Thong Yong Quan, Danny

+65 98456880 | dannyyqthong@gmail.com | [LinkedIn](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

AI/ML Engineer with a unique blend of expertise in Python, Deep Learning, LLMs (RAG), and MLOps, coupled with a strong background in financial compliance and product analytics. I leverage deep analytical rigor and practical experience to develop and deploy data-driven solutions, particularly in regulated environments, as demonstrated by a 15% improvement in clinical decision-making efficiency through end-to-end ML pipelines.

TECHNICAL SKILLS

Programming & ML: Python, SQL, PyTorch, Scikit-learn, FastAPI, LangChain, Deep Learning, NLP, Transformers
LLM Ecosystem: OpenAI, OLLAMA, GROQ, LangGraph, CrewAI, Hugging Face, RAG
MLOps, Cloud: AWS (EC2, ECR), Azure (Container registry, Web App), Docker, CI/CD (GitHub Actions), MLflow, DagsHub
Data Tools & Databases: Power BI, FAISS, Chroma, Astra DB
Others: [Financial Modeling (Netflix forecast)](#), Financial Analysis, Compliance

PROFESSIONAL EXPERIENCE

AI Singapore – AI Apprentice **| Jan 2024 - Oct 2024**
- Enhanced clinician efficiency by 15% through an **end-to-end ML pipeline** (20 data sources, 4 engineers) that enabled prioritization of low-confidence daily predictions.
- Eliminated the need for clinicians to manually extract symptoms from medical reports, enhancing efficiency by collaborating on fine-tuning a **biomedical BERT LLM** that improved feature F1 score by 25%.
- Developed a **customized date aggregation module** using the **Least Squares method** to handle irregularities in **time-series data**, resulting in a 10% improvement in **explainability boosting model** F1-score.
- **Containerized** the ML pipeline with **Docker** and integrated **FastAPI** for testing, ensuring OS-agnostic deployment and robust validation across Windows, Linux, and Mac.
- Mentored 8 junior apprentices in Neural Networks and NLP, fostering collaboration and skill development.

CriAT (Credit Risk Analytics Startup) – Product Analyst | **Jan 2022 - Sep 2022**
- Benchmarked internal Probability of Default (PD) models against competitors (Moody's/S&P) standards using **Power BI**, delivering actionable insights to sales teams to improve competitive positioning.
- Engineered Python scripts to automate data quality assurance for credit risk irregular spikes alerts, enhancing daily QA efficiency by 25% and eliminating over 8 hours of manual review weekly.
- Started data migration analysis from Bloomberg to FactSet, validating at least 90% company coverage within product.

Various Top Tier Banks & Deloitte – Private Banking Support Manager | **2016 - 2021**
- Reduced client onboarding (KYC) turnaround from 5 days to 4 days (20%) by streamlining document verification workflows
- Managed complex regulatory compliance processes (KYC, CDD, transaction monitoring, fraud prevention) across diverse financial products, ensuring adherence to stringent industry standards.

Freelance – TMT Equity Research Assistant | **Jun 2018 - Aug 2019**
- Built financial models and evaluated SEC reports to forecast earnings and profitability for listed and pre-IPO companies.

PROJECTS

[RAG-Powered Chatbot for PDF Search](#)
- Designed and deployed a **RAG LLM** chatbot on **Streamlit Cloud**, saving 70% of user time with instant PDF/web answers, expanding search by 2x with **Arxiv**, **Wikipedia**, **DuckDuckGo** tooling.
- Utilized **LangChain**, **GROQ**, PyPDFLoader, **FAISS vectorstore**, and HuggingFace embeddings for efficient querying.

[Network Security Phishing Project End to End Deployment](#)
- Implemented a traffic classifier with a 90% F1-score, which reduced analyst time by 50% for prioritizing low-confidence predictions.
- Deployed on **AWS (EC2, ECR)** and **Azure(Azure Registry, Web App)** with **Docker**, ensuring scalability and reliability.

[NLP Kindle Review Sentiment Classification](#)
- Engineered an NLP system to classify Kindle review sentiment using **Scikit-learn**, **NLTK**, and **Gensim** and used **Docker** to debug environment issues, enabling successful deployment on **Streamlit Cloud.**
- Achieved 64.2% accuracy on AvgWord2Vec, outperforming BoW and TF-IDF techniques by 10%.
- Engineered a pipeline with tokenization, stopword removal & vectorization (**BoW**, **TF-IDF**, **Word2Vec**, **AvgWord2Vec**).

EDUCATION AND CERTIFICATIONS

AI/ML & Analytics: [Associate AI Engineer](#) (AI Singapore), [Business & Data Analytics](#) (BCG)
Corporate Finance Institute® (CFI): Financial Modeling & Valuation Analyst
University of London: Bachelor's Degree in Banking and Finance