Anthony Chen

anthony.chen@uci.edu
anthonywchen.github.io

Education

2018-Present **PhD in Computer Science**, *University of California, Irvine*,
Advised by Sameer Singh.

2016-2018 **M.S. in Computer Science**, *University of California, Irvine*.

2012-2016 **B.S. in Computer Science**, *University of California, Davis*.

Publications

2020 **MOCHA: A Dataset for Training & Evaluating Generative Reading Comprehension Metrics**,
Anthony Chen, Gabriel Stanovsky, Sameer Singh, and Matt Gardner.
Empirical Methods in Natural Language Processing (EMNLP)

2019 **Evaluating Question Answering Evaluation**,
Anthony Chen, Gabriel Stanovsky, Sameer Singh, and Matt Gardner.
Machine Reading for Question Answering Workshop @ EMNLP Best Paper

Industry Experience

Summer 2020 **Machine Learning Research Intern**, *Apple*, Cupertino, CA,
Worked with Dr. Xiao Ling, Shayne Longpre, and Pallavi Gudipati.

- Worked on the Siri query understanding team, developing an evaluation set to assess the robustness of the entity linking system on tail entities.

Summer 2018 **Data Scientist Intern**, *Ancestry*, San Francisco, CA.

- Developed feature extraction pipelines in Java and ranking models in Python for historical records recommendation in family trees, leading to a significant gain in ranking performance.

Summer 2017 **Data Scientist Intern**, *Allstate*, Menlo Park, CA.

- Developed feature extractors and machine learning models on billion row datasets towards predicting customer defection, leveraging PyTorch, XGboost, Hadoop, and Spark.
- Provided actionable steps in which customer retention could be improved.

2014 **Software Engineering and DevOps Intern**, *Intel*, Folsom, CA.

Projects

2017 **Improving Sequence to Sequence Video Captioning**,
Approaches to automatic video captioning are limited by dataset sizes. We attempt to circumvent this by leveraging captioned images to bolster the training set as well as fusing a language model into the decoder for more syntactically correct captions. We show that using a language model provides a drastic improvement of the generated video captions.

2017 **Speech Modeling for Parkinson's Detection**.

2015 **Feature Learning of fMRI Data via Deep Autoencoders**.

Skills

ML Frameworks PyTorch, AllenNLP

Storage Frameworks Hadoop, SQL

Languages Python, Java, C/C++