Jackson Yang

Email: zhuoxin.yang@duke.edu | **LinkedIn**: Jackson_Yang | **GitHub**: Personal_Website

EDUCATION

Duke University | Durham, NC   **Expected May 2023**

- M.S. in Business Analytics   GPA: 3.85/4.00
- Coursework: Machine Learning, Data Visualization, Data Infrastructure, Applied Statistics, Market Research, CRM

University of Washington | Seattle, WA   **March 2022**

- B.A. in Business Administration and Information Systems   GPA: 3.83/4.00
- Coursework: Data Mining, Statistical Methods, Predictive Modeling, Database Management, Business Communications

TECHNICAL SKILLS

Tools: Python (Pandas, Numpy, Sklearn), SQL (SQL Server, MySQL), R (Tidyverse), Tableau, Git, Spark, Google Analytics

Statistics Analysis: A/B Testing, Statistical Test (Z-test, T-test, F-test, Chi-square test), ANOVA, Time Series Analysis (ARIMA)

Machine Learning: Classification, Regression, Clustering, NLP, PCA, Feature Engineering, Monte Carlo Simulation

WORK EXPERIENCE

Avidian Technologies, Seattle, WA   Apr 2021 - Present
Data Scientist

- Reduced customer churn rate to a **3%** per month by creating a data pipeline to track and clean 7K + customer usage data and then training a churn rate prediction model on the cleaned data using Random Forests in **Python**
- Reached a **5%** increase in click-through rates for email marketing campaigns through the design and implementation of **A/B tests** on email elements such as the subject line, layout, personalization, and content
- Built a data processing pipeline to supply categorized lead data for email marketing campaigns by extracting 1K lead data from ZoomInfo weekly and cleaning and partitioning the data into batches using **Python**
- Facilitated faster customer information access and lowered customer response time by **30%** using **SQL** to clean, update, and segment 18K+ subscription data from 3 databases (Zuora, Stripe, and Outlook)
- Improved customer experience and increased website conversion rate to **8%** monthly by designing and creating various visualizations in 7 **Tableau** dashboards to monitor web traffic data and customer journeys
- Achieved a **15%** boost in demo sign-up rate using **Python** to wrangle 120K + lead data from 2020 to 2022 and train a lead scoring model using XG Boost to generate prime prospect lead data for targeted outreach

Duke Fuqua, Durham, NC   Aug 2022 – March 2023
Marketing Intelligence Analyst

- Gathered customer intelligence by writing queries in Brandwatch to scrape 500+ textual data from websites monthly
- Recognized customer preferences and identified trendy topics by building a text pre-processing pipeline to translate and clean textual data and then using Topic Modeling technique in **Python** to extract insights and topics
- Optimized marketing content and uncovered valuable customer insights by cleaning and classifying 2K+ online customer comments from popular media platforms with Spacy and Naive-Bayes classifier in **Python**

PROJECTS

Optimizing Email Marketing Campaign: Analysis and Click Rate Prediction | *Report , Website*

- Cleaned and wrangled 100k+ customer data in **Python** by encoding categorical data, tackling imbalanced datasets, etc.
- Uncovered click-prone behaviors and groups by performing PCA and K-Means clustering in Python to identify segments
- Developed ideas for A/B Tests using Logistic Regression, Decision Tree, Partial Dependence Plot, and RuleFit analysis
- Constructed a click probability prediction model in Python by fine-tuning logistic regression with grid search tuning

Fraud Detection At Self-Checkouts In Retail | *Report , Website*

- Prepared data for visualization and fraud prediction model building by wrangling 2k + customer check-out data in **Python**
- Reduced sampling bias by resampling (SMOTETomek) unbalanced datasets and using cost-sensitive algorithms
- Identified fraud-prone behaviors and suspicious groups by visualizing data using pair plots, PCA, and clustering techniques
- Developed a fraud detection model using XG Boost, improving fraud prevention and resulting in a cost savings of 102K