

Sachith Abeywickrama

Singapore | +65 9351 0851 | e240203@e.ntu.edu.sg | [LinkedIn](#) | [GitHub](#)

RESEARCH PROFILE

Ph.D. researcher developing efficient, trustworthy, and adaptive AI systems for sequential data. Expertise in foundation models, continual learning, interpretable AI, LLM applications, semantic search, and distributed model training.

EDUCATION

Nanyang Technological University

Singapore

Ph.D. Student in Electrical and Electronic Engineering

Aug. 2024 – Present

- Conducting research with A*STAR on efficient, interpretable, and continually adaptive AI systems for non-stationary sequential data.

University of Colombo

Sri Lanka

B.Sc. (Hons) in Industrial Statistics and Mathematical Finance

Jan. 2019 – Apr. 2023

HONORS & AWARDS

Second Runner-up

July 2026

China-ASEAN Student Innovation & Entrepreneurship Challenge

Guizhou, China

- Awarded second runner-up at the 2026 Challenge for the project “Precision Drive Coming” (autonomous core humanoid robotics joint solution), representing Nanyang Technological University.
- Selected as one of the elite 32 finalist projects out of 1,007 registered entries from 149 universities, involving 7,656 student competitors.

Singapore International Graduate Award (SINGA), A*STAR

2024

- Competitive Ph.D. scholarship supporting doctoral study at NTU and research at A*STAR’s Institute for Infocomm Research.

SELECTED PH.D. RESEARCH: *Under Review*

ConceptTime: Predictive Surprise as a Self-Grounding Concept Bottleneck

- Addressed the lack of interpretable and reliability-aware time-series classifiers by developing a self-supervised model that automatically discovers meaningful temporal segments and human-inspectable concepts.
- Introduced a built-in trust score that identifies unfamiliar or unreliable inputs; evaluated across 17 benchmarks and achieved 97.9% accuracy on Epilepsy using only 1% labeled data.

EntroPE: Entropy-Guided Dynamic Patch Segmentation for Time-Series Transformers

- Addressed accuracy and efficiency losses caused by fixed-length input windows that split meaningful temporal patterns at arbitrary locations.
- Developed an entropy-guided segmentation method and adaptive encoder that place boundaries at natural changes, reducing average forecasting error by approximately 8–10% compared with static patching.

PACT: Identity-Preserving Domain-Incremental Adaptation of Frozen Time-Series Foundation Models

- Addressed catastrophic forgetting when deployed foundation models encounter new operating conditions sequentially and receive only limited labeled data.
- Developed lightweight local adapters with identity-preserving routing, maintaining 0.984 average macro- F_1 across four sequential conditions using only 5% labeled calibration data and no condition identity at test time.

NeurInferno: Field Inference for Unknown Binary Protocols Without Priors or Binaries

- Addressed the need to recover the structure of unknown network protocols when specifications, source code, executables, and expert-provided information are unavailable.
- Developed a pretrained byte model and cross-message transformer that achieved 0.807 average boundary F_1 and 0.815 recall across 12 unseen protocols, improving recall from 0.54 to 0.81 over strong trace-only baselines.

RESEARCH AND INDUSTRY EXPERIENCE

Institute for Infocomm Research, A*STAR

Singapore

Ph.D. Researcher / Research Scholar

Aug. 2024 – Present

- Develop efficient, trustworthy, and continual-learning methods for foundation models operating under distribution shifts, limited labels, and changing real-world conditions.
- Build reproducible PyTorch pipelines for multi-dataset benchmarking, few-shot learning, continual adaptation, out-of-distribution detection, and model ablations.
- Train and evaluate time-series foundation models and open-source LLMs on SLURM-managed distributed multi-GPU clusters.

Rootcode

Sri Lanka

Associate Machine Learning Engineer

May 2023 – Jun. 2024

- Solved enterprise information-access problems by building document question-answering, semantic search, named-entity recognition, and open-source LLM applications.
- Productionized machine-learning models through Python and FastAPI services, scalable data pipelines, automated evaluation workflows, Docker, and cloud infrastructure.
- Delivered technical demonstrations to customers, performed statistical model validation, mentored interns, and investigated memory-efficient generative-model deployment.

Rootcode

Sri Lanka

Artificial Intelligence Engineer Intern

Nov. 2022 – Apr. 2023

- Built conversational AI and semantic-search systems to improve product discovery and customer support for e-commerce applications.
- Developed FastAPI endpoints and deployed containerized AI services using Docker, AWS, and cloud-based NoSQL databases.

Nanyang Technological University

Remote

Research Intern

Jun. 2022 – Sep. 2022

- Addressed data-efficient polymer property prediction by extracting informative topological features from molecular structures.
- Trained convolutional, transformer, and ensemble models using persistent spectral and molecular fingerprint representations.

TECHNICAL SKILLS

Programming: Python

Machine Learning: PyTorch, TensorFlow, scikit-learn, Transformers, Foundation Models, Large Language Models, Self-Supervised Learning, Continual Learning, Generative AI

Trustworthy and Efficient AI: Interpretable AI, Out-of-Distribution Detection, Few-Shot Learning, Model Adaptation, Efficient Sequence Modeling, Uncertainty Estimation

Applications: Time-Series Modeling, Natural Language Processing, Semantic Search, Information Retrieval, Document Question Answering, Named-Entity Recognition, Anomaly Detection

Systems and Deployment: SLURM, Distributed Multi-GPU Training, Docker, Git, FastAPI, AWS, Google Cloud Platform, Microsoft Azure