

William Huang

225 Rector Pl. Apt 12C, New York, NY 10280

☎ (203) 675-4977 · ✉ will.huang@nyu.edu · [in linkedin.com/in/williamhuang93](https://www.linkedin.com/in/williamhuang93) · 🌐 <https://wh629.github.io/>

Education

New York University, Courant Institute, New York, NY

Anticipated: May 2021

M.S., Scientific Computing

Cumulative GPA: 3.83

Relevant Coursework: Team Project Seminar in Crowdsourcing (Current), Natural Language Processing (Current), Natural Language Understanding, Deep Learning, Fundamental Algorithms, Programming Languages, Numerical Methods I

Cornell University, College of Engineering, Ithaca, NY

May 2016

B.S. with Honors, Operations Research Engineering, Dyson Business Minor for Engineers

Major GPA: 3.76 | Cumulative GPA: 3.60 | Cum Laude

Relevant Coursework: Machine Learning, Optimization, Stochastic Processes, Engineering Probability and Statistics

Publications (*Equal contribution)

2020 “Precise Task Formalization Matters in Winograd Schema Evaluations.” Haokun Liu,* **William Huang**,* Dhara A. Mungra, and Samuel R. Bowman. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

2020 “Counterfactually-Augmented SNLI Training Data Does Not Yield Better Generalization Than Unaugmented Data.” **William Huang**, Haokun Liu, and Samuel R. Bowman. In *Proceedings of the 2020 EMNLP Workshop on Insights from Negative Results in NLP*.

Research & Professional Experience

Machine Learning for Language (ML²), New York, NY

Advised by Prof. Samuel R. Bowman

Jun 2020 — Present

- Researched the use of Item Response Theory (IRT) as a method to evaluate the relative difficulties of NLP benchmarks
- Developed a script using Pyro to apply black-box Variational Inference to fit an IRT model to a set of machine responses by experimenting with different prior and posterior approximation assumptions
- Selected as set of NLP benchmarks based on reasoning criteria for a comparison of evaluation tasks
- Researched properties of counterfactually-augmented natural language inference (NLI) training data by evaluating RoBERTa models on out-of-domain examples and challenge examples to test generalization and robustness
- Analyzed the lexical diversity of counterfactually-augmented training sets to understand the effectiveness of minimal edit constraints and the role of dataset size
- Coordinated several hundred training runs on a high performance computer cluster to train RoBERTa and BERT models on seven Winograd Schema Challenge (WSC) task formalizations
- Analyzed and visualized results using Pandas and Matplotlib for an ablation analysis between two popular WSC formalizations to attribute an 11-21% difference in validation accuracy
- Built a custom task sampler with PyTorch to study the effect of batching similar examples during deep learning training

American International Group, Inc (AIG), New York, NY

Enterprise Risk Management Analyst II

Aug 2016 — Jun 2019

- Managed quarterly capital estimation by coordinating with a multinational team and presenting risk profile changes from business activities and market movements to key stakeholders
- Led a team of multinational analysts to migrate portions of AIG’s capital model to Python in collaboration with IT

Skills & Interests

Technical Skills Python • PyTorch • Pandas • Numpy • Transformers • Matplotlib • Pyro • Microsoft Office • L^AT_EX
Interests Board Games • Skiing • Fantasy Football • Cooking