Navapat Nananukul

Email: nananuku@usc.edu
Website: navapatn.github.io

Education

University of Southern California — Los Angeles, CA
M.S. in Computer Science — *Aug. 2021 – Aug. 2023*

Chulalongkorn Univesity — Bangkok, Thailand
B.S. in Mathematics and Computer Science — *May 2013 – May 2017*

Research Experience

Information Science Institute (ISI), AICS lab — Los Angeles, CA
*Knowledge Graph and Large Language Model with **Prof. Mayank Kejriwal*** — *Aug. 2023 – Present*

- **Research Direction:** Large language model evaluations and applications, Knowledge graph construction
- **Computational Social Science:** Analyzing and visualizing health inequality in the USA

University of Southern California — Los Angeles, CA
*Domain Adaptation and Computer Vision with **Prof. Mohammad Rostami*** — *Feb. 2023 – May 2023*

- **Research Direction:** Multi-source domain adaptation for medical image segmentation

Publications and Preprints

Balancing Efficiency and Quality in LLM-Based Entity Resolution on Structured Data [pdf]
Navapat Nananukul, Mayank Kejriwal
The 2024 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2024)

What if Red Can Talk? Dynamic Dialogue Generation Using Large Language Models [pdf]
Navapat Nananukul, Wichayaporn Wongkamyan
Wordplay: When Language Meets Games, ACL Workshop 2024

Multi-Source Data Integration for Segmentation of Unannotated MRI Images [pdf]
Navapat Nananukul, Hamid Soltanian-Zadeh, Mohammad Rostami
IEEE Journal of Biomedical and Health Informatics

Cost-Efficient Prompt Engineering for Unsupervised Entity Resolution in the Product Matching [pdf]
Navapat Nananukul, Khanin Sisaengsuwancha, Mayank Kejriwal
Discover Artificial Intelligence

HALO: An Ontology for Representing and Categorizing Hallucinations in Large Language Models [pdf]
Navapat Nananukul, Mayank Kejriwal
SPIE 2024: Disruptive Technologies in Information Sciences

Demos and Dataset Articles

The Plausibility Machine Commonsense (PMC) Dataset: A Massively Crowdsourced Human-Annotated Dataset for Studying Plausibility in Large Language Models [pdf]
Navapat Nananukul, Ke Shen, Mayank Kejriwal
Elsevier Data in Brief

ISAC: An Interactive Hierarchical Interface for Efficient Structural Analysis and Vertex Search in Complex Networks (Demo) [pdf]
Navapat Nananukul, Khanin Sisaengsuwanchai, Mayank Kejriwal
The 2024 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2024)

Work Experience

Agoda International (USA) New York City, NY
Senior Data Analyst/Scientist *Sep. 2018 – July 2021*

- **Automation:** Built and automated internal data preprocessing tools, including EDA, ETL processes, and data visualization for the management team.
- **Machine Learning:** Developed statistical and machine learning models to predict sales performance, market management targets, and potential opportunities.
- **Experimentation:** Evaluated and measured the impact of strategic initiatives through experimentation and A/B testing. Built success metrics and data models that forecast market trends.

Agoda Bangkok, Thailand
Business Intelligence Developer *Jun. 2017 – Sep. 2018*

- **BI Product:** Provided end-to-end BI product and data engineer solutions, including ETL processes, analysis, OLAP, and visualization using Tableau.

Other Projects

Socpify: Knowledge graph for European Soccer with News Clarification Feature [slides]

- Performed end-to-end knowledge graph building, including crawling, entity resolution, ontology design, visualization, and UI development for a knowledge graph application.
- Built a European soccer player knowledge graph for our web application, which suggests and clarifies soccer jargon in news articles.

Parkinson's Disease Detection using CNN-LSTM Model for Time-series Keystroke Data [slides]

- Proposed a CNN-LSTM model that outperforms baseline models, including SqueezeNet, MobileNet, and AlexNet, in predicting Parkinson's disease.
- Proposed a solution for imbalanced data by performing time-series subsequence undersampling, achieving better performance compared to SMOTE.

Technical Skills

Languages: R, Python, SQL, C/C++, R, Git, HTML, CSS
Frameworks: PyTorch, TensorFlow, CUDA, Pandas, NumPy
Annotation: Amazon Mechanical Turk, Prolific
Libraries: Pandas, NumPy, Matplotlib, SciPy, spaCY, NLTK, KGTK