# Chenwei Wu

chenweiwu@g.harvard.edu | (310) 733-0864 | [github.com/ChenweWu](github.com/ChenweWu)

## EDUCATION

**Harvard University, Cambridge, MA** *Expected May 2023*
- *M.S. in Data Science,* **GPA** *4.00/4.00* — *OPT 3 years*

**University of Rochester, Rochester, NY** *May 2020*
- *B.S. in Data Science (Highest Distinction), B.A. in Financial Economics (Distinction), Minor in Japanese,* **GPA** *3.92/4.00*
- TA: CS Department Tutor; Intermediate Statistical Methods TA; Data Mining TA; Globalization Economics TA
- Honors and Awards: Phi Beta Kappa; Dean's List for all semesters

## WORK EXPERIENCE

**Credit Suisse** **New York, NY**
*Technology Analyst in Credit Data Analytics* *July 2020 - June 2021*
*Summer Analyst in Core Technology Data Analytics* *June 2019 - August 2019*
- Served as developer & business analyst for Investment Banking Division, building Airflow automated data ETL pipelines and constructing a centralized Azure cloud data platform for bonds and credit default swaps.
- Served as project manager for the PoC of a firm-wide chat platform that leverages NLP to assist sales & trading team to a competitive edge.
- Collected big data streams and performed deep learning time series predictions on stock trends.

**PwC Strategy&** **Shanghai, China**
*Winter PTA* *Dec 2018 - Mar 2019*
- Organized surveys on Beijing's business and investing environment and performed data preprocessing for the survey to provide the Government with a better view of the Beijing market and help policy making.

**Commonwealth of Massachusetts| Operational Services Division** **Boston, MA**
*Data Analyst* *June 2018 - Aug 2018*
- Analyzed certification and accounting data via R and Excel VBA to identify opportunities and make strategies for the recruitment, engagement, and retention of minorities, women, LGBTQ, physically disadvantaged and veteran vendors. Actively attended off-line events and make calls to promote the government help programs and re-establish connections with disadvantaged vendors.
- Utilized BI reporting tool (Sugar) for data gathering, analysis and visualization; Researched and utilized queries via MS Access to quantify relevant data within the Commonwealth's Data Warehouse.

## RESEARCH EXPERIENCE

**Rajpurkar Lab | Harvard Medical School, Dr. Pranav Rajpurkar** **Cambridge, MA**
*Research Assistant* *Nov 2021 - Present*
- Develop diversity viewmaker networks, which are generative models with stochastic boundaries for data augmentations, via Pytorch Lightning, to adversarially auto learn and generate augmentations on 12-lead electrocardiogram (ECG) sensor data for self-supervised learning tasks, so as to reduce the rigorous trial and error by human experts.
- Develop self-distillation with no labels algorithms for 12-lead ECG data using Convolutional Neural Networks and Vision Transformers.
- Investigate and compare the performance of viewmaker networks to those of other previous contrastive methods, in particular whether viewmaker networks learned views that are medically sensible, and whether they are more robust to corruptions commonly observed in ECG data collection settings.

**Learning, Information &Technology Lab | Harvard School of Education, Dr. Bertrand Schneider** **Cambridge, MA**
*Research Assistant* *Sep 2021 - Present*
- Develop real time 3D gaze detection (Gaze360) and facial recognition algorithms via Pytorch for Harvard Makerspace, reconstruct gaze predictions in a 3D space, and integrate the tracking system into the Multimodal learning analytics cloud data pipeline.
- Collect data of student behavior in Makerspace lab sessions by leveraging the Multimodal learning analytics pipeline, and perform analysis to understand social learning aspects like collaboration and student attention.

**University of Rochester Medical Center | Surgery, Cancer Control, Dr. Huiwen Xu** **Rochester, NY**
*Research Assistant* *May 2019 - Aug 2021*
- Utilized unsupervised machine learning algorithms to cluster nursing homes based on the percentage of residents with dementia, depression, and serious mental illness, and detected previously unknown patterns of resident case-mix and staffing in nursing homes. [1] Predicted deficiency scores of nursing homes on a longitudinal basis with supervised learning algorithms.
- Developed text association and pattern mining algorithms to classify cancer therapies. Combined synthetic minority oversampling Technique (SMOTE) with supervised learning techniques to deal with unbalanced caregiver datasets and help clinicians identify potential mental and physical health risk factors for caregivers of the elder people. Implemented Local Interpretable Model-Agnostic Explanations (LIME) to give a non-black-box explanation for ML results in a clinical setting.

**Social Cognition & Psychopathology Lab** | UR Psychology, Dr. David Dodell-Feder         **Rochester, NY**
*Research Assistant*         *April 2019 - Present*
• Performed sentiment and topic analysis on text behavioral data in a functional neuroimaging study to better understand the relationship between brain functional connectivity and social anhedonia. [2]
• Applied machine vision techniques including semi-supervised CNNs, optical flow and open face, on video datasets to analyze group differences in nonverbal synchrony during social interactions among people with schizophrenia and controls. [3] Currently studying patterns of combinations of facial actions to evaluate the effect of oxytocin on patients' social abilities.

**Youth Risk and Resilience Lab** | UR Psychology, Dr. Catherine Glenn         **Rochester, NY**
*Research Assistant*         *Feb 2019 - May 2020*
• Processed data from anagram tests among suicidal adolescents to examine level of anxiety and cognitive flexibility.
• Applied text processing techniques via Python to extract patient information from electronic medical record notes of adolescents admitted for suicidal crises across a variety of care units at University of Rochester Medical Center.
• Developed scripts in Python to consolidate inclusion/exclusion criteria data across multiple reviewers for a set of 1700 empirical articles examining self-injury and suicidal thoughts and behaviors among gender minority adolescents.

**Active Perception Lab** | UR Cognitive Sciences Department         **Rochester, NY**
*Research Assistant*         *Aug 2018 - Dec 2018*
• Studied the optical components of a digital eye tracker and conducted experiments to test influence of eye movement on visual perception.
• Performed data mining via MATLAB on eye movement data and conducted experiments for data acquisition.

## CONFERENCE PAPERS & POSTER PRESENTATIONS

[1] Huiwen Xu, Orna Intrator, Zhengqing Ye, Eva Culakova, Chenwei Wu (2020).
"Effective Clustering of Nursing Homes Using Unsupervised Machine Learning Focusing on Dementia and Mental Illness" (Abstract, Academy Health poster presentation ARM 2020)
https://academyhealth.confex.com/academyhealth/2020arm/general/papers/viewonly.cgi?password=869456&username=38863
[2] Shovestul, B.J., Dudek, E., Wu, C., Dodell-Feder, D. (2019).
"Does functional connectivity within the DMN predict individual differences in social pleasure in schizophrenia?" (Abstract, Poster Presentation. Society for Research in Psychopathology Annual Meeting 2019, Buffalo, New York; Cognitive Neuroscience Society 2020, Boston, Massachusetts)
https://www.researchgate.net/publication/336316178_Does_the_DMN_Predict_Differences_in_Anticipatory_and_Consummatory_Pleasure
[3] Lin, L., Matvey, M., Bradley, E., Wu, C., Lai, J., Harshman, I., Dodell-Feder, D. & Woolley, J.D. (2020).
"Using Optical Flow to Quantify Movement Differences in Responses to Emotional Stimuli Among People with Schizophrenia and Controls". Poster presented at the annual Cognitive Neuroscience Society Meeting: Boston, MA
https://www.brainlatam.com/blog/cns-2020-cognitive-neuroscience-society-poster-session-c1-c69-1268

## LEADERSHIP AND AFFILIATIONS

**Rochester Data Science Society** | *University of Rochester*         **Rochester, NY**
*President*         *Sep 2018 - Jan 2020*
• Organized career talks and networking events involving industrial leaders from financial, health care and public sectors (https://rdssur.wixsite.com/rdss)
• Led manage journal clubs bi-weekly regarding newest technologies in machine learning and deep learning
• Initiated an interdisciplinary consortium between Data Science and Economics Department to award and present students' self-conducted researches applying big data analysis to the field of social sciences.

**Gapper International Volunteer Organization**         **Tioman Island, Malaysia**
*Volunteer*         *July 2017 - Aug 2017*
• Received 2 weeks' dive training and earned Open Water diving Permit.
• Observed and protected creatures and coral reefs, and helped reduce the level of pollution on the island.

**Changzhou Tian'ai Children Rehabilitation Center**         **Jiangsu, China**
*Volunteer Activity Leader*         *Jan 2015 - Jan 2020*
• Provide interest-based volunteer education for children diagnosed with autism.
• Organized public shows and fund-raising events to get clothes, books, and financial support from companies and local citizens for children, in addition to recruiting volunteers from local high schools and colleges.

## SKILLS

**Technical:** Python, Java, R, SQL, Pytorch, Pytorch-lightning, Sklearn, Tensorflow, Keras, OpenCV, Azure Databricks, PySpark, Scala, Airflow, Tableau, Docker, HTML, Flask, Excel, AWS, Linux
**Interests:** Powerlifting (600kg+), Diving (OW Permit), Rowing (varsity), Chinese Painting (16 yrs), Apex Legends (Predator)