Qinzhe Wang

(919) 214 - 2441 | Email: wqz0102@outlook.com

GitHub: www.github.com/wangqinz | Personal Website: https://wangqinz.github.io

EDUCATION

2019.08-2022.05 Duke University Master of Science Durham, NC

Statistics (STEM) GPA 4.0/4.0

Relevant 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

2014.08-2019.05 Indiana University Bachelor of Science Bloomington, IN

Triple major: Statistics, Mathematics, and Economics, with a Business Minor ***GPA 3.9/4.0***

Honors: Phi Beta Kappa Honor Society, Founders Scholar for Sustained Academic Excellence 2018 & 2019, Dean's List (recipient for 7 semesters)

Relevant Courses: Machine Learning, Statistical, Time Series Analysis, Big Data, Bayesian Statistics, Statistical Consulting, Linear algebra, Game Theory

WORK EXPERIENCE

2021.04-2021.08 Huatai Securities Quantitative Research Intern Shanghai, China

- Assisted to apply multi-factor analysis and feature selection based on over 200,000 credit debt market data; validated and enhanced statistical soundness and conceptual stability of existing high & low-frequency bond trading strategies
- Simulated different portfolio constructions with Monte-Carlo algorithm and portfolio optimization techniques to predict the future movement of returns; assessed the influence of COVID-19 by time series analysis
- investigated empirical relationships between bonds' yield and other features including macroeconomic environment, industry status, and company's non-systematic risks

2020.09-2020.12 NetEase Data Science Intern Beijing, China

- Analyzed relationship between the Cost Per Million (CPM) and users' tag (general information, consumption pattern, and searching behavior) based on the industry research; integrated social science (potential social influence & homogeneity) with statistical science (Distance Correlation & Spearman Correlation) to cluster existing customers, develop potential markets, and set up recommendation system
- Trained linear Support Vector Machine, Random Forest, Naïve Bayes, GBDT, XGBoost models based on over 4 million observations to improve the precision of advertising technology, increased the average conversion rate of video ads by 27.4%
- Identified target audiences to match the personas; participated in creating optimal strategies for over 20 advertisers

2018.06-2018.08 Flickering AI Artificial Intelligence Intern Shenzhen, China

- Identified predictors consisted of each pixel using convolution and pooling techniques and applied the principal component analysis and feature selection technique
- Utilized CNN Deep Learning Method, Support Vector Machine, and Random Forest to build the artificial intelligence models for handwritten numeral recognition based on over 60000 observations; trained and cross-validated the performance of machine grading; performed root-cause analysis and adjusted parameters to improve the recognition precision by about 19%
- Developed a WeChat application named Math UFO in conjunction with a small team of senior developers: the private beta test successfully attracted more than 1500 users. The full-scale application launched application market in 2019

PROJECT EXPERIENCE

2020.09-2021.04 Football Search Engine APP (R shiny app) Durham, NC

Application designed to serve football fans: provide game information, display the data of teams & players with download function.

- Developed an application by R shiny app function. Accessed and collected data through football API, presented game results and the future match arrangements, supported fuzzy inquiry for teams & players, 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.)

2020.08-2020.12 News Popularity Project (Python) Durham, NC

Project designed to help authors on Mashable to predict the click rates and to determine whether their articles will be popular.

- Performed exploratory data analysis with basic statistics plots (scatterplots and boxplots) and feature selection methodology in R to examine the most influential predictors
- Adopted multiple basic regression and machine learning methodologies like the Logistic Regression, Generalized Additive Model, 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: R, Python, Jupyter, MATLAB, MySQL, LaTeX, Tableau

Languages: TOEFL: 118/120, GRE 333/340