

Yifan (Eva) Zhong

+1 206-612-3010 | zhongyifan.eva@gmail.com | <https://www.linkedin.com/in/yifan-eva-zhong>

EDUCATION

Carnegie Mellon University, Pittsburgh, PA

May 2021

- Bachelor of Science: Statistics and Machine Learning Major
- Minor in Computer Science

Honors: Dean's List (2018 Fall, 2019 Spring)
(GPA: 3.77/4)

Relevant Coursework:

- **Statistics and Data Analysis:** Data Mining; Advanced Data Analysis; Data Over Space and Time (F'20); Statistical Modeling
- **Computer Science and Machine Learning:** Intro to Machine Learning; Computer Systems; Parallel and Sequential Algorithms and Data Structures; Intro to Deep Learning(F'20); Artificial Intelligence(F'20)

PROFESSIONAL EXPERIENCES

Summer Technology Analyst, Credit Suisse, New York, NY

June 2019 - Aug. 2019

- Teamed with three interns to implement classification Machine Learning models to detect anomalies in financial reports and to optimize the bank's internal control processes. Presented to 50+ analysts at Credit Suisse
- Worked as a part of the product-control team and collaborated with other business units under CIO to optimize the auditing process by automating financial account classification, using natural language processing and machine learning models

Teaching Assistant, Girls Who Code Inc. (NGO), New York

June 2017 - Aug. 2017

- Taught 40 high school girls fundamental computer science topics including Python, fundamental algorithms, Web development frameworks, and JavaScript
- Supported our partnership with our corporate host Pfizer, and managed the classroom to maintain an inclusive learning environment

RESEARCH EXPERIENCES

Undergraduate Researcher, Language Technology Institute, Carnegie Mellon University

Project 1: A Deep Learning approach to multilingual low-resource language studies

April. 2020 - Present

- Apply text-to-speech models (Tacotron) and meta-learning (REPTILE) to perform language clustering on multilingual datasets of both textual and audio forms of inputs
- Develop novel techniques to generate and evaluate embeddings for low-resource languages, broadly aiming to bridge linguistics studies with deep learning
- Co-authored and submitted a research paper to ICASSP

Project 2: Robust machine learning for natural language processing

June. 2020 - Present

- Implement a novel character-based encoding method and evaluate its robustness against challenging medical data which contain misspellings and rarely-seen terminologies

Undergraduate Researcher, GroupLens, University of Minnesota

Jan. 2018 - May 2018

- Conducted data analysis (using Python, R, SQL) to evaluate different recommender systems algorithms deployed on a MovieLens, a movie website extensively used for ML research
- Authored a four-page research paper (accepted by ACM RecSys) based on analysis of algorithms and presented
- Presented published research findings to 200+ researchers at a Recommender System conference in October 2018

RELEVANT PROJECTS

Student Consultant, Corporate Capstone Project (36497), NPD Market Research Group

Jan. 2020 - May 2020

- Developed data science solutions which improved a baseline receipt classification model's recall accuracy from 77% to 98%, contributing to one of NPD's product line; methods include BERT, Mutual Information, TFIDF, sen2vec and tree-based methods
- Collaborated directly with a corporate data scientist and delivered final presentation to NPD's project stakeholders and leaderships
- Awarded first prize in the department-wide semester project showcase

Developer, Data Science Club, Carnegie Mellon University, Pittsburgh, PA

Feb. 2019 - May 2020

- Improved the performance of deep learning models (CNN & RNN) on time series data to measure sleepiness in speech, applying feature engineering such as PCA and Fourier Transformation

TECHNICAL SKILLS

- **Data Science:** R(Proficient); Scikit-Learn, PyTorch, Jupyter, Pandas (Proficient); SQL (Elementary)
- **Programming:** Python (Proficient), C(Intermediate), Java(Elementary)