Zhehao (Kenny) Zhang

Personal Website: zhehao-zhang.org

Email : zhehaoz@uw.edu

Mobile : +1-805-837-9616

Education

- **Univeristy of Washington, Seattle** — Seattle, WA
  PhD Student in Statistics (Machine Learning and Big Data Track) — *Sept. 2020 – Current*
- **University of California, Santa Barbara** — Santa Barbara, CA
  B.S. in Mathematics; B.S. in Statistics; GPA: 3.94 (Highest Honor) — *Sept. 2016 – June. 2020*

Publication

Z. Zhang, C. Segal, B. T. Karras, D. Revere, G. Zane, J. G. Baseman. "Early Epidemiological Evidence of Public Health Value of WA Notify, a Smartphone-based Exposure Notification Tool: Modeling COVID-19 Cases Averted in Washington State". Link.

Z. Zhang, N. LaPierre, B. Hill, C. Cinelli. "PySensemakr: Sensitivity Analysis Tools for Regression Models in Python ". Link.

Industrial Experience

- **Google Cloud Platform** — Sunnyvale, CA
  Software Engineer (PhD) Intern — *Jun.2022 - Sept.2022*
- **Comcast Applied AI Team** — Washington, DC
  Graduate Research Intern — *Jun.2021 - Sept .2021*

Research Experience

- **Causal Inference and Sensitivity Analysis** — Washington, WA
  Research Assistant in Department of Statistics — *Sept.2021 - Present*
  - **Research Topics**: Supervised by Professor Carlos Cinelli and Professor Thomas Richardson. Studies estimating causal effects from generic machine learning tools and quantify sensitivity analysis from OLS regression.
  - **Others**: Currently working on developing sensitivity analysis tool PySensemakr package in Python. Github.
- **WA Notify Data Analysis and Evaluation Team; University of Washington** — Washington, WA
  Research Assistant with Department of Public Health — *Feb.2021 - Jun.2021*
  - **Modeling and Analysis**: Worked with Washington State Department of Health, Apple and Google to evaluate the effectiveness of Bluetooth notification technology to eliminate the transmission of COVID-19.
  - **Others**: Large-scale data manipulation; Time series analysis; Privacy and fairness of data.
- **Toronto Western Hospital (Fujitsu); University of Toronto** — Toronto, ON
  Summer IMS Researcher (with Dr. Mojgan Hodaie's group) — *Apr.2020 - Sept 2020*
  - **Automatic Brain Tumor Segmentation**: Use a U-net based deep learning architecture to perform image semantic segmentation on brain MRI images with a focus on Trigeminal Neuralgia.
  - **Others**: Gamma Knife optimization algorithm based on sphere packing. Multi-Modality MRI image generation by conditional General Adversarial Networks.
- **Statistics Department, UCSB** — Santa Barbara, CA
  Researcher, Senior Thesis (with Prof. Alex Shkolnik) — *Jan.2019 - Apr.2020*
  - **Optimal James-Stein Shrinkage for Regression**: Develop a new James-Stein type estimator for cross-sectional ordinary least square regression with asymptotic optimization guarantee on dispersion bias. Provide theoretical guarantees and numerical experiments. Thesis.
- **Fields Institute for Mathematical Science** — Toronto, ON
  Summer REU Researcher (with Prof. Andreas Hilfinger) — *June.2019 - Sept.2019*

- **Inverse Problem for Stochastic Models**: Inferred rate functions for complex stochastic models in biological processes. Used supercomputer to simulate large-scale continuous time Markov Chain and developed algorithms for solving linear network topology models based on Hill functions. Preprint.
- **Others**: Presented at Pacific Math Alliance Conference and Undergraduate Mathematics Symposium.

- **Vitality Group** Santa Barbara, CA; Chicago, IL
  Data Analyst (with Prof. Ian Duncan and Dr. Xiyue Liao) *Sept.2018 - Apr.2019*
  - **Multi-Year Longitudinal Diabetes Analysis**: Used Vitality Group dataset to find factors of changing from pre-diabetes to healthy. Used R language to implement elastic net, T-SNE, General Additive Model. Report.
  - **Others**: Graduate level classes PSTAT 296AB. Presented at URCA and InsurTech Summit.

Projects

- **Option Trading Visualization App**: Write a Python Dash App for option trading visualization. Link.
- **Recommender Systems Against Shilling Attacks**: Evaluated the Robustness of Collaborative Filtering Recommender Systems against Shilling Attacks. Implemented 4 algorithms and more than 20 attacks. Link.
- **Hull Tactical ERP prediction contest**: Investigated stock return prediction using Long Short Term Memory (LSTM) models. Won most creative category with $1000 in contest. Link.
- **Time Series Analysis (PSTAT 274)**: Built a Time series model to predict on 5-year break-even inflation rate. Link.
- **Statistical Machine Learning (PSTAT 231)**: Used the dataset of 2016 president election and US census to build a model for election prediction. Data organization, visualization. Link.
- **Statistical Learning and Computer Vision (STAT 535)**: Used the CIFAR-10 data set to build a convolutional neural network model for classification (PyTorch). Link.
- **Statistical Learning with Sparsity (STAT 538)**: A detailed study on a Variational Inference paper used for posterior sampling with implementation (Julia, Turing.jl). Link.

Teaching

Teaching Assistant: Statistics, UW
STAT 311 Elements of Statistical Methods. . . . . Sep 2020 — Mar 2021
STAT 504 Applied Regression. . . . . Jan 2022 — Mar 2022
STAT 396 Finite Markov Chains and Monte-Carlo Methods. . . . . Mar 2022 — Jun 2022
Directed Reading Program: Statistics, UW
Teach and guide project on Deep Learning and Computer Vision. . . . . Jan 2021 — Mar 2021
Teach and guide project on Causal Inference. . . . . Sep 2021 — Dec 2021

Honor and Awards

REU fellowship at Fields Institute in University of Toronto . . . . . Jun 2019
Most creative category in Hull Tactical ERP Prediction Contest . . . . . May 2019
Putnam Mathematical Competition, top 5 in UCSB . . . . . Dec 2017
College of Creative Studies Honor . . . . . Jan 2017 — Mar 2020
Dean's Honor in College of Letters and Science . . . . . Sept 2016 — Mar 2020
American Math Competition top 5% worldwide . . . . . Mar 2016

Coursework

Graduate-level: Probability Theory (A), Stochastic Calculus (A), Statistical Machine Learning (A-), Statistical Learning (STAT 535, 538: A), Machine Learning and Big Data (A-), Time Series (A), Matrix Analysis (A), Regression Methods (A), Advanced Theory for Statistical Inference (A-), Causal Inference (A)
Relevant Undergraduate Classes: Real Analysis (measure theory) (A), Linear Algebra (A+), Stochastic Process (A+), Probability Theory (A+), Bayesian Analysis (A), PDE (A), Abstract Algebra (A+), Graph Theory (A), Regression Analysis (A), Linear Programming (A).

Programming Skills and Research

- **Languages**: Python, R, Julia, C++, LaTeX **Research**: Causal Inference, Stochastic, Machine Learning