Bo Zhang, Ph.D.

Austin, TX, 78717 https://bozhangecon.weebly.com/
505.717.6780 ▪zhangbo.eco@gmail.com https://www.linkedin.com/in/bo-zhang-33188b71/

EXPERIENCE

Machine Learning Scientist II
Overstock, remotely work at Austin, TX Jun 2021 – Present

- Led and worked on the new promotion and pricing algorithm; built the demand elasticity model at product level using machine learning and deep learning models for 2 millions of products; developed optimization methods to maximize revenue and profit under budget constraints; results from A/B testing show that the revenue increased by 6% (~$0.15B annual) and profit increased by 4% (~$10M annual) with the new promotion algorithm
- Led the analysis on the return of company's marketing spending using statistical and machine learning models on both individual and aggregated data; provided data-driven strategies on maximizing the marketing spending on site sale, coupon, and advertisement; the strategies were implemented company wide and resulted in an about 5% increase in the revenue
- Productionized machine learning algorithm; wrote production code; deployed the model using Docker, Jenkins, and Airflow at both local server and cloud
- Involved in the A/B testing for new algorithm test, including the test design, method development, sample splitting, metrics selection, power calculation, and results analysis and interpretation
- Worked with professors from top tier universities on pricing and personalized promotion algorithm using statistical, machine learning, and experimental methods
- Mentored three junior machine learning scientists and data scientists

Data Scientist
Plymouth Rock Assurance, Boston, MA Oct 2019 – Jun 2021

- Initiated and led new pricing models in home insurance; conducted data cleaning and sample design on various big data sets with over 10m records; selected features from over 20k variables and conducted dimension reduction; conducted model design, algorithm selection, and parameter tuning; improved the performance by 20%; communicated with the product team on model results and implementations
- Evaluated new solar radiation data on the pricing model performance; applied regression, random forest, XGBoost, and LightGBM to predict the property loss based on millions of records using Python; improved the accuracy of the prediction by 4% and provided better prediction on the large losses
- Cleaned and organized large and unstructured quote/underwriting data (over 100millon record) using Python Dask; generated over1000 variables related to quotes, coverages and policies

PhD Researcher
Penn State University, State College, PA Aug 2014 - Aug 2018
Estimating the Effect of House Prices on Food Consumer Behaviors Using Nielsen Scanner Data

- Analyzed over 10 million households' grocery food shopping records; estimated consume demand models using R and Stata; discovered that food expenditure increases by 0.7%, but the diet quality does not change for a 1% increase in house prices

House Prices and Marriage Behaviors: A Survival Analysis

- Analyzed over 1 million individuals' marriage behaviors in China using census data; merged individual-level data with city-level economic conditions and house price data; estimated the discrete-time duration model using R and Stata and found that the hazard rate of marriage entry decreased by 0.3% for a 1% increase in house prices

SKILLS

Programming and Software: Python, SQL, Spark, R, Git, Docker, Airflow
Causal inference: difference-in-differences, propensity score matching, A/B testing, panel data analysis, causal forest, double machine learning
Machine learning: regression, time series, decision tree, random forests, gradient boosting machine, XGboost, feed-forward neural network, recurrent neural network

EDUCATION

PhD in Applied Economics
Pennsylvania State University, State College, PA Aug 2014 - Aug 2019
MA in Economics
University of New Mexico**,** Albuquerque, NM Aug 2012 - May 2014