Shuaihang Zhou

Mobile:(+86)18520767627| Email: shuaihang.zhou@gmail.com

EDUCATION

Beijing Normal University (BNU) Beijing, China
B.S in Statistics (GPA: 3.8/4.0) Sept. 2019 - July 2023 (Expected)

- Relevant Courses: Probability Theory, Advanced Algebra, Mathematical Analysis, Mathematical Statistics, Econometrics, Applied Multivariate Statistical Analysis, Applied Stochastic Process, Sampling Survey, Statistical Learning, Linear Model, Experimental Design, Applied Time Series Analysis

Awards and Honors:
The Second Prize Scholarship of Beijing Normal University, BNU, 2022&2021&2020;
Honorable Mention, Mathematical Contest in Modeling, COMAP, 2022
The First Prize Competition Scholarship of Beijing Normal University, BNU, 2021

WORK EXPERIENCE

Apollo Intelligent Driving Technology (Beijing) Co., Ltd Beijing, China
Data analyst Intern, Autonomous Driving Business Operation Department July 2022 – Sept. 2022

- Built a new online taxi order quantity prediction model based on several variables (i.e. weather conditions, day of the week, holidays, district label) using the LSTM network. The new model decreased the monthly MSE by 5 percent.
- Found the reasons why the order completion rate differs between the orders formed in different channels by analyzing the completion rate extracted using SQL query for each hour and each region.
- Made graphs and dashboards to analyze changes in key metrics on a daily, weekly, and monthly basis.

China Construction Bank Corporation Guangzhou, China
Fintech Intern, Fintech Department July 2021– Aug. 2021

- Collected relevant document files from websites by building a web crawler using Python and extracted the keywords by cutting the text into words and computing the TF-IDF of a word. The efficiency of filtering out important documents has been greatly improved.
- Analyzed the conversion rate of advertising in different marketing channels for different products and recommended possible improvement solutions which have improved the conversion rate in certain channels.

RESEARCH EXPERIENCE

Iowa Home Sales Price Forecast using Integrated Learning
Advisor: A.P. Yonghong Du, Statistics Department July 2022

- Analyzed the data set and performed Box-Cox transformation, generated new features, and selected the features.
- Built three forecasting models: a LightGBM model, a KNN regression model, and a ridge regression model.
- Compared results from each model and produced the final prediction model combining models using the stack method.

The relationship between economic indicators and air quality in China: Based on optimized robust statistical methods
Advisor: Prof. Jiao Jin, Statistics Department May 2022

- Optimized K-means clustering by estimating the covariance matrix in Mahalanobis Distance with Kendall's τ coefficient and Median absolute deviation of the sample, and selecting the initial centroids with a new method.
- Analyzed and concluded the relationship between economic indicators and air quality by applying robust canonical correlation analysis to the data. Factors that have a significant impact on air quality were evaluated by robust PCA.
- Applied the improved robust K-means clustering to the data so that further exploration and more advice can be made.

Variable selection in classification problems: Based on red wine quality data
Advisor: Prof. Gaorong Li, Statistics Department Jan. 2022

- Fitted and optimized the penalized logistic regression model in R using the red wine data. A classification tree with pruning and an SVM model was built for comparison.
- Summarized that penalized logistic regression was likely to be affected by the correlation between variables but the classification tree was sensitive to changes in individual variables by comparing the variable selection results.

A comparison of four variable selection methods in linear models: Based on simulated data
Advisor: Prof. Xinwei Tong, Statistics Department Nov. 2021

- Summarized the theory of four variable selection methods: Lasso, adaptive Lasso, best subset selection using AIC, and best subset selection using BIC. Derived and summarized the algorithm, which is based on the local quadric approximation proposed by Fan and Li(2001), for solving the Lasso and adaptive Lasso.
- Applied four methods to the data generated by a multiple linear regression given the parameters including the beta coefficients, the variance of the normal error term, and the distribution of the dependent variables.

批注 [M1]: 感觉这个好像也有点水

批注 [M2]: 原来的中文标题是：《空气质量与经济发展、产业结构的关系探究—基于稳健统计方法》
英文标题是：**Relationship between air quality, economic development and industrial structure-Based on robust statistical methods**

批注 [M3]: 这一个点不知道要不要说，感觉好像没啥实际内容

- Analyzed and graphed the variation of the correct selection rate of the four methods with the change of parameters including the variance of the normal error term, and the covariance matrix of the dependent variables.

EXTRACURRICULAR ACTIVITIES AND LEADERSHIP

Beijing Normal University Baige Youth Volunteers Association Beijing, China

Deputy of Web Department August 2020 – Sept. 2021

- Streamlined the video editing software training process for the association's media department and other clubs on campus by introducing the procedure of preparing and distributing training documents in advance.
- Planned, organized, and coordinated campus activities that were attended by more than 100 people.

批注 [M4]: 不知道这种志愿者经历要不要加进来

SKILLS AND OTHER

- Languages: Mandarin (native), English (fluent)
- Programming: Python, R, C, SQL
- Interests: Tennis, Hiking