Ou Stella Liang

Ph.D. Candidate in Information Science

ol54@drexel.edu New York, NY

EDUCATION

Drexel University
Ph.D. Information Science 2022
GPA: 4.0/4.0
Graduate College Doctoral Fellowship - Gold (1/4 awardees)
Upsilon Pi Epsilon Honor Society
Coursework: Data Acquisition and Preprocessing (A+), Machine Learning (A+), Data Analysis at Scale (A+), Social Network Analysis (A+), NLP with Deep Learning (A+), Deep Learning Specialization Certificate (Coursera Credential ID YH3G7VXLNV2U)

Johns Hopkins University
Master of Health Administration 2011
GPA: 3.8/4.0
Upsilon Phi Delta Honor Society

Peking University
Bachelor of Science Applied Pharmacy 2009
GPA: 3.6/4.0
AstraZeneca Pharmacoeconomics Scholarship

RESEARCH PROJECTS

Road Injury Prevention Using Naturalistic Driving Data

- Developed generalized linear mixed models to evaluate risk exposure of at-fault accidents accounting for individual random effects (R). (Publication #1)
- Applied machine learning algorithms to determine feature importance for predicting road injuries using high frequency multichannel sensor data. Corrected for data imbalance and improved model recall by 4% with grid search (scikit-learn and Tensorflow). (Publication #2)

Technology and Substance Use

- Wrote a web crawler to download online discussions and performed qualitative coding and statistical analyses to study issues related to unsupervised drug use recovery during pregnancy (Python, R). (Publication #5,6)
- Performed social network analyses and visualization to identify distinct participation patterns in peer support groups (Python, NetworkX, Gephi).
- Served as co-principal investigator of a randomized control study to evaluate the effect of online peer support on substance use recovery. Won Drexel's COVID-19 fast response grant award.

Visualize Medication Use During Pregnancy

- Iteratively queried 250 million records of commercial insurance claims data for cohort identification (SAS, bash) .
- Transformed transactional data to temporal data structure aligned with clinical treatment episodes (R).
- Developed time sequence visualization scheme to analyze episodic drug use patterns during pregnancy (R). Won the AMIA best student paper 2nd place. (Publication #7)

SUMMARY

Experienced in data wrangling, data visualization and time series pattern recognition using machine learning and statistical models.

SKILLS

TECHNICAL: Python, R, SQL, D3.js, Java
FINANCIAL: Passed the CFA Level 1 Exam (2016), Financial Modeling

EMPLOYMENT

Goldman Sachs, New York, NY
Quantitative Strategist Intern, Liquidity Risk
6/2021 to 8/2021

- Developed a data visualization application that allows user inputs and animation playback to interactively investigate the drivers of key events using D3.js (front-end), Python (data preprocessing) and Slang (data extraction). Formed understanding of the firm's liquidity risk management framework.

Merck & Co., West Point, PA
Data Science Intern, Global Vaccine Technical Operations
4/2020 to 6/2020

- Developed a DTW-KNN based clustering algorithm prototype in Python to identify abnormal vaccine production yields using multivariate (3000+ features), temporal (~80 timesteps) spectroscopy data, achieving an 80% accuracy. Proactively reached out to stakeholder departments to improve model performance with domain knowledge.

Johns Hopkins Medicine, Baltimore, MD
Manager, Financial Analysis
6/2016 to 5/2018

- Led team effort to automate departmental financial analysis pipelines with SQL macros and advanced Excel modeling.
- Managed concurrent projects with 100% on-time deliveries by anticipating rate-limiting requirements.

Project Administrator, Financial Analysis
6/2014 to 5/2016

- Built financial models and performed financial due diligence on institutional investment opportunities. Presented to senior executives.
- Sample project: valuation modelling of an international hospital joint venture, resulting in a $300M operation agreement.

Junior & Senior Analyst, Financial Analysis
6/2011 to 5/2014

- Queried and analyzed historical clinical data using SQL to predict multiyear patient volumes and revenues.
- Developed interactive Tableau dashboards with multilevel drill-down on operating room supply usage trends using Epic and SAP data. Uncovered 5% cost saving opportunity in the neurosurgical product lines.

VOLUNTEERING

IEEE International Conference on Healthcare Informatics
Tutorial Co-Chair (2021), Publication Co-Chair (2019)

7 Cups Online Therapy and Emotional Support
Trained Active Listener

Ou Stella Liang

Ph.D. Candidate in Information Science

PUBLICATIONS

Road Safety

1. **Liang OS**, Yang CC. Determining the risk of driver-at-fault events associated with common distraction types using naturalistic driving data. (accepted)
2. **Liang OS**, Yang CC. How Are Different Sources of Distraction Associated with At-Fault Accidents among Drivers of Different Age Gender Groups? (under revision)
3. Monselise M, **Liang OS,** Yang CC. Identifying important risk factors associated with vehicle injuries using driving behavior data and predictive analytics. In Proceedings of the 7th IEEE International Conference on Health Informatics. Xi'an, China, 2019.
4. Yang CC, **Liang OS**, Ontanon S, Ke W. Predictive Modeling with Vehicle Sensor Data and IoT for Injury Prevention. In Proceedings of the 4th IEEE International Conference on Collaboration and Internet Computing. Philadelphia, PA, USA, 2018.

Substance Use and Technology

5. **Liang OS**, Chen Y, Bennett D, Yang CC. Identifying self-management support needs for pregnant women with opioid misuse in online health communities: mixed-method analysis of web posts. Journal of Medical Internet Research. 20/12/2020:18296 (forthcoming/in press).
6. **Liang OS**, Richitt H, Yang CC, Bennett DS, Online Social Support Seeking among Pregnant Women Coping with Opioid Misuse, IEEE International Conference on Health Informatics 2020. **Best Poster Award.**
7. **Liang OS**, Sheffield JS, Taylor CO. Detecting Patterns of Prescription Drug Use During Pregnancy and Lactation with Visualization Techniques. AMIA Summits on Translational Science Proceedings 2019 (May 6, 2019): 478–87. **Best Student Paper, Second Place.**
8. Meacham MC, **Liang OS**, Zhao M, Yang CC, Thrul J, Ramo DE. Connectedness Based on Shared Engagement Predicts Remote Biochemically Verified Quit Status Within Smoking Cessation Treatment Groups on Facebook. Nicotine Tob Res. 2021 Jan 7;23(1):71-76.

Other Collaboration

9. Jazayeri A, **Liang OS**, Yang CC. Imputation of Missing Data in Electronic Health Records Based on Patients' Similarities. J Healthc Inform Res 4, 295–307 (2020).
10. Monselise M, Greenberg J, **Liang OS**, Pascua S, et al. An automatic approach to extending the consumer health vocabulary. Journal of Data and Information Science (2020).
11. Voong KR, **Liang OS**, Dugan P, et al. Thoracic Oncology Multidisciplinary Clinic Reduces Unnecessary Health Care Expenditure Used in the Workup of Patients With Non-small-cell Lung Cancer. Clin Lung Cancer. 2019 Jul;20(4):e430-e441.
12. Rao K, **Liang OS**, Cardamone M, et al. Cost implications of PSA screening differ by age. BMC Urol 18, 38 (2018).
13. Tekes A, Jackson E, Ogborn J, **Liang OS**, Bledsoe M, et al, How to Reduce Head Computerized Tomography Orders in Children with Hydrocephalus using the Lean Six Sigma Methodology: Experience at a Major Quaternary Care Academic Children's Center, American Journal of Neuroradiology (AJNR), 2016 Jun, 37(6), 990–996.

STANDARDIZED TESTS

GRE
2017
Verbal Reasoning 96%, Quantitative Reasoning 94%, Analytic Writing 93%

GMAT
2013
Overall 94%, Integrated Reasoning 93%, Analytical Writing 91%