Bo Zhang, Ph.D.

Austin, TX, 78717 https://Zhang1987.github.io/
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 –

- Led and worked on the new promotion and pricing algorithm; Built the demand and elasticity model at product level using machine learning and neural network models for millions of products; Developed optimization method to maximize revenue and profit under budget constraints; The A/B testing results show the new promotion algorithm increased the revenue by 6% (~$0.15B annual) and profit by 4% (~$10 million annual)
- Led the analysis on the return of company's marketing spending using statistical and machine learning model on both individual and aggregated data; Provided data-driven recommendations on marketing money spending on site sale, coupon, and paid search; The recommendations were implemented at company wide and resulted in a significant 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 the 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 predictive 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, R, SQL, Git, Docker, Airflow, PySpark
Causal Inference: difference-in-differences, propensity score matching, A/B testing, panel data analysis, causal forest, double machine learning
Machine Learning and Deep Learning: regression, time series, decision tree, random forests, gradient boosting machine, XGboost, MLP, RNN, CNN, Transformer

EDUCATION

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