William Huang

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

(203) 675-4977 · wh629@nyu.edu · 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: Seminar in Semantics (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

Research & Professional Experience

Machine Learning for Language (ML^2); New York, NY

Research Assistant advised by Prof. Samuel R. Bowman — Sep 2020 — Present

- Researched the use of Item Response Theory (IRT) as a method to evaluate the relative difficulties of NLP benchmarks
- Developed a program 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

Collaborator — Jun 2020 — Sep 2020

- Built a custom task sampler with PyTorch to investigate the effect of batching highly similar data during training
- Curated a set of annotated natural language datasets based on reasoning criteria for a comparison of evaluation tasks

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

Enterprise Risk Management Analyst — Aug 2016 — Jun 2019

- Analyzed AIG's loss distribution to estimate each subsidiary's required capital to buffer against financial downturn
- Managed quarterly capital estimation and firm wide reporting to key stakeholders including the Chief Risk Officer

Publications (*Equal contribution)

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

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

Projects

Counterfactually-Augmented Training Data; New York, NY — Jun 2020 — Aug 2020

- Researched the use of counterfactually-augmented natural language inference (NLI) training data to improve generalization to out-of-domain problems
- Analyzed properties of counterfactually-augmented NLI training data by evaluating RoBERTa-based models on a set of diagnostic examples to test both linguistic ability and robustness to distractors

Task Sensitivity to Problem Formalization; New York, NY — Feb 2020 — Jun 2020

- Coordinated several hundred training runs on New York University's high performance computer cluster to train RoBERTa-based models on seven Winograd Schema Challenge (WSC) task formalizations
- Analyzed and visualized results using Pandas and Matplotlib to perform an ablation analysis between two popular WSC formalizations and attribute the 6% difference in test accuracy

Skills & Interests

Technical Skills Python • PyTorch • Pandas • Numpy • Huggingface • Matplotlib • Pyro • Microsoft Office • LaTeX

Interests Board Games • Skiing • Fantasy Football • Cooking