JONGMIN MUN

jongminm@usc.edu / jong-min.org / (+1) 984 317 2288

PhD candidate specializing in Machine Learning, Optimization, and Causal Inference. 5+ years of experience developing algorithms for large-scale data problems, dynamic pricing, and policy learning. Passionate about leveraging mathematical programming (MIP/SDP) and reinforcement learning to solve complex problems.

TECHNICAL SKILLS

- **Core Competencies:** Causal Inference, A/B Testing, Experimental Design, Revenue Management, Optimization (MIP/SDP, Gurobi), Differential Privacy, Heterogeneous Treatment Effects (HTE).
- **Programming:** Python (PyTorch, Pandas, NumPy), R (Tidyverse, Caret), SQL, PySpark, Bash.
- **Packages Built: privateAB** (pytorch; pypi.org/project/privateAB/), **FShybridPLS** (R; github.com/Jong-Min-Moon/FShybridPLS), **spkmeans** (matlab;github.com/Jong-Min-Moon/sparse_kmeans)

RESEARCH EXPERIENCE

Offline Dynamic Pricing & Policy Learning under Distribution Shift (First author, NeurIPS MLxOR workshop) openreview.net/forum?id=ZL748l6oQG
- Designed **off-policy transfer learning** algorithm for **continuous treatment allocation (pricing)**, directly applicable to revenue management and offer targeting.
- Addressed the challenge of **market shift** (covariate shift) and privacy constraints by implementing a twofold pessimistic estimation of counterfactual response

Iterative Exploration-Driven Sparse SDP Clustering via Thompson Sampling (In prepration for Journal of American Statistical Association) arxiv.org/abs/2505.20478
- Developed **mixed integer semidefinite programming**-based sparse **clustering** framework and solved it via **reinforcement learning** to escape local optima and improve robustness in high-dimensional settings
- Implemented reproducible, object-oriented MATLAB package with rigorous benchmarking

Privacy-Preserving A/B Testing (first author, Journal of Machine Learning Research and package `privateAB`) jmlr.org/papers/v26/24-2016.html
- Developed a **private A/B testing** algorithm using Google's RAPPOR mechanism to enhance the trustworthiness of **federated analytics**, achieving **2x higher testing power** compared to prior approaches

Heterogeneous Treatment Effect Estimation & Subgroup Discovery (co-first author, NeuroImage) doi: 10.1016/j.neuroimage.2024.120956
- Designed cluster-aware generative oversampling **causal inference** method to address severe class imbalance, improving estimation of **heterogeneous treatment effects** in high-dimensional data
- Applied machine learning and statistical modeling to quantify subgroup effects and uncover interpretable patterns
- Identified **4x more brain regions** exhibiting significant heterogeneous treatment effects than conventional methods

WORK IN PROGRESS

ML-Accelerated Mixed-Integer Optimization for Large-Scale Time Series
Project for USC ISE-617: Machine Learning for/with Mixed-Integer Optimization; [paper link]
- Formulated a Mixed-Integer Quadratic Programming (MIQP) framework to achieve globally optimal exact L0-penalized inference, outperforming traditional L1 heuristic relaxations for complex, noisy time-series data
- Currently developing edspecialized mathematical reformulations and advanced branch-and-bound techniques, targeting computational scalability to process massive sequences of 50,000+ data points efficiently
- Developing "learning-to-optimize" architecture utilizing Graph Neural Networks (GNNs) trained on large-scale databases to significantly accelerate exact MIP solver times.

WORK EXPERIENCE

Applied Scientist, **Center for Army Analysis and Simulations**, Gyeryong, South Korea   2020.01 – 2021.07
Project: **Modeling wildfire risk based on Army artillery training data**
- Engineered a data pipeline using MS SQL to integrate massive, disparate databases via complex spatial-temporal variables, a technique highly relevant to geographic routing and supply chain network modeling
- Applied cluster-aware **generative oversampling** to mitigate severe class imbalance in rare-event prediction models
- Published two first-authored papers on methodology (**Journal of Classification**; doi: 10.1007/s00357-024-09467-1) and theory (**Computational Statistics & Data Analysis**; doi:10.1016/j.csda.2024.108078)

EDUCATION

Ph.D. in Data Sciences and Operations, University of Southern California, Los Angeles, CA   2023.08 – 2027.05
M.A. and B.A. in Statistics, Yonsei University, Seoul, South Korea   2014.03 – 2023.02