KUANG-TING CHANG

New York City · 718-568-4129 · www.ktchang.nyc · ktchang@pm.me

Build and deploy machine learning models to production.

PROFESSIONAL EXPERIENCE

Senior Data Scientist
Penguin Random House - largest trade book publisher the world — Jan, 2020 - Present

- Expanded personalization to prh.com via APIs
  - Led a cross-team collaboration to design the recommender system's API for website and email usages. Responsible for scope definition, development and deployment.
  - Automated the deployment with AWS CloudFormation and Gitlab CI.
- Users onboarding recommendation experience
  - Designed a quiz to collect users' explicit feedback to provide personalized experience.
  - Built a dynamic list of products that model identity as having the most explainability of one's taste.
- Full stack web development for internal tools
  - Print capacity forecasting tool for supply chain team. Bulk pricing tool for corporate sales team.
  - Visualization with JavaScript, built APIs with the Django REST framework.
  - Deployed using docker-compose to serve the app built with Django and Nginx.

Data Scientist
Penguin Random House - largest trade book publisher the world — Apr, 2018 - Dec, 2019

- Recommender system
  - Optimized and Containerized the recommender system of 15 million users, 70,000 products, increased reliability, and productivity. Reduced the pipeline execution time by 50%.
  - Developed the A/B testing framework. Regular tests increased click-through rate by 30%.
  - Migrated the system from Netezza to Snowflake database, and from on-premise server to AWS EKS.
- E-mail targeting system
  - Developed and deployed a "lookalike" email waves system using TensorFlow to achieve smaller but high precision targeted sends, later email waves often have >50% click-through rate.
  - Increased the click-through rate by 200%, reduced the churn by 90%. While only making up 7% of total delivery volume in 2019, 25% of revenue was generated by model-driven campaigns.

Data Analyst
Genesis Research - international healthcare consulting firm — Feb, 2016 - Aug, 2017

- Designed and implemented real-world data-oriented pharmaceutical studies using SAS and Python with various datasets typically over 100 million unique patients' medical records.
- Provided strategies that optimize statistical efficiency and quality, ranging from descriptive to univariate statistical analyses to regression modeling.
- Developed a Python web app that facilitates identification and retrieval of medical codes, reduced need for data entry personnel by 90%.

TECHNICAL SKILLS

- **Data Wrangling:** SQL, numpy, pandas
- **Data Pipeline:** luigi, dbt
- **ML:** TensorFlow, scikit-learn, xgboost
- **Infrastructure:** Linux, bash, Git, Gitlab CI/CD, Docker
- **AWS:** API Gateway, Lambda, Dynamodb, CloudFormation
- **Web:** Django, Nginx, HTML, CSS, JavaScript

EDUCATION

New York University, Graduate School of Arts and Sciences — New York, NY
Master of Arts, Applied Quantitative Research — Dec, 2015

Yuan-Ze University, Department of Social & Policy Sciences — Taiwan
Bachelor of Arts, Sociology — June, 2012

LANGUAGES

Native Mandarin Chinese, Intermediate Japanese (JLPT N4)