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**: Deep Learning, NLP, Data Mining, Machine Learning, Python, SQL, Excel VBA, Tableau, R, Business Analytics, Integer Programming (Optimization), Data Visualization, Project Management, HTML
- **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
- 1st place in SAS Optimization Challenge
- Google Data Analytics Certificate
- Google Digital Marketing Certificate
- 3rd place in 2021 KGSA Case Competition
- 3rd place in STAMINA Analytics Case Competition
- 3rd place in Kaggle Crossroads Classic Datathon 2022

PROFESSIONAL EXPERIENCE

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

- Established a method that optimizes the sequence of training datasets for machine translation models via **JoeyNMT** and provides a clear strategy for choosing the semantic domains that minimize the translation cost and time
- This method can be used for translating between different languages, especially practical for rare languages that require human translation

Shoetown Group **Qingyuan, China**
Data Analyst Intern **Jul 2019 – Aug 2019**

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

PROJECTS

Prescriptive Analytics – Selected by Pycon Conference 2022 Poster Session **Dec 2021 – Present**

- Provided end users decision confidence with a framework that integrated predictive model using the python library **PyCaret** into an optimization model using **Pyomo**
- This project is selected by **Pycon** poster session which is the biggest python conference in the US

Kaggle in-class competition - Credit Card Prediction Project **Nov 2021 – Dec 2021**

- Developed an ensemble predictive model that utilizes historical payment status and demographical data to help credit card firms evaluate the likeliness of credit default via **SAS EM** and won 10th place in the Kaggle competition

Improving the Craigslist User Experience - Video Gaming Classification **Nov 2021 – Dec 2021**

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

Heart Disease Classification Project **Aug 2021 – Aug 2021**

- Created an application that performs two roles via R Shiny. One provides an interactive dashboard to visualize exploratory data analysis, while the other facilitates speedy risk factor identification by allowing users to input 14 body variables and forecast their likelihood of heart failure levels using a random forest model

Special Topic Group Study **Feb 2019 – Feb 2020**

- Led a group to define the assembly line problem, aggregate and cleaned datasets with **Excel VBA,** and construct 3 models with **Python** and **Gurobi** that allow managers to find optimized workstation solutions within 1 minute
- This project allows managers to make optimized machine layout decisions without relying on human

EDUCATION

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

- GPA: 3.71/4.0

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

- Top 30%