Thodsaporn Chay-intr

Machine Learning Engineer, Software Engineer, and Researcher
Bangkok, Thailand • +66-084-568-1166 • t.chayintr@gmail.com • linkedin.com/in/tchayintr • github.com/tchayintr • https://boated.zip

Machine learning engineer and researcher with 5 years of hands-on experience in academic and commercial applications, specializing in NLP, ML, and software development. Directed the design, development, and deployment of applications. Led pioneering projects that contributed to the Thai NLP community. Achieved state-of-the-art performance in text segmentation, with F1-scores ranging from 97.7% to 99.4% across Asian languages. Accomplished improvements in multimodal sentiment analysis and text generation. Demonstrated expertise in Python, graph techniques, and toolkits, notably PyTorch, across various projects, underlined by a record of multidisciplinary collaboration and leadership.

Relevant Work Experience

Artificial Intelligence Association of Thailand, Pathum Thani, Thailand Jan 2018 – Aug 2023
Machine Learning Researcher and Lecturer (Ad Hoc and Project-based)

- Designed and delivered Python courses on data analysis, AI, and ML. Covered tools (e.g., pandas, scikit-learn, spacy) and algorithms (e.g., regression, SVM, neural nets) in text and image analysis to educate hundreds of students, cultivating ML enthusiasts in Thailand.
- Consulted industries and students on theoretical and practical ML concepts, guiding project plans, assisting in the development of projects and software, and contributing to the publication of over 10 international research papers and projects.

Tokyo Institute of Technology, Tokyo, Japan Sep 2019 – Mar 2020
Research Assistant

- Collaborated with a multidisciplinary team to implement tailored APIs for NLU/NLG units for Japanese conversational dialogs.
- Designed, developed, and applied LSTM-based sequence-to-sequence models using PyTorch for text generation and refinement, resulting in enhanced fluency and a more natural linguistic output, validated through human evaluations.

iApp Technology Limited, Pathum Thani, Thailand Mar 2017 – Feb 2018
Machine Learning Engineer, Software Engineer, and Researcher

- Spearheaded and led a cross-disciplinary team of linguists and software developers to construct the first Thai Treebank, comprising over 5,000 entries, contributing Thai NLP resources, enhancing research, innovations, and applications in the NLP community.
- Led the development and deployment of native and web applications in React, Python, and GCP, as a pipeline product, for syntactic extraction stored in Firestore, with an emphasis on intuitive UX, bolstering ongoing developments in the Thai NLP community.

Selected Relevant Projects

A Unification-based Knowledge Graph Construction for Thai Profile Generation from Online Resources Sep 2023

- Constructed a knowledge graph for Thai researchers, using 6+ million entries crawled from online research databases.
- Designed a semi-supervised method with multi-task learning to extract entities/relations, improving F1-score by 8% over baseline.

Simple2In1: A Simple Method for Fusing Two Sequences from Different Captioning Systems into One Sequence Sep 2023

- Developed a T5-based generative model for Thai captions fusion, outperforming baselines by 5.2% in sBLEU and ROUGE-L scores.
- Accomplished a sBLEU score of 79% and a ROUGE-L score of 90% for a small captioning dataset comprising 3,168 samples.

LATTE: Lattice ATTentive Encoding for Character-based Word Segmentation (Journal of NLP) Jun 2023

- Proposed, and implemented a sequence labelling method that integrates multi-granularity linguistic units, Lattices, GNNs, PTMs, and Attention Mechanism using PyTorch and PyG to generate and refine text representations for word segmentation.
- Achieved state-of-the-art performance (97.7% to 99.4% of F1-score) across Asian languages: Japanese, Chinese, and Thai.

Multimodal Sentiment Analysis Using Multiple Labels from Different Modalities Mar 2023

- Collaborated with students to design and implement a sentiment analysis model for social network data, leveraging text, image, and multimodal labels using CLIP, BERT, and RoBERTa. Yielded up to 2% improvement in F1-score over recent models.
- Attained F1-scores of 74.1% for MVSA-single and 62.0% MVSA-multiple datasets.

Public Budget Usage Monitoring System (Bronze Medal - The 47th International Exhibition of Inventions Geneva) Feb 2019

- Cooperated with an interdisciplinary team to develop a monitoring system that utilizes Scrapy to crawl large-scale unstructured data from government sites, such as procurement and budget portals, for corruption detection in text data. Deployed by two organizations.
- Developed a text classification method in TensorFlow, with rule-based enhancements, for corruption detection, validated by experts.

Education

Tokyo Institute of Technology, Tokyo, Japan Apr 2019 – Sep 2023
Doctor of Engineering — Information and Communications Engineering

Sirindhorn International Institute of Technology, Pathum Thani, Thailand Jul 2015 – Aug 2018
Master of Engineering — Information Communication and Technology for Embedded Systems

Thammasat University, Pathum Thani, Thailand Jun 2011 – Aug 2015
Bachelor of Science — Computer Science

Skills

Technical Skills

- **Programming Languages**: Python, C/C++, Java
- **ML Toolkits**: PyTorch/Lightning, TensorFlow/Keras, Hugging Face, PyG, DGL, OpenCV, Scikit-learn, Spacy, NLTK
- **Tools & Technology**: Linux, Hadoop, Spark, SQL, NoSQL (MongoDB, Firestore), Docker, Jupyter, Neo4j, GCP, AWS, Git

Languages: Thai (Native), English (Advanced), Japanese (Intermediate)