

Awards and Scholarships

Best Project Award, Stanford CS 224R (Deep Reinforcement Learning) Spring 2023
Instructor: Chelsea Finn

Outstanding Student Award, University of Toronto 2015 - 2020
Awarded to the top 3 students in Electrical Engineering

W.S. Wilson Medal, University of Toronto 2020
Awarded to student with highest academic standing in the graduating year

Baptie Scholarship, University of Toronto 2016
Awarded to students with high academic standing (top 1%)

Technical Skills

Languages: Python, C/C++, JavaScript, HTML/CSS, R, Bash, MATLAB, CUDA
Frameworks: LaTeX, Git, RStudio, Jupyter Lab/Notebook, Gdb, Valgrind
Libraries: PyTorch, TensorFlow, Keras, Scikit-learn, pandas, NumPy, Matplotlib