

A 3rd year PhD student majoring in Systems and Information Engineering with interdisciplinary backgrounds in Business Administration, Industrial Engineering, and Machine Learning. My research concentrates on Health Informatics, Machine Learning in Healthcare, Causal Inference for Medical Decision-Making, and Fairness of AI Applications. I've compelling passion for teaching and have had sufficient amount of teaching experience in both business and engineering subjects. My teaching philosophy is to explain complex concepts using the simplest language.

SKILLS & COURSES TAKEN

Programming	Python, R, Julia, MySQL, InfluxDB
Communication	Native and Highly Proficient in English, Mandarin
Statistics	Applied Probability & Statistics, Modeling & Simulation, Linear Models, Probability & Stochastic Processes, Introduction to Bayesian Methods, and Causality Inference
Machine Learning	Mobile Sensing & Machine Learning, Data Mining, Time Series & Applications, Reinforcement Learning, and Machine Learning
Systems-Related	Systems Analysis & Design, Decision Analysis, Business Decision Models, Systems Thinking in Public Health, Business Info. Systems Development, Electronic Commerce Strategy, Networks & Distributed Systems, Supply Chain Strategy, Systems Design, Electronic Commerce, Business Database Systems, Production & Schedule Control, Operations Research, Integer Programming, and Behavioral Sciences

TECHNICAL EXPERIENCE & HIGHLIGHTED PROJECTS

Research Project: *Organ Match-Run Optimization and Pattern Detection* Jan 2021 — Present

- Develop data processing pipelines to turn the raw data into structured and cleaned formats.
- Select and implement appropriate methods to impute missing data.
- Conduct literature review to explore innovative machine learning algorithms for healthcare domain.
- Develop statistical and machine learning models to capture patterns and practical data-driven solutions for current transplantation practices.
- Explore and implement other cutting-edge machine learning algorithms to achieve more efficient and interpretable model performance.

Research Project: *Predict Racial Bias Using Physiological Data* Jan 2020 — Dec 2020

- Built data collection pipelines to collect data from mobile sensing devices(Cellphone and Empatica E4 wristband).
- Built MySQL and InfluxDB databases to support data storage.
- Built a data feature extraction pipeline to extract possible features from physiology and survey data.
- Conducted correlation and importance analysis for features used for modeling.
- Tested popular ensemble meta-algorithms, and eventually implemented XGBoost, facilitated by the framework of Bootstrap Aggregation (Bagging) and Boosting mechanisms to predict if someone is biased or not based on their physiology data under certain circumstances.
- Investigated feature importance, as well as how much each feature contributed to the algorithm's decision process.

Class Project as Student / Predictions for Market Sales Data Using Time Series Techniques Oct 2020 — Dec 2020

- Implemented various models to predict retail sales using data from a supermarket.
- Performed data cleaning, timestamp conversion and indexing.
- Decomposed the time series data into three distinct components: trend, seasonality and noise to identify the unstable retail items(e.g. furniture, rugs, water faucets etc.).
- Implemented ARIMA models with "grid-search" for optimal sets of parameters to make predictions on store sales.
- Validated the model performance using metrics of MSE and RMSE.
- Explored the practicability of using PROPHET library developed by Facebook on the supermarket sales data.

Class Project as Student / Predicting Flight Delays Using Basic Data Mining Techniques Mar 2020 — May 2020

- Implemented a data cleaning pipeline for a nation-wide airline flights dataset(cleaning, imputation, aggregation etc.).
- Performed a series of Exploratory Data Analysis(EDA) to create intuitive and accessible visualizations.
- Built linear and polynomial models for univariate or multivariate regressions, and utilized cross-validation to determine accurate model parameters.
- Explored how fundamental regularization techniques ($l1$, $l2$, and *elasticnet*) could improve model performance.
- Successfully predicted flight delays with a satisfactory MSE, and identified airline companies and airports which produced the most and least delays.

Class Project as Student/ Customer Churn Prediction

Mar 2021 — May 2021

- Conducted a series of customer churn analysis using machine learning methods to identify the actual problems in the business's services (e.g. poor quality product/service, poor customer satisfaction, wrong management strategies and goals, etc.).
- Explored the practicability of making correct strategic decisions that would lead to higher customer satisfaction and consequently higher customer retention using machine learning techniques.
- Investigated and implemented possible methods to take care of imbalanced datasets, and compared the models' performance with balanced datasets.
- Implemented Naive Bayes and Logistics Regression models as baseline models, then built other models using Support Vector Classifier, Random Forest Classifier, Gradient Boosting Classifier, and Xtreme Gradient Boosting Classifier to search for the best approach to address the churn problem.
- Improved the models by conducting hyperparameter tuning and feature selections for all the models listed above.

Class Project as Student/ Using Deep Neural Network to Detect Credit Fraud

Oct 2021 — Dec 2021

- Implemented Decision Tree, Random Forest as baselines to detect possible credit card fraud.
- Resampled the highly imbalanced dataset using oversampling and SMOTE, which later greatly improved model performance.
- Investigated the difference between Rectified, Sigmoid and Hyperbolic Tangent functions in terms of their impact on final model performance. Rectified(ReLU) outperformed others.
- Implemented a 5-layer Deep Neural Network(w/ Keras) with an activation function of Rectified Linear Unit(ReLU) for the hidden layers, Sigmoid function for the outer layer.
- Explored the interpretability of DNN models, specifically for the case of financial fraud detection, with the help of this book: [Interpretable Machine Learning](#).

Class Project as T.A./ National Weather Disasters Analysis

Oct 2021 — Dec 2021

- Helped students address the overfitting issue by explaining and implementing penalization methods(*l1, l2, elasticnet*) for them.
- Introduced other tree-based classification methods to them, and asked them to implement Decision Tree and Random Forest to compare the prediction results with the original Logistic Regression model.
- Asked them to distinguish the difference between common metrics such as accuracy, recall, precision, f1 etc.

Class Project as T.A./ COVID-19 Data Visualization and Analysis

Mar 2022 — May 2022

- Helped students to build reusable and well-packaged code for data cleaning.
- Guided the students to perform dummy-encoding for multi-class categorical features.
- Explained the beauty of logistic regression, and informed them why and then the classic statistical modeling approaches are better than the famous, but black-boxed approaches.
- Introduced RShiny and asked them to build a data dashboard using RShiny.
- Facilitated them to build and interpret linear and logistic regression models to tackle different problems defined in advance.

EDUCATION & AWARDS

PhD in Systems and Information Engineering, *University of Virginia*

Aug 2019 — Present

M.S. in Industrial and Systems Engineering, *SUNY - Binghamton*

Aug 2017 — May 2019

B.S. in Supply Chain Management, *Arizona State University*

Aug 2013 — May 2017

B.S. in Computer Information Systems, *Arizona State University*

Aug 2013 — May 2017

PhD Fellowship(Stipend, Tuition and Insurance Coverage), *University of Virginia*

Aug 2019 — Present

New American Scholarship (\$4,000/Year), *Arizona State University*

2013 — 2017

ACTIVITIES & EMPLOYMENT (DETAILED CV AVAILABLE UPON REQUEST)

Graduate Research Assistant, *UVa*

Aug 2019 — Present

Graduate Teaching Assistant, *UVa*

Aug 2019 — Present

Global Operations Intern, *Dell Technologies*

Jun 2018 — Aug 2018

Business Solutions Development Intern, *Sentinel Medical Devices LLC*.

Jan 2016 — Aug 2017

PAC-12 Lead Interpreter, *ASU*

Sept 2016 — May 2017

ASU101 Course Instructor, *ASU*

Aug 2015 — Dec 2016

Marketing Specialist, *Inition Education Solution*

Oct 2016 — May 2017

Vice President, *Coalition of Intl' Students at ASU*

Mar 2014 — May 2016

Marketing Rep., *ASU*

Oct 2014 — Jul 2015

Marketing Rep., *Airdog LLC*.

May 2014 — Aug 2014