Yadong Liu

+1 6468918962 | JeremyColumbiaAPAN@outlook.com | 250 W 100 St, NYC, USA 100025

EDUCATION

Columbia University — New York, USA
M.S. in Applied Analytics (STEM), GPA:3.92 — Sep 2021 - Dec 2022

Hubei University of Technology — Wuhan, China
B.S. in Electronic Information Science and Technology — Sep 2016 - June 2020
Awards: Distinction Graduate (TOP 1%), 1st Prize Scholarship, 2017/2019 (TOP 3%)

Vaasa University of Applied Sciences — Vaasa, Finland
Exchange Student in Computer Science — Jan 2019 - June 2019

EMPLOYMENT

Lianwei Data Tech — Shanghai, China
Junior Data Scientist (Strategy & Analytics concentration) — Jul 2020 - Jul 2021
Role: *Provided data-driven service to help 6 brands in P&G and Estee Lauder groups grow revenue in Alibaba E-com Platform*

- **Detective:** Using Python, SQL and Tableau to construct data dashboards from Ali Databank (10TB OLAP data engine) to track metrics and KPIs to provide actionable insights by diagnosing user cohort's portrait, traffic, trading, behavior and ad placement
- **Strategist:** Reporting by building data models of Machine Learning (regression, collaborative filtering, clustering, feature engineering, NLP) to review performance of campaign, new product launches, shopping carnival and make predictive strategy
- **Innovator:** Cooperated with the engineering team and Ali partner to develop Business Intelligence tools for internal teams to improve efficiency and for external stakeholders to better track and review their performance

Achievement: *Assisted Safeguard achieve the GMV goal of 100 million in Ali TMALL during "Double-11" Shopping Carnival; Built a SEM optimization model based on LTR and XGBoost that improved the ROI of Safeguard's new product launch by 23%; Helped Gillette achieve the highest ROI in P&G China during the "618" Grand Promotion.*

INTERNSHIP

Columbia University — New York, USA
Part-time Research Assistant (Showcase: https://columbia.shinyapps.io/youtube_jeremy/) — Dec 2021 - In progress

- Implemented functions of content-based YouTube music video recommender website with R, PostgreSQL and Rshiny.

Nielsen — Shanghai, China
Data Analyst Intern, Finance Vertical — Sep 2019 - Dec 2019

- Collaborated with programmers to deploy online questionnaires on brand performance of Union Pay in 18 countries
- Conducted data cleaning with VBA and quantitative analysis with SPSS to complete Brand Tracking reports for UnionPay
- Monitored social trends by integrating data with SQL from the Social Listening database and performed a customer sentiment analysis for Vivo with 1.75 million user-generated records in R Studio and delivered presentation in slides to manager

RESEARCH EXPERIENCES

Kaggle x Airbnb Coursework Competition — Oct 2021 - Dec 2021

- Analyze and predict the price for Airbnb rental based on 96 variables regarding its property, host, geography and past reviews.
- Monitored psychographic data exploration, cleaning, preprocessing, feature selection and model training with R
- Continually using linear regression, best-subset selection, random forest, ranger and SVM to train the model, and finally using XGBoost to obtain the optimal RMSE and scored TOP 20% ranking.

Application Research on Key Opinion Leader (KOL) Group Identification Based on Improved Particle Swarm Algorithm
Graduate Capstone (awarded Excellent Graduation Thesis) — Jan 2020 - May 2020

- Established a Weibo user influence evaluation model and key opinion leader group identification model based on the improved PSO and K-means Fusion Algorithm
- Web Crawled large-scaled records of user data that participated in the topic of "U.S. Epidemic" and conducted data tidying
- Used pandas and scikit-learn to conduct statistical test of the improved Fusion algorithm (68% higher than the ordinary K-means algorithm in DBI)

Prediction Team, Innovation Center in Hubei University of Technology — Wuhan, China
Researcher on Weather Prediction Systems — Mar 2018 - May 2019

- Designed and built a data assimilation model leveraging MM5 (weather model) by optimizing the gradient descent algorithm and particle swarm optimization (PSO) algorithm with time-varying constrict factor
- Visualized the accuracy and robustness of the model using MATLAB with massive assimilation data, improved the system accuracy by two orders of a magnitude of 10^{-11} to 10^{-13}
- Analyzed the assimilation results and improved efficiency by 54% with distributed parallel computing

Liu Y D, *Zhou M R, Xie J Y, Tong Y L. "Research on Simulation Model for Data Assimilation in Solar Radiation Prediction Based on Particle Swarm Optimizer with Time-Varying Constrict Factor" Published on Acta Energiae Solaris Sinica, 2021, Vol.42, No.4*

SKILLS

Programming and Tools: Python, R, C/C++, MATLAB, SPSS, SQL, AWS, Spark, Hadoop, Tableau, Microsoft Office
Techniques: Statistics, Machine Learning, Optimization, A/B testing, Consulting, Web Scraping, Web/APP development