Xuechun Wang

(734)546-6457 | hesterwang0514@gmail.com | 618 S Main St, Apt#527, Ann Arbor, MI

EDUCATION

University of Michigan, Ann Arbor, MI **Aug. 2021 – May. 2023**
M.S in Applied Statistics

Purdue University, West Lafayette, IN **Aug. 2017 – Dec. 2020**
B.S in Statistics and Mathematics, Minor in Economics

- Dean's list in Spring 2018 and Fall 2020

INTERNSHIPS

MasterCard Analytics Consulting, China, Shanghai, *part-time intern* **June. 2021 – July. 2021**

- Utilized Logistic Regression and K-means Clustering in interpreting 4-year debit/credit card promotions campaigns datasets, categorized by demographics, education, martial status and sex, for exploring and ameliorating the effectiveness of yearly/monthly/quarterly promotions
- Captured multicollinearity and heteroscedasticity in data preprocessing, and applied manipulation
- Profiled distribution of enrolling in instalment plan or full payment, and identified new classifiers in categorizing customers on social status scores
- Tuned the hyperparamters to the optimized performance, and supported the reliability of story line

RESEARCH EXPERIENCES

COVID-19 Predication and Control, West Lafayette, IN, *Research Assistant* **May. 2020 – Mar. 2021**

- Established an infectious disease modeling (SEIR model) based on ordinary differential equations to predict COVID-19 spread rate at Purdue campus for fall reopening
- Modeled contact matrix between students and school faculties based on heterogeneous patterns to simulate the possible routes of infection
- Leveraged the model with time-dependent variable to parametrize real-time infectious rates
- Validated the robustness of the model by evaluating it on 6 other universities' COVID-19 data, and reported corresponding precaution measures to school officials

Nursing Home Ranking System, West Lafayette, IN, *Research Assistant* **Sept. 2020 – Nov. 2020**

- Collected over 67 million CMS government data points with rating information by web scraping
- Developed a nursing home ranking system with customer reviews using HTML, CSS, and JavaScript
- Delivered a personalized recommendation function using Item-Based Collaborative Filtering in Python

PROJECTS

New York City Taxi Fare Prediction, West Lafayette, IN **Jan. 2020**

- Performed exploratory data analysis with summary statistics, correlation matrix, histograms, etc.
- Created a geographic map that showed the overall distribution of pick up and drop off locations
- Preprocessed the dataset with over 55 million rows by invalid data removal and missing imputation
- Conducted feature engineering and trained multiple machine learning models including Linear Regression, LASSO regression, Random Forest, XGBoost to predict taxi fare based on 32 variables
- Selected XGBoost as the final model by measuring Root Mean Square Error (RMSE) on 5-fold cross validation and fine-tuned hyperparameters through Grid Search, which achieved RMSE of $5

SKILLS

- Programming: Python (Sklearn, TensorFlow, Keras, Pytorch), SQL, R, MATLAB, C++
- Data Visualization: Python (Matplot, Seaborn), R (Shiny, ggplot2), Excel, Tableau