Zhiyan (Kylie) Huang

(608) 358 5132   hzykelly@gmail.com   linkedin.com/in/zhiyan-kylie-huang   github.com/KylieH1314

EDUCATION

Carnegie Mellon University   Pittsburgh, PA
- Master of Information System Management – Business Intelligence and Data Analytics   Dec 2021
- Master of Statistical Practice   May 2020
- GPA: 3.9/4.0

University of Wisconsin-Madison   Madison, WI
- Bachelor of Industrial Engineering   Dec 2018
- Double majored in Applied Mathematics with Computer Science Certificate
- GPA: 3.6/4.0

SKILLS

Programming: Python (w/ framework of scikit-learn, Pandas, PyTorch, SpaCy), Java (w/FX), R, Bash (Linux)
Database & Big Data: SQL (Oracle, MySQL, PostgreSQL), Hadoop (w/ Hive, Spark)
Analytics / Machine Learning: Classification, Clustering, Experimental Design, GLM, Time Series
Applications: Tableau, Excel, Git, Visual Studio Code

PROJECT EXPERIENCE

[machine learning] Waiting Time Prediction Analysis with Yelp   Fall 2019
- Optimized the existing algorithms to better predict restaurant waiting time based on 5GB records in 2016 via Python
- Performed Data Cleaning, Feature Engineering and Exploratory Data Analysis to have a complete view of data
- Trained supervised machine learning models like Random Forest and Gradient Boosting, and applied cross validation to tune parameters and overcome overfitting
- Improved 15% waiting time prediction accuracy with a new estimation algorithm

[object oriented programming/python] Classification and Regression Tree   Fall 2019
- Designed algorithms to build, prune trees and generate random forest from scratch in Python.
- Implemented interface to support SQL and data frame as input using psycogy2 and pandas package
- Applied unit-testing, benchmarking, and profiling to evaluate computation performance.

[natural language processing] Deceptive Online Hotel Reviews Detection   Spring 2020
- Developed algorithms to classify deceptive online hotel reviews based on the labeled text data by Python
- Preprocessed data by constructing bag of words, conducting part of speech analysis and sentiment analysis
- Categorized reviews by performing K-means clustering using the keywords extracted from TF-IDF
- Improved 10% classification accuracy by using the decision trees compared with manual detection.

[python | data visualization] Aerie Brand Expansion   Fall 2020
- Provided brand expansion strategy recommendations based on the customer behavior and sales data in 2018-2020
- Trained and tested Linear Regression and Random Forest to quantify the relationship between behaviors and values
- Visualized the results from acquisition analysis and machine learning models using Tableau, matplotlib and seaborn
- Won 1st in the Aerie Brand Expansion Case Competition on campus

WORK EXPERIENCE

Horizon Research Consultancy Group   Guangzhou, China
Assistant Analyst   February - June 2019
- Collaborated with software engineers, project managers and stakeholders to improve customer satisfaction rate of a local call center by optimizing its serving process
- Designed satisfaction surveys and analyzed the results to find the bottleneck in the process
- Provided a digital transition plan which can handle up to 5% demand growth each quarter without hiring more staff.