

# Charles Dickens

Email: cadicken@ucsc.edu

GitHub: [github.com/dickensc](https://github.com/dickensc)

Phone: (831)-419-7656

---

## PROFILE

I am a Ph.D. computer science and engineering candidate at the University of California Santa Cruz. My general research interests are in machine learning, modeling richly structured graph data, and empowering AI with domain knowledge and symbolic reasoning capabilities. I have developed models for a variety of applications, including computer vision, language modeling, social networks, demand forecasting, and recommendation. I am seeking a research engineer role where I can contribute my skills to produce impactful research and products.

---

## EDUCATION

**Ph.D. in Computer Science and Engineering** | University of California Santa Cruz      Expected Summer 2024  
Advanced to Candidacy with Honors      GPA: 4.0  
Advisor: Dr. Lise Getoor  
Thesis: Convex and Bilevel Optimization for Neuro-Symbolic Inference and Learning

**B.S. in Electrical and Computer Engineering, Minor in Mathematics** | University of Hawai'i Manoa      2019  
Summa Cum Laude Honors      GPA: 3.9

---

## EXPERIENCE

**Graduate Research Assistant** | University of California, Santa Cruz      Fall 2019–Present

LINQS Statistical Relational Learning Group: [linqs.org](https://linqs.org)

During my Ph.D. I have led and worked on multiple industry and academic collaborative research efforts. Specifically, funded by a Google research grant focused on building practical tools and applications, I worked with Dr. William Yang Wang and his lab at the University of California Santa Barbara to create NeuPSL, a general Neuro-Symbolic (NeSy) framework extending on the Probabilistic Soft Logic (PSL) open-source machine learning library ([github.com/linqs/psl](https://github.com/linqs/psl)). Further, under the NSF-funded TRIPODS initiative focused on theoretical research, I collaborated with Dr. Stephen Wright and his lab at the University of Wisconsin Madison to develop general NeSy inference and learning algorithms. Additionally, I worked with Dr. Barbara Block and her lab at Stanford University to create a shark identification tool to assist with their analysis of the great white shark population in the San Francisco Bay Area. These research directions have led to numerous contributions, including 16 academic articles, a unifying theoretical framework for NeSy modeling, and a new open-source machine learning library.

**Applied Scientist Intern** | Amazon      June 2022–December 2022

Amazon Search Graph Mining Team

During my internship with the Amazon search graph mining team, I worked on the scalable graph neural network (GNN) training initiative. I led research on scaling GNN training via graph coarsening and produced a robust GNN training experiment infrastructure with DGL ([github.com/amazon-science/convolution-matching](https://github.com/amazon-science/convolution-matching)) and an academic paper. I also contributed to research on distributed GNN training and modeling with massive-scale heterogeneous graphs ([github.com/amazon-science/random-tma](https://github.com/amazon-science/random-tma)). I gained experience working with real-world data and training deep models on AWS servers.

**Data Science Intern** | Clari      Summer 2019

I implemented and demonstrated applications of a sales opportunity similarity metric to department executives. Additionally, I created a UI framework for future intern demos with REACT, SemanticUI, and D3.js.

**Data Science Intern** | University Health Partners of Hawai'i      Spring 2019

I utilized H2O, MapR, and SQL to develop a pipeline for triaging health documents for review with a human interpretable explanation derived from a decision tree. Furthermore, I improved model accuracy by implementing procedures for hyperparameter optimization, model selection, and feature space dimensionality reduction.

**Undergraduate Researcher** | Collaborative Software Development Lab      Summer 2018–Summer 2019

I worked on the Open Power Quality project, an open-source solution for distributed power quality data collection, analysis, and visualization ([openpowerquality.org](https://openpowerquality.org)). Using signal processing and machine learning, I developed plugins to classify electrical disturbances and transients ([github.com/openpowerquality/opq](https://github.com/openpowerquality/opq)).

**Research Fellow** | Hawai'i Data Science Institute      Summer 2018–Summer 2019

I was in the first cohort of research fellows with the Hawai'i Data Science Institute. We worked with academic and industry partners on projects, including the development of a tsunami detection system, which was deployed onboard cargo ships, and building Raspberry Pi sensors to monitor volcanic activity.

**Undergraduate Researcher** | University of Hawai'i Big Data Lab      Winter 2017–Spring 2019

I developed a semi-supervised clustering algorithm for measuring the polarity between schools of thought on Google news and Twitter. This project was rewarded funding from the NSF Center for Science of Information.

## PROJECTS & PUBLICATIONS

---

**Neuro-Symbolic AI** | University of California, Santa Cruz

Fall 2021–Present

I specialize in building principled theory and tools for integrating neural models (to process low-level data) and symbolic frameworks (to reason over high-level symbolic structures). To achieve this, I introduced Neuro-Symbolic Energy-Based Models (NeSy-EBMs), a unifying framework for NeSy learning and inference. Further, in collaboration with researchers from Google, the University of California Santa Barbara, and the University of Southern California, I created NeuPSL, a NeSy-EBM implementation extending PSL. Additionally, with researchers from the University of Wisconsin Madison, I developed general NeSy-EBM learning algorithms and established the theoretical basis for new inference methods in NeuPSL. I have built NeSy models for numerous applications, including computer vision, large-scale graph node classification, link prediction, recommendation, and demand forecasting.

- (In Progress) **Charles Dickens**, Connor Pryor, Changyu Gao, Stephen Wright, and Lise Getoor. Neuro-symbolic energy-based models. *JMLR*, 2024c
- (In Submission) **Charles Dickens**, Changyu Gao, Connor Pryor, Stephen Wright, and Lise Getoor. Convex and bilevel optimization for neuro-symbolic inference and learning. In *ICML*, 2024a
- **Charles Dickens**, Connor Pryor, and Lise Getoor. Modeling patterns for neural-symbolic reasoning using energy-based models. In *AAAI-MAKE*, 2024d
- **Charles Dickens**, Connor Pryor, and Lise Getoor. Top-5 finisher in the road event awareness in autonomous driving with requirements (RoadR) challenge at NeurIPS 2023.
- Connor Pryor, **Charles Dickens**, and Lise Getoor. Deep neuro-symbolic weight learning in neural probabilistic soft logic. In *KLREDL Workshop at ICML*, 2023
- **Charles Dickens**, Connor Pryor, Eriq Augustine, Alon Albalak, William Yang Wang, and Lise Getoor. NeuPSL: Neural probabilistic soft logic. In *IJCAI*, 2023b
- **Charles Dickens**, Connor Pryor, Eriq Augustine, Alon Albalak, and Lise Getoor. Efficient learning losses for deep hinge-loss markov random fields. In *TPM Workshop at UAI*, 2022
- Eriq Augustine, Connor Pryor, **Charles Dickens**, Jay Pujara, William Yang Wang, and Lise Getoor. Visual sudoku puzzle classification: A suite of collective neuro-symbolic tasks. In *NeSy Workshop*, 2022b
- Eriq Augustine, Pegah Jandaghi, Alon Albalak, Connor Pryor, **Charles Dickens**, William Wang, and Lise Getoor. Emotion recognition in conversation using probabilistic soft logic. In *Arxiv*, 2022a

**Scaling Graph Neural Networks** | Amazon Search Graph Mining Team

Summer 2022–Winter 2022

I led research on scalable graph neural network (GNN) training via coarsening, achieving a training graph of just 1% of its original size while maintaining up to 95% of the GNN’s performance. Additionally, I made substantial contributions to work on massive graph partitioning and distributed GNN training techniques, resulting in state-of-the-art performance and a 2.31x speedup in training time on a 1.3 billion edge e-commerce network.

- **Charles Dickens**, Eddie Huang, Aishwarya Reganti, Jiong Zhu, Karthik Subbian, and Danai Kutra. Graph coarsening via convolution matching for scalable graph neural network training. In *DCAI Workshop at WWW*, 2024b
- Jiong Zhu, Aishwarya Naresh Reganti, Eddie Huang, **Charles Dickens**, Nikhil Rao, Karthik Subbian, and Danai Koutra. Simplifying distributed neural network training on massive graphs: Randomized partitions improve model aggregation. In *LL Workshop at ICML*, 2023

**Shark Fin Identifier** | University of California, Santa Cruz and Stanford University

Winter 2023–Present

I am building an encoder that embeds images of great white shark fins to support image-based searches to assist Dr. Barbara Block and her lab at Stanford University with population counting and behavior analysis. Specifically, the encoder is a vision transformer that is fine-tuned using data augmentation with self-supervised learning.

**Probabilistic Soft Logic** | University of California, Santa Cruz

Fall 2019–Present

I am an admin and developer of Probabilistic Soft Logic (PSL), an open-source machine learning framework for structured prediction ([github.com/linqs/psl](https://github.com/linqs/psl)). Notably, I brought PSL’s symbolic reasoning capabilities to deep models implemented in major frameworks, including PyTorch, TensorFlow, and DGL. Furthermore, I have advanced the theory underlying PSL and created new principled learning and inference algorithms, yielding over 100× learning runtime improvements and significant prediction performance gains. Moreover, I led the OnlinePSL project, an extension of PSL for performing collective inference over evolving graphical models. I achieved this by implementing a client-server architecture for issuing model updates and extending an out-of-core inference algorithm. Finally, I have reintroduced the power of mixed-integer programming to PSL with Gurobi, strengthening its modeling and reasoning capabilities.

- Sriram Srinivasan, **Charles Dickens**, Eriq Augustine, Golnoosh Farnadi, and Lise Getoor. A taxonomy of weight learning methods for statistical relational learning. *Machine Learning*, 2021
- **Charles Dickens**, Connor Pryor, Eriq Augustine, Alex Miller, and Lise Getoor. Context-aware online collective inference for templated graphical models. In *ICML*, 2021b
- **Charles Dickens**, Eriq Augustine, Connor Pryor, and Lise Getoor. Negative weights in hinge-loss markov random fields. In *TPM Workshop at UAI*, 2021a
- **Charles Dickens**, Alex Miller, and Lise Getoor. Online collective demand forecasting for bike sharing services. In *HICSS*, 2023a

**HyperFair** | University of California, Santa Cruz

Fall 2019

I developed HyperFair, a framework for enforcing soft fairness constraints on the HyPER recommender system.

- **Charles Dickens**, Rishika Singh, and Lise Getoor. Hyperfair: A soft approach to integrating fairness criteria. In *FACCTRec Workshop at RecSys*, 2020

**Open Power Quality** | Collaborative Software Development Lab

Summer 2018–Spring 2019

I built a transient classification system in an electrical grid monitoring system.

- **Charles Dickens**, Anthony J. Christe, and Phillip M. Johnson. A transient classification system implementation on an open source distributed power quality network. In *IARIA*, 2019

## TEACHING & LEADERSHIP EXPERIENCE

---

**Introduction to Artificial Intelligence Teaching Assistant** | University of California, Santa Cruz Winter 2021  
I led discussion sections, hosted office hours, created assignments and exams, and graded for a course of 100+ students.

**Introduction to Algorithms Teaching Assistant** | University of California, Santa Cruz Fall 2019  
I led discussion sections, hosted office hours, and graded for a course of 50+ students.

**Data Wrangling with Pandas** | Hawai'i Data Science Institute Spring 2019  
I wrote a course on data wrangling with Python and Pandas and led a corresponding workshop at the University of Hawai'i ([github.com/hawaiidatascience/pandas\\_data\\_wrangling](https://github.com/hawaiidatascience/pandas_data_wrangling)).

**Research Team Leader** | Purdue and NSF CSOI Data Science Workshop Summer 2018  
I was invited to lead a group of student researchers to make progress on and write a proposal for the Polarization of Information project. Our project received NSF funding to continue the collaboration.

**Student Research Mentor** | University of Hawai'i College of Engineering June 2017–August 2017  
I led a group of high school students participating in the University of Hawai'i College of Engineering summer research program to build a Python graph algorithms library.

**Math and Science Tutor** | Online Learning Academy, Honolulu Hawai'i Winter 2016–Summer 2018  
I tutored students across the state of Hawai'i in STEM courses at the grade school to sophomore college level.

## RELEVANT COURSES

---

### **Artificial Intelligence and Machine Learning:**

UCSC CSE: Machine Learning (A+), Responsible Data Science (A), Computing for Society(A+)

UHM EE: Machine Learning (A), Signals & Systems (A+)

### **Mathematics:**

UCSC AM: Convex Optimization (A), Numerical Optimization (A)

UCSC STAT: Applied Bayesian (A+)

UHM MATH: Linear Algebra I, II (A), Real Analysis(A+), Statistical Inference (A), Calculus I-IV (A)

### **Computer Science:**

UCSC CSE: Algorithms and Analysis (A), Programming Languages (A), Information Theory (A)

UHM EE: Operating Systems (A), Computer and Network Security (A+), Signals and Systems (A)

UHM ICS: Software Engineering I, II (A)

## SKILLS

---

**Research:** Machine Learning, Deep Learning, Structured Prediction, Graph Mining, Graph Neural Networks, Neuro-Symbolic Artificial Intelligence, Statistical Relational Learning, Nonlinear Programming, Bilevel Optimization

**Frameworks** PSL, PyTorch, TensorFlow, DGL, Gurobi

**Languages:** Python, Java, MATLAB, C, Bash, SQL, L<sup>A</sup>T<sub>E</sub>X.

**Tools:** Git, GitHub, AWS, JetBrains IDEs, Jupyter Notebooks

## AWARDS

---

UCSC Regents Fellowship, NSF CSOI Channels Scholar REU Award, UHM Dean's List 10 Semesters