Lin **Xiao Ya** | Mobile Phone: (+65) 8039 4280 | Email: linx0070@e.ntu.edu.sg
LinkedIn: www.linkedin.com/in/xiaoya-lin

EDUCATION

Nanyang Technological University   Aug 2023 – May 2027
- Bachelor of Electrical and Electronic Engineering, CGPA: 4.86/5.00, Honors (Highest Distinction)
- Specialization in Data Science and Machine Learning | **Dean's List (2024)**
- **NTU Science & Engineering Undergraduate Scholarship Recipient**
- **NTU President Research Scholar**
- Presenter at **The International Conference of Undergraduate Research (ICUR) 2025**
- Relevant Modules: (1) *Introduction to Data Science and Artificial Intelligence* [Python]
  (2) *Data Structures & Algorithms* [Python] (3) *Designing and Developing Database* [MySQL, NoSQL]

SKILLS

- **Software Programming:** Python, R, C Language
- **Database:** MySQL, NoSQL, MongoDB
- **Software Applications**: Figma, Jupyter Lab
- **Cloud & Big Data:** AWS S3, Sagemaker, Boto3, PySpark

RESEARCH & INTERNSHIP EXPERIENCE

GlobalFoundries, *Data Scientist Intern*   May 2025 – Dec 2025
Project 1: Scalable Data Pipeline & Compression
- Developed modular data pipelines for multi-month trace data processing with AWS S3, Boto3, PySpark, optimizing query performance.
- Implemented Parquet and Snappy compression techniques to improve retrieval speed and efficiency for high-volume semiconductor datasets.
- Reduced extraction runtime from hours to 4–8 mins; cut storage by 90%; eliminated session timeouts.

Project 2: Trace-to-Image Machine Learning Pipeline for Fault Detection
- Developed a novel trace-to-image machine learning pipeline, converting raw process data into image representations to address limitations in traditional Fault Detection and Classification (FDC) methods.
- Achieved baseline model accuracies of 97.0%, 97.0%, and 94.6% across three distinct tool types, outperforming the conventional guard band method by 19.3%.

A*STAR, *Healthcare Data Pre-Processing Research Intern*   Jan 2025 – Apr 2025
- Supported the project "National Healthcare Group (NHG): Validating an AI Facial Health Screening Product with A*STAR", developing data preprocessing workflows to anonymize sensitive healthcare datasets while ensuring compliance with regulations.
- Conducted feature preservation analysis and collaborated with the AI research team to optimize preprocessing methods for model performance.

ACADEMIC PROJECTS / MODULE PROJECTS

AI & Machine Learning Projects:
URECA Project: *AlzCare Smart Watch* (**Team Leader** of 3 Members)   Aug 2024 – May 2025
- Spearheaded a team in the conceptual design and prototyping of an AI-powered wearable device for dementia patients, with features for real-time health monitoring, fall detection, and geofencing.
- Collaborated with SG Jamiyah Nursing Home to define project requirements and evaluate initial design concepts, gaining experience in translating real-world needs into a technical solution.

NTU Module: *Introduction to Data Science and Artificial Intelligence*   Feb 2024 – Apr 2024
Module Project: Hotel Booking Trends and Cancellation Forecast (**Team Leader** of 4 Members)
- Led a team of 4 in applying Python with 4 Machine Learning Models to analyze a Kaggle dataset, achieving an 86.6% accuracy with Random Forest.
- Applied optimization techniques including TensorFlow, L2 regularization, and Batch Normalization, boosting model accuracy to a final 91%.

LEADERSHIP / CO-CURRICULAR ACTIVITIES / COMMUNITY INVOLVEMENT / VOLUNTEERISM

NTU Institution of Engineering and Technology (IET), *Liaison Manager*   Jul 2024 – Present
NTU Investment Interactive Club (IIC) Investment Academy, *Member*   Jul 2024 – Present
CFLS Model United Nation Club (CFLS-MUN), *Academic Director*   Dec 2019 – Dec 2021
- Hosted the North-East MUN Conference (**1,000+** participants) as **Chairperson**.
- Led delegations to World MUN and Yale MUN, proposing draft resolutions.
- Instructed **100+** members in CFLS-MUN about conference processing and document writing.