JINGLIN ZHAO

New Brunswick, NJ | 747-217-5887 | jlinzhao@outlook.com | https://www.linkedin.com/in/jinglin-zhao-146b0b1ba/

SKILLS

Technologies: Python (Pandas, Numpy, Scikit-Learn, PyTorch, Keras, TensorFlow), SQL, R, SAS, SPSS, Tableau, Power BI, Excel, Hadoop, Spark, AWS
Specialties: A/B Testing, Statistical Modeling, Hypothesis Testing, Data Visualization, Machine Learning, NLP, Marketing Analytics

EDUCATION

Master of Business and Science: Analytics and Data Science, Rutgers University *Jan 2024 (Expected)*
Courses: Applied Artificial Intelligence into Market, Project Management, Data Structure and Algorithm, Regression Analysis, Database and Data Warehousing, Cloud Computing, Market Assessment, Business Intelligence with Visual Analytic
GPA: 3.9/4.0
Bachelor of Science, Psychology (Hons.), Hong Kong Baptist University *Jun 2016*

PROFESSIONAL EXPERIENCE

ECPAT USA **New Brunswick, NJ**
Data Science Extern Lead *May 2022 - Present*
- Utilized **spatial analytics** in **ArcGIS** to predict trafficking probabilities across U.S. regions, considering income levels, crime rates, communication trends,and distribution of transit stops/routes
- Collaborated with a team of 9 to apply **natural language processing (NLP)** techniques to analyze unstructured data, specifically traffickers' hashtags and emoji usage, obtained through web scraping using **Python**
- Unraveled communication trends in human trafficking and identified high-risk regions by employing effective **metrics** (economic status, the number of postings); visualized findings in a **Tableau Dashboard**
- Provided actionable insights for enhanced intervention against human trafficking during critical time frames, such as weekends and holidays, as well as in high-risk regions

CGI - New Jersey Big Data Alliance (NJBDA) **New Brunswick, NJ**
Data Science Extern Lead *Aug 2020 - Jan 2023*
- Offered **statistical consulting** services to the government with the objective of mitigating financial risk resulting from flood losses, achieving by developing precise flood prediction and warning systems
- Built an **end-to-end ETL pipeline** using **SQL** to preprocess and consolidate 3 years of daily precipitation data from two stations located near Trenton, New Jersey
- Employed **machine learning** model (**XGBoost**, LSTM) with 20+ features and statistical time series models (**SARIMA**) to forecast water levels using **Python**; visualized the flood susceptibility area in Trenton via ArcGIS
- Presented financial loss estimates associated with different levels of flood events to **executive-level stakeholders** and developed flood mitigation strategies

U-haul International, Inc **Piscataway, NJ**
Data Science Extern *Jun 2021 - Jun 2022*
- Managed 4+ years, **64M+** rows of historical order data in **SQL**, comprising daily number of orders at each U-haul store
- Conducted time-series seasonality analysis and rental **demand forecasting** using **PySpark**; employed SARIMA, XGBoost, and Facebook Prophet models on **Azure Machine Learning Studio**
- Achieved a **25%** reduction in Mean Squared Error (MSE) compared to the benchmark, resulting in significantly improved demand forecasting accuracy and facilitating more effective resource allocation

Decomil, LLC **New Brunswick, NJ**
Project Manager *Jan 2021 - Aug 2021*
- Performed market research and **competitor analysis** for an online store, focusing on sales trends and social media traffic
- Gathered customer data from online retail platforms (Amazon, eBay, Etsy, and Wayfair) through **web scraping** using **Python**
- Developed **Tableau dashboards** to offer insights and recommendations regarding website design, holiday sales plans, and marketing strategies tailored to specific vendor target groups, enhancing the understanding of weekly customer search trends
- Designed and analyzed **A/B testing** to measure the impact of product image and description changes, leading to **stakeholder** decisions to revise these elements, which improved the **conversion rate** by 20%
- Analyzed the impact of reviews on conversion and performance matrix **click-through rates** using **Propensity Score Matching** for **causal inference**, allowing for the rationalization of relevant market campaigns without the costs of online experiments

PROJECTS

Healthcare Project: Disease Detection System with Treatment Recommendation **Jun 2021 – Present**
- Oversaw R&D efforts and Employed **feature engineering** such as categorical feature encoding, bucketing numeric features on extensive disease-symptoms dataset (Heart Disease, Skin Disease, Retinal Disease)
- Reduced feature dimension by **PCA** to save computational cost along with **regularization** to prevent overfitting
- Sampled 20% data as holdout and conducted 5-fold **cross-validation** on the training-validation dataset; reached an overall accuracy of 83.94% in detecting disease by **Multinomial Naïve Bayes model**