

# Shang Gao

Phone: 470 301 7095

Email: emailshang@gmail.com

Github: <https://github.com/iamshang1/Projects>

=====

## **Oak Ridge National Laboratory – Graduate Research Assistant, Oak Ridge, TN Jan 2017 – Present**

- Develop and deploy deep learning architectures for automated information extraction from cancer pathology reports; techniques include CNN, RNN, and Transformer-based approaches
- Developed new state-of-the-art text classification model for cancer pathology reports based on neural self-attention that achieves better accuracy and trains 10x faster than the previous state-of-the-art method
- Develop and apply multi-task, semi-supervised, and uncertainty quantification methods for deep learning models
- Scale deep learning algorithms across multiple GPUs and nodes on Oak Ridge supercomputer clusters

## **University of Georgia – Graduate Research Assistant, Athens GA Aug 2016 – Dec 2016**

- Work with interdisciplinary team on human activity recognition project that attempts to classify activity type based on hip-worn accelerometer device
- Developed convolutional-LSTM model that achieves competitive performance on human activity recognition tasks without requiring manual engineering of features

## **Noble Systems – Technical Writer, Atlanta GA Mar 2012 – Jul 2016**

- Produce customer-facing online training for a wide range of contact center products, including campaign management software, IVR scripting interfaces, and more
- Maintain and develop structure, templates, procedures, and single-sourcing guidelines for internal, VAR, and customer knowledge bases—content includes product technical specifications, client connectivity information, troubleshooting and configuration guides, database reference tables, and best practices
- Troubleshoot all technical problems related to internal and customer knowledge bases, including issues with HTML/CSS formatting, Team Foundation Server version control, and nightly auto-build and publishing process

=====

## **Ph.D. Data Science – University of Tennessee, Knoxville Aug 2017 – Dec 2019**

- Research focus: deep learning for natural language processing

## **B.S. Economics – Duke University Aug 2005 – May 2009**

- Major in Economics, minor in Film, second minor in Markets and Management

=====

## **Notable Publications and Awards**

- 2019 Gao S, et. al. *Classifying Cancer Pathology Reports with Hierarchical Self-Attention Networks*. Accepted for publication in Artificial Intelligence for Medicine.
- 2019 Patent for Live Call Debugging and Monitoring Tool for an Interactive Voice Response Unit (10212283)
- 2018 Gao S, et. al. *Hierarchical Convolutional Attention Networks for Text Classification*. Representation Learning for Natural Language Processing, Proceedings of ACL 2018.
- 2018 Bredesen Center Entrepreneurship Award
- 2018 Fall NSF I-Corp South Regional Cohort Alumni
- 2017 Gao S, et. al. *Hierarchical attention networks for information extraction from cancer pathology reports*. Journal of the American Medical Informatics Association.
- 2017 Smoky Mountain Data Challenge Most Novel Solution
- 2016 Patent for Utilizing Predictive Models to Improve Predictive Dialer Pacing Capabilities (US9723144B1)

## **Programming Languages and Skills**

- I am experienced with Python, Numpy, Pandas, SciKitLearn, TensorFlow, PyTorch, and PySpark
- I am comfortable working in Linux, SQL, Java, R, Matlab, and MPI
- I can speak conversational Chinese