

Scott Geng

scott.geng@columbia.edu | scottgeng.com | [in](#) scottgeng00 | [G](#) scottgeng00

EDUCATION

Columbia University

Expected May 2023

B.A. Computer Science

New York, NY

- GPA: 4.1/4.0
- Relevant Coursework: Advanced Programming, Making and Breaking Codes, Data Structures, Statistics
- Selected as a Rabi Scholar, a distinction given annually to the top ~10 incoming freshman STEM majors

TECHNICAL SKILLS

Languages: Python, Java, C, Bash, Perl, HTML/CSS, R, LaTeX

Libraries: NumPy, SciPy, pandas, scikit-learn, spaCy, Matplotlib, seaborn, H2O

EXPERIENCE

Columbia University Medical Center

March 2020 – Present

Software Developer Intern

New York, NY

- Designed and developed a Python analysis pipeline for cerebellar electroencephalogram data using Fourier and machine-learning methods, saving 200+ hrs. of physician time and increasing number of analysis metrics by 500%
- Applied pipeline to real-world data from healthy and diseased patients, yielding significant publishable results
- Drafted a manuscript detailing software methods and results for submission to a peer-reviewed academic journal
- Presented project progress using self-designed figures to an interdisciplinary team at weekly lab-wide meetings
- Communicated work to the wider academic community at Columbia's annual undergraduate research conference

National Institutes of Health

June 2018 – Aug. 2018

Data Science Intern

Bethesda, MD

- Developed and evaluated gradient-boosting and random forest machine learning models in R to predict N-glycan occupancy within any protein, leading to an overall prediction accuracy of 95%, the highest obtained in the field
- Implemented minimum redundancy pre-training feature selection, further increasing prediction accuracy by 7%
- Designed Perl scripts in a Linux remote shell to scrape data from previously published scientific research, automating dataset generation and resulting in over 5000 unique features per data entry

PROJECTS

Natural Language Interpreter | *Python, Google Cloud, Bash, C*

October 2020 – Present

- Developed an app in a team of 4 that automatically generates outline notes of Zoom meetings via speech recognition and natural language processing, winning 1st place in the education track at Columbia's annual hackathon
- Implemented a TextRank algorithm and a neural network to detect important moments in Zoom lecture recordings

CPU Architecture and Assembler | *HDL, Python*

July 2020 – Present

- Designed a theoretical 8-bit CPU in a hardware description language with only NAND and D-flip-flop logic gates
- Developed an assembler in Python to convert user-written low-level instructions to CPU-readable bytecode
- Implemented instruction jumping, random memory access, conditional statements, variable declaration, and labels

... and more on [Github](#)

LEADERSHIP & ACTIVITIES

Columbia Bach Society

Feb. 2020 – Present

Webmaster and Board Member

New York, NY

- Designed club website to better align it with the board's vision, improving our communication with the wider community and resulting in an increase of unique website traffic by 20% (columbiabachsociety.weebly.com)
- Organized and executed social bonding events such as parties and club dinners to improve intraclub relationships

Frontiers of Science

Sept. 2019 – Dec. 2019

Teaching Assistant

New York, NY

- Held weekly office hours to help other students in a ~500 student class with lecture concepts and homework