# Jai-Hua (Kevin) Yen

jaihuayen@gmail.com
https://jaihuayen.github.io/

## EDUCATION

**The State University of New York at Buffalo**  New York, USA
*Doctor of Philosophy in Biostatistics*  Aug. 2024 - Present

**National Taiwan University**  Taipei, Taiwan
*Master of Science in Biometry*  Sep. 2017 - Aug. 2019

**National Tsing Hua University**  Hsinchu, Taiwan
*Bachelor of Science in Quantitative Finance*  Sep. 2013 - Jun. 2017

## SKILLS

- **Programming:** Python, R, SQL, Golang, Matlab, SAS, C, C++, Linux, Docker, Git, LaTex

## RESEARCH EXPERIENCE

**National Taiwan University**  Taipei, Taiwan
*Graduate Student Research Assistant*  Sep. 2017 - Aug. 2019

- Derived an adjusted Chao2 species richness estimator, reducing underestimation by 15% and mitigating a 10% overestimation in scenarios with very low corrected denominators while mitigating identity errors common in species surveys
- Conducted data cleaning on weed and plant cover species surveys conducted in Soft Bridge County, Taiwan, and Grand St. Bernard Pass, Switzerland

## TEACHING EXPERIENCE

**National Taiwan University & Taipei Medical University**  Taipei, Taiwan
*Teaching Assistant Conference Lecturer*  Feb. 2019

- Conducted teaching sessions for 100+ students
- Presented teaching methodologies, crafting materials, and becoming a proficient teaching assistant

**National Taiwan University**  Taipei, Taiwan
*Teaching Assistant*  Sep. 2017 - Jan. 2019

- Taught statistics and linear algebra to accounting, agricultural chemistry, agronomy, and veterinary medicine majors
- Designed and delivered TA courses covering R and Excel fundamentals
- Achieved a 4.58 rating out of 5.0 from 54 students; received the Excellent Teaching Assistant Award (Top 1% evaluation score)

## INDUSTRY EXPERIENCE

**QNAP** Taipei, Taiwan
*Senior AI Software Engineer* Apr. 2024 - Jun. 2024
*AI Software Engineer* Feb. 2022 - Apr. 2024

- Trained a ResNet50 image classification model with a 70% F1-Score across 250 image classes and optical character recognition (OCR) models using vision transformers with 90% accuracy to enable efficient customer photo searches
- Leveraged CLIP (Contrastive Language-Image Pre-Training), a multimodal model, to integrate semantic search into the NAS search engine, empowering customers to search for photos and files using human-readable descriptions instead of conventional keywords
- Deployed models on edge devices, including customer NAS systems with Tensor Processing Unit (TPU) and Neural Processing Unit (NPU) acceleration, improving data security and inference speed within constrained computing resources
- Mentored junior personnel on programming, project organization, and training OCR models

**TutorABC** Taipei, Taiwan
*AI Algorithm Engineer* Jun. 2020 - Feb. 2022

- Created an 81% accurate customer churn detection model using random forest to decrease churn rates by 5%
- Developed an XGBoost customer purchase detection model with a 78% accuracy rate; utilized explainable machine learning method (SHAP value) to boost purchase rates and formulate targeted sales strategies
- Extracted key customer feedback using computer vision (YOLO, OCR) and natural language processing (TFIDF, BERT) models, enhancing data analysis for additional product features and improved user experience
- Accelerated the extraction, transformation, and loading (ETL) processes by 50%; consolidated by implementing Airflow for machine learning operations (MLOps)
- Mentored a junior engineer in programming, project structuring, and MLOps

## PUBLICATION

**Yen, J.-H.**, Chiu, C.-H. (2020). Richness estimation with species identity error. *Proceedings 62nd International Statistical Institute World Statistics Congress*, **Volume 6**, 401-408.