Hyoin An

614-619-2267 · an.355@osu.edu · hyoin-an.github.io

EDUCATION

The Ohio State University Columbus, Ohio
Ph.D. in Statistics (Dissertation Advisor: Dr. Steven MacEachern)   Aug. 2019 - May. 2024 [anticipated]

Ewha Womans University Seoul, South Korea
M.S. in Statistics (Thesis Advisor: Dr. Jae Keun Yoo)   Mar. 2017 - Feb. 2019

Ewha Womans University Seoul, South Korea
B.A in Economics & B.S. in Statistics, *Magna Cum Laude*   Mar. 2012 - Feb. 2017

INTERNSHIP/TEAMWORK EXPERIENCE

Analytics & Innovation group, United Airlines Chicago, Illinois
Intern Jun. 2022 - Aug. 2022
- Developed a prototype model for the subscription value and the unsubscription cost of marketing emails applying propensity score matching framework
- Handled and transformed big data with 13 million rows using SQL, SAS, and PySpark while running the models on Python
- Extended the model to estimation of revenue incrementality of email campaigns and evaluated the results

Consumer Behavior Prediction using Airbnb Web Log data Seoul, South Korea
Final project in Graduate Data Mining course Mar. 2018 - Aug. 2018
- Predicted binary decision-making time (fast/slow) of Airbnb customers until the first reservation, using web log data provided by Kaggle and published the results
- Created derived variables that helped improve prediction performance and used various machine learning methods such as Support vector machine, Gradient Boosting, and XGBoost on R

L Point Big Data Competition, Lotte Members Seoul, South Korea
Grand Prize (Second place) Feb. 2018
- Participated as a team of three students to provide customized contents for customers, given demographic information and purchase logs of 18 thousands of customers
- Constructed a recommender system on R which suggests items based on demographic information, purchase log, and seasonal/weather effect

RESEARCH EXPERIENCE

Pulmonary Arterial Hypertension Analysis, The Ohio State University Columbus, Ohio
Graduate Research Associate (Supervisor: Dr. Shili Lin) May. 2021 - Aug. 2021
- Predicted 1 year survival of Pulmonary Arterial Hypertension patients using genetic pathways using survival deep neural network models on R and Python
- Fit Bayesian network models using clinical variables on Yale Open Data Access platform to analyze quality of life of Pulmonary Arterial Hypertension patients

Cancer operation rate & SES Analysis, Ewha Womans University Seoul, Korea
Research Assistant (Supervisor: Dr. Won Bok Lee) Dec. 2018 - July. 2019
- Analyzed National Health Insurance data at Korean National Health Insurance Service and studied relation between operation rate of cancer and socioeconomic status
- Transformed and cleaned data using SAS, applied logistic regression, visualized results using R and ggplot2

Chronic diseases and Nutrition Analysis, Mokdong Ewha W. Univ. Hospital Seoul, Korea
Research Assistant (Supervisor: Dr. Hye Ah Lee) Apr. 2018 - Feb.2019
- Analyzed nutrition intake patterns to find linkage between nutrition intake patterns and chronic diseases using cohort data from Korean Genome and Epidemiology Study

Sufficient Dimension Reduction, Ewha Womans University   Seoul, Korea
Research Assistant (Supervisor: Dr. Jae Keun Yoo)   Mar. 2017 - Aug. 2018
- Studied sufficient dimension reduction(SDR) methodologies, focused on drawback of sliced average variance estimation(SAVE), and presented fused approach to SAVE
- Researched SDR literature, conducted numerical studies on dimension test, won bronze prize in poster competition at 2017 fall conference of the Korean Statistical Society, and published paper based on work

PUBLICATIONS

Lee, H. A., **An, H.**, & Lee, E. (2020) Dietary patterns related to cardiovascular disease based on reduced rank regression analysis of healthy middle-aged Koreans: data from the community-based Korean Genome and Epidemiology Study (KoGES) cohort, *The American Journal of Clinical Nutrition*, Volume 111, Issue 6, Pages 1159-1169.

Lee, H. A. & **An, H.**, (2020) The Effect of High Carbohydrate-to-fat Intake Ratios on Hypo-HDL-cholesterolemia Risk and HDL-cholesterol Levels over a 12-year Follow-up, *Scientific Reports*, 10, 913.

Lee, H. A., **An, H.**, & Park, H. (2019) Identification of Dietary Patterns Related to Metabolic Diseases and Their Association with Cardiovascular Disease: From the Korean Genome and Epidemiology Study, *Nutrients*, Volume 11, Issue 10, 2434.

An, H., Choi, Y., Oh, R. E., & Song, J. W. (2019) Consumer Behavior Prediction using Airbnb Web Log data, *The Korean Journal of Applied Statistics*, Volume 32, Issue 3, Pages 391-404.

An, H., Won, S., & Yoo., J. K. (2017). Fused Sliced Average Variance Estimation, *Journal of the Korean Statistical Society*, Volume 46, Issue 4, Pages 623-628.

HONORS AND AWARDS

Travel Award, *13th International Conference on Bayesian Nonparametrics*	October 2022
University Fellowship, *The Ohio State University, Columbus, Ohio*	Fall 2019 - Spring 2020
Lubrizol Start-up Fellowship, *The Lubrizol Foundation, Wickliffe, Ohio*	August 2019
Brain Korea 21 Plus Fellowship, *Korean Ministry of Education*	Fall 2017 - Fall 2018
4th Place in Poster Competition, *Korean Statistical Society*	Fall 2018
Grand Prize in 4th L. Point Big Data Competition, *Lotte Members, Seoul, Korea*	February 2018
Bronze Prize in Poster Competition, *Korean Statistical Society*	Fall 2017
Admission Scholarship, *Ewha Womans University*	Spring, Fall 2017
Dean's List, *Ewha Womans University, Seoul, Korea*	Fall 2012 - Spring 2014, Fall 2015, Fall 2016
Korea WEST Fellowship, *Korean National Institute for International Education*	Spring 2016
Global Internship Scholarship, *Ewha Womans University, Seoul, Korea*	Spring 2016
Creative Social Science Academy Scholarship, *Ewha Womans University*	Fall 2015
Placement & Work Experience Scholarship, *Ewha Womans University*	Summer 2015
International Exchange Scholarship, *Mirae Asset Park Hyeon Joo Foundation, Seoul, Korea*	Spring 2015

TEACHING EXPERIENCE

Graduate Teaching Assistant, The Ohio State University, Columbus, Ohio   Aug. 2020 - Present
- Statistical Inference for Data Analytics (Spring 2022): Hold three in-person computer lab sessions on weekly basis
- Statistics for the Business sciences (Fall 2021): Led and assisted four online recitation sessions on weekly basis for about 80 students in each session
- Introduction to Statistical Analysis I (Fall 2020 - Spring 2021): Held two online recitation sessions on weekly basis for about 40 students in each session

SKILLS

Software: Python, R, Rshiny, SQL, SAS   **Language**: Korean (Native), English (Fluent)