# KYONG JU LEE

Chicago, IL lkj6113@gmail.com (872) 262-2197 LinkedIn/KyongJu

## EDUCATION

**The University of Chicago**, Chicago, IL *Sep. 2023 - Mar. 2025*
*Master of Science in Statistics*
- Courses: Generalized Linear Models, Applied Bayesian Modeling, Machine Learning, Applied Analysis
- 25% Merit-based Scholarships

**Yonsei University**, Seoul, Republic of Korea *Mar. 2017 - Feb. 2023*
*Bachelor of Arts in Applied Statistics*
- Courses: Deep Learning, Data Mining, Generalized Mixed Models, Exploratory Data Analysis, Sampling Theory

## SKILLS

**Technical:** Python (NumPy, Pandas, Scikit-Learn, TensorFlow, Matplotlib, PyTorch, Scanpy), R, Rcpp, SQL, Tableau, SAS, Slurm, LaTeX
**Languages:** English (Fluent), Korean (Native), Chinese (Bilingual), Japanese (Fluent)

## WORK EXPERIENCE

**BayeSoft Inc.**, Chicago, IL *May. 2025 - Present*
Intern - Marketing Team
- Wrote blog articles to introduce Bayesian adaptive clinical trial methods and published on the company website and social media account.

## RESEARCH EXPERIENCE

**The University of Chicago**
Precision Dose-Finding Design for Phase I Oncology Trials || *R, RJAGS, Slurm* *Jun. 2024 - Present*
- Integrated individual pharmacology data in establishing patient-level toxicity probabilities under a Bayesian framework.
- Proposed precision medicine-based dose-finding algorithm, and compared MTD estimation and safety in patient assignment against Continual Reassessment Method (CRM).

**The University of Chicago**
Doublet Detection in Single-Cell RNA Seq via Noisy PU Learning || *Python (PyTorch), R* *Apr. 2024 - Mar. 2025*
- Proposed new method for synthetic doublet generation by incorporating random coefficient, adapted Noisy Positive-Unlabeled learning for classification.
- Compared AUPRC socres with established methods across 16 benchmarking datasets and outperformed 4 existing methods.

## PROJECT EXPERIENCE

**Impact of Abatacept Treatment on FOXP3 Gene Expression in T Cells** || *R (ggplot2, glmmTMB)* *Winter 2024*
- Analyzed gene expression data to assess the effects of Abatacept treatment across different cell types using Poisson regression models.
- Suggested the best model by minimizing AIC and BIC values using a mixed zero-inflated negative binomial model.

**Housing Values in Suburbs of Boston Modeling** || *R* *Winter 2024*
- Evaluated the impact of various socio-economic factors on median values of owner-occupied homes.
- Enhanced model accuracy by applying gamma generalized linear model and conducting both-direction stepwise variable selection methods to identify significant predictors.

**Waste Classification** || *Python (TensorFlow)* *Spring 2021*
- Achieved 83.72% classification accuracy by building a CNN model to classify waste images as recyclable or non-recyclable, evaluated and compared different models by varying parameters.

**Estimation of Cases of Rotavirus Disease** || *Python (Keras)* *Spring 2021*
- Improved MSE accuracy by 21% by optimizing LSTM models with varying input window sizes for time-series forecasting of reported cases.