

C.J. DUAN

Data Scientist. Ph.D

@ Data.Scientist@dulun.com
in cj-duan-3ab27a191

757-742-3896
hublun

DRC Lab, Virginia

www.dulun.com



EXPERIENCE

Data Scientist (contract)

Fortune 500 Company (CPG), Research Data Science

2021 - 7.2022

Remote

- Input Data (Snowflake on AWS): Time-series of weekly sales data and weekly media (advertising) spend and GRP (Impression)s
- Model: State-Space Bayesian Media Mix Model using R and RStan
- Output Results: ROI effectiveness in terms of MC (media contribution) over MS (media spend)

Chief Research Data Scientist

DRC Lab @ Dulun Research & Consulting

2016 - Now

Freelance

Completed Projects:

- (2020-2021) Algorithmic Trading Robot, coded in Python, digests recent (commodity future) prices and execute order composition and routing automatically.
FFT, Risk Control, Trading
- (2016-2017) Develop the best distribution routing model, which saved an Alabama flower-growing farmer thousands fuel dollars on monthly basis.
Transportation Scheduling, Constraint Programming, OR

Replication & Learning Projects (2016 - now):

- Analysis of 6.4 million SARS-CoV-2 genomes identifies mutations associated with fitness (Obermeyer, et. al., 2022, Science)
Bayesian multinomial logistic regression model using Pyro
- Survival analysis in Stan (Filipe S. Dias, 2022, Web Blog)
Survival Analysis
- Deep learning recommendation model for personalization and recommendation systems (Naumov, et. al. 2019, ArXiv)
TorchRec
- Bayesian Inference for a Generative Model of Transcriptome Profiles from Single-cell RNA Sequencing (Lopez, et. al., 2018, Nature Methods)
Variational Inference for Generative Model of scRNA seq using PyTorch and Pyro
- Bayesian aggregation of average data: An application in drug development (Weber, et. al., 2018, Annals of Applied Statistics)
Pharmacometric Model using RStan

Assistant Professor of Quantitative Methods

Troy University

March 2009 - May 2017

Troy, AL

Research Projects:

- Research Paper (2021): Team Contingent or Sport Native? A Bayesian Analysis of Home Field Advantage,
- Research Paper (2021): Exposing model bias in machine learning revisiting the boy who cried wolf in the context of phishing detection

WORK PHILOSOPHY

"Modeling in the sciences is in fact almost always generative modeling"

Kingma and Welling (2019)

MOST PROUD OF



Hidden Markov Model (HMM)

state space, transition, and emission



Markov Decision Process (MDP)

states, actions, rewards



Pyro & PyTorch

SVI (stochastic variational inference) based Variational Autoencoder



Multi-Arm Bandit (MAB) and RL

π policy mapping actions to contexts



Media Mix Modeling

ROAS, mROAS, Adstock, Hill function



A/B Testing and MCI

multiple causal (treatment) inference



LDA / Mixture / Clustering

Latent Dirichlet Allocation, Topic Models



Stan & Bayesian Modeling

data, parameters, model, and prediction



Servire Est Vivere

helping Braydon Farm LLC in Alabama optimize their daily flower delivery routes using Transportation Model

EDUCATION

Ph.D in Industrial Management

Clemson University

2000 - 2007

REMARKS

"... On behalf of my wife and I, we would like to sincerely thank Dr. Duan, Mrs. Amy Hu, and you for the time you spent developing a distribution plan for our new floral operation. We have incorporated it into our operation and it appears to have significantly reduced our travel time and expenses in distribution our products to the Marvin Stores chain. Hopefully this will be

PUBLICATIONS

Journal Articles

- (Chaojie), D. C. J., & Chakravarty, A. (2021). Team contingent or sport native? a bayesian analysis of home field advantage in professional soccer. *Journal of Business Analytics*, 4(1), 67–75. doi:10.1080/2573234X.2020.1854625
- (Chaojie), D. C., & Gaurav, A. (2021). Exposing model bias in machine learning revisiting the boy who cried wolf in the context of phishing detection. *Journal of Business Analytics*, 4(2), 171–178. doi:10.1080/2573234X.2021.1934128. eprint: <https://doi.org/10.1080/2573234X.2021.1934128>
- Duan, C. [C.J.], & Pierce, G. S. (2017). Starting operations management with a 'Big Bang': Using sitcom to introduce OM concepts to students. *Operations Management Education Review*, 11, 115–122.
- Duan, C. [CJ], Hu, J., Garrott, S., et al. (2016). Using excel solver to solve Braydon Farms' truck routing problem: A case study. *South Asian Journal of Management Sciences (SAJMS)*, 10(1), 38–47. doi:10.21621/sajms.2016101.04
- Duan, C. [Chaojie], Grover, V., Roberts, N., & Balakrishnan, N. (2014). Firm valuation effects of the decision to adopt relationally governed business process outsourcing arrangements. *International Journal of Production Research*, 52(15), 4673–4694. doi:10.1080/00207543.2014.884289. eprint: <https://doi.org/10.1080/00207543.2014.884289>
- Kronenburg, M. A., Shetterly, D. R., Duan, C., Krishnamoorthy, A., & Loutzenhiser, K. (2013c). The impact of contract design on contract performance satisfaction. *International Journal of Management Research and Reviews*, 3(4), 2673.
- Duan, C. [Chaojie], Grover, V., & Balakrishnan, N. (2009). Business process outsourcing: An event study on the nature of processes and firm valuation. *European Journal of Information Systems*, 18(5), 442–457. doi:10.1057/ejis.2009.38

Conference Proceedings

- Duan, C. [C.J.]. (2013). Retrieving hypothesis parameters from real life events. In *The 43rd southeast decision sciences institute conference proceedings*. SE DSI.
- Kronenburg, M. A., Shetterly, D. R., Duan, C., Krishnamoorthy, A., & Loutzenhiser, K. (2013a). Public management skills needed in contract development. In *The southeastern conference for public administration*. SECoPA.
- Kronenburg, M. A., Shetterly, D. R., Duan, C., Krishnamoorthy, A., & Loutzenhiser, K. (2013b). Public management skills needed in contract development. In *The southeastern conference for public administration*. SECoPA.
- Kronenburg, M. A., Shetterly, D. R., Duan, C., Krishnamoorthy, A., & Loutzenhiser, K. (2012). The impact of contract design on contractor performance - a second look. In *5th international public procurement conference*. IPPC.
- Duan, C. [C.J.]. (2004). Teams in TQM: An knowledge based view. In *The 35th dsi*.
- Duan, C. [C.J.]. (2001). One machine scheduling of jobs to minimize total weighted tardiness. In *The 37th southeast informs*. SE INFORMS.

an on-going study and we will be able to "tweak" or alter the system based on your recommendations as our business grows..."

Stephen (Brad) Bradshaw Garrott
Owner, Braydon Farms, 2015

"...I work on projects focused on educating high school students in Mississippi in entrepreneurship opportunities and what resources are available to start-ups. I use Excel frequently and I am thankful for being familiar with excel formula's like vlookup and IF. ... I wish there was an Operations Management class because that was my favorite in undergrad and where I find myself passionate. I am considering entering the workforce aiming for an Operations Management position based on my success in your class and how much I enjoyed it. I think your class is very important and I'm glad I had the opportunity to learn from you. I am also open to entrepreneurship opportunities as I am so active in the entrepreneurship ..."

Russell, QM 3345, 2016