Qinzhe Wang

Cell: (919) 214 - 2441 | Email: wqz0102@gmail.com
LinkedIn: https://www.linkedin.com/in/wangqz/ | Personal Website: https://wangqinz.github.io/

EDUCATION

DUKE UNIVERSITY, Durham, NC **Aug. 2020-May 2022**
Master of Science in **Statistical Science** (GPA: 4.0)

- **Courses**: Machine Learning, Bayesian Statistics, Natural Language Processing, Data Consulting, Hierarchical Models, Time Series Analysis, Statistical Decision Theory, Missing Data, High-dimension Data, Mathematical Analysis, APP Develop

INDIANA UNIVERSITY, BLOOMINGTON, IN **Jan. 2015-May 2019**
Triple major: Bachelor of Science in **Statistics**, Bachelor of Arts in **Mathematics** and **Economics**, Business Minor (GPA: 3.9)

- **Honors**: Founders Scholar for Sustained Academic Excellence 2018 & 2019

WORK EXPERIENCE

Quantitative Research Intern **Huatai Securities** **Shanghai, China** **Apr. 2021-Aug. 2021**

- Pre-processed 20M+ credit debt market data; investigated relationships between bond's return and empirical features including macroeconomic environment, industry status, and company's fundamentals
- Performed multi-factor analysis and visualized the results with histogram, boxplot, and other basic statistical plots
- Simulated future return of selected bonds with Monte-Carlo algorithm, combined the data support with Markowitz theory, increased the total return of bond investment (ROI) by 4% in the third quarter

Data Science Intern **NetEase** **Beijing, China** **Sept. 2020-Dec. 2020**

- Performed correlation analysis on 4M+ existing users' tag (personal information, consumption pattern, etc.) and the target personas; selected highly correlated tags based on the Distance Correlation and Spearman Correlation
- Integrated social science theory (potential social influence and homogeneity) and selected tags, trained and enhanced Logistic Regression, Support Vector Machine and Random Forest model to predict the ads clicks with the F1 score
- Designed A/B testing framework to improve the ranking of advertisements: selected matrices, determined sample size, and conducted A/A testing; increased the average conversion rate of ads by 27.4%
- Assisted in constructing a tag recommendation system to raise efficiency, mainly responsible for the Content-Based-Recommendations and the collaborative filtering

Artificial Intelligence Intern **Flickering AI** **Shenzhen, China** **June 2018-Aug. 2018**

- Conducted pixel analysis with convolution and pooling techniques, applied principal component analysis and feature selection to reduce the dimension
- Utilized Support Vector Machine and Random Forest to build artificial intelligence models for handwritten numeral recognition based on 650K+ observations; trained and cross-validated the performance of machine grading; performed root-cause analysis and adjusted parameters to improve the recognition precision to 87%

PROJECT EXPERIENCE

Football Search Engine APP (R shiny APP) **Sept. 2020-Apr. 2021**
Application designed for football fans: provide game information, display the data of teams & players with download function.

- Developed an application with R shiny APP function. Collected data through football API, presented game results, supported fuzzy inquiry for teams and players, and visualized player statistics by radar charts and tag clouds in the APP
- Applied K-means clustering for players with dynamic diagrams based on the similarity (goals scored, assist numbers, etc.)

News Popularity Project (Python) **Aug. 2020- Dec. 2020**
Project designed to help authors on Mashable predict the click rates and determine whether their articles will be popular.

- Performed exploratory data analysis on 20K+ observations with basic statistics plots (scatterplots and boxplots) and feature selection methodology to examine the most influential predictors
- Adopted multiple basic regression and machine learning methodologies like the Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine to build classification methods
- Analyzed the accuracy of each model and selected the best model for prediction using cross-validation

ADDITIONAL INFORMATION

Technical Skills: Python (NumPy, Pandas, Matplotlib, Scikit-learn), R (ggplot2, tidyverse), MySQL, LaTeX, Tableau
Languages: Mandarin (native), English (fluent)