YUEH-HUAN (KEVIN) HO

619-953-8494 | yueh-huan.com | Yueh-Huan.Ho@rady.ucsd.edu | linkedin.com/in/y6ho | github.com/yue87

EDUCATION

University of California San Diego — San Diego, CA
Master of Science in Business Analytics — *Jul. 2022 – Dec. 2023*

Fu Jen Catholic University — Taipei, Taiwan
B.B.A in Finance and International Business — *Aug. 2017 – Jul. 2021*

SKILLS

Languages: Python, R, SQL (Postgres, SQL Server), SAS
DS Tools: Tableau, Spark, A/B Testing, AWS (SageMaker, Glue, Athena, EMR), Git
Libraries: Pandas, NumPy, Scikit-learn, SQL Alchemy, Keras, Statsmodels, Pytorch, Matplotlib
Other Skills: Marketing Optimization, Recommender Systems, Fraud Analytics, Customer Analytics
Certification: AWS Certified Machine Learning Specialty

EXPERIENCE

Data Scientist Intern — Jun. 2023 – Aug. 2023
Rocket Companies - Client Growth — *Detroit, MI*
- Created a **retention model** leveraging client's journey in mortgage milestones and identified decile with clients having a **twice close rate** than average, contributing to optimizing B2B and B2C marketing strategies
- Built a **lead forecasting model** (Beta Regression) that identified the 5 most effective marketing channels and affiliate partners, **reducing lead forecasting error rate by 22%**
- Constructed a data and ML model pipeline using **SQL Server** and **SageMaker**, generating table for client journey and saving 30% of time on Market Mix Modeling (MMM) training
- Collaborated with marketing teams to design **A/B testing** framework for model production and marketing strategy experiments

Research Data Scientist — Oct. 2021 – Jul. 2022
Fu Jen Catholic University — *Taipei, Taiwan*
- Designed the data pipeline, used Python to conduct feature engineering, and analyzed relationships between CSR decoupling and firm's performance indicators
- **Discovered the significant impact of CSR decoupling on companies' financial metrics** by building regression models and conducting multiple hypothesis tests on 607 Taiwanese companies' CSR indexes
- Built Random Forest models that predicted loan status for 900K+ customer records with 68% recall, leveraging techniques including K-fold **cross validation**, **SMOTE data resampling**, and **feature engineering**

Business Analyst Intern — Jun. 2020 – Sep. 2020
Metro Walk Shopping Center — *Taoyuan, Taiwan*
- Identified top 10 product categories and customers' spending behaviors using excel modeling (Linear Regression), and conducted customer analytic by designing and collecting 500 questionnaires
- Developed market strategies that **increased company revenue by 11%** through analyzing over 20 product metrics and customer behavior indicators

PROJECTS

Guidewire Software - Capstone | *Home Insurance Internal Water Loss Analysis* — Feb. 2023 – Jun. 2023
- Developed machine learning models **(Logit & LightGBM)** that analyzed 970 variables across 797K property records, identifying properties with **100% higher flood risk** in P&C insurance underwriting process
- Conducted feature engineering and **ensemble learning** to increase models' AUC by 19%, generating actionable insights into high risk properties and insurance premium optimization strategies
- Delivered new premium pricing strategies and client segmentation approachs by structuring model output into deciles, supporting an **9.7% increase in underwriting revenue**

Credit Card and Property Fraud Analytics | *Python, PCA, Fraud Model* — Feb. 2023 – Jun. 2023
- Built ML-based fraud models (**MLP, RF**) that predicts fraudulent credit card behaviors, generating an estimated **$20M in additional profit** through reduced fraud and optimized lending strategies
- Used **unsupervised model (PCA)** and **auto-encoder (MLP)** to calculate fraud scores from 1M+ New York property records with 32 variables, enabling clients to identify potentially fraudulent entries