Nelson F. Liu

E-mail: nfliu@cs.washington.edu
Github: github.com/nelson-liu
Web: homes.cs.washington.edu/~nfliu

Research Interests

Natural Language Processing, Machine Learning, Statistical Inference, Deep Learning

Education

University of Washington, Seattle, Washington
B.S, Computer Science, B.S., Statistics, B.A., Linguistics. Cumulative GPA: 3.81
Expected graduation date: June 2019
Relevant Coursework: CSE490U (Natural Language Processing), CSE590 (Graduate Artificial Intelligence Seminar), LING461 (Syntax), LING461 (Phonetics), MATH308 (Matrix Algebra)

Publications

Johannes Welbl, **Nelson F. Liu**, Matt Gardner. "Crowdsourcing Multiple Choice Science Questions." In review, submitted to EMNLP 2017.

Research Experience

University of Washington Computer Science and Engineering Seattle, Washington
Noah's ARK Group Undergraduate Researcher **September 2015 - present**
Working on various machine learning and natural language processing research problems.

- Currently investigating the extent of arbitrariness in language by creating models for discovery of non-arbitrary form meaning associations in language, known as phonesthemes.
- Previously built a model to predict Supreme Court case outcome from the text of amicus briefs and other court documents.
- Advisor: Professor Noah A. Smith

Allen Institute for Artificial Intelligence (AI2) Seattle, Washington
Research Intern **January - March 2016**
Worked on deep learning methods for reading comprehension as part of Project Aristo.

- Implemented and reproduced the results of two neural reading comprehension models (Attentive Sum Reader and the Gated Attention Reader) in the `deep_qa` library.
- Conducted transfer learning experiments to improve generalization across domains in multiple choice and span-predicting reading comprehension models.
- **Awarded an "AI3 award" for outstanding contributions to `deep_qa`, research productivity, and more.**
- Mentor: Dr. Matt Gardner

University of Washington eScience Institute Seattle, Washington
Undergraduate Researcher **September - December 2016**
Contributed to pomegranate, a library for probabilistic programming and probabilistic graphical modelling in Python.

- Writing code, usage examples, documentation, unit tests, and improving testing infrastructure.

University of Washington Computer Science and Engineering Seattle, Washington
Networks Lab Undergraduate Researcher **June - August 2015**

- Worked on wireless backscatter of ambient signals to power various embedded devices with the use of batteries.
  - Wired and programmed a low-power temperature sensor for use with wireless backscatter.
- Advisor: Professor Shyam Gollakota

Massachusetts Institute of Technology CSAIL Cambridge, Massachusetts
Research Intern **June 2013 - September 2014**

- Worked on the START natural language question answering system and its relational knowledge database backend, Omnibase.
  - Wrote Scheme scripts to dynamically extract JPL and U.S. Election Atlas data
- Advisor: Professor Boris Katz

Professional Experience

scikit-learn
Google Summer of Code Developer **May - August 2016**
Contributed to the popular Python machine learning library scikit-learn (18k+ Github stars) by implementing various enhancements to decision tree module (adding new impurity splitting criterion and tree pre-pruning).

Selected Personal / Open Source Projects

paraphrase-id-tensorflow
Creator (100+ Github stars) **March 2017 - Present**

Implemented three paraphrase identification models from the literature in TensorFlow. Adhered to TensorFlow best practices, and the codebase is extensively documented with thorough unit tests (98% coverage + CI).

- Baseline Manhattan BiLSTM (Mueller et. al 2016)
- Siamese BiLSTM with Matching Layer (Liu et. al 2016)
- State of the Art Bilateral Multi-Perspective Matching model (Wang et. al 2017)

deep_qa
Contributor (100+ Github stars) **January 2017 - Present**

A software library for deep learning for NLP written on top of Keras and Tensorflow. Focused on enabling a workflow for quickly experimenting with and developing models for question answering and reading comprehension tasks.

scikit-learn
Contributor (18.6k+ Github stars) **November 2015 - Present**

Popular machine learning toolkit for Python. I contribute Python and Cython patches to the project, answer questions on the project mailing list and issue tracker, and review contributions from other developers.

Honors and Awards

Mary Gates Research Scholarship, 2016
SciPy Scholarship, 2016
Best Use of Microsoft Technology / Best Use of Data, DubHacks 2015
Winner, Hack The Dot Seattle 2015
Finalist, AngelHack Seattle 2015
Winner, Lockheed Martin Code Quest Sunnyvale 2015
Winner, Dave Wittry Memorial Programming Competition 2015
Winner, California State University Los Angeles ProgFest 2015
Winner, Improving MIT Award, MIT BitComp 2014

Leadership

Machines Who Learn
President **2016 - Present**
Officer **2015 - 2016**

- Organizing and leading a machine learning and data science student organization at the University of Washington.
- Giving weekly talks about various topics of interest to club members of all skill levels. Previous topics include subjects like sequence-to-sequence recurrent neural networks and the Python language and its tools for data science.
- Helping other students gain real-world experience with data science projects by collaboratively working on machine learning competitions and other data science/machine learning-related projects.

Technical Skills

- Languages / Tools: Python, Cython, Java, git, Unix shell scripts