KAREN ZHOU

kz265@cornell.edu • karen-zhou.com • (781) 266-7188

EDUCATION

Cornell University *Expected May 2021*
Bachelor of Science in Computer Science GPA: 3.91

RESEARCH EXPERIENCE

CS 6741 - Topics in NLP and ML Project *September - December 2020*
- Advised by Professor Yoav Artzi.
- Exploring the roles of language and cognition in the making of moral judgments, using datasets of real-life and hypothetical ethical dilemmas.

Cornell Natural Language Processing Group *September 2019 - Present*
- Advised by Professor Lillian Lee.
- Annotated corpus of oral conversation alignments with expert analysis.
- Trained a baseline Transformer-Ranker model for SCOTUS Justice identification.
- Tracked US Supreme Court reporting activity in news streams.

Cornell Computational Physiology Lab *February 2018 - May 2019*
- Advised by Professor Thomas Cleland.
- Attended weekly journal meetings discussing neuroscience and machine learning papers.
- Created Python GUI to perform biochemical calculations and visualizations and distributed the executable application to lab.

PROFESSIONAL EXPERIENCE

Facebook, Software Engineer Intern *June - August 2020*
- Built PHP backend and React frontend (30,000+ lines of code) for a suggestions engine on Facebook Pages that surfaces attributes from connected accounts.
- Tested extensively with unit and integration test frameworks for backend and Jest tests for frontend.

Cisco, Data Science Intern *May - August 2019*
- Performed ETL and analytics work using SQL and PySpark on 200 million rows of Cisco.com session activity data for integration with causal inference pipeline.
- Processed unlabeled text data and built Named Entity Recognition model to extract top Cisco security product names with 95% accuracy.

Sanofi, Data Science Intern *June - August 2018*
- Parallelized the execution of a disease prediction algorithm in R on a Computer Cluster to optimize target/disease prioritization.
- Developed scripts to automate verification of predictions for 1000 genes using various pharmaceutical data sources.

PROJECTS

ROC Story Cloze Task *December 2019*
Final Project for *CS 4740: Natural Language Processing*
- Implemented BiLSTM RNN and fine-tuned BERT model for reading comprehension task, achieving over 70% baseline accuracy.

Tweet Classifier *December 2019*
Final Project for *CS 4780: Machine Learning for Intelligent Systems*
- Created classifier and engineered features to identify whether tweets were written by President Trump or his staff, which achieved 6th place for accuracy out of over 100 teams.

TEACHING EXPERIENCE

Teaching Assistant *September 2018 - December 2020*
Held weekly office hours (in-person and remote), graded course work, developed solutions to problem sets and exams.
- CS 4740: Natural Language Processing (Fall 2020)
- CS 4850: Mathematical Foundations for the Information Age (Spring 2020)
- CS 2800: Discrete Structures (Fall 2018, Spring 2019)

Academic Excellence Workshop Facilitator *August 2018 - December 2019*
Created lectures and interactive worksheets/activities each week to help students understand course topics.
- CS 1110: Intro to Python (Fall 2018, Fall 2019)

PROFESSIONAL SERVICE, AFFILIATIONS, AND HONORS

Tau Beta Pi *2020 - Present*
Cornell chapter of the engineering honor society recognizing distinguished scholarship and exemplary character as students in engineering.

Alpha Phi Omega - Gamma Chapter *2018 - Present*
Cornell chapter of the co-ed national service fraternity founded on the principles of Leadership, Friendship and Service.

Cornell Tradition *2017 - Present*
Fellowship awarded to Cornell students who demonstrate significant work experience, a commitment to campus and/or community service, and academic achievement.

SKILLS

Programming Languages: Python, Java, R, Bash, SQL/PySpark, OCaml, C/C++, HTML & CSS, PHP (Hack), React.

Libraries: PyTorch, Pandas, Scikit-learn, NumPy.

Tools: Git, LaTeX, Jupyter Notebook. Eclipse. RStudio. Cluster Computing.

Languages: Spanish (Proficient), Mandarin (Advanced).