WENHAO ZHANG

3320 Sawtelle Blvd ◇ Los Angeles, CA 90066 ◇ +1 (213) · 806 · 0186 ◇ wenhaoz@ucla.edu

EDUCATION

University of California, Los Angeles • Ph.D. in Computer Science *2017-2023*

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*

SELECTED PUBLICATIONS AND PROFESSIONAL CONFERENCE

Temporal convolutional networks and data rebalancing for clinical length of stay and mortality prediction
Scientific Reports Dec. 2022

Range of Motion Sensors for Monitoring Recovery of Total Knee Arthroplasty(Best paper honorable mention)
The 17^{th} *IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks* Sep. 2022

ECG Heartbeat classification using deep transfer learning with Convolutional Neural Network and STFT technique [C]
The 4^{th} *International Conference on Computing and Data Science* **(Best paper award)** July 2022

Causal Inference in medicine and in health policy, a summary [Book chapter]
second edition of the handbook of computational intelligence *Aug. 2022*

Physical Activity Behavior of Patients at a Skilled Nursing Facility: Longitudinal Cohort Study [J]
Journal of Medical Internet Research May 2022

The derivation of an ICD-10-based trauma-related mortality model utilizing machine learning [J]
Journal of trauma and acute care surgery *Sep. 2021*

Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning[C]
The Web Conference 2020 *Apr. 2020*

GenSample: A Genetic Algorithm for Oversampling in Imbalanced Datasets
arXiv *preprint*

WOTBoost: Weighted Oversampling Technique in Boosting for imbalanced learning[C]
IEEE BigData 2019 Special session: 5th Special Session on Intelligent Data Mining *Dec. 2019*

Combination of Indoor Localization and Wearable Sensor-Based Physical Activity Recognition to Assess Older Patients Undergoing Subacute rehabilitation: Baseline Study Results[J]
Journal of Medical Internet Research *July, 2019*

Using Smart Watch Sensing in At-Risk Populations (SARP) in a Sub-Acute Rehabilitation Center[A]
Archives of Physical Medicine and Rehabilitation *Dec. 2018.*

INTERNSHIP

Research Intern @ Yahoo Research Center *July, 2020 - Sep, 2020*
Highlights: Recommendation, Rein-enforcement Learning, Causal Inference

· Researched the cold-start recommendation issue in a dynamic online environment using counterfactual reasoning and contextual multi-armed bandit.
· Implement the Causal Thompson Sampling algorithm.

Research Intern @ Alibaba Group *July, 2019 - Sep, 2019*
Highlights: Conversion rate estimation, Causal inference, Multi-task learning, Selection bias

- 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.
- Drafted paper "Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning".
- Paper submission has been published as a short paper to *The Web Conference 2020*
- Pre-print version: "https://arxiv.org/pdf/1910.09337.pdf"

SKILL HIGHLIGHTS

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

OPEN-SOURCE CONTRIBUTION

Contributions to Scikit-learn *June, 2018 - July, 2018*
Highlights: Python, open-source contribution, model selection, Scikit-learn

- Solved the compatibility issue with python 3.7.0b5 in version 0.19.2 (Merged pull request #11256)
- 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 requestion #11354)

Contributions to wkdict *Jan, 2019 - Feb, 2019*
Highlights: Python, open-source contribution, translation tool

- Published a dictionary app that sits in CLI enviroment, *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

Risk Stratification Model for Faculty Practice Group at UCLA *Jan. 2019 - Present*
Highlights: Machine learning, Healthcare data analytic, Risk analysis

- This model aims to support care management activities by identifying risk of future hospitalization and Emergency Dept. visits in individual ambulatory patients.
- The model utilizes machine learning methods along with the information of patient characteristics/demographics, economic indicator, prior utilization/exposure, test results, medical conditions to predict risk of hospital admission or ED visit in individual patients over next year.
- This model is also later being deployed and integrated in the hospitals on most University California Campuses .

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.

MAJOR AWARDS AND HONORS

Best paper honorable mention at 2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN) *2022/09/27*

Best paper award at The 4^{th} International Conference on Computing and Data Science *2022/07/16*

Outstanding Students of Harbin Engineering University *2012/09/19*

Zhongji Social Scholarship by Zhongji Company *2012/09/19*

Sino-Pacific Social Shcolarship 2011/09/29

Outstanding Volunteer in Harbin Engineering Universitys 2011/04/06

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