

# **Expertise**

Tools: Python, SQL, STATA, R, Matlab, Git, Tableau, Plotly, Mode Analytics

Packages: pandas, numpy, matplotlib, scipy, statsmodels, sklearn, tensorflow, sqlalchemy

**Statistical Modeling:** Causal Inference, Experimental Design, A/B Testing, Difference-in-Differences, Propensity Score Matching, Synthetic Control Methods, Survival Analysis, Time Series Analysis, Bayesian Inference

**Machine Learning:** Linear and Logistic Regressions, Random Forests, Boosting, Support Vector Machines, K-nearest Neighbors, Clustering(K-means, Hierarchical, Kernel), Feature Engineering, Model Selection, Cross-validation

## **Education**

### Ph.D., Applied Economics

Dec 2019

Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, USA

B.S., Economics Jun 2014

Nanjing Audit University, Nanjing, China

# **Experience**

### Data Science Fellow, Insight Data Science Fellows Program

Sep 2019-Jan 2020

- Consulted with an online learning start-up company to evaluate their referral program to improve user acquisition.
- Queried over 200K+ granular data across 20+ tables in relational database using Mode Analytics and created 30+ feature metrics relevant to site performance, course property, and user actions using SQL and Python.
- Built a supervised classification model to predict successful user referrals using logistic and random forests with SMOTE and recursive feature elimination cross-validation to reduce overfitting (+85% model accuracy on test set).
- Delivered a set of suggestions to boost their user acquisition and engagement by nudging specific users to act, optimizing class formats, and providing a great initial course experience, which has been adopted by the company to upgrade their products and gained substantial revenue increase.

#### Postdoc Research Associate, Virginia Tech

Jan 2020-present

- Developed a global simulation model to assess the impacts of trade agreements, projected billions of losses to U.S.
   livestock industry due to withdrawal from CPTPP, where over 90% are recovered from U.S.-Japan trade agreement.
- Developed a survival model to estimate the impact of non-tariff measures (NTMs) on US and global agricultural
  exports using millions of product-line trade data over 20 years, which quantified a 3%-8% increase in the probability
  of failure of trade relationships owning to the presence of NTMs.
- Built a data pipeline to extract public data sources using Python and API and bulk preprocessed these documents and excel data used in statistical modeling and economic impact analysis.

#### **Graduate Researcher**, Virginia Tech

Aug 2015-Sep 2019

- Applied a pseudo-Poisson maximum likelihood model to estimate the trade import elasticities, which was cited by USDA to estimate the 2018 Trade Damage Assessment and Farm Market Facilitation Program.
- Proposed a transitional demand system to examine market shocks caused by food safety outbreaks using panel data, which detected significant drifts in consumer preferences over imported products.
- Built a demand model to assess the impact of China's trade diversification strategy on U.S. farmers and agribusinesses, which found declining competitiveness of U.S. exports against Australia, Argentina and Brazil in China.

### **Instructor**, Virginia Tech

Aug 2017-May 2019

- Designed the undergraduate level course principle of microeconomics and organized paired group activities to improve students' learning experience (full-semester).
- Lectured *math*, *statistics* and *econometrics* review sessions for first-year Ph.D. students' qualifying exams and *international trade and finance* for senior Ph.D. students' field research (multiple sessions).
- Facilitated the department workshops and seminar events to improve collaborations across research fields.