IN SON ZENG

105 S State St,
Ann Arbor, MI 48109

insonz@umich.edu
https://son520804.github.io
University of Michigan, Ann Arbor

Education

BS, Mathematics and Applications, The University of Macau Sep.2013 - Jun.2017
Minor: Music Composition and Computing Sep.2014 - Jun.2017

MS, Applied Statistics, The University of Michigan - Ann Arbor Sep.2017 - Dec.2018
Courses: Applied Bayesian Inference (BIOSTAT 668), Bayesian Modeling and Inference (STATS 551)

Faculty-reviewed Massive Open Online Course

Instructor of Applied Probabilistic Data Science in Python Specialization Jan.2021 - May.2022

- Platform: Coursera; Collaborator: Christopher Brooks
- Course 1: Fundamentals of uncertainty and Bayesian statistics Jan.2021 - May.2021
- Course 2: Probabilistic Data Analysis: Bayesian Estimation Jun.2021 - Sep.2021
- Course 3: Probabilistic Experimentation: Bandits Algorithms in Bayesian Experimentation Sep.2021 - May.2022

  Abstract: Coursera's Applied Probabilistic Data Science with Python specialization takes the advantage of the ever-increasing significance of Bayesian methods applied to data science questions in both academic and commercial scenarios. This series of courses is particularly suitable for novice programmers who want to improve programming skills, learn statistics and grow a new analytic mindset. The learning objectives of the first two courses are to explain what Bayesian thinking is and go through fundamental Bayesian estimation techniques using PyMC3. The learning objectives of the last course are to go deeper to demystify online predictions based on unknown outcomes using Bandit algorithms and Bayesian reinforcement learning techniques. Taking advantage of high-performance computing, this specialization will develop students with real-life applications of Bayesian methods in experiments.
- Video Notebook Samples: Video Notebook Samples

Research Experience

Research Fellow, School of Information, University of Michigan Mar.2020 - present

- Single-handedly wrote and recorded 80+ programming lecture videos covering **probability theory**, **Bayesian estimation techniques**, and novel **Bayesian experimentations** using Deepnote, built 8+ autograded Python assignment and structured 4-week basis courses in the **Applied Probabilistic Data Science in Python** Coursera specialization (Lecture Sample)
- Designed quasi-experiments to improve the effectiveness of instructional videos by analyzing playback speed options, video composition conditions and student's click-stream behaviors
- Predicted early student attrition for courses, student's affective responses, engagement, knowledge transfer and information retention among various student demographic groups using **mixture models** and **Bayesian Estimation Supersedes the t Test**
- Deployed BinderHub on Google Kubernetes Engine to continuously integrate UMSI educational materials and reproducible scientific analysis, allowing sharing publicly at low cost and with interactivity

Research Assistant, School of Public Health, University of Michigan Jun.2018 - Mar.2019

- Simulated, debugged and calibrated the Michigan Model of Diabetes (MMD) with 350 variables and functions in six sub-models through **agent-based modeling** in Java-based Anylogic
- Calibrated the risk factors and transition probabilities of type 2 diabetes to estimate disease progression for 30 years using **survival analysis**
- Documented the tutorial of MMD through a 50-page MMD user manual for UM medical experts to support the clinical evaluation in University of Michigan Health System and disease prevention in Public Health departments

Team Lead, NeurIPS 2020 Education Challenge Aug.2020 - Oct.2020

- Conducted learning analytics for a log data involving 1.3M+ rows containing interactions from 4.9K+ students
- Implemented **Light Gradient Boosting Machine** to rank the quality of 948 diagnostic questions, achieving 0.80 agreement fraction (rank: 5/182)

Teaching Experience

Instructional Aide, School of Information, University of Michigan Aug.2020 - Dec.2020

- Created 30+ lecture videos on Bayesian statistical thinking and data visualization in Python to Coursera to support the Master of Applied Data Science program
- Maintained the Coursera Shell and Slack channels for 250+ student classes (SIADS 505 and SIADS 521)
- Prebuilt 10+ Python data manipulation and animated visualization projects to give weekly office hours (link)

Graduate Student Instructor, University of Michigan, Ann Arbor Sep.2018 - Dec.2018

- Organized the STATS 412 courseware and summarized each lecture by writing class notes in GitHub (link)
- Clarified misconceived statistical concepts via an integration of technical and non-technical explanations (example)

Teaching Assistant

- **SIADS 505: Data Manipulation** Sep.2020 - Oct.2020
  School of Information, University of Michigan
  Instructor: Christopher Brooks
- **SIADS 521: Visual Exploration of Data** Nov.2020 - Dec.2020
  School of Information, University of Michigan
  Instructor: Christopher Brooks

Conference Tutorials

PyCon 2021 Workshop Presentation May.2021

Topic: Effective Bayesian Estimation using PyMC3 and Arviz

Honors and Awards

Meritorious Winner at 2017 COMAP Interdisciplinary Contest in Modeling Apr.2017

- Simulated scalable smart-growth city models using **quantile regression** and **agent-based modeling**

The Second Prize at China Undergraduate Mathematical Contest in Modeling Nov.2016

Technical Skills

Programming Languages: Python, R, Anylogic(Java based), SQL, Netlogo(Java based), HTML5, JavaScript, Stata, SAS, Matlab, CSS, Latex
Frameworks: React, Vue, Django, Overleaf, GKE, Git, Hadoop, Docker, Helm, Vissim
Language: English (Fluent), Cantonese (Native), Mandarin (Fluent), Portuguese (Fluent)