

# Shichang Zhang

825 Weyburn Terrace, Los Angeles, CA 90024 · Tel: +1(626) 325-4675 · Email: [shichang@cs.ucla.edu](mailto:shichang@cs.ucla.edu)

## EDUCATION

UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 09/2019 – Present

### Ph.D. in Computer Science

STANFORD UNIVERSITY, STANFORD, CA 09/2017 – 04/2019

### M.S. in Statistics, GPA 3.95

UNIVERSITY OF CALIFORNIA, BERKELEY, CA 08/2015 – 05/2017

**B.A. in Statistics, GPA 3.92**, Honors in Statistics, High Distinction in General Scholarship

## EXPERIENCE

**Data Scientist Intern**, WeWork, Palo Alto, CA 06/2019 – 08/2019

- Implemented a data pipeline in SQL and Python for querying and cleaning raw data, feature engineering, and KNN feature generation.
- Trained an XGBoost model with 2 million data points with 30 features to predict occupancy rate for WeWork buildings and achieved 0.093 MAE on test set.
- Applied data augmentation and prediction calibration techniques to the unbalanced dataset and drew insightful price elasticity.
- Presented my project as a selected outstanding project to over 40 employees of the Research and Applied Science team including the VP. [Poster](#)

**Research Assistant**, Stanford University, Stanford, CA 09/2018 – 12/2018

- Wrote R programs for automatically constructing bootstrap confidence intervals (BCI).
- Performed BCI analysis to evaluate the relationship between initial diagnoses and survival rate of organ transplant patients.

## RELEVANT PROJECTS

**Machine Reading Comprehension on SQuAD 2.0**, Stanford, CA 01/2019 – 03/2019

- Implemented an end-to-end neural network question answering model on SQuAD 2.0 using PyTorch.
- Incorporated the Hierarchical Attention Fusions to the baseline model to combine representations from multi-level granularity and boost model performance.
- Fine-tuned the model to achieve competitive 66.6 F1 score and 63.7 EM score on the SQuAD 2.0 test set without pretraining (BERT/ELMo).

**Predicting Conference Paper Acceptance**, Stanford, CA 09/2018 – 12/2018

- Designed and extracted vector representations of research papers including word embedding of paper abstracts, number of math equations, number of references, and others.
- Trained an RBF-kernelized SVM model to increase the paper acceptance prediction accuracy by 7% compared to the logistic baseline.
- Fitted the data with neural network models, selected the best topological structures and meta-parameters through grid search and on validation set.

## SKILLS

Python (Pandas, XGBoost, PyTorch), SQL, C++, C, Java, R, Jupyter Notebook, LaTeX