

Yijun Wu

Email: yijunwu2021@u.northwestern.edu

EDUCATION

Northwestern University

Evanston, IL

Master of Science in Analytics

Sep 2020 - Dec 2021(Expected)

- **Expected coursework:** Predictive Analytics, Java and Python Programming, Data Visualization, Data Mining, Deep Learning, Analytics Value Chain

University of Illinois at Urbana Champaign

Urbana, IL

Bachelor of Science in Statistics; Bachelor of Science in Civil Engineering

Aug 2016 - May 2020

- **Honors:** Dean's List (Top 20%); Wilson H. Tang CEE Risk, Reliability, and Decision Analysis Scholarship
- **Coursework:** Statistics & Probability, Applied Linear Algebra, Statistical Analysis, Matrix Theory, Differential Equations, Microeconomic Principles, Applied Statistics with R, Applied Regression and Design, Data Science Foundations, Basics of Statistical Learning, Stochastic Process

PROFESSIONAL EXPERIENCE

31 Jiu Trading (Shanghai) Co., Ltd.

Shanghai, China

Data Analyst Intern

Jul 2019 - Aug 2019

- *Collated* sales data of the soda beverage from 2017 to 2019, calculated the monthly and annual growth rate, drew the variation curve via Excel and summarized the seasonal changes in the sales volume
- *Conducted* regression prediction model for sales performance 2017-2019 through R using Holt-winters method.
- *Calculated* safety stock, offered data support to the subsequent optimization of order placement and sales program.

Shanghai Wangren Co.

Shanghai, China

Data Analyst Intern

Jun 2018 - Jul 2018

- *Proposed* supply chain optimization for this traditional Chinese food wholesaler via inventory to sales ratio and inventory resource prioritization, drafted coping strategies in advance of the possible deviation from the prediction
- *Built* data visualization for the historical and the current year sales data of each retailer via drawing corresponding curves, performed year-over-year comparison, and created rolling forecast for better supply distribution

SELECTED RESEARCH AND PROJECT EXPERIENCE

Household Income Prediction Based on Housing Features (R)

Urbana, IL

Individual Research

Feb 2020 – May 2020

- *Developed* a feasible model form that uses housing features to predict household income.
- *Fitted* a multiple regression model using data collected; checked the Box-cox plot and performed a log transformation on the response; performed stepwise variable selection based on AIC.
- *Checked* the assumptions of the model via diagnostic plots, conducted a train-test validation, and calculated the average percentage of the difference between the predicted value and the actual value, 27.8%.

Google Play Store App Rating Prediction (Python)

Urbana, IL

Team Leader

Nov 2019 - Dec 2019

- *Predicted* the rating of apps in Google play store using Kaggle data source; preprocessed data that includes missing data refilling, one-hot encoding and train:test data split.
- *Fit* four models: lasso, k-nearest neighbor, random forest and support-vector machine, cross validated each model with training data; compared and selected the model with the lowest RMSE; used the model selected and random forest model to make prediction on the test set and calculated test RMSE, 0.3085.

Regression Model for Glucose Prediction

Urbana, IL

Individual Project

Mar 2019 - Apr 2019

- *Conducted* outlier reduction Bonferroni-adjusted critical value from given *Pima* Dataset from the *Faraway* package, checked if the assumption of linearity, normality and constant variance still satisfied
- *Selected* best performing model from forward, backward, stepwise, regulation and regular models and reached mean square error at the lowest score
- *Applied* the model to the test set to check the feasibility; checked the influential point by looking at Cook's distance

SKILLS

Computer: R, Python, Matlab, Sql, Java, Autodesk Revit, Photoshop

Language: English(Fluent), Mandarin(Native)