Jasper Yat Shun Lee

Hong Kong | +852 9304 9832 | yatshunlee@gmail.com | github.com/yatshunlee | linkedin.com/in/yatshunlee | Portfolio

PROFESSIONAL EXPERIENCE

Blue Cross (Asia-Pacific) Insurance Limited, AIA Group   **Hong Kong**

Data Analyst, Full-Time   September 2024 – Present

- Automated 10+ ETL pipelines handling 60+ tables from databases such as Azure SQL MI and MySQL, cutting manual processing time by 10% daily through SQL, Spark, Qlik and Pandas.
- Collaborated with 10+ stakeholders to translate business requirements into 20+ actionable Power BI reports, streamlining workflows and saving 10+ manual hours weekly.
- Migrated 10+ dashboards and 5+ datamarts from Tableau to Power BI for sales, claims and marketing departments, resolving 100% issues ahead of the deadline in 2 months.

ASRC at the Hong Kong Polytechnic University   **Hong Kong**

Data Scientist, Full-Time   February 2021 – June 2023

- Developed and deployed an AI software by Python with YOLOv5 deep learning model, generating $1.4M revenue and achieving an F1 score of 0.93 (a 130% enhancement) via fine-tuning with 2k+ images and data augmentation.
- Improved the text extraction accuracy by 200% in an industrial project, by recommending the use of PaddleOCR, a state-of-the-art OCR model, and synthetic data generation with generative AI to train the model.
- Enhanced the performance of U2Net image segmentation model by nonlinear brightness scaling and image augmentation, ensuring its ability remains effective 99% of the times across varying lighting conditions.
- Achieved a 75% reduction in aerostructure distortion via analysis of material properties and machining parameters in a $5M project, employing A/B testing, visualization, and regression with Matplotlib, SciPy and Scikit-Learn.

EDUCATION

City University of Hong Kong   **Hong Kong**

Master of Science in Data Science   2021 – 2023

- GPA: 3.83 (Distinction); Selected as the graduation representative of Class 2023 from 100 graduates.

The Hong Kong Polytechnic University   **Hong Kong**

Bachelor of Engineering (Honors) in Mechanical Engineering   2015 – 2019

DATA SCIENCE PROJECTS

LendingClub Loan Analysis – Link   Pandas, AWS RDS (MySQL), Power BI, Scikit-Learn, MLflow

- Modelled the probability of default for credit risk analysis, materially boosting ROI of a portfolio by 58%.

Quantitative Risk Modelling – Link | Ranked 1st in class, 10+ Stars GitHub Repo   NumPy, SciPy, AWS EC2

- Developed a value-at-risk model based on Quantile Regression, improving the original computational speed by 90%.

Training a Gaming Agent by Reinforcement Learning – Link | Ranked 1st in class   PyTorch, OpenAI Gym

- Led a team of 4 to train an agent with DQN and customized reward function, reducing 3x completion time.

NLP for Fact Checking – Link   Requests-HTML, PyTorch, Transformers, Langchain, OpenAI API

- Engineered a fact-checking solution utilizing web scraping, vector similarity search and Gemma-2b LLM with RAG.

SKILLS

Programming: Python (NumPy, Pandas, Tensorflow, PyTorch), SQL   **Languages**: Cantonese, English, Mandarin

Technology: Power BI, Tableau, Linux, Git, Docker, PostgreSQL, MS SQL Server, MongoDB, AWS, GCP, Azure