WANMENG LIU

(678)-575-8748 | wanmengliu@gatech.edu |
www.linkedin.com/in/wanmengliu/

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY, *Atlanta, Georgia* Jan. 2021 - Expected Graduation: May 2022

Master of Science, Health Systems

Relevant courses: Regression Analysis, Computational Statistics, Simulation, Deterministic Optimization, Engineering Economy, Healthcare Design of Future, Healthcare Delivery

THE GEORGE WASHINGTON UNIVERSITY, *Washington, D.C.* Sept. 2017 - Dec. 2019

Bachelor of Arts, International Affairs (Honors: cum laude)

SKILLS

Python | R | SQL | Spark | MATLAB | Numpy | Pandas | Scikit-learn | Seaborn | Plotly | Arena | Mandarin Chinese (fluent)

PROFESSIONAL EXPERIENCE

GEORGIA INSTITUTE OF TECHNOLOGY, *Atlanta, Georgia* August 2021 - Present

Graduate Assistant, ISyE

- Extracted millions of rows of Medicaid data for 50 U.S. states with SQL for further analysis on pediatric mental health treatment pathway
- Built data manipulation and processing pipeline to handle large amounts of patient data to ensure data security and data quality
- Continuously refined data extraction process with testing and performance evaluation using Python and Pyspark
- Remedied ad hoc data extraction issues
- Mentored new team members and helped them get on board with projects
- Built large scale medical event table for further analysis
- Wrote code to deal with new data format of Medicaid data

GEORGIA INSTITUTE OF TECHNOLOGY, *Atlanta, Georgia* July 2021 - August 2021

Graduate Assistant, ISyE

- Conducted exploratory data analysis and regression analysis on state of Georgia Covid-19 data from March 2020 to June 2021
- Determined associations between socioeconomic factors and covid-19 cases, deaths, and hospitalizations with regression models and non-parametric hypothesis tests in R
- Investigated correlations between political factors and vaccination rates in each county with one-way ANOVA in R
- Developed 20+ interactive maps to show Covid-19 hotspots in 159 counties in Georgia effectively with Python Plotly Library
- Created temporal plots to compare Covid-19 cases, deaths, and hospitalizations in each county with Python Matplotlib and Seaborn
- Delivered scientific reports and presentations to communicate results of statistical modeling and their policy implications with audience from both technical and non-technical backgrounds

NATIONAL LGBT CHAMBER OF COMMERCE (NGLCC), *Washington, D.C.* June 2018 - July 2018

Intern, NGLCC Global

- Produced social media content, newsletters, invitation letters for donors, and blog posts
- Recommended over ten grant programs to fund NGLCC Global projects
- Delivered presentations on virtual reality, employee training, and social inclusion

PROJECT EXPERIENCE

GOOGLE CUSTOMER REVENUE PREDICTION *language: Python*

- Preprocessed high-dimensional Google customer data with Principal Component Analysis
- Predicted future transaction revenue with linear regression model and tree-based models (random forest and gradient boosting)
- Optimized predictive models with parameter tuning and cross-validation

STROKE PREDICTION *language: Python, R*

- Analyzed a healthcare dataset with 10 clinical features
- Built a logistic regression model with AIC model selection criterion to predict the probability of a patient having a stroke
- Conducted hypothesis tests to determine the factors that are significantly correlated with the occurrence of stroke

NBA GAME OUTCOME PREDICTION *language: Python, R*

- Analyzed NBA game data from 2004 season to 2020 season to predict competition outcomes
- Employed various predicative machine learning models inlacing random forest, logistic regression, support vector machine