Scott Geng

scott.geng@columbia.edu | www.scottgeng.com | 443-939-9246

Education

Columbia University Expected May 2023
B.A. Computer Science and Mathematics (GPA: 4.15/4.0) *New York, NY*

- Co-advised by Prof. Carl Vondrick and Prof. Junfeng Yang

Research Interests

Computer vision, vision-language reasoning, self-supervised representation learning, multi-modal learning, video understanding, social intelligence, robustness.

Publications

Scott Geng*, Revant Teotia*, Purva Tendulkar, Sachit Menon, Carl Vondrick. "Affective Faces for Goal-Driven Dyadic Communication." *In submission*.

Chengzhi Mao*, **Scott Geng***, Junfeng Yang, Xin Wang, Carl Vondrick. "Understanding Zero-shot Adversarial Robustness for Large-Scale Models." *In submission*.

Kexin Pei, Dongdong She*, Michael Wang*, **Scott Geng***, Zhou Xuan, Yaniv David, Junfeng Yang, Suman Jana, Baishakhi Ray. "NeuDep: Neural Binary Memory Dependence Analysis." *ESEC/FSE* 2022.

Shi-Bing Wong, Yi-Mei Wang, Chih-Chun Lin, **Scott Geng**, Nora Vanegas-Arroyave, Seth Pullman, Sheng-Han Kuo, Ming-Kai Pan. "Cerebellar Oscillations in Familial and Sporadic Essential Tremor." *Cerebellum* 2021.

Note: equal contributions are denoted by *.

Research Experience

Columbia Computer Vision Lab Apr 2022 – Present
Advisor: Prof. Carl Vondrick *New York, NY*

- Defined and explored a new AI social intelligence task, listening with goals: given video of a speaker, output video of someone listening with a *specified goal*; created RealTalk video dataset
- Investigating language models for learning social intelligence and solving listening task
- Introduced novel problem of adapting models for zero-shot adversarial robustness and proposed language grounded contrastive finetuning as a first solution; method achieves 2.5x accuracy of baselines

Columbia Software Systems Lab Nov 2020 – Present
Advisor: Prof. Junfeng Yang *New York, NY*

- Investigating graph embeddings for learning robust representations of code that are provably equivariant to natural symmetries; designing adversarial attacks to evaluate limits of this robustness
- Designed a multi-modal masked autoencoder to learn representations of binary programs by modeling their execution semantics; conducted experiments evaluating the transferability of the representation

Columbia Ataxia and Tremor Lab Jan 2020 – Mar 2021
Advisor: Prof. Sheng-Han Kuo *New York, NY*

- Designed a data analysis framework for cerebellar electroencephalogram data (i.e., brain wave scans) to quantitatively characterize neurological movement disorders such as Parkinson's and essential tremor

Presentations

Poster presentation, Columbia Undergraduate Research Symposium (2022)
"Zero-Shot Adversarial Robustness for AI Vision."

Poster presentation, Columbia Undergraduate Research Symposium (2021)
"TMod: Learning Operational Semantics for Binary Program Dependence Analysis."

Poster presentation, Columbia Undergraduate Research Symposium (2020)
"Reading Brainwaves: Differential Electrophysiological Activity in Patients with Cerebellar Ataxia."

Honors and Awards

Junior Phi Beta Kappa Nov 2022

- Awarded to top 2% of graduating class (25 students total).

Rabi Fellowship Sep 2019

- Columbia College's most prestigious undergraduate research fellowship, awarded to the top ∼10/1600 incoming freshman STEM students. Fellowship provides research funding every semester and summer.

Teaching and Leadership

Columbia University

Teaching Assistant *New York, NY*

- Graduate Machine Learning (COMS 4771) — Fall 2021, Spring 2022. Topics include decision trees, SVMs, dimensionality reduction, PAC learnability, and optimality bounds for common learning algorithms.
- Frontiers of Science (SCNC 1000) — Fall 2019.

Scientists and Engineers for a Better Society Sep 2019 – Present

Treasurer, Executive Board Member *New York, NY*

- Lead and organize outreach events (tutoring, experiment demonstrations, etc.) targeting disadvantaged elementary and middle school students in local schools to promote scientific engagement and interest.

Work Experience

Regeneron Pharmaceuticals Jun 2021 – Aug 2021

Software Engineering Intern *Albany, NY*

- Engineered large-scaled, automated data analysis pipeline for results of genomic sequencing experiments

Personal Skills

Coding Languages. Python, C, OCaml, Java, HTML/CSS
Technical Tools. fairseq, PyTorch, TensorFlow/Keras, Numpy, Scipy, Scikit-Learn, Matplotlib, Seaborn, Django
Spoken Languages. English: Native; Mandarin Chinese: Professional Working Proficiency
Other. Amateur clarinet and guitar player