Yeo Meng Han

+65-98150541 | yeo_menghan@u.nus.edu | Linkedin | Github | Portfolio

WORK EXPERIENCE

AI Engineer Intern, NUS Office of Internal Audit **Sep 25 - Present**

- Designed and deployed an end-to-end fraud detection system to flag anomalous vendor invoices, combining OCR (spaCy-Layout), PDF metadata forensics (PyMuPDF), and document similarity (TF-IDF + cosine similarity); containerized with Docker and served via Gradio, reducing manual audit review time by ~30%.
- Developed an explainable unsupervised anomaly detection module using Isolation Forest to identify suspicious staff/vendor claims, with SHAP values providing interpretable risk scores for auditors - enabling transparent, actionable alerts without labeled fraud data.

Machine Learning Engineer Intern, Changi Airport Group **May 25 - Aug 25**

- Improved taxi passenger demand forecasting at Terminal 3 by applying time series decomposition and feature engineering to XGBoost, increasing accuracy from 48% to 62%; deployed via CI/CD (Bitbucket, AWS Lambda/ECR/CloudFormation).
- Analysed 18,000+ flight turnarounds to pinpoint ground-crew bottlenecks, delivering reports and recommendations to reduce departure delays.
- Built an automated SageMaker data pipeline (Bash/Python) integrating flight, baggage, and passenger flow data - cut monthly processing from 12h to 8.5h with improved logging.

Research Assistant, Cisco-NUS Accelerated Digital Economy Corporate Laboratory **Jan 25 - May 25**

- Built a real-time human detection and tracking system using YOLOv8 + DeepSORT on overhead fisheye video feeds, achieving 84% mAP at 15 FPS on a standard workstation.
- Developed a low-latency pipeline to convert calibrated video streams into standardised JSON bounding boxes, enabling near real-time motion analysis for digital-twin applications.

AI & Computer Vision Engineer Intern, Rockship (NUS Overseas College Vietnam) **Jan 24 - Jun 24**

- Built a RAG pipeline using Langchain and ingest 50+ publicly available n8n workflows to generate simple n8n workflows; added observability via LangSmith to monitor agent performance and reduce hallucination.
- Developed and deployed robust, scalable LlamaIndex RAG APIs for text and image embedding with built-in OCR safeguards, adopted as boilerplate by internal engineering team to accelerate prototype development.
- Improved YOLOv8 object detection mAP by 20 percentage points through targeted labeling and data augmentation; integrated model into an ARKit iOS app for sub-centimeter nail alignment and distance measurement in construction scenarios.

PROJECTS

eKYC Facial Liveness Detection System | Github Link | Website **Jan 26 - Jan 26**

- Built and deployed an end-to-end facial liveness detection pipeline (live vs spoof) with PyTorch, ONNX, FastAPI, and AWS, including image quality checks and automated deployment.

Smart Delivery Allocator (Grains) | Github Link **Oct 25 - Oct 25**

- Built an AI agent using GPT-4.1 to auto-assign catering orders to drivers under real-world constraints (geography, time windows, capacity, VIP priority, driver's event expertise).
- Used iterative LLM prompting with deterministic validation to refine allocations and reduce errors.
- Delivered a full-stack prototype: Python backend + Streamlit/Folium frontend for visualizing route assignments.

EDUCATION

National University of Singapore **Jul 25 - Present**

Master of Science, Data Science and Machine Learning

- Relevant Coursework: Scalable Distributed Computing for Data Science, Deep Learning, Cloud Computing
- Expected Graduation: December 2026

National University of Singapore **Aug 22 - Jul 25**

Bachelor of Engineering, Major in Computer Engineering

- GPA: 4.51 / 5.00 (First Class Honours)
- NUS Engineering Scholarship Programme (awarded to top engineering students)
- NUS Computing Club - Project Director, LifeHack 2023 | Residential College 4 - Photographer & Videographer, Freshmen Orientation 2023 | Web Developer, E-Scholar Community

TECHNICAL SKILLS

- **Programming Languages:** Python (primary), SQL, C/C++, Java
- **AI / ML:** PyTorch, TensorFlow, scikit-learn, OpenCV, PySpark
- **Generative AI:** Diffusion Models (Stable Diffusion, AnimateDiff), LangChain, LlamaIndex, RAG, Multimodal AI
- **MLOps & Deployment:** Docker, FastAPI, CI/CD (GitHub Actions, Bitbucket Pipelines), Model Monitoring (LangSmith, LangFuse), Model Optimization (ONNX, Lightweight Edge Inference)
- **Cloud**: AWS, Google Cloud Platform (Professional ML Engineer Certified)