Jackson Yang

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

SKILLS

Tools: Python (Numpy, Pandas, Sklearn, Spacy), SQL (SQL Server), 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

EXPERIENCE

Avidian Technologies, Seattle, WA — Apr 2021 – May 2023

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

Purple Wave, Manhattan, KS — Mar 2023 – May 2023

Data Scientist Student Consultant

- Communicated with different stakeholders to gather data and insights in order to understand the company's sales goals
- Ensured correct decision-making for future growth by wrangling 42M user browsing data from 2018-2023, fitting an L1 logistic model to determine the North Star metric, and identifying important features for improving key metrics
- Created territory-specific growth threshold metrics by fitting a Decision Tree model to 208K transactions in **Python**
- Detected fraudulent activities by fitting an XGBoost model on transaction data in combination with k-means clustering

Duke Fuqua, Durham, NC — Aug 2022 – Mar 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**

Data On Demand, Taipei**,** Taiwan — Dec 2020 – Mar 2021

Business Intelligence Analyst Intern

- Collaborated with stakeholders to understand data requirements and prepared visualization dashboards by wrangling 24K+ insurance data from 6 tables using **SQL**, including data type transformations, data aggregation, and feature selection
- Built 4 **MicroStrategy** dashboards to evaluate loss ratio, sales, and profit margins by channel and product type
- Identified an abnormally high loss ratio for property insurance in Q3 and conducted various analysis to find the root causes

EDUCATION

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

- M.S. in **Business Analytics** — GPA: 3.82/4.00

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

- B.A. in **Information Systems and Supply Chain Management** — GPA: 3.83/4.00