Helen Jin

560 Riverside Drive, New York, NY 10027
hj2368@columbia.edu • 646-643-1348 • https://helenjin.github.io/
CURRICULUM VITAE - PREPARATION DATE: Jun 2020

EDUCATION

University of Pennsylvania, Philadelphia, PA — Expected: May 2026
Ph.D. Candidate in Computer and Information Science (CIS) — GPA: -----
- **Advisor: Dan Roth**, Eduardo D. Glandt Distinguished Professor
- Cognitive Computation Group
- Areas: Machine Learning, Natural Language Processing, Artificial Intelligence

Columbia College, Columbia University, New York, NY — Sep 2016 - May 2020
B.A. – Double Major in Mathematics and Computer Science, Concentration in East Asian Studies — GPA: 3.72 / 4.00
- **Computer Science Track: Intelligent Systems**
- **Relevant Coursework**: Data Structures in Java, Advanced Programming in C/C++, Discrete Mathematics, Fundamentals of Computer Systems, Computer Science Theory, Artificial Intelligence, Machine Learning, Natural Language Processing, Computer Vision, Computation and the Brain, Advanced Spoken Language Processing, Databases, Linear Algebra, Modern Algebra I + II, Modern Analysis I + II, Topology, Calculus-Based Statistics, Probability Theory
- **Honors/Awards**: Dean's List

Girls Who Invest, The Wharton School, University of Pennsylvania, Philadelphia, PA — May 2018 – Aug 2018
Summer Intensive Program Scholar
- Completed highly selective, rigorous 4-week training program focused on core investment concepts and skills taught by leading business school professors and investment professionals; upon completion, earned a paid 6-week asset management internship
- Coursework and case studies included accounting, valuation, financial modeling, asset allocation and presentation skills

Stuyvesant High School, New York, NY — Sep 2012 – Jun 2016
Advanced Regents Diploma — GPA: 4.0 / 4.0
- **Activities**: Girls' Varsity Swimming, Lifeguard, Stuyvesant Spectator, Math Team, CS Dojo and Writing Center Tutor
- **Honors/Awards**: AP National Scholar; National Merit Finalist; ARISTA National Honor Society; National Latin Society; National Junior Classical League; National Latin Exam Awards: Maxima Cum Laude; Award of Distinction in Mathematics

PROFESSIONAL EXPERIENCE

Department of Statistics, Columbia University, New York, NY — Jun 2019 – Present
Undergraduate Researcher
- PI: Tian Zheng, Professor
- Worked on a research project that seeks to develop an artificial perceptual learning framework that can aid semi-supervised learning problems, specifically with a focus on image classification on large images of tree canopies with a small labeled tree species data set
- Used various statistical and deep learning methods primarily in Python, including TensorFlow and Keras packages

LionBase, LLC, New York, NY — Dec 2018 – Sep 2019
Client-Facing Team Member
- LionBase is a new student-led data science and product development group at Columbia University
- Worked in a team of six to solve real-world industry problems related to data analytics, ML, NLP, and statistical analysis
- Spring 2019: Recruitment Platform for Early Stage Data Science Talent – my team collaborated with an executive search firm to design, develop, and implemented from scratch a platform that matches data science talent and company based on various technical and non-technical tests.

Data Science Institute, Columbia University, New York, NY — Mar 2019 – May 2019
Data for Good (DFG) Scholar
- Worked in a team in collaboration with Okimo, a startup using eye tracking technology to help identify and aid individuals' reading skills especially targeting younger age groups in low resource communities in developing countries such as Paraguay

RippleMatch, Remote Location Feb 2019

Growth Marketing Intern

- Selected from a pool of thousands of candidates to work closely with the leaders of RippleMatch's Marketing Team.
- Leveraged various growth strategies and tools including social media, email marketing, presentations, and peer & faculty member networking to grow the user base and awareness on campus.
- Strategically assessed growth and performance metrics to improve, change and/or help design new growth strategies

Grantham, Mayo, van Otterloo & Co. LLC (GMO), Boston, MA Jul – Aug 2018

Quantitative Equity Analyst Intern

- Using MATLAB and SQL, created a working tool to measure the relationship between institutional investor concentration and downside risk in equities, and modeled risk in global equity markets

Computational and Systems Biology, Memorial Sloan Kettering Cancer Center, New York, NY May 2017 – Aug 2018

Undergraduate Researcher

- PI: Dana Pe'er, Professor
- Developed computational methods and packages to analyze single cell RNA-seq data using Python and R

Finance Office, Graduate School of Architecture, Preservation, and Planning (GSAPP), New York, NY Jan – Nov 2017

Administrative Assistant

- Tasks included: data entry; creating Excel spreadsheets and Word documents for office projects; keeping various paper files up to date; scanning and uploading invoices, purchase orders, and other supporting material; shredding confidential documents; organizing paychecks; creating folders for new hires, casual employees and TA's

SELECTED TEACHING EXPERIENCE

- **Private Tutor** for Mathematics and Chemistry subjects, high school level (Feb 2017 – Jun 2017)
- **Teaching Assistant** for Calculus-Based Intro to Statistics (Fall 2017)
- **Teaching Assistant** for Calculus I (Fall 2017)
- **Teaching Assistant** for Calculus III (Fall 2018, Summer 2019)
- **Teaching Assistant** for Calculus IV (Spring 2018, Summer 2019)
- **Teaching Assistant** for Linear Algebra (Spring 2019)
- **Teaching Assistant** for Natural Language Processing (Fall 2019, Spring 2020, Summer 2020)

PUBLICATIONS

- Chengliang Tang, Helen Jin, Maria Uriarte, Douglas Morton, and Tian Zheng. **Artificial Perceptual Learning: Image Categorization with Weak Supervision.** *Submitted to AISTATS 2020 Conference*, September 2019. <https://drive.google.com/open?id=1O9LPIo2dMgi-qzDIy1lDHp1pi0jH4EPF>

SELECTED LEADERSHIP

Girls Who Code (GWC) at Columbia University, New York, NY Feb 2017 – Mar 2020

President

- Oversaw high school outreach and recruitment for program that holds weekly CS classes for ~50 high school girls each semester
- Planned on-site visits, created fundraisers, and reached out to companies and local businesses for sponsorship
- Worked with other Executive board members to better improve program
- Managed ~30 people on the Managing Board (Finance, Programming, Publicity, HS Recruitment)
- Previously served as Vice President of External Affairs, and before that, Sponsors and Finance Team

One-to-One Tutoring at Columbia University, New York, NY Oct 2016 – Sep 2019

Treasury Coordinator

- Allocated and managed annual $2,000 budget to various on-campus events from weekly meetings to fundraisers
- Supervised and oversaw volunteers during tutoring sessions each semester to ensure attendance
- Worked with other coordinators and Community Impact staff to discuss issues and implement solutions to improve program
- Volunteered weekly to tutor and mentor 6-12 year old individual throughout academic year

Columbia Table Tennis Club, New York, NY Sep 2017 – Dec 2018

Secretary

Smart Woman Securities (SWS), Columbia Chapter, New York, NY Sep 2017 – May 2018

Senior Research Analyst

- Led a team of SWS Analysts researching the Consumer industry
- Conducted background research, analyzed and formulated investment opinions on companies to make real-life investment recommendations

SELECTED PROJECTS

Persistent Homology

- Created a visualization tool for the persistent homology in Python for my Math seminar on Elementary Applied Topology

SCAnalysis

- Built a Python package and Jupyter notebook to analyze single cell RNA-seq data that incorporates MAGIC and Wishbone technologies, with additional features (including Palantir). Wishbone is an algorithm to align single cells from differentiation systems with bifurcating branches. MAGIC (Markov-Affinity Based Graph Imputation of Cells) is an interactive tool to impute missing values in single-cell data and restore the structure of the data.

Tetris

- Created a modified version of the classic game Tetris in Processing
- There are two different modes of play: levels and infinite. In Levels mode, there are twenty successive levels you can play. To reach the next level, you must clear a certain number of Tetrominos. The required number increases as the level increases. In Infinite mode, you can play forever and ever

HONORS & AWARDS

- **Dean's List**, Columbia University
- **Rewriting the Code (RTC) Fellow** (2017- present)
- **Science Technology Engineering Program (STEP)**, Columbia University (2017)

SKILLS & INTERESTS

Technologies: Java, Python, MATLAB, SQL, R, C/C++, Git, Jupyter Notebook, LaTeX , HTML/CSS/JavaScript, Processing, Racket, NetLogo, Microsoft Office (Excel, Word, PowerPoint), G Suite (Gmail, Drive, etc.), Training the Street (Financial Modeling)

Languages: English (native), Korean (intermediate), Mandarin Chinese (intermediate)

Personal Interests: Psychology, Philosophy, Music, Visual Arts, Swimming, Yoga, Traveling