TE-KAI CHOU

Kelvin Chou | bestkelvin08@gmail.com | El Cerrito, CA | (470) 601-4801 | www.linkedin.com/in/tekaichou/ | https://github.com/TeKaiChou

WORK EXPERIENCE

Taelor (AI-Powered Menswear Rental Service)   **San Francisco, CA**
Data Scientist   *Jul 2023 – Present*

- Developed an automatic **multitask predictive model** for product image labeling in **PyTorch**, reducing manual labeling time by 50%
- Conducted analysis of customer buying behavior using **Python** and **Matplotlib**, resulting in a 60% increase in resale revenue
- Built data pipelines leveraging **SQL**, **dbt**, and **BigQuery** data warehouse, cutting data cleaning time by 1 hour per colleague weekly and reducing misreported data incidents by 10%
- Created automated reports and dashboards with **Python**, **SQL**, and **Looker Studio**, accelerating reports generation efficiency by 20%
- Enhanced inventory utilization by analyzing product sales and rentals with **SQL**, **Excel**, and **Pivot Tables**, recovering 10% of unused stock and reducing merchandising costs by 15%

BlueX Trade (Logistics Payment Fintech Platform)   **Taipei, TW**
Data Scientist Intern   *Jul 2022 – Aug 2022*

- Developed a **time series SARIMA model** for future bank account balances forecast with **PyCaret** and **Plotly**, resulting in **1 day acceleration** in loan approval response time for underwriting team decision-making
- Built data pipelines to transform transaction data using **SQL**, **dbt**, **Airflow**, and **Snowflake**, cutting 50% report generation time
- Initiated the automation and deployment of custom machine learning models in **AWS** using **SageMaker** and **ECR** with data engineers, boosting model-building efficiency by 80%

National Chiao Tung University (Supply Chain Management Lab)   **Hsinchu, TW**
Undergraduate Research Assistant   *Sep 2018 – Jun 2019*

- Constructed **decision tree models** for panel pixel defect detection using **Python**, achieving 95% yield
- Engineered image features with **GLCM**, **NumPy**, and **OpenCV**, enhancing defect detection accuracy by 20%

PROJECTS

Intelligent PDF Document Assistant App

- Developed a **Q&A Conversational AI** for PDF information summarization and retrieval by leveraging **Retrieval Augmented Generation (RAG)** solution powered by **Google Gemini-Pro API**, **LlamaIndex**, **Streamlit**, and Custom Prompts

Next Meme Stocks Forecast

- Built **logistic regression**, **random forest**, and **artificial neural network models** for meme stock classification using **Sklearn**, integrating social media data and stock prices from Reddit and Twitter APIs
- Implemented **sentiment scoring** and extracted stock tickers from posts with **NLTK** for trend and correlation analysis (EDA)

News Recommender Engine

- Developed a **content-based** news recommender engine using **TF-IDF** and **Word2Vec** to match articles with user interests
- Collected and processed data from 4 independent news sources using **Pandas** and **NLTK**, and stored the aggregated data in the **SQLite relational database** for efficient querying and recommendation

CAPTCHA Recognition System

- Developed an automatic CAPTCHA Recognition System **PyTorch** Model by leveraging **CNN** and **RNN** algorithms
- Engineered image noise reduction with **NumPy** and **OpenCV**, enhancing CAPTCHA recognition accuracy by 15%

EDUCATION

Georgia Institute of Technology   Atlanta, GA
Master of Science in Operations Research (***Data Science Concentration***)   *Dec 2022*
National Chiao Tung University (NCTU)   Hsinchu, Taiwan
Bachelor of Science in Industrial Engineering & Management   *Jun 2020*

SKILLS

- **Programming Languages:** Python, R, SQL
- **Data Analysis / Machine Learning:** Regression Analysis, Statistics, Time Series Analysis, KNN, K-means, Naive Bayes, A/B Testing, Natural Language Processing (NLP)
- **Data Visualization:** Google Looker Studio, Tableau, Power BI, MS Office, Jupyter Notebook
- **Data Engineering and Databases:** BigQuery, dbt, Airflow, MySQL, PostgreSQL, SQLite, Snowflake, Spark, PySpark, ETL
- **Cloud Platforms:** Google Cloud Platform (GCP), Amazon Web Service (AWS), S3, SageMaker, ECR
- **Tools and Technologies:** PyTorch, Tensorflow, Docker, Git, Scikit-Learn
- **Language:** English (Fluent), Mandarin (Fluent)

CERTIFICATIONS

- **Data Scientist Associate** by DataCamp   Nov 2024
- **Career Essentials in Generative AI** by Microsoft and LinkedIn   Nov 2024
- **AWS Certified Cloud Practitioner** by Amazon Web Services (AWS)   Jan 2023