

Junghyun Min

Seongnam, Korea | jmin10@jhu.edu | <https://github.com/Aatlantise> | <http://www.aatlantise.science/>

EDUCATION

Johns Hopkins University

Master of Arts in Cognitive Science

Baltimore, MD

Aug 2019 – Dec 2020

- Research track; concentration in computational approaches to linguistics; 3.91 GPA

Bachelor of Science in Physics with a second major in Mathematics

Sep 2014 – Dec 2017

- Dean's List Spring 2015, Fall 2017; General Honors; KGSA Baseball Captain; 3.53 GPA

Wolfram Summer School

Waltham, MA

Science Track

Jun 2016 – Jul 2016

- Courses on computational topics such as deep learning, image processing, optimization in Wolfram language

PUBLICATIONS & PROJECTS

Master's thesis: **Junghyun Min**. 2020. The roots and effects of heuristics in natural language inference and question answering models.

ACL 2020 paper: **Junghyun Min**, R. Thomas McCoy, Dipanjan Das, Emily Pitler, and Tal Linzen. 2020. [Syntactic data augmentation increases robustness to inference heuristics.](#)

BlackboxNLP 2020 paper: R. Thomas McCoy, **Junghyun Min**, and Tal Linzen. 2020. [BERTs of a feather do not generalize together: Large variability in generalization across models with similar test set performance.](#)

Internship project at Wolfram Research: **Junghyun Min** and Giorgia Fortuna. 2017. [What on Earth is this Map?](#)

WORK & RESEARCH EXPERIENCE

NCSoft Language AI

Seongnam, Korea

Research Engineer

Jan 2021 – Current

- Design predicate range optimization framework to improve DL-based open information extraction models.
- Outperform state-of-the-art open information extraction (2%p more accurate), chunking (25% faster) models.
- Establish event and relation extraction prototype pipeline for KB population and value chain construction.
- Deploy improved models internally. Optimize chunking model for scaling as requested (10 calls per ms).

Johns Hopkins University Computation and Psycholinguistics Laboratory

Baltimore, MD

Graduate Research Assistant

Jul 2019 – Oct 2020

- BERT fine-tuned on MNLI and is unstable and vulnerable to syntactic heuristics (McCoy, **Min**, Linzen 2020).
- Adversarial data augmentation via syntactic manipulation of training set data significantly increases robustness to augmentation-like examples and general syntactic sensitivity too (**Min**, McCoy, Das, Pitler, Linzen 2020).
- Heuristics likely arise from both the pre-training and the fine-tuning dataset. Currently popular fine-tuning and evaluation paradigm has drawbacks that can be patched with longer fine-tuning on unbiased datasets, multi-seed out-of-distribution evaluation, and syntactic adversarial augmentation (Master's thesis).

Harford Community College Department of Analytics and Planning

Bel Air, MD

Research Associate

Mar 2019 – Jul 2019

- Improved student retention, success prediction by developing machine learning based predictive models.
- Increased data request processing volume by 20%, by automating edit checks and recurring data requests.
- Launched Tableau implementation project for internal use, handled data definitions and validated mappings.

Research Assistant

Apr 2018 – Mar 2019

- Self-taught SAS, SQL, and SPSS to query and respond to internal requests, external compliance reports.
- Facilitate data-driven engagement for non-technical departments by redesigning enrollment data reports.

Wolfram Research

Waltham, MA; Baltimore, MD

Student Researcher

Jun 2016 – Jul 2017

- Self-trained in Wolfram language and its machine learning package before, during Wolfram Summer School.
- Built an Inception-based classifier with 94% accuracy in continent identification with map image as input.
- Expanded the model by extracting geometric properties from the map and comparing to real-world data.

SKILLS & INTERESTS

Natural & Computer Language

Korean, English, German, Mandarin Chinese, Python, SAS, SQL, Wolfram language, R, Unix

Interests

Geography, cartography; baseball analytics, sabermetrics; urban planning, public transport; low-cost travel.