Jonas Lim Hong Xiang

Block 625 Choa Chu Kang Street 62 S(680625) | +65 9663 2152

259 Concord Road, NSW 2138 | +61 0439 716 134

jonaslimhx@gmail.com | linkedin.com/in/jonaslimhx | github.com/jonaslinhx

EDUCATION & AWARDS

Master of Computer Science, Data Science & AI | University of Sydney, Australia   *Aug 2024 – June 2026*

Singapore IMDA SG Digital Scholar (Infocomm Media Development Authority)

• COMP5318 Machine Learning: Implemented ViT from scratch to train a model which predicts human age.
• COMP5046 Natural Language Processing: Implemented encoders and decoders. LLM prompting.
• COMP5425 Multimedia Retrieval: Built a RAG based chatbot.
• COMP5427 Usability Engineering: Analyze, design and evaluate an E-book.
• COMP5426 Parallel & Distributed Programming: Designed C programs using pthreads, OpenMP, OpenMPI and Cuda.
• COMP5216 Mobile Computing: Built a food wastage mobile app using TypeScript and React Native.
• ELEC5514 IoT Wireless Sensing & Network: Built real-time chat function with STM32 MCU and LoRa Transceiver.

Bachelor of Engineering, Aerospace | Nanyang Technological University, Singapore   *Aug 2015 – May 2019*

First Class Honors, Highest Distinction

PROFESSIONAL EXPERIENCE

Software Architect | NockHub, Singapore (Part Time)   *May 2024 – Present*
• Translated business requirements into technical features through direct customer engagement.
• Architect AWS backend in Python & Terraform, designed 40+ PostgreSQL schemas & 30+ RESTful APIs.
• Designed & deployed the full-stack infrastructure using Python (Backend), Flutter (Frontend), and AWS (Terraform)
• Led a team of 3 engineers, overseeing technical planning, code reviews, and API documentation..
• Developed 5 cross-platform applications serving 4 archery ranges and 6,000+ users.
• Designed application interfaces (Figma) and workflows based on Jakob Nielsen's UX heuristics.

Senior Machine Learning Engineer | Carro, Singapore   *Jun 2022 – Jul 2024*
• Worked on ML projects spanning across predictive pricing, vehicle damage detection and engine acoustics.
• Built scalable and tolerant ML system and microservices on GPU cluster; serving $\geq$ 12,000 requests/day.
• Evaluated and replaced production Mask-R CNN model with YOLOv5; reducing inference time from 500ms to 50ms.
• Built data preparation pipelines, feature store, metadata store which improved training data quality by 50%.
• Collaborated with cross-functional stakeholders from product management, operations and data science across
   Singapore, Malaysia, Indonesia and Taiwan to increase ML service user adoption by 40%.
• Optimized AWS network architectures, databases and model training servers; saving US$120k in 2023.
• Replaced monolithic web crawler with a distributed system which reduced processing time from 7days to 9hours.
• Reimagined team's ML infrastructure and work processes which saved 75 man hrs/month/pax.
   • Designed data engineering frameworks and mentored team on automating data pipelines and best coding practices..
   • Refined project repository structures and optimized system designs for current production systems.

Data Scientist | Micron Technology, Singapore   *Jun 2019 – May 2022*
• Built defect classification model which saved inspection tool capacity by 10%.
• Optimized SQL queries in model data preparation pipeline which reduced query time by 50%.
• Built a predictive model for product yield and deployed to on-premises servers which achieved $11M savings.
• Developed end-to-end ML solution; from problem formulation, solution development, to productionalization.

TECHNICAL EXPERTISE

Frontend Technologies: React Native, Flutter
Backend Technologies: Pytorch, TensorFlow, PySpark, FastAPI, Docker
Programming Languages: Python, SQL, Dart, Bash, C, TypeScript, Terraform
Databases: MySQL, PostgreSQL, Redis, MongoDB, DynamoDB, DocumentDB, RDS
Cloud: AWS, GCP, Firebase

HACKATHONS

2025 Global Health Systems Hackathon Hub, "Building High-Value Health Systems: Leveraging AI"
Team *I Know You Know* won [2nd Place in Sydney, Australia](#) and [top 20 in the Global round](#).

GITHUB PROJECTS

pyspark_project [Data Engineering]
Pyspark application which aggregates data and publishes to Kafka topic for downstream service systems consumption.

producer-consumer-ml-system [MLOps]
Asynchronous instance segmentation microservice implemented in Terraform.

CERTIFICATION

Data Engineering

Apache Spark 3, Udemy	*December*	*2023*

Cloud Solutions

AWS Certified Solutions Architect Professional, Amazon Web Services	*January*	*2024*
AWS Certified Advanced Networking Specialty, Amazon Web Services	*July*	*2023*
AWS Certified Solutions Architect Associate, Amazon Web Services	*August*	*2022*
Terraform for AWS, Udemy	*June*	*2022*