

YUE (Lynnc) LIANG

liangyue291@outlook.com | www.linkedin.com/in/yue-lynn-liang | (206) 822-1022 | Saint Paul, MN 55108

EDUCATION

- UNIVERSITY OF MINNESOTA, Twin Cities, Minneapolis, MN** 2019 – 2023
PhD in Data Science and Health Informatics, Institute for Health Informatics, Fully Funded. GPA 3.9
Master of Science in Data Science, Department of Computer Science & Engineering. GPA 3.9
Coursework: Advanced Machine Learning, Advanced Causal Inference, Bayesian Data Analysis, Data Mining, Database, Artificial Intelligence, Deep Learning Theory, Advanced Optimization.
- DUKE UNIVERSITY, School of Medicine, Durham, NC** 2017 - 2019
Master of Science in Biostatistics, Partial Tuition Scholarship. GPA 3.7
Coursework: Statistics Theory and Method, Applied Biostatistics Methods, Survival Analysis, Statistical Programming for Big Data, Statistical Learning, Observational Studies (Causal Inference), Analysis of Longitudinal Data.
Master's Thesis (Awarded with Outstanding Data Science Master's Project)
Performance of deep learning model in modeling longitudinal Electronic Health Records (EHR)
- UNIVERSITY OF WASHINGTON, School of Art and Science, Seattle, WA** 2013 - 2017
Bachelor of Science in Mathematics & Economics. GPA 3.7

SKILLS

Proficient: R, Python (NumPy, pandas, TensorFlow, PyTorch, scikit-learn, XGBoost, CatBoost, etc.), SQL, SAS, Spark.
Familiar: Git/GitHub, Tableau, Java, MATLAB, Unix Shell

HIGHLIGHTED EXPERIENCE

- Counterfactual Prediction for Statin Users using Causal Inference Methods and Machine Learning Model** 2019 - Present
Graduate Research Assistant, University of Minnesota, Twin Cities, Minneapolis, MN
- **Integrated Inverse Probability Weighting with Machine learning for doing the counterfactual prediction for statin user LDL lab-value reduction and cardiovascular risk reduction.**
 - Performed the data cleaning for 5.5 million patient's data from OptumLabs claims data for Statin users.
 - Built up the data pipeline for missing data handling, data preprocessing, and tidy data generation with different data source.
 - Conducted deep learning model building using PyTorch for side-effect risk prediction and LDL reduction prediction.
 - Built causal based model using propensity weighting methods incorporating deep learning for counterfactual risk prediction.
 - Re-evaluated model prediction result that causal based model shows better alignment with clinical insights.
- Causal Analysis for Clinical Team Combination Effects using EHR** 2020 - 2021
Research Assistant, Arizona State University, Phoenix, AZ
- **Performed the causal inference analysis to determine the best combination of clinical teams on Patient Activation Score (PAS) and generate the ranking for the clinical team combinations.**
 - Conducted patient's cohort definition and cohort data extraction from Fairview Hospital database using SQL.
 - Performed data wrangling and data preparation for PAS and relative demographics.
 - Used the mixed effect model of ANOVA to determine PAS is statistically significant among different clinical teams when adjusting for other covariates.
 - Computed the propensity score using logistic regression and calculated the Average Treatment Effect (ATE) of PAS using overlap weights where teams are regarded as treatments.
 - Ranked the teams based on the result and determined that clinical teams including Physician Assistant achieves the highest PAS.
- Deep Learning Modeling for Longitudinal Time-to-event EHR for ICU Patients** 2018 - 2019
Research Assistant Intern, Duke University, Durham, NC
- **Utilized several machine learning models for time-to-event risk prediction to determine the best clinical history look-back window for ICU patients.**
 - Conducted complex "time to event" dynamic data cleaning for raw EHR data (around 10GB) using R.
 - Performed LASSO regularization and the basic deep learning model for predicting patients' event using TensorFlow.
 - Applied Recurrent Neural Network (RNN) model and validated the importance of the time order of EHR of patients, then determined the best length of longitudinal history to look back for ICU patients' event prediction.
 - Achieved around 10% increase in performance of time-to-event data modeling using RNN compared to traditional multilayer perceptron.
- ### PUBLICATIONS
- **Yue Liang**, Tatiana Lenskaia, Matt Loth, Prajwal Mani Pradhan, Pui ying Yew, Jennifer G. Robinson, Terrence J Adam, Peter Tonellato, Chih-Lin Chi. Patient Centered Prescription using Optimization Algorithms for Statin Users (09/2021)
 - **Yue Liang**, Zidi Xiu, Cara O'Brien, MD, Armando Bedoya, MD, Ricardo Henao, PhD, Benjamin A. Goldstein, PhD. Predictive Performance of Deep Learning Models using Longitudinal Electronic Health Records Data (03/2019).
 - Zidi Xiu, **Yue Liang**, Cara O'Brien, MD, Armando Bedoya, MD, Benjamin A. Goldstein, PhD, Ricardo Henao, PhD. Adversarial Learning in Time-to-event Prediction (03/2019).