

PAUL CHEN

463-210-6665 | chen3876@purdue.edu | linkedin.com/in/paiweichen | https://paul46186.github.io/portfolio/

PROFILE AND SKILLS

Data enthusiast with proven technical qualities, analytical mindset, strong communication, and actionable insights

- **Technical Skills:** Python, SQL, Excel VBA, Tableau, R, Business Analytics, Integer Programming, Data Visualization, Project Management, Machine Learning, Data Mining, HTML, Deep Learning, NLP
- **Power Skills:** Leadership, Active Listener, Change Management, Agile, Self-Learner, Empathy, Presentation
- **Language:** Mandarin Chinese (Native), English (Professional), Taiwanese (Native), German (Beginner)

CERTIFICATIONS AND AWARDS

- Data Camp Data Analyst with SQL Server Track
- Data Camp Data Scientist with Python Track
- Data Camp Data Scientist with R Track
- Codecademy Deep Learning Tensor Flow Path
- Google Data Analytics Certificate
- Google Digital Marketing Certificate
- 3rd place in 2021 KGSA Case Competition
- 3rd place in STAMINA Analytics Case Competition

PROJECTS

Industry Practicum – SIL Optimizing Seed Data for Machine Translation West Lafayette, IN
Analyst Jan 2022 – Present

- Created a method that optimizes datasets for the training of machine translation models via **JoeyNMT**, **PyTorch nltk**, **SpaCy**, and **polyglot** and a clear strategy for sequencing a translation project. This method can be used for translating between different languages with minimum cost and time

Prescriptive Analytics – Window Breakage Optimization West Lafayette, IN
Analyst Dec 2021 – Present

- Demonstrates a code design that integrates a trained predictive model using the python library **PyCaret** into an optimization model using **Pyomo**. efficiently integrate the predicted model into the optimized model and then show an outcome distribution based on how the user might want to change an input parameter

Kaggle in-class competition - Credit Card Prediction Project West Lafayette, IN
Analyst Nov 2021 – Dec 2021

- Created and tuned an ensemble model that combined neural network, adaptive lasso and random forest through **SAS EM** to predict the likeliness of credit default and won 10th place in the Kaggle in-class competition

Improving the Craigslist User Experience - Video Gaming Classification West Lafayette, IN
Analyst Nov 2021 – Dec 2021

- Scrapped, tokenized, and lemmatized the Craigslist video game data via **beautifulsoup**, **nltk** and **sklearn**, then built 6 models with **keras**. The best model has 88.45% accuracy which can increase information gain by 34% that will improve the retention rate and customer experience in Craigslist website

Heart Disease Classification Project West Lafayette, IN
Team Leader Aug 2021 – Aug 2021

- Constructed a random forest model and created a website via **R Shiny** that presents interactive Exploratory Data Analysis, allows users to input 14 body factors and predict the potential heart failure level

Coursera Project West Lafayette, IN
Team Leader Jul 2021 – Jul 2021

- Cleaned, integrated, manipulated, and visualized Coursera datasets to find critical elements and offer business recommendations to enhance course enrolment using statistical models like **ANOVA**, **OLS**, and **regression**

Special Topic Group Study Tainan, Taiwan
Team Leader Feb 2019 – Feb 2020

- Led a group to define the assembly line problem, aggregate datasets with **Excel VBA**, and construct 3 models with **Python** and **Gurobi** that allow manager to find optimized workstation solutions within 1 minute

PROFESSIONAL EXPERIENCE

Shootown Group Qingyuan, China
Operation Intern Jul 2019 – Aug 2019

- Reduced workers' daily input time from 6 hours to 1 minute by using **Excel VBA**
- Built an **OEE** system to enable the company to capture their machine performance

EDUCATION

Purdue University, Krannert School of Management West Lafayette, IN
Master of Science in Business Analytics and Information Management Jun 2021 – May 2022

National Cheng Kung University Tainan, Taiwan
Bachelor of Management Science, Industrial and Information Management Sep 2016 – Jun 2020