

# Jason Wei

jason.weng.wei@gmail.com    jasonwei20.github.io

Experience	<b>Google AI Language</b> <i>AI Resident</i> New York, New York	Oct. 2020 - Present
Education	<b>Dartmouth College</b> <i>Bachelor of Arts, Computer Science. GPA: 3.91/4.00</i> Thesis Advisor: Lorenzo Torresani	2016 - 2020
Selected Papers	<b>EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks</b> <b>Jason Wei</b> , Kai Zou <i>Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP) 2019</i>  <b>What Are People Asking about COVID-19? A Question Classification Dataset</b> Jerry Wei, Chengyu Huang, Soroush Vosoughi, <b>Jason Wei</b> <i>Workshop on NLP for COVID-19 at ACL (NLP-COVID) 2020</i>  <b>Label Noise Reduction Without Assumptions</b> <b>Jason Wei</b> <i>Undergraduate Thesis 2020</i>  <b>Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathology Slides</b> [Medical Imaging] <b>Jason Wei</b> , Arief Suriawinata, Louis Vaickus, Bing Ren, Xiaoying Liu, Mikhail Lisovsky, Naofumi Tomita, Behnaz Abdollahi, Adam Kim, Dale Snover, John Baron, Elizabeth Barry, Saeed Hassanpour <i>JAMA Network Open 2020</i>  <b>Pathologist-Level Classification of Histologic Patterns on Resected Lung Adenocarcinoma Slides with Deep Neural Networks</b> <b>Jason Wei</b> , Laura Tafe, Yevgeniy Linnik, Louis Vaickus, Naofumi Tomita, Saeed Hassanpour <i>Scientific Reports 2019 (#3 most downloaded of 1024 cancer papers in Scientific Reports in 2019)</i>  <b>Automated Celiac Disease Detection on Duodenal Biopsy Slides: A Deep Learning Approach</b> <b>Jason Wei</b> , Jerry Wei, Christopher Jackson, Bing Ren, Arief Suriawinata, Saeed Hassanpour <i>Journal of Pathology Informatics 2019</i>	
Code	<b>EDA: Easy Data Augmentation for NLP</b> Github, 800+ stars, 100+ forks, <a href="https://github.com/jasonwei20/eda_nlp">https://github.com/jasonwei20/eda_nlp</a>  <b>DeepSlide: histopathology image classification</b> Github, 200+ stars, 50+ forks, <a href="https://github.com/BMIRDS/deepslide">https://github.com/BMIRDS/deepslide</a>	
Honors	Phi Beta Kappa, Tau Beta Pi, Gamma Sigma Alpha Neukom Prize for Outstanding Undergraduate Research NSF GRF Honorable Mention Barry Goldwater Scholarship (\$7,500 award) Neukom Computational Research Scholar Rufus Choate Scholar (top 5% of class) Intel Science Talent Search Semifinalist	2020    2019 2018  2016
Service	Teaching Assistant, Dartmouth College: Machine Learning and Statistical Data Analysis (Spring 2018) Reviewer: EACL (2020), NeurIPS ML4H Workshop (2020, 2019)	

### ***Text Data Augmentation***

#### **Curriculum Data Augmentation for Highly Multiclass Text Classification**

**Jason Wei**, Chengyu Huang, Soroush Vosoughi, Yu Cheng, Shiqi Xu

*Under Review*

#### **Don't Forget Where You Come From: A Multi-Task View of Data Augmentation**

**Jason Wei**, Chengyu Huang, Soroush Vosoughi

*Under Review*

#### **EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks**

**Jason Wei**, Kai Zou

*Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP) 2019*

### ***NLP for Social Science***

#### **Anxiety in the Digital Era: Linguistic Complexity Correlates of Mental Health in Text-Based Therapy**

**Jason Wei**, Kelly Finn, Emma Templeton, Thalia Wheatley, Soroush Vosoughi

*Under Review*

#### **Mitigating Political Bias in Language Models through Reinforced Calibration**

Ruibo Liu, Chenyan Jia, **Jason Wei**, Guangxuan Xu, Lili Wang, Soroush Vosoughi

*AAAI Conference on Artificial Intelligence (AAAI) 2021*

#### **What Are People Asking about COVID-19? A Question Classification Dataset**

Jerry Wei, Chengyu Huang, Soroush Vosoughi, **Jason Wei**

*Workshop on NLP for COVID-19 at ACL (NLP-COVID) 2020*

#### **Narrative Origin Classification of Israeli-Palestinian Texts**

**Jason Wei**, Eugene Santos Jr.

*Florida Artificial Intelligence Research Society Conference (FLAIRS) 2020*

### ***Deep Learning for Histopathology Image Analysis***

#### **Learn like a Pathologist: Curriculum Learning by Annotator Agreement for Histopathology Image Classification**

Jerry Wei, Arief Suriawinata, Bing Ren, Xiaoying Liu, Mikhail Lisovsky, Louis Vaickus, Charles Brown, Michael Baker, Mustafa Nasir-Moin, Naofumi Tomita, Lorenzo Torresani, **Jason Wei**, Saeed Hassanpour

*Winter Conference on Applications of Computer Vision (WACV) 2021*

#### **Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathology Slides**

**Jason Wei**, Arief Suriawinata, Louis Vaickus, Bing Ren, Xiaoying Liu, Mikhail Lisovsky, Naofumi Tomita, Behnaz Abdollahi, Adam Kim, Dale Snover, John Baron, Elizabeth Barry, Saeed Hassanpour

*JAMA Network Open 2020 (Neukom Prize for Outstanding Undergraduate Research, 3rd place)*

#### **Generative Image Translation for Data Augmentation in Colorectal Histopathology Images**

Jerry Wei, Arief Suriawinata, Louis Vaickus, Bing Ren, Xiaoying Liu, **Jason Wei**, Saeed Hassanpour

*Machine Learning for Health Workshop at NeurIPS (ML4H) 2019 [Spotlight]*

#### **Pathologist-Level Classification of Histologic Patterns on Resected Lung Adenocarcinoma Slides with Deep Neural Networks**

**Jason Wei**, Laura Tafe, Yevgeniy Linnik, Louis Vaickus, Naofumi Tomita, Saeed Hassanpour

*Scientific Reports 2019 (#3 most downloaded of 1024 cancer papers in Scientific Reports in 2019)*

#### **Automated Celiac Disease Detection on Duodenal Biopsy Slides: A Deep Learning Approach**

**Jason Wei**, Jerry Wei, Christopher Jackson, Bing Ren, Arief Suriawinata, Saeed Hassanpour

*Journal of Pathology Informatics 2019*