

# Will Lim Wen Long

Data Scientist

IMDA SG Digital Scholar

wenlonglim96@hotmail.com | +65 98437531 | [github.com/wenlong96](https://github.com/wenlong96) | [linkedin.com/in/limwenlong/](https://linkedin.com/in/limwenlong/)

## RELEVANT WORK EXPERIENCE

### PRICEWATERHOUSECOOPERS

Senior Data Scientist

Jan 2025–Present

- Led development of Agentic Graph RAG solutions using LangChain, LlamaIndex, and PyTesseract, generating US\$1.2M in revenue.
- Developed and custom-trained lightweight U-Net, improving accuracy by 10% in image-based tasks.
- Designed hyperparameter-tuned splitting strategies pre-embedding, improving splitting accuracy by 20% over base chunking strategies in vector DB retrieval.

Data Scientist

Singapore  
Jun 2023–Jan 2025

- Developed time-series models for inventory management using ARIMAX and decision tree boosting models, reducing MAPE by 20%.
- Refactored RAG pipelines with asyncio and multiprocessing, reducing runtime by 80% and streamlining performance bottlenecks.
- Built automated web scraping and data mining pipelines with Selenium, SQL, and factor analysis, enabling market mapping and whitespace identification.

### DINERS CLUB

Data Analyst

Singapore  
Jan 2023–Jun 2023

- Deployed automation watchers for data crawling, tidying, ingestion, and report generation, saving 30 minutes daily.
- Developed and maintained complex PostgreSQL/MySQL queries to support stakeholders, and deployed Cron Jobs for automated report generation.
- Streamlined processes by redefining logic from legacy systems to improve data quality and optimize query performance.

## SKILL & LANGUAGE

### LANGUAGES / DATABASE

R, Python, SQL

### TECHNOLOGIES / FRAMEWORKS

LangChain, PyTorch, scikit-learn, LlamaIndex, Azure DevOps, PowerBI, Tableau, Git, asyncio, Selenium, Power Automate, Copilot Studio, Pytesseract, mlr3, plotly, Markdown, NetworkX, DGL

### MACHINE LEARNING / DEEP LEARNING

Time Series Analysis, RNN, CNN, GNN, Transformer, Autoencoders, Regression, Classification, Probabilistic Clustering, Decision Tree Models, Gradient Boosting Models, Bayesian Networks, Bayesian Inference

## AWARD & CERTIFICATION

PwC Spot Award (2024) - Top Performer in Division

IMDA Singapore Digital Scholarship (Postgraduate) - Artificial Intelligence

AWS Certified Machine Learning - Specialty

Singapore Management University - School of Economics Dean's List

## EDUCATION

### NATIONAL UNIVERSITY OF SINGAPORE

Master of Computing - Artificial Intelligence Specialisation (cGPA: 4.75/5.0)

Singapore  
Jan 2025-Jun 2027

### SINGAPORE MANAGEMENT UNIVERSITY

Bachelor of Science in Economics, Data Science & Analytics  
Honours: Magna cum laude (GPA: 3.71/4.0)

Singapore  
Aug 2019-May 2023

### SINGAPORE POLYTECHNIC

Diploma in Banking and Finance (GPA: 3.52/4.0)

Singapore  
Apr 2013-Apr 2016