

Fabrice Harel-Canada

SOFTWARE ENGINEER · COMPUTER SCIENCE RESEARCHER

☎ (+1) 954-805-5735 | ✉ fabricehc@cs.ucla.edu | 🏠 fabrice.harel-canada.com | 📱 fabriceyh

Biography

I am a computer science PhD student at UCLA where I am part of the Software Evolution and Analysis Lab and advised by Miryung Kim. I also work with Violet Peng in the PLUS Lab and Quanquan Gu in the Statistical Machine Learning Lab. My research aims at developing new techniques for training, testing, and evaluating ML systems. Previously, this has entailed devising novel adversarial attacks, investigating test metrics like neuron coverage for deep learning models, and creating *sibylvariant transformations* to improve model generalization and robustness. My current projects span three distinct areas: a) developing fine-grained data lineage to support the creation of automated data augmentation policies and data exploration; b) producing reference-free evaluation metrics to quantify the diversity of natural language; c) analyzing android apps containing ML models to generate tests that expand post-inference coverage.

Education

University of California, Los Angeles (UCLA)

PHD | COMPUTER SCIENCE

- 3.92 / 4.00 GPA
- Majors: Software Engineering + Artificial Intelligence
- Minor: Data Science Computing

Advisor: Miryung Kim

Jan 2020 - Jun 2023

University of California, Los Angeles (UCLA)

MS | COMPUTER SCIENCE

- Major: Artificial Intelligence
- Thesis: *Is Neuron Coverage a Meaningful Measure for Testing Deep Neural Networks?*

Advisor: Miryung Kim

Sept 2018 - Dec 2019

University of Florida (UF)

BS | BUSINESS ADMINISTRATION

- Majors: Management + Information Systems and Operations Management
- Minors: Entrepreneurship + Finance
- Thesis: *The Viability of the Student-Run Housing Cooperative: A Case Study of the Cooperative Living Organization*

Aug 2008 - May 2012

Publications

ENDEX: EVALUATION OF DIALOGUE ENGAGINGNESS AT SCALE

Guangxuan Xu, RuiBo Liu, **Fabrice Harel-Canada**, Nischal Reddy Chandra and Nanyun Peng, in *Findings of EMNLP 2022, Abu Dhabi, United Arab Emirates. Empirical Methods in Natural Language Processing*.

Findings of EMNLP 2022

Acceptance Rate: 29 %

SIBYLVARIANT TRANSFORMATIONS FOR ROBUST TEXT CLASSIFICATION

Fabrice Harel-Canada, Muhammad Ali Gulzar, Nanyun Peng and Miryung Kim, in *Findings of the Association for Computational Linguistics: ACL 2022, pages 1771–1788, Dublin, Ireland. Association for Computational Linguistics*.

Findings of ACL 2022

Acceptance Rate: 37 %

IS NEURON COVERAGE A MEANINGFUL MEASURE FOR TESTING DEEP NEURAL NETWORKS?

Fabrice Harel-Canada, Lingxiao Wang, Muhammad Ali Gulzar, Quanquan Gu and Miryung Kim, in *Proc of ACM SIGSOFT International Symposium on the Foundations of Software Engineering (ESEC/FSE)*, Sacramento, California, USA, 2020.

ESEC/FSE 2020

Acceptance Rate: 28 %

Research Experience

University of California, Los Angeles

GRADUATE STUDENT RESEARCHER

- Developing automated evaluation metrics for natural language generation systems - e.g. diversity, engagement
- Designed a new class of sibylvariant data transformations, releasing an open source tool for the community [Sibyl]
- Evaluated the reliability of testing metrics for deep learning
- Developed diversity-promoting extensions to adversarial attack algorithms

Supervisor: Miryung Kim

June 2019 - present

Professional Experience

Axenzi Consultants

CEO

Oct 2018 - present

- Helping clients to plan and implement technology projects - specs, devops, application development, and database development

University of California, Los Angeles

Supervisor: *Miryung Kim*

TEACHING ASSISTANT, CS 130 : SOFTWARE ENGINEERING

Sept 2020 - Dec 2020

- Developed teaching materials, activities, and exam questions that reinforced the lectures and gauged conceptual proficiency.

University of California, Los Angeles

Supervisor: *Judson Caskey*

RESEARCH ASSISTANT

July 2020 - Oct 2020

- Conducted LDA topic modeling and phrase-level sentiment analysis using BERT transformers on millions of Amazon and Glassdoor product / employer reviews

University of California, Los Angeles

Supervisor: *Carlo Zaniolo*

READER, CS 143 : DATABASE SYSTEMS

Jan 2019 - Mar 2019

- Developed test cases for a Spark project in Scala, actively addressed student questions, and assisted students with project work

Verox Tech

Supervisor: *Michael Simhai*

PROJECT MANAGER + DEVELOPER TEAM LEAD

July 2015 - Sept 2018

- Designed and successfully launched a smarter insurance app by building, motivating, and guiding a team of 23
- Contributed 100s of check-ins, primarily consisting of C# Web APIs and database tables, triggers, sprocs, and T-SQL scripts
- Designed and implemented a 3-system legacy data migration
- Responsible for producing, coordinating and reviewing nearly every aspect of the SDLC:
 - code reviews and code optimization
 - requirements and spec generation
 - database design + SSIS + SSRS
 - technical documentation
 - UI design standards + product design
 - build and deployment scripts

People's Trust Insurance

Supervisor: *Scott Thompson*

BUSINESS + DATA ANALYST

Dec 2012 - Sept 2018

- Go-to data expert for executives to produce hundreds of reports and analyses for pivotal decision making at the executive level
- Developed numerous front-end / back-end application features that were successfully deployed into production to automate tedious and time-intensive workflows
- Performed continuous requirements gathering & demos with business stakeholders to support 60+ sprints of development for both applications and business intelligence teams, including the transition to a new policy and claim management system

Cooperative Living Organization

PRESIDENT

Dec 2010 - May 2012

- Appointed and oversaw the performance of 17 managers who were responsible for executing the policies and procedures I developed
- Increased house occupancy from 60 residents to full capacity (80) and increased net income 57% while expanding services & activities for residents

Skills

Programming	sql, python, java, c#, linq, html, css, js, scala, prolog, prolog, \LaTeX
Frameworks	pytorch, tensorflow, keras, spark, numpy, pandas, scipy, sklearn, opencv, matplotlib, seaborn
Tools	ssms, ssis, ssrs, tfs, azure devops, visual studio, excel, aws, gcp, exchange server, moqups

Selected Coursework

2022	Natural Language Generation , CS 269	<i>UCLA</i>
2022	Automated Testing 4 Data Systems , CS 239	<i>UCLA</i>
2021	Large-Scale Social and Complex Networks , EC 232E	<i>UCLA</i>
2020	Machine Learning Algorithms , CS 260	<i>UCLA</i>
2020	Probabilistic Programming , CS 267A	<i>UCLA</i>
2019	Neural Networks and Deep Learning , EC 239AS	<i>UCLA</i>
2019	Natural Language Processing , CS 269	<i>UCLA</i>
2019	ML Testing and Debugging , CS 239	<i>UCLA</i>
2018	Advanced Data + Knowledge Bases , CS 240B	<i>UCLA</i>

Honors & Awards

2019	Magna Cum Laude , 3.90 / 4.00 GPA	<i>UCLA</i>
2012	Magna Cum Laude , 3.85 / 4.00 GPA	<i>University of Florida</i>
2008 - 2012	Recipient , Bright Futures Academic Scholarship	<i>University of Florida</i>
2011 - 2012	Recipient , Anderson Scholar with Distinction	<i>University of Florida</i>

Extracurricular Activity

2020-2021	Mentor , Queers in STEM	<i>Los Angeles, CA</i>
2019	Volunteer , TreePeople	<i>Los Angeles, CA</i>
2018	Volunteer , Gay for Good (G4G)	<i>Los Angeles, CA</i>
2017	Volunteer , Los Angeles LGBT Center	<i>Los Angeles, CA</i>
2014	Volunteer , Stonewall Pride	<i>Wilton Manors, FL</i>
2009	Volunteer Counselor , CDS Interfaith Youth Shelter	<i>Gainesville, FL</i>
2009	Co-founder, Editor, Contributor , Florida Inquiry Journal	<i>Gainesville, FL</i>
2009 - 2010	Volunteer , Habitat for Humanity	<i>Fort Lauderdale, FL</i>
2008 - 2012	Member , UF Entrepreneurship Club	<i>Gainesville, FL</i>