Sarah Wiegreffe

Atlanta, GA
Email: swiegreffe6@gatech.edu
Website: sarahwie.github.io

EDUCATION

Georgia Institute of Technology (Georgia Tech) **August 2017 - present**
PhD in Computer Science
Passed qualifying exams: Spring 2019

Research interest in natural language processing and machine learning with applications in interpretability & explainability, bias & fairness, and healthcare.

Honors College at the College of Charleston **August 2013 - May 2017**
Bachelor of Science in Data Science, Summa Cum Laude
Minors in Mathematics and International Studies

University of Tartu, Estonia **January - June 2015**
Visiting Student, Faculty of Mathematics and Computer Science

PUBLICATIONS

Wiegreffe, S. and Pinter, Y. (equal contribution). *Attention is not not Explanation*. EMNLP 2019. Long paper. **Oral presentation.**

Wiegreffe, S., Flores, G., Choi, E., and Dai, A. *Learning Bi-Directional Clinical Event Representations: a Comparison of Architectures*. Preprint. 2019.

Wiegreffe, S., Choi, E., Yan, S., Sun, J. and Eisenstein, J. *Clinical Concept Extraction for Document-Level Coding*. ACL BioNLP Workshop 2019. Long paper.

Mullenbach, J., **Wiegreffe, S.**, Duke, J., Sun, J. and Eisenstein, J. *Explainable Prediction of Medical Codes from Clinical Text*. NAACL 2018. Long paper. **Oral Presentation.**

Wiegreffe, S., Anderson, P. and Obeid, J. *Can Classifications of Publications by Translational Categories be Automated?*. American Medical Informatics Association (AMIA) Joint Summits on Translational Science 2017.

RESEARCH

Research Intern, Google AI Health (formerly Medical Brain) **May - August 2019**

Worked with Dr. Edward Choi and Gerardo Flores on the Medical Records team to improve on outcome prediction for clinical time-series using unsupervised pretraining techniques.

Computational Linguistics Lab at Georgia Tech **August 2017 - present**

Projects include tying convex optimization to word embeddings to learn under hierarchical constraints as well as domain-knowledge grounding for end-to-end learning of effective representations of clinical text via deep learning.

Research and Development Intern, Sutter Heath **May 2018 - December 2018**

Worked with the Research, Development, and Dissemination group and Professor Jimeng Sun (Georgia Tech) to develop deep learning methodology for disease prediction from clinical text.

Anderson Lab at the College of Charleston **January 2016 - May 2017**

Researched extensions to Google's Word2Vec algorithm used to generate word embeddings for variable-length documents. Investigated performance of the algorithm when used directly as a classifier, and whether this technique, along with similarly created ensemble methods, could outperform benchmark preprocessing and machine learning pipelines on topic recognition tasks.

PROFICIENCIES

Daily Use	**Languages:** Python (**Pytorch/Tensorflow**/sklearn/pandas/nltk/numpy/multiproc), Bash. **Tools:** Git, TeX.
Past Use	**Languages:** R, Java, SQL, SAS, Octave. Fluency in French. **Tools:** Dynet, Oracle RDBMS, MongoDB, Tableau.

REVIEWING

EMNLP 2019
ACL 2018 (subreviewer), 2019
NeurIPS Machine Learning for Healthcare Workshop 2017, 2018, 2019
AMIA Informatics Summit 2018, 2019
Student Volunteer EMNLP 2019, ACM FAT* 2019, NAACL 2018

AWARDS

School of Interactive Computing Travel Grant (2019). Georgia Tech. Funded EMNLP attendance.

Graduate Student Government Association Travel Grant (2019). Georgia Tech. Funded EMNLP attendance.

NeurIPS Travel Grant (2019). Funded attendance.

NeurIPS Women in Machine Learning Workshop Travel Grant (2019). Funded attendance. Invited to present poster.

Computing Research Association (CRA-W) Grad Cohort for Women Attendee (2018, declined 2019). Funded attendance.

Phi Kappa Phi Graduate Fellowship (2017). $5,000.

Data Science Major of the Year, Departmental Honors (2017). College of Charleston.

Grace Hopper Scholar, the Anita Borg Institute (2016). Funded attendance. Grace Hopper Celebration of Women in Computing Attendee (2015, 2016).

William Aiken Fellow (2013-2017). A fellowship representing the top 1% of students at the College of Charleston. $92,000.

Crosby Computer Science Award (2014). Awarded by professorial nomination to the most promising student in an introductory computer science course at the College of Charleston.

TEACHING AND SERVICE

Graduate Teaching Assistant, Georgia Institute of Technology

Machine Learning (advanced undergraduate). ~100 students. **Spring 2019**
Deep Learning (graduate). ~200 students. **Fall 2019**
Guest Lecture: Transformers and Natural Language Applications.

Graduate Research Assistant, Georgia Institute of Technology **Fall 2017 - Fall 2019**

Teaching Assistant, Charleston Digital Corridor **Fall 2014**
MongoDB database course for adult professionals.

Women in Computing Club at the College of Charleston **2014 - 2017**
President (2017), Vice-president (2016) and Treasurer (2015-16). Worked to promote diversity in computer science, host professional development workshops, and conduct community outreach.