

# WENHAO ZHANG

725 Weyburn Terrace; APT 055 ◇ Los Angeles, CA 90024 ◇ +1 (213) · 806 · 0186 ◇ wenhaoz@ucla.edu

## EDUCATION

---

University of California, Los Angeles • Ph.D. in Computer Science	2017-now
University of Southern California • M.S. in Computer Science	2016-2017
University of Southern California • M.S. in Electrical Engineering	2013-2015
Harbin Engineering University • B.S. in Electrical Engineering	2009-2013

## PUBLICATIONS AND PROFESSIONAL CONFERENCE

---

<b>Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning</b> <i>The Web Conference 2020</i>	submitted
<b>GenSample: A Genetic Algorithm for Oversampling in Imbalanced Datasets</b> <i>arXiv</i>	preprint
<b>WOTBoost: Weighted Oversampling Technique in Boosting for imbalanced learning[C]</b> <i>IEEE BigData 2019 Special session: 5th Special Session on Intelligent Data Mining</i>	accepted
<b>Combination of Indoor Localization and Wearable Sensor-Based Physical Activity Recognition to Assess Older Patients Undergoing Subacute rehabilitation: Baseline Study Results[J]</b> <i>Journal of Medical Internet Research</i>	July, 2019
<b>Using Smart Watch Sensing in At-Risk Populations (SARP) in a Sub-Acute Rehabilitation Center[A]</b> <i>Archives of Physical Medicine and Rehabilitation</i>	Dec. 2018.

## INTERNSHIP

---

<b>Research Intern @ Alibaba Group</b> <i>Highlights: Conversion rate estimation, Causal inference, Multi-task learning, Selection bias</i>	July, 2019 - Sep, 2019
<ul style="list-style-type: none"><li>Identified the <i>selection bias</i> and <i>data sparsity</i> issues in conventional conversion rate (CVR) estimation</li><li>Proposed two theoretically unbiased CVR estimators, i.e., Multi-IPW, and Multi-DR, which solves these issues from a causal perspective.</li><li>Evaluated the proposed models on a public dataset and a production dataset (with 10 Billion data samples), and the results reveal that the proposed method outperform the state-of-the-art CVR models.</li><li>Drafted paper “Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning”, and submitted it to <i>The Web Conference 2020</i>.</li><li>Submitted a pre-print version: “<a href="https://arxiv.org/pdf/1910.09337.pdf">https://arxiv.org/pdf/1910.09337.pdf</a>”</li></ul>	

## SKILL HIGHLIGHTS

---

<b>Development Languages</b>	Python(Proficient), Java(Proficient), C, C++, OCaml, Scheme, Prolog, SQL, JavaScript
<b>Development Platform</b>	Google Cloud Platform, Amazon Web Service, Tensorflow, Pytorch, Hadoop
<b>Tools</b>	Emacs, Vim, Matlab, Eclipse, Android Studio, Linux, Node.js, Git, Unix, Visio

## OPEN-SOURCE CONTRIBUTION

---

<b>Contributions to Scikit-learn</b> <i>Highlights: Python, open-source contribution, model selection, Scikit-learn</i>	June, 2018 - July, 2018
<ul style="list-style-type: none"><li>Solved the compatibility issue with python 3.7.0b5 in version 0.19.2 (Merged pull request #11256)</li><li>Added a new interface in model selection module (sklearn.model.selection) in version 0.21.0. This feature adds more flexibility in identifying the best estimator. (Merged pull request #11354)</li></ul>	

- Published a dictionary app that sits in CLI environment, <https://pypi.org/project/wkdict/>

## TEACHING ASSISTANT SERVICE

---

### TA services at University of California, Los Angeles (UCLA)

Course “**Programming Languages**”(CS131) with Prof. Paul Eggert in Spring, 2019

Course “**Programming Languages**”(CS131) with Prof. Paul Eggert in Winter, 2019

Course “**Intro to Algorithms and Complexity**”(CS 180) with Prof. Majid Sarrafzadeh in Fall, 2018

### TA service at University of Southern California (USC)

Course “**Internet and Cloud Computing**”(EE 542) with Prof. Kai Hwang in Summer, 2017

Course “**Wireless Internet and Pervasive Computing**”(EE 532) with Prof. Kai Hwang in Spring, 2017

## SELECTED RESEARCHES & PROJECTS

---

### Large-scale causal approaches to debiasing post-click conversion rate estimation Jul. - Oct., 2019

*Highlights: CVR estimation, selection bias, causal inference, tensorflow*

- Identified the *selection bias* and *data sparsity* issues in conventional conversion rate (CVR) estimation
- Proposed two theoretically unbiased CVR estimators, i.e., Multi-IPW, and Multi-DR, which solves these issues from a causal perspective.
- Evaluated the proposed models on a public dataset and a production dataset (with 10 Billion data samples), and the results reveal that the proposed method outperform the state-of-the-art CVR models.

### Data Analytic in Sensing at Risk Population (SARP) Project

Oct., 2017 - now

*Highlights: Data Analytic, Machine Learning, Data Visualization, Python, R*

- Conducted a baseline analysis of combining indoor localization and wearable sensor-based physical activity recognition to assess older patients in Berkeley East rehab.
- Conducted a longitudinal analysis to understand the improvement pattern of the geriatric population with sensor-based physical recognition and clinical records.

### Machine learning with imbalanced data

Apr. -Aug., 2018

*Highlights: Ensemble learning, SMOTE, oversampling, undersampling*

- Proposed an ensemble learning algorithm with a combination of oversampling and undersampling technique for learning from imbalanced dataset.
- Tested the proposed algorithm on 18 imbalanced datasets, and compared the classification results with other well-known algorithms.

### Speaker Recognition for Dialogue system

Oct, 2016 - Nov, 2016

*Highlights: Natural Language Processing, Machine Learning, Multi Layer Perceptron Classifier, Speaker Recognition*

- Goal: Identify the speaker from the dialogue text in TV drama transcript corpus
- Preprocessed the corpus of “The Big Bang Theory” transcripts by removing stop words, lemmatization, splitting each episode into sequence of scenes.
- From the preprocessed dataset, extracted several useful features such as, Bag of words, POS tags, and so on.
- Created input feature vector for Multi Layer Perceptron classifier.
- Trained the classifier and tuned the parameters to improve the performance.

## MAJOR AWARDS AND HONORS

---

Outstanding Students of Harbin Engineering University 2012/09/19

Zhongji Social Scholarship by Zhongji Company 2012/09/19

Sino-Pacific Social Scholarship 2011/09/29

Outstanding Volunteer in Harbin Engineering University 2011/04/06

1st-Level scholarship of Harbin Engineering University 2011/03/18

