

EDUCATION

The University of Illinois Urbana-Champaign

August 2022 - Present

Master of Science in Computer Science

3.7 Cumulative GPA, May 2024 Expected Graduation

Coursework: Statistical RL, Transfer Learning, Algorithmic Market Microstructures**Activities:** Quant @ Illinois - Research Division, Data Structures TA, iDEA Lab

The University of North Carolina at Chapel Hill

August 2019 - May 2022

Bachelor of Science in Computer Science, Bachelor of Science in Statistics

3.7 Cumulative GPA

Coursework: Stochastic Modeling, Probability, Data Structures, Algorithms, Discrete Math, Linear Algebra**Activities:** Carolina Analytics and Data Science (CADS), UNC Finance Society, LUPA Lab

EXPERIENCE

Cutler Group

July 2023 - August 2023

Quantitative Trading Intern, LAG Team

Practiced mock trading, market-making, and options theory.

Reconstructed time decay profiles to improve theta decay models using historical data.

Automated skew analysis to alert traders to imbalanced hyperparameters.

Google

April 2023 - July 2023

Software Engineer Intern, YouTube Ads

Improved ad quality by applying conditional multimodal generative AI models CTRL and MMUM to automate advertiser-friendly campaigns, improving ad evaluation by X% using ad attributes as features.

Collaborated with Google Research to finetune internal models; a publication is under review at WWW '24.

Deployed models on internal platform to reach XXXmm customers.

Capital One

May 2022 - August 2022

Software Engineer Intern, Customer Experience Team

Constructed a sentiment analysis pipeline in Python using a roBERTa-Large with Self-Explaining model.

Used for managing real-time customer feedback; improves proprietary satisfaction index by X%.

Deployed pipeline on Amazon Web Services with SQS, Lambda, and DynamoDB microservices.

IQVIA

May 2021 - August 2021

Artificial Intelligence Intern, Internal Vendors Team

Designed an invoice parser to automate information extraction through modeling in Python, saving over \$0.XM annually by directly billing vendors with parsed invoices.

Utilized the PyTesseract library for optical character recognition of invoices.

Implemented a graph convolutional neural network to incorporate both spatial and semantic information.

PROJECTS

Melanoma Classification

August 2019

Finetuned and utilized VGG, ResNet, and ViT to classify skin cell images as malignant or benign with UNC

Hospitals' patient data; obtained 87% accuracy.

Transformed data with rotations and flips to augment training dataset.

RESEARCH

Generative Reasoning on Knowledge Graphs

August 2022 - Present

Dr. Hanghang Tong, Department of Computer Science, UIUC

Currently researching nascent extensions of diffusion with knowledge graphs to solve various tasks such as knowledge graph completion, policy optimization over graphically structured data, and other reasoning tasks.

Two publications are under review at WSDM '24.

Hypothesis-Test Driven Coordinate Ascent

August 2021 - May 2022

Dr. Junier Oliva, Department of Computer Science, UNC-CH

Explored black-box optimization via Hypothesis-Test Driven Coordinate Ascent (HDCA) to decompose the policy space and make statistically guided updates to parameters to avoid costly gradient calculations.

PUBLICATIONS

Ginkgo-P: General Illustrations of Knowledge Graphs for Openness as a Platform

Blaine Hill, Lihui Liu, Hanghang Tong - Under Review, WSDM '24

A demo paper to both automate infrastructure for KG visualization and to codify several important KG reasoning categories: KG completion, KG question answering, KG subgraph extraction, and KG with RL optimization.

Conversational Question Answering with Reformulations over Knowledge Graphs

Lihui Liu, Blaine Hill, Boxin Du, Hanghang Tong - Under Review, SDM '24

A long paper to tackle the challenge of conversation question answering by using a learned RL policy to both reformulate natural language questions and answer them using graph data.

Amalgamation of Predictive Modeling and LLMs in YouTube Ads

Poorva Potdar, Blaine Hill, Shobha Diwakar - Under Review, WWW '24

An industry paper to combine different multimodal large language models towards improving ads in YouTube by learning the underlying characteristics of well-performing ads and using them as features.

PROJECTS

Arcane

Constructed a web application employing Spotify user data to generate unique personalized discographies; creates listening sessions by over 2 hours on average.

Utilized the React, Express.js, and Node.js stacks as well as the Spotify Web API.

Programmed in JavaScript, HTML and styled with the Tailwind CSS and Bootstrap frameworks

Melanoma Classification

Finetuned and utilized VGG, ResNet, and ViT to classify skin cell images as malignant or benign with UNC Hospitals' patient data; obtained 87% accuracy.

Transformed data with rotations and flips to augment training dataset.

Written in Python, utilizing the PyTorch package for fluidity.

HONORS / AWARDS

3rd Place 2019 Pokemon Trading Card Game World Championship

August 2019

Was invited to compete after the 2018-2019 tournament circuit.

Finished as the best placing American in the 2019 season.

SKILLS

Languages

Python, R, C++, C, Java, JavaScript, HTML, SQL

Tools / Frameworks

AWS, Git, Snowflake, Keras, Tensorflow, pandas, NumPy, nltk, scikit-learn, React, Node.js, Express.js, MySQL, MongoDB, Material-UI, Tailwind CSS