

Zuobei (Joe) Zhou

Willing to relocate | (757) 367-3934 | zzhou13@email.wm.edu | linkedin.com/in/zuobeizhou/

EDUCATION

COLLEGE OF WILLIAM AND MARY

Master of Science in Business Analytics / STEM / GPA: 4.0/4.0

Williamsburg, Virginia

August 2019 – May 2020

GUANGDONG UNIVERSITY OF FOREIGN STUDIES

Bachelor of Economics / GPA: 3.91/4.00 (Top 1)

Guangzhou, China

September 2015 – June 2019

ACHIEVEMENTS AND CERTIFICATES

- Beta Gamma Honor Society (Top 10%)
- Udacity Data Scientist Nanodegree (Data Engineering, Software Engineering, Experimental Design)

TECHNICAL SKILLS

Tools: Python, R, SQL, Spark, Tableau, Alteryx, Gurobi, AWS, Git, TensorFlow

Techniques: Predictive Modeling, Data Warehousing, Big Data, A/B Testing, Statistical Inference, Optimization

EXPERIENCE

Capstone (Major telecommunication company in the U.S.)

Williamsburg, Virginia

Data Science Intern

April 2020 – May 2020

- Collected census and geographic data through web scraping and ArcGIS, identified areas with rapid population growth in four years by tree-based models and Recurrent Neural Networks with weighted MAPE of 7.10%
- Spatially optimized future cell towers placement utilizing Gurobi with total revenue retained of \$44,900,000 based on available investment, existing cell towers localities and predicted customer demands
- Prepared developer documentation, presented actionable findings and technical methodology to non-technical audiences in an understandable way

Ipsos (Top 5 of the Market Research Sector in 2019)

Shanghai, China

Data Analyst Intern

March 2019 – June 2019

- Tracked 18 products performance worth \$1,500,000 through compiling 30+ monthly reports based on sentiment analysis of customer reviews, provided product innovation and target marketing advice
- Independently automated data aggregation ETL and ad-hoc data requests using SQL and Python to save manual work by 80%, maintained a Tableau dashboard for metrics monitoring

Danone (Food-products corporation among Fortune 500)

Guangzhou, China

Sustainability Analyst Intern

August 2018 – December 2018

- Revised data integration system by assembling multiple data sources (supplier, logistic and retailers) and automating carbon footprint computation to enhance analytics efficiency using SQL and Excel
- Minimized logistics cost by 30% through reassigning supplies distribution using linear programming via Excel Solver

DATA SCIENCE PROJECTS

Home Credit Default Risk Prediction (R)

- Joined 7 data tables with over 200 features from different sources by summarizing them at a consistent level
- Constructed creative features such as ratios and percentages based on domain knowledge, assessed client default risk using Random Forest, GBM and XGBoost with an 17 % improvement on AUC from 0.672 to 0.783

Customer Conversion Prediction (Python, Spark)

- Forecasted customer conversion using GBM with PySpark RDD and DataFrame API in Spark ML
- Formulated Sprak ML Pipelines to connect feature engineering (vector assembler/indexer), hyperparameter tuning and model evaluation

Data Warehouse for Kiva (SQL, Alteryx, Tableau)

- Enabled analysis around loan distributions by creating data warehouse based on Kiva's loan data with more than 670,000 records and 20 features from 2013-2017 using Alteryx and MySQL
- Created dashboards presenting demographic, geographic and time series analysis using Tableau