Xintong Li

xli2224@wisc.edu | (608)-949-4642 | https://kaylee0501.github.io/

EDUCATION

University of Wisconsin-Madison Madison, WI

Bachelor of Science in Computer Sciences and Data Science Sep 2019 – May 2023

GPA: 3.981/4.0

RESEARCH INTERESTS

- Weak Supervision, Representation learning
- Foundation models, Automated Machine Learning
- Optimization for machine learning, Adaptive gradient descent

PUBLICATIONS

*equal contribution

CONFERENCE PUBLICATIONS

Nicholas Roberts*, **Xintong Li***, Tzu-Heng Huang, Dyah Adila, Spencer Schoenberg, Cheng-Yu Liu, Lauren Pick, Haotian Ma, Aws Albarghouthi, Frederic Sala. *AutoWS-Bench-101: Benchmarking Automated Weak Supervision with 100 Labels.* Neural Information Processing Systems (NeurIPS), 2022.

UNDER REVIEW

Nicholas Roberts, **Xintong Li**, Dyah Adila, Sonia Cromp, Tzu-Heng Huang, Jitian Zhao, Frederic Sala. (2022). *Escaping Label Subspaces via Label Geometry.*

RESEARCH EXPERIENCE

UW-Madison: Department of Computer Science Feb 2022 – Present

Undergraduate Research Assistant for Optimization | Advisor: Prof. Jelena Diakonikolas

- Used potential function-based framework to study the convergence of adaptive gradient descent methods
- Learned the tradeoff between minimizing the optimality gap and the norm of the gradient under different functions
- Extended to non-convex and local smoothness case to search for better convergence rates

UW-Madison: Department of Computer Science Dec 2020 – Present

Undergraduate Research Assistant for Machine Learning | Advisor: Prof. Fred Sala

- Implemented plug-and-play combinations of feature representations and automatic label function generation and selection framework using Scikit-Learn and PyTorch to enable weak supervision for diverse learning tasks
- Fused multiple sources of signal, including foundation models, into automated weak supervision pipelines to maximize performance
- Incorporated the geometric relationship of label spaces in order to learn in partially observed label spaces of extremely high cardinality

SELECTED PROJECTS

Filtering Rules on Deer Data in Linear Prediction by Land Cover Features Sep 2021 – May 2022

Team Lead | Advisor: Prof. Tyler Caraza-Harter, University of Wisconsin-Madison

- Structured observations and built new columns to map professional and volunteer datasets together
- Compared p-value with uncertainty using Fisher model to eliminate outliers in the datasets
- Coordinated the division of labor in a team of 4 people and communicated with researchers to report progress

Forecasting Airport Waiting Time using Machine Learning Models Sep 2021 – Dec 2021

Team Lead

- Separated time intervals and created new datasets using Pandas and R to obtain factors that affect waiting time
- Selected variables by feature importance and found the best hyperparameter for SVM, Random Forest, XGBoost, and Multilayer Perceptron using GridSearchCV to generate the highest test accuracy for each model
- Calculated the p-value using McNemar's test to compare the performance of models with similar test accuracy

Restaurant Information Desk Oct 2020 – Dec 2020

CS 400: Class Project Team Member

- Stored each restaurant's menu in a Red Black Tree and generated the shortest path from different restaurants using Dijkstra's shortest path algorithm to create a restaurant map
- Consulted with Front end developers to design a functional JavaFX application for web visualization
- Collaborated with a group of 8 members to build a feasible restaurant information desk

HONORS & AWARDS:

NeurIPS Scholar Award 2022

Wisconsin Science and Computing Emerging Research Stars (exploreCSR award) 2022

UW–Madison Dean's List 6 consecutive semesters

SELECTED COURSES:

Machine Learning, Deep Learning, Nonlinear Optimization, Operating Systems, Database, Data Modeling, Probability, Linear Algebra, Computer Programing, Algorithm, Natural Language Processing, Foundation of Data Science

LEADERSHIP, ACTIVITIES, AND EXTRACURRICULAR:

Data Science Research Group Team Member

University Housing Team Member

UW-Madison Society of Women Engineers Team Member

Red Cross Club Team Member

SKILLS

Programming Language: Java, C++, Python, C, R, Assembly language

Web Programming Knowledge: JavaScript, HTML, CSS, JavaFX

Tools: Pandas, Pytorch, TensorFlow, Scikit Learn, SQL, Scipy, Numpy