Cheng-Pang (Tego) Chang

Email: tegochang@gmail.com

Phone: 919-945-6512

Website: tego-chang.github.io

LinkedIn: tego-chang

*A data scientist with **5+ years of product analytics and management experience** in the **high-tech/fintech** industry, dedicated to extracting insights from data to **inform product directions and formulate business strategies***.

Experience

Data Scientist – Explainable Machine Learning   Sep. 2022 – Present

2nd Order Solutions *(Duke MIDS Capstone Project)*   USA

- Evaluated the performance of <u>Explainable Boosting Machine (EBM) and Gradient Boost Machine (GBM) in execution duration and interpretability</u> by applying them to a classification prediction problem.
- Demonstrated the feasibility of <u>applying EBM to credit advisory topics in the fintech industry</u> by generalizing the model pipeline and interpretation methods.

Product Data Scientist Intern – IoT Risk Prediction   Jun. 2022 - Aug. 2022

TeleSign   USA

- <u>Proposed to the CTO a business opportunity in IoT security</u> by presenting a project – *Mobile Auth for Wi-Fi Access*.
- Realized risk evaluation of every connected Wi-Fi device by <u>utilizing unsupervised/supervised machine learning and TeleSign's strength in two-factor authentication</u>.

Senior Product Manager – Document AI   Sep. 2020 - May 2021

Foxit Software *(US company based in Fremont, CA)*   Taiwan

- Built and released an AI SaaS for contract review, ***iDox.ai***, by leading a team to <u>collaborate with external stakeholders to integrate legal domain knowledge into the outcomes from NLP models</u>.
- Stimulated user growth by 1.2% by <u>leveraging Google Analytics to design A/B testing for email campaigns.</u>
- <u>Contributed to the design of DevOps and MLOps processes</u> to ensure production requirements, such as logging and alerting, were met.

Software Engineering Product Manager – Wi-Fi Router Analytics   Aug. 2017 - Apr. 2020

NETGEAR *(US company based in San Jose, CA)*   Taiwan

- <u>Expanded the company's data dictionary by 30% (400+ attributes)</u> by collecting requirements from diverse stakeholders, designing corresponding technical specs, and leading internal and external teams to complete implementation.
- Verified the data accuracy by <u>collaborating with data engineers to troubleshoot issues during the cloud data processing pipeline</u>.
- Conducted the company's data-driven decision-making process by visualizing data outcomes, identifying key metrics, and <u>extracting business insights for the executives of the engineering and marketing teams</u>.

Software Product Manager - Wi-Fi Router   Apr. 2015 - Jun. 2017

SYNOLOGY   Taiwan

- Managed software releases by defining development plans and <u>driving cross-functional teams to meet schedules.</u>
- <u>Delivered 30+ technical presentations to a diverse range of audiences, including executives, technical or non-technical stakeholders, and reseller partners</u>, to secure product selling points.

Software Engineer - Wi-Fi Router | *REALTEK Semiconductor*   Jun. 2012 - Oct. 2013, Taiwan

Leadership and Awards

- Promoted among managers of 20+ at NETGEAR for leading 10+ software engineers to develop data analytics projects, deploying them to production, and providing insights for several teams at the company. Key projects include:
  - *Wi-Fi Connection Analytics:* accelerated the support team's remote troubleshooting process for wireless connection failures through the visualization of wireless data.
  - *Onboarding Method Analytics*: guided the campaign strategies for the marketing teams by identifying customers' preferred engagement channels and frequencies.
- Managed teams of 15+ members, including R&D, QA, UI/UX, PM, and Marketing, at Foxit Software, bringing the AI contract review service, *iDox.ai*, to market in six months.

Projects (GitHub)

Explainable Machine Learning – Predicting Donations upon Receiving Mail Offers

- Compared the performance of EBM and GBM in different feature selection settings for predicting donors in the KDD Cup 1998 dataset.
- Investigated the monotonic relationships between EBM's top influencing variables and the response for interpretability.

Impact of Superhost on the Annual Revenue of Listings on Airbnb

- Analyzed the causal effects between a listing's annual revenue and its host, as well as other factors in the hospitality industry.
- Performed data cleaning, wrangling, statistical modeling, and visualization on a geospatial dataset using Python.

Customer Segmentation and Churn Prediction for Payment Transaction Service

- Classified customers into four categories by applying an RFM framework based on engagement and lifetime value (LTV).
- Predicted the potential churn according to business context by utilizing Semi-supervised Learning methods.

Personalized Product Recommender

- Built an online recommendation system and achieved a 74% accuracy on a 1M-observation like-unlike dataset.
- Utilized collaborative filtering with Pandas to calculate similarity scores among users.

IoT Risk Identification and Prevention Framework – Mobile Auth for Wi-Fi Access

- Developed a risk score profile for each connected IoT device using its 2FA phone number and machine learning.
- Synthesized a Wi-Fi access dataset with annotated labels using clustering and outlier detection algorithms.

Education

Duke University, Durham, NC, ***Master of Interdisciplinary Data Science (MIDS)***; **GPA: 3.8/4.0**   **Expected May 2023**

National Taiwan University, Taiwan, *Master of Computer Networking & Telecommunication*   Jan. 2011

Knowledge and Skills

Data Science Machine Learning, Statistical Modeling/Inference, Hypothesis Testing, Causal Inference, Algorithmic Trading, Deep Learning, Computer Vision, Natural Language Processing, AWS, Azure, CI/CD, DevOps, MLOps.

Product SaaS, Experimentation and A/B Testing, Data Visualization and Storytelling, Product Management, Agile/Scrum Development, API Development, Tech Spec Design, User Growth, Customer Lifetime Value (LTV), SEO, UI/UX.

IoT and Cybersecurity Wireless Communication, Network Protocols (OSI Model 7 Layers), Digital Identification, Two/Multi-factor Authentication (2FA/MFA).

Programming Python (scikit-learn, PyTorch), SQL, R, C++/C.

Tools Tableau, Databricks AutoML, Docker, Google Analytics, Jira, Confluence, Microsoft Office Suite.