Intro. Of Chris and some highlights
AI4BI Workshop Part 2

Prof. Wen-Chih Peng
Dept. of Computer Science
National Ying Ming Chiao Tung University

Bio-Sketch - Wen-Chih Peng

- **Ph.D., Department of Electrical Engineering, National Taiwan University**
- **Professional Experiences**
  - Chairman of Dept. of Computer Science, NYCU (2019/8 - 2022/7)
  - Director of E-SUN-NCTU AI/Fintech Center (2019/8 - now)
  - Vice Director of Digital Medicine Center (2021/2 - 2022/7)
  - Director of Institute of Multimedia Engineering (2016/8 - 2019/7)
  - Visiting professor at LARC, SMU, Signapore (2018/1 - 2018/2)
  - Visiting professor at ITEE, University of Queensland, Australia (2014/7-2015/7)
  - Visiting professor at CS of UIUC, USA (2009/7 - 2009/11)
- **Research Interests**
  - Data mining, Machine learning/Deep learning, Databases/Data management, AI

Honors and Awards

- Receipt of Y. Z. Hsu Scientific Paper Award (有庠科技論文獎) in the category of Artificial Intelligence , 2023.
- Receipt of Future Tech. MOST, 2021
  - **Design and Implementation of Mining Purchasing Datasets for Purchasing Behavior Prediction and EDM Subject Generation**
- Receipt of Outstanding professor of Electrical Engineering (Chinese Institute of Electrical Engineering), 2020
- Receipt of **K. T. Li Breakthrough Award, Institute of Information & Computing Machinery, 2019 (李國鼎穿石獎)**
- **Receipt of MSRA Collaborative Research Grant Awards, 2018**
  - **Crowd-flow: Discovering and Inferring Crowd Transportation Modes via Cellular Data**
- Receipt of Ta-You Wu Memorial Award, NSC (Taiwan), 2012
- Receipt of Outstanding Electrical Engineers (Chinese Institute of Electrical Engineering), 2011

Some Selected Projects (Led by Wen-Chih Peng)

Project titles	Time	Sponsors
Exploring AI in Fintech Services	2018-12 - Now	E-Sun Bank
Design and Implementation of Mining Multi-dimensional Fintech Cubes from Fintech Datasets	2017/6 - 2021/7	MOST
Next-Generation Smartphone: Intelligent User Behavior Sensing, Mining, and Visualization	2013/01 - 2014/12	hTC
Mining Application Usage Patterns on Smart Phones for Application Recommendation	2011/01 - 2011/12	hTC
Mining User Latent Intention from Browsing Trajectories for Recommendation in Mobile EC App	2013/01 - 2013/12	Yahoo

About Our Lab (2018-2024, >20 papers in AAAI/KDD/EMNLP/ICDE/CIKM)

Badminton Coach AI

Intent Detection

Real Estate Appraisal

- AI Detection and Causality Inferences for Fab Operations
- AI Causality Inferences for TSMC Operations

Fake News Detection

Intelligent Medical Service System

FinTech - Real Estate Appraisal

Why do we need Real Estate Appraisal?

Customer

Apply for
House Mortgage

House Mortgage

Bank/Financial Institution

Ask for
Real Estate Appraisal

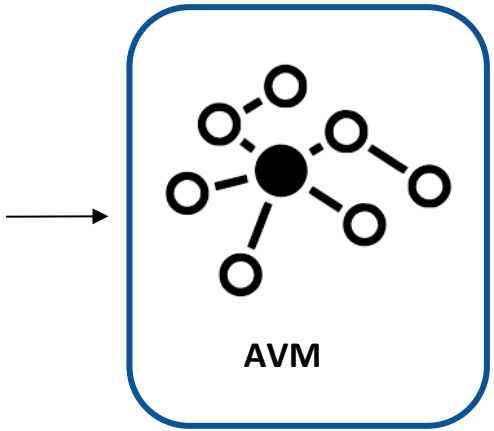
Real Estate Value

Appraiser

Real Estate Appraisal

Input

- Real Estate Transactions
  - house size
  - house floor
  - house age
  - city name/town name
  - …
- Point of Interest (PoI)
  - convenient store
  - school
  - supermarket



Output: **Unit Price**

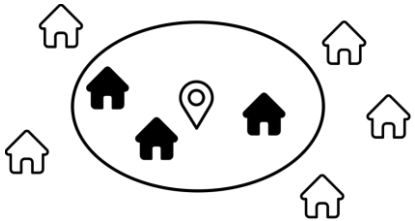
Previous AVMs for Real Estate Appraisal

Machine Learning-based Method

Linear Regression, Support Vector Regression, Boosted Regression Trees

Graph-based Method

MugRep (KDD 2021) [1], ReGram (AAAI 2023) [2]



[1] Zhang, Weijia, et al. "MugRep: A Multi-Task Hierarchical Graph Representation Learning Framework for Real Estate Appraisal." SIGKDD. 2021.
[2] Chih-Chia Li, Wei-Yao Wang, Wei-Wei Du, and Wen-Chih Peng. "Look Around! A Neighbor Relation Graph Learning Framework for Real Estate Appraisal." AAAI workshop 2023 and PAKDD 2024.

Research Direction

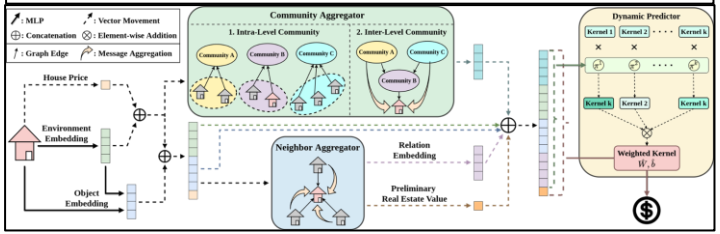
Microsoft, 2017

LightGBM

Using GNN to leverage **neighbor samples** for better accuracy and interpretability

2022.Jul （ Previous Goal ）

ReGram (AAAI'23 MUFin Best Paper)



Integrate previously developed models to create a comprehensive **house valuation system**

2023.Jul

Express

Combining Self-Supervised Learning and Few-Shot learning to improve model performance with **sparse neighbor samples**

2023.Mar

DoRA (ACM CIKM 2023)

2023.Jun

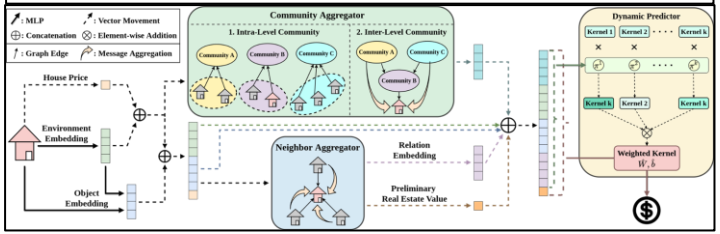
Transformer-Based x SSL

Combining **SSL** and **Transformer** to enhance model representation learning on tabular data

Research Direction

2022.Jul （ Previous Goal ）

ReGram (AAAI'23 MUFin Best Paper)



Microsoft, 2017

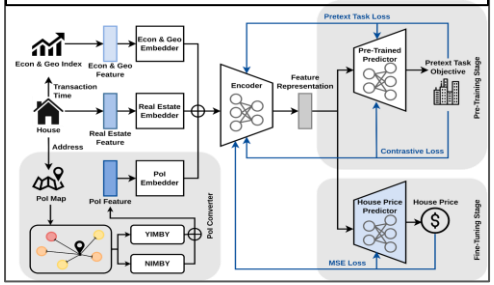
LightGBM

Using GNN to leverage **neighbor samples** for better accuracy and interpretability

Combining Self-Supervised Learning and Few-Shot learning to improve model performance with **sparse neighbor samples**

2023.Mar

DoRA (ACM CIKM 2023)



2023.Jul

Integrate previously developed models to create a comprehensive **house valuation system**

2023.Jun

Combining **SSL** and **Transformer** to enhance model representation learning on tabular data

Research Direction

Microsoft, 2017

LightGBM

Using GNN to leverage **neighbor samples** for better accuracy and interpretability

2022.Jul (Previous Goal)

ReGram (AAAI'23 MUFin Best Paper)

Integrate previously developed models to create a comprehensive **house valuation system**

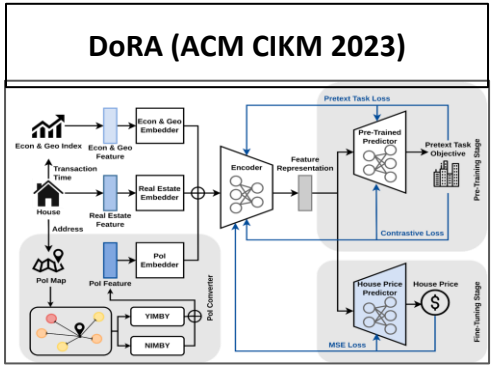
2023.Jul

Express

Combining Self-Supervised Learning and Few-Shot learning to improve model performance with **sparse neighbor samples**

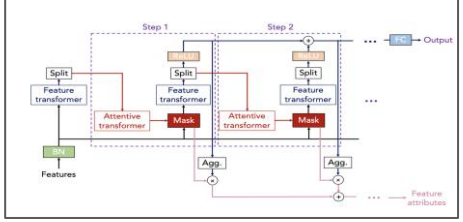
2023.Mar

DoRA (ACM CIKM 2023)



2023.Jun

Transformer-Based x SSL



Combining **SSL** and **Transformer** to enhance model representation learning on tabular data

Research Direction

Microsoft, 2017

LightGBM

Using GNN to leverage **neighbor samples** for better accuracy and interpretability

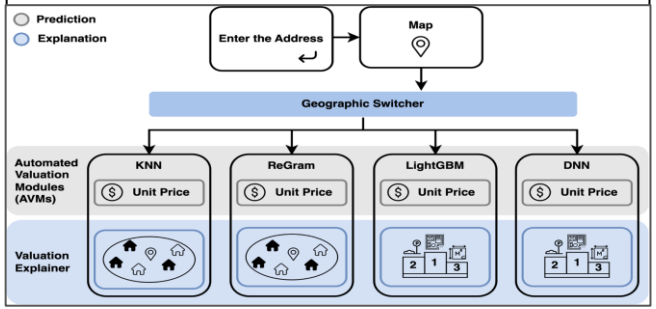
2022.Jul (Previous Goal)

ReGram (AAAI'23 MUFin Best Paper)

Integrate previously developed models to create a comprehensive **house valuation system**

2023.Jul

Express



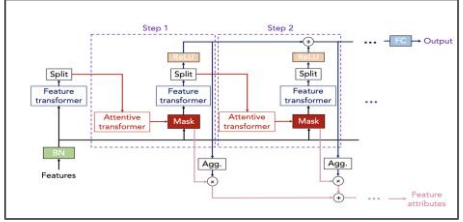
Combining Self-Supervised Learning and Few-Shot learning to improve model performance with **sparse neighbor samples**

2023.Mar

DoRA (ACM CIKM 2023)

2023.Jun

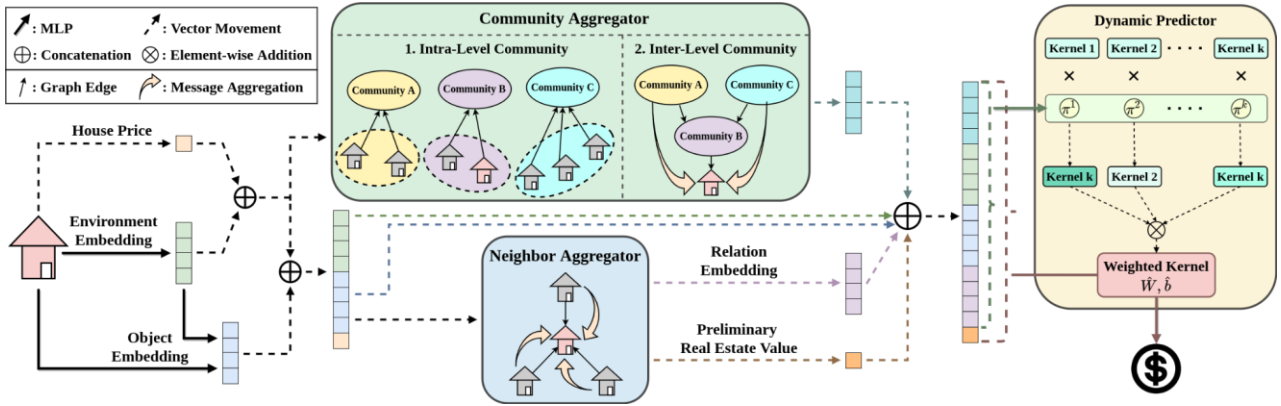
Transformer-Based x SSL



Combining **SSL** and **Transformer** to enhance model representation learning on tabular data

Real Estate Appraisal

- Our paper is accepted by AAAI-23 MUFin workshop and the extended version is accepted by PAKDD 2024.
- **Best Paper Award**



Reference: Li, Chih-Chia, et al. Look Around! A Neighbor Relation Graph Learning Framework for Real Estate Appraisal

Previous AVMs for Real Estate Appraisal

- Performance are **sensitive to the number of labeled data** and cannot tackle with low-resource scenario
- For graph-based method, **what if there are no available neighbors?**

Low-Resource Real Estate Appraisal

01. Rural Area

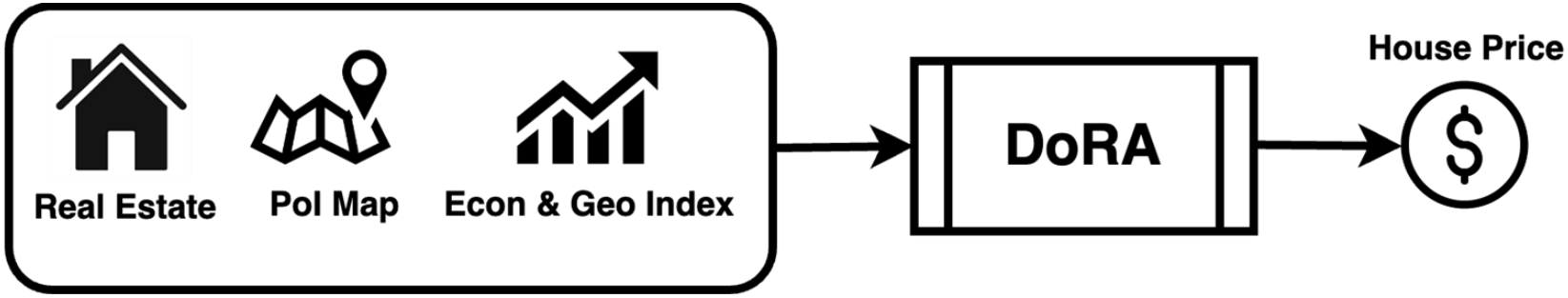
There are few neighbors in the area outside towns and cities.

02. New Developing Area

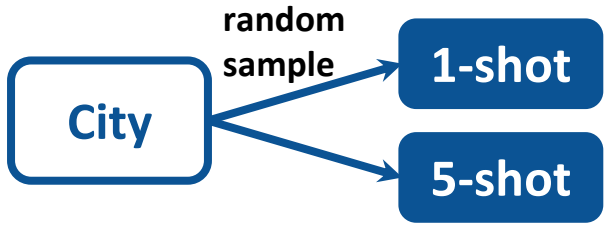
An urban planning or land readjustment area lacks historical records.

Problem Definition

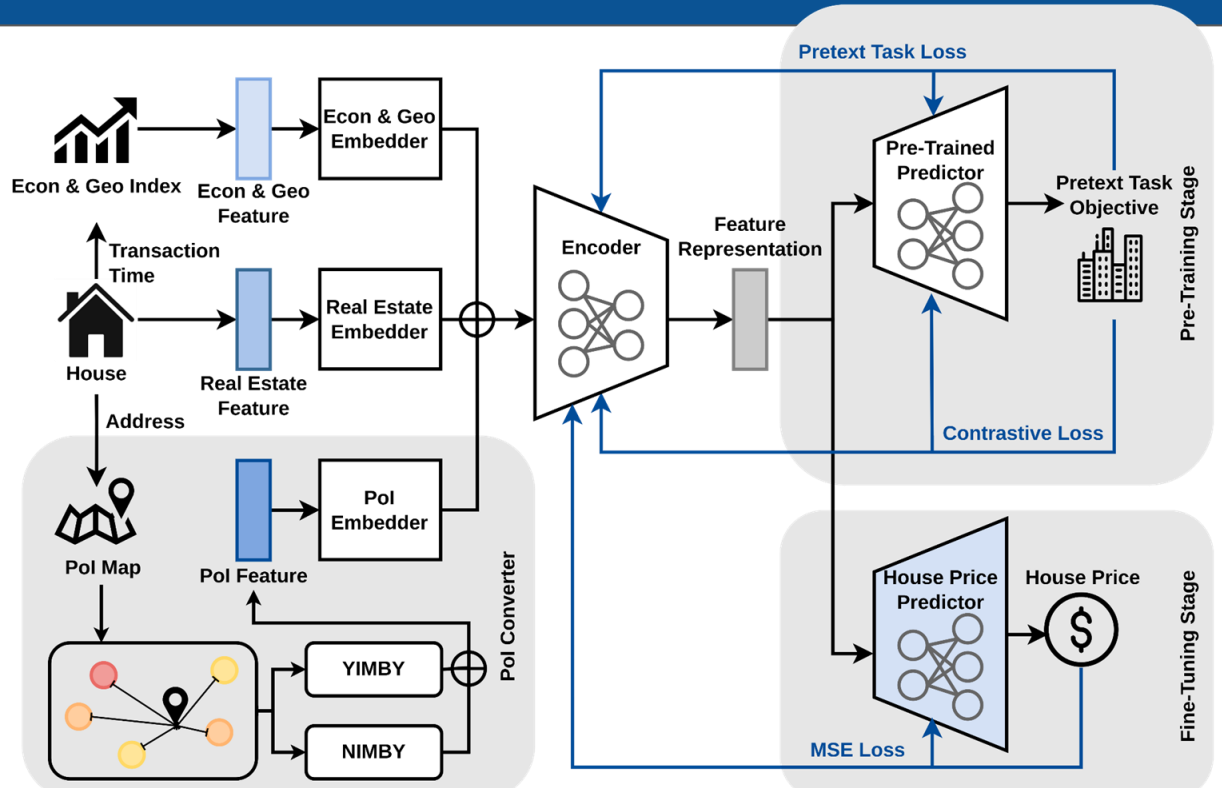
- Input: **Real Estate Features, PoI Features, Economic and Geographical Index**
- Output: **Unit Price $**



- Few-shot Setting: randomly sampled 1 and 5 shots for each city from the labeled training set as annotated examples



Proposed Framework: DoRA



Reference:

W.-W. Du. W.-Y. Wang, and W.-C. Peng, ``DoRA: Domain-Based Self-Supervised Learning Framework for Low-Resource Real Estate Appraisal," Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (ACM CIKM 2023), Birmingham United Kingdom, Oct. 21 - 25, 2023.

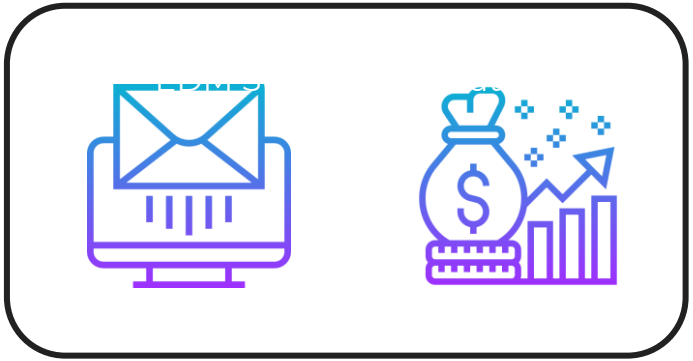
FinTech - Understand User Purchasing

Know Your Customers

Previous Project - Understand User Purchasing

1. Off-Line Behavior Mining from Purchasing data

2. Collaborative Inspired Headline Generation



Reference:

Y.-T. Wen, P.-W. Yeh, W.-C. Peng and H.-H. Shuai,``Customer Purchase Behavior Prediction from Payment Datasets," Proceedings of the 11th ACM International Conference on Web Search and Data Mining (ACM WSDM), Los Angeles, California, USA, Feb. 6-8, 2018.

Y.-H. Chen, P.-Y. Chen, H.-H. Shuai, and W.-C. Peng, "TemPEST: Soft Template-based Personalized EDM Subject Generation Through Collaborative Summarization," Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI), 2020.

Yu Ting Wen, Hui-Kuo Yang, Wen-Chih Peng, ``Mining Willing-to-Pay Behavior Patterns from Payment Datasets.," ACM Trans. on Intelligent Systems and Technology, 2022.

Payment Datasets

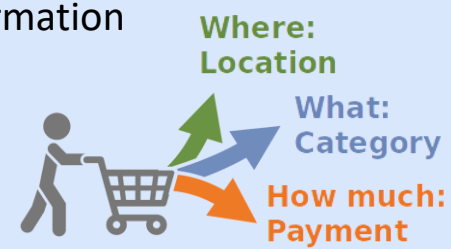
Payment record

- Time
- Category
- Store/Places
- Amount
- Rewards

Mining Purchasing Behavior of Users

Step 1
User Profiling

Understand the customer by profiling the time, location, category, payment information



Step 2
Next Step Prediction
(Where, What, How much)

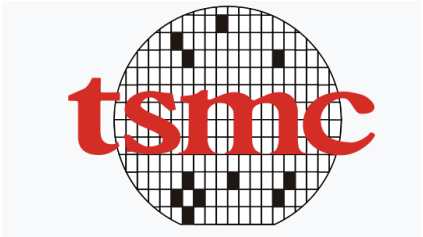
Propose a **probabilistic graphical model** that explores credit card's characteristics to predict the next step (location, category, payment) of customers

Reference:
Y.-T. Wen, P.-W. Yeh, W.-C. Peng and H.-H. Shuai,``Customer Purchase Behavior Prediction from Payment Datasets," Proceedings of the 11th ACM International Conference on Web Search and Data Mining (ACM WSDM), Los Angeles, California, USA, Feb. 6-8, 2018.
Yu Ting Wen, Hui-Kuo Yang, Wen-Chih Peng, ``Mining Willing-to-Pay Behavior Patterns from Payment Datasets.," ACM Trans. on Intelligent Systems and Technology, 2022.

AI for Industry Applications

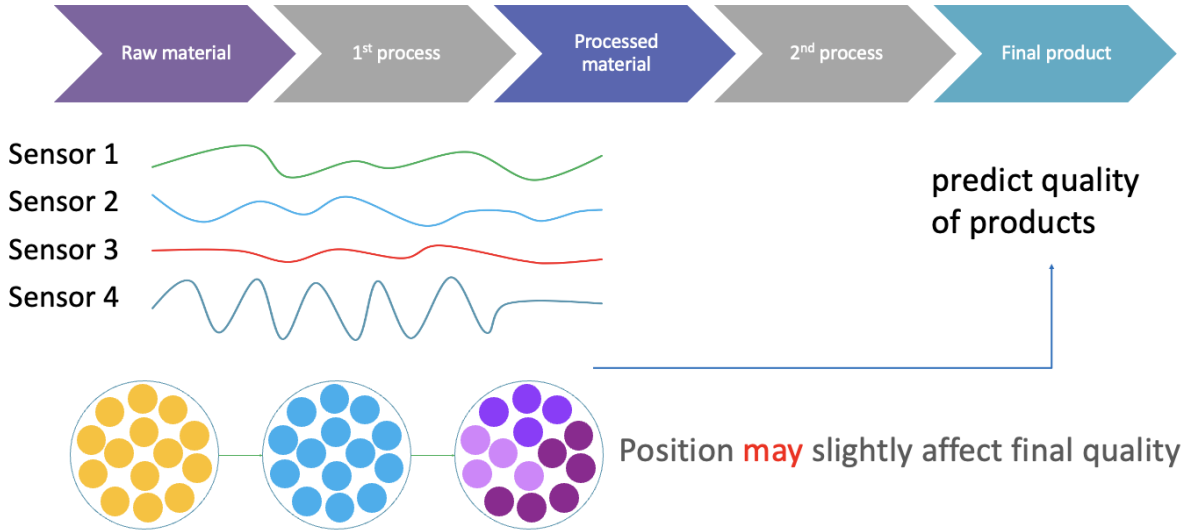
Previous Projects-related to industries

- Anomaly detection in manufacturing datasets
- Product quality prediction
- Root cause analysis in microservices



Product quality prediction

- **Early stage quality prediction**
- **Interpretable model and intermediate output**



C.-H. Yen, Y.-C. Fan and W.-C. Peng,``Interpretable Multi-Task Learning for Product Quality Prediction with Attention Mechanism," Proceedings of the 35th IEEE International Conference on Data Engineering (ICDE), Macau SAR, China, April 8-12, 2019.
H.-Y. Chih, Y.-C. Fan, W.-C. Peng and H.-Y. Kuo, "Product Quality Prediction with Convolutional Encoder-Decoder Architecture and Transfer Learning," Proceedings of the 29th ACM International Conference on Information and Knowledge Management, Oct.19-23, 2020.

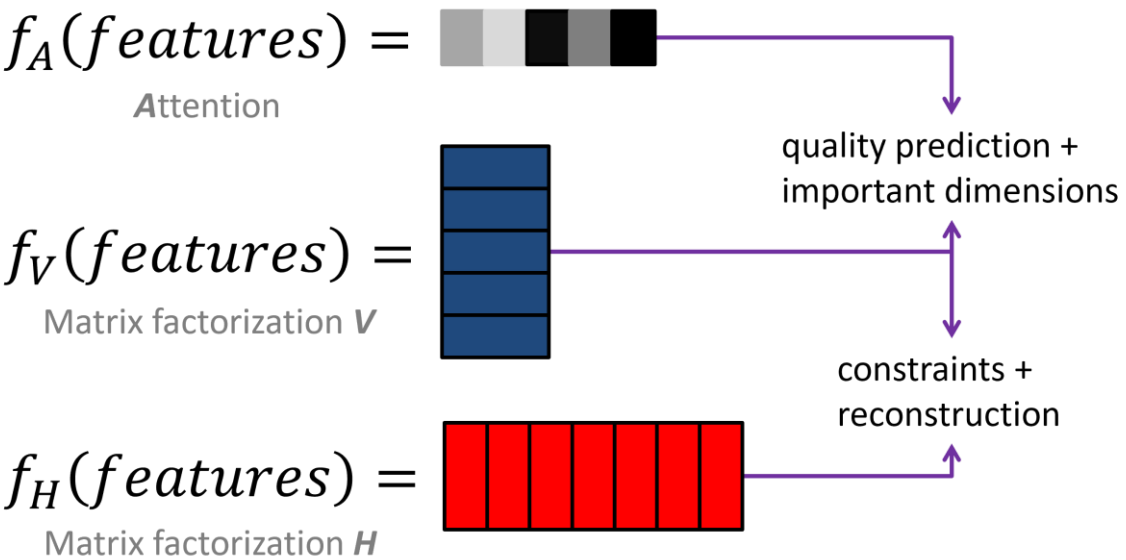
Product quality prediction

- **Interpretable multi-task learning**

Matrix Factorization	Reconstruction Quality Prediction

Similar to **Time-Series** Imputation

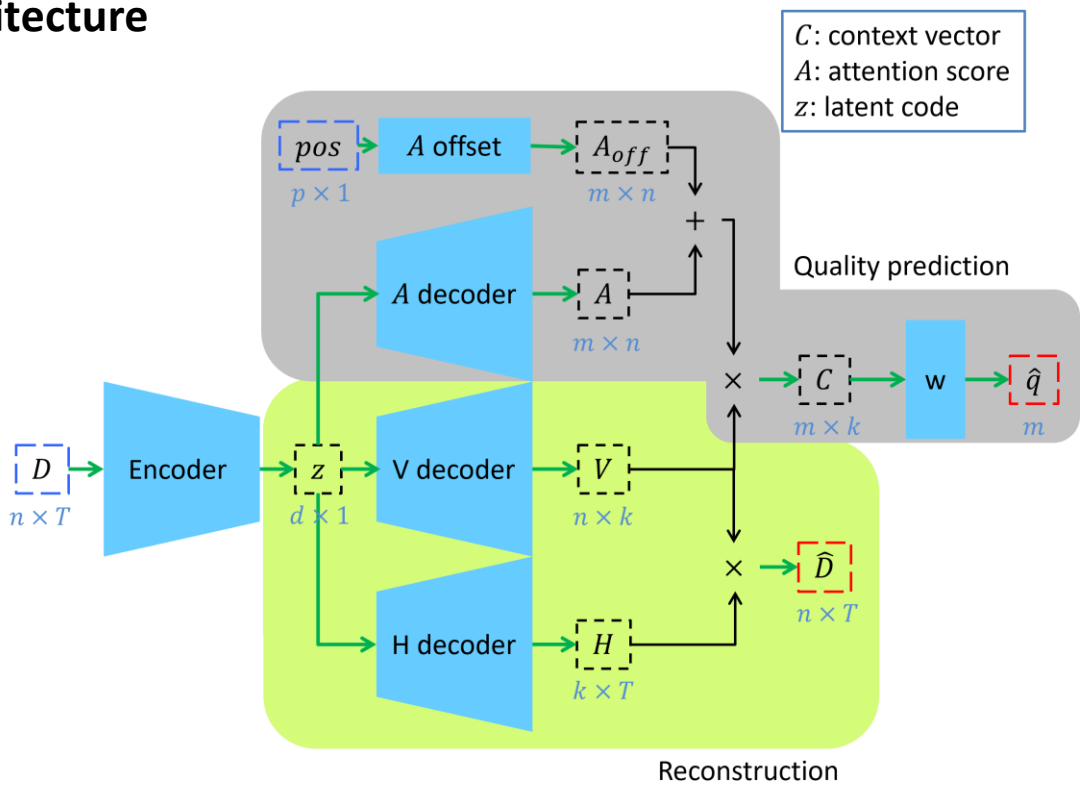
f_A, f_V, f_H are neural networks



C.-H. Yen, Y.-C. Fan and W.-C. Peng,``Interpretable Multi-Task Learning for Product Quality Prediction with Attention Mechanism," Proceedings of the 35th IEEE International Conference on Data Engineering (ICDE), Macau SAR, China, April 8-12, 2019.

Product quality prediction

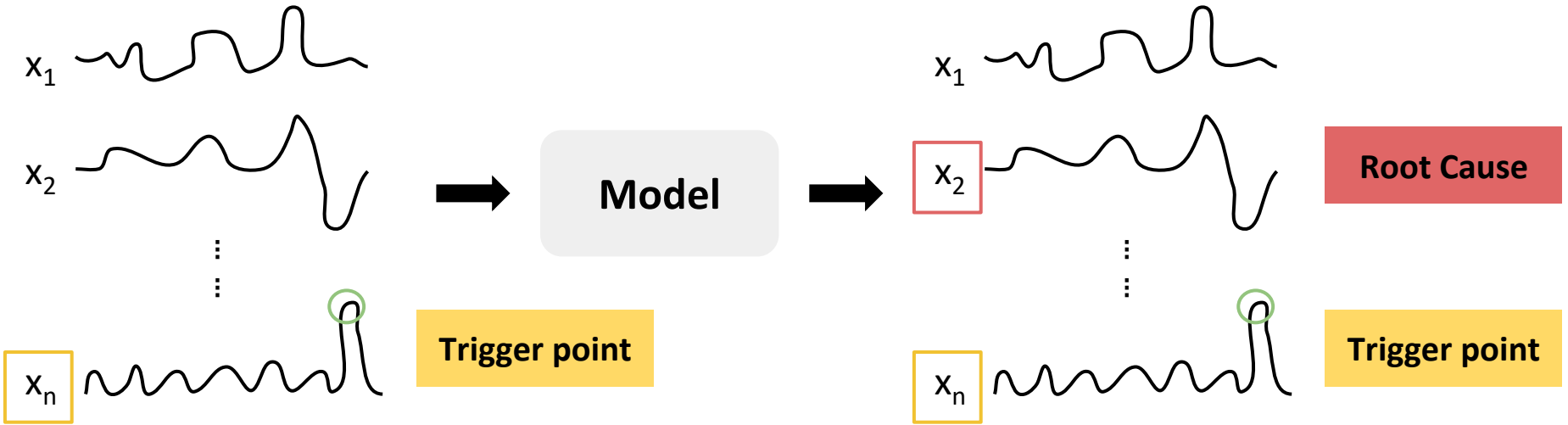
- **Model Architecture**



C.-H. Yen, Y.-C. Fan and W.-C. Peng,``Interpretable Multi-Task Learning for Product Quality Prediction with Attention Mechanism," Proceedings of the 35th IEEE International Conference on Data Engineering (ICDE), Macau SAR, China, April 8-12, 2019.

Root cause analysis

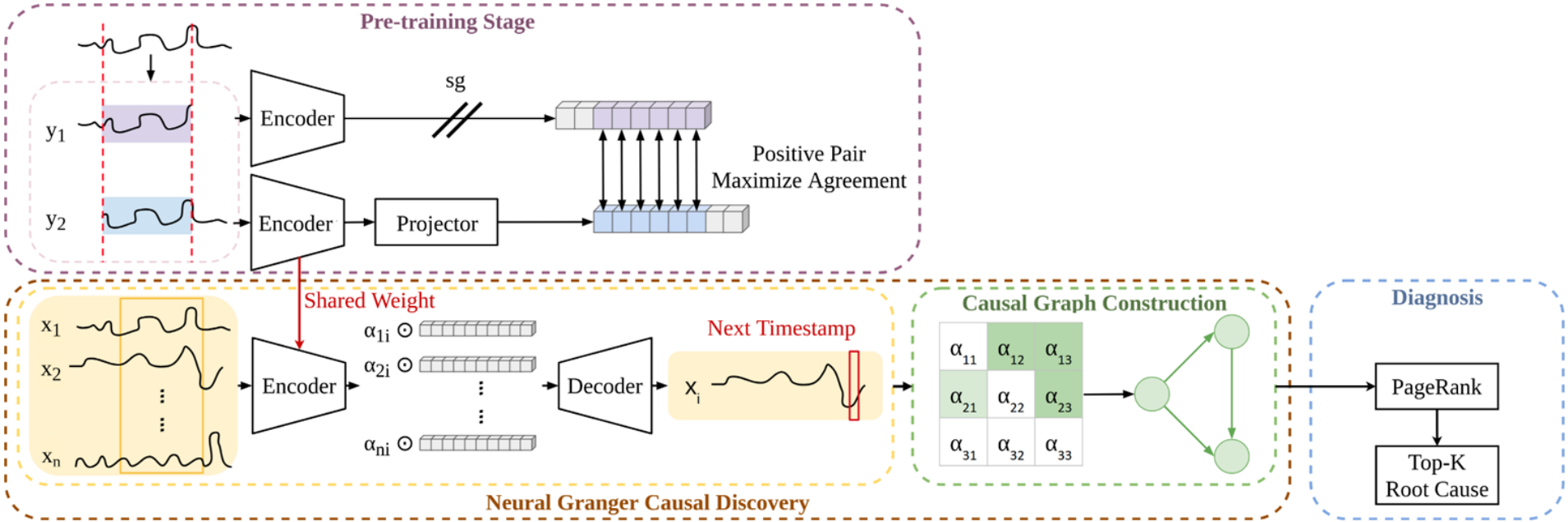
- Find the **root causes** of **trigger point** based on **multivariate time series**
  - Trigger point: the value of the variable is found to deviate from the standard



C.-M. Lin, C.-Chang, W.-Y. Wang, K.-D. Wang and W.-C. Peng, `` Root Cause Analysis In Microservice Using Neural Granger Causal Disocvery," *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, 2024.

Root cause analysis

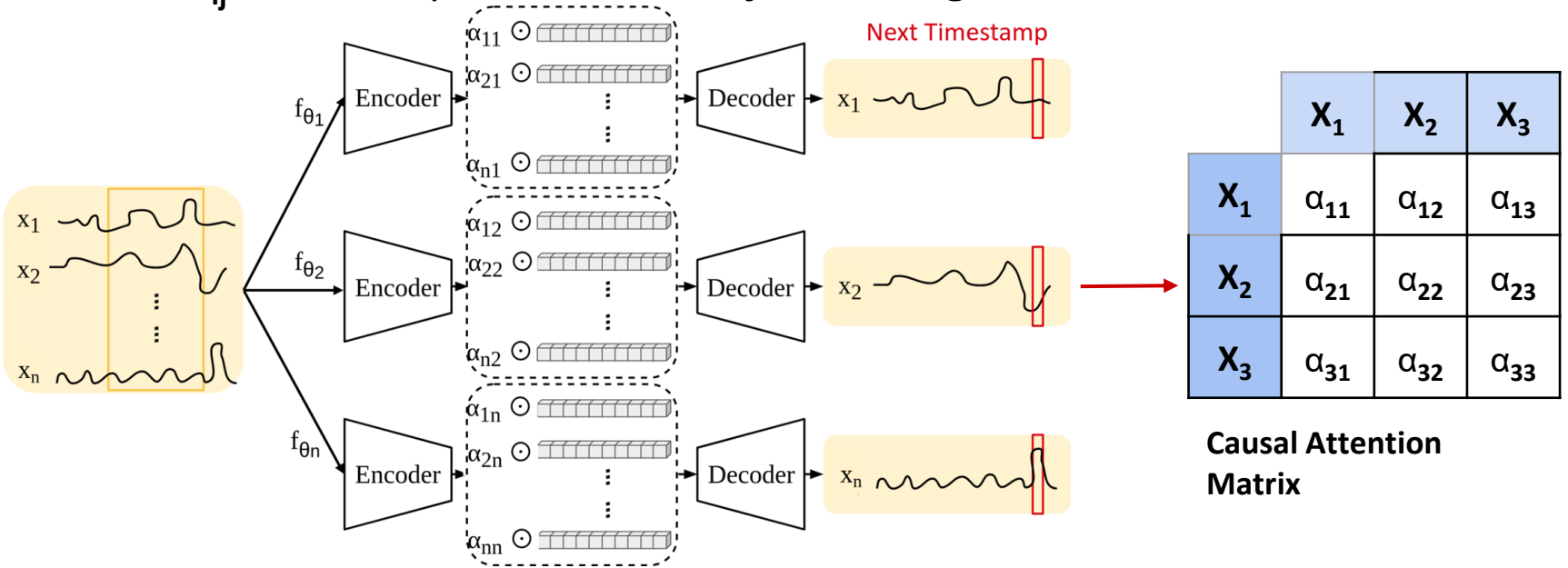
- Framework



C.-M. Lin, C.-Chang, W.-Y. Wang, K.-D. Wang and W.-C. Peng, `` Root Cause Analysis In Microservice Using Neural Granger Causal Disocvery," *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, 2024.

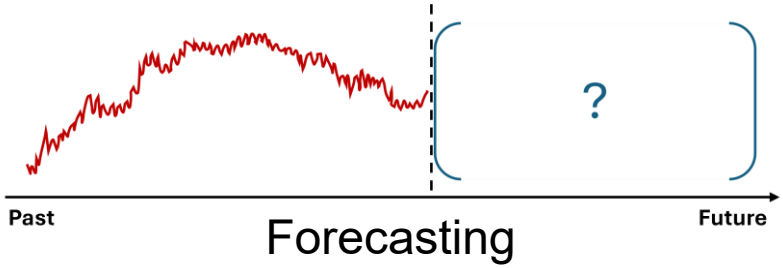
Root cause analysis – Time-Series Forecasting

- There are **N** independent neural networks with the same architecture
  - α_{ij}: **i** is the input time series, **j** is the target time series

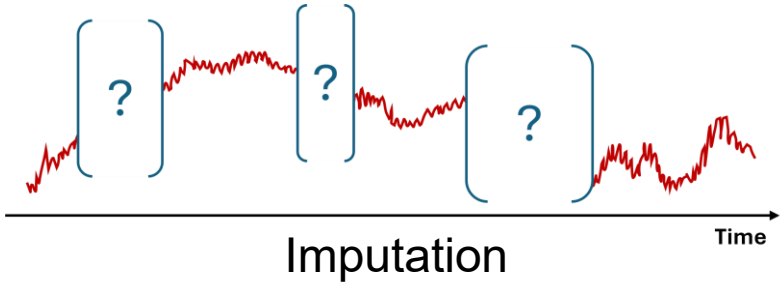


C.-M. Lin, C.-Chang, W.-Y. Wang, K.-D. Wang and W.-C. Peng, `` Root Cause Analysis In Microservice Using Neural Granger Causal Disocvery," *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, 2024.

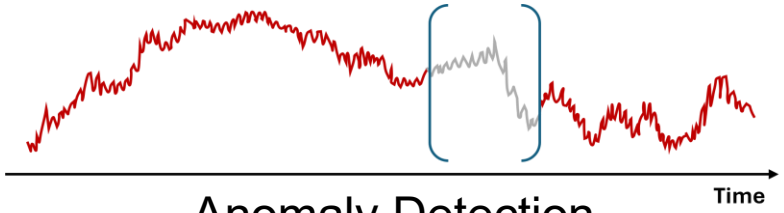
Time-Series Analysis



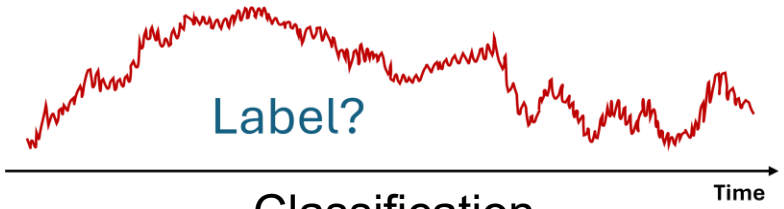
Forecasting



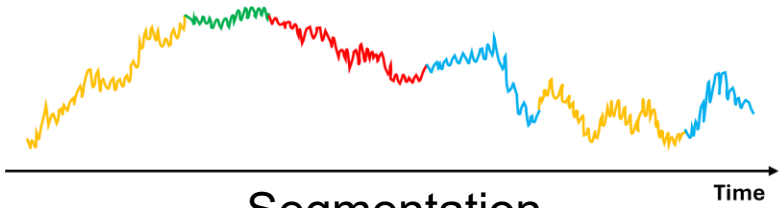
Imputation



Anomaly Detection



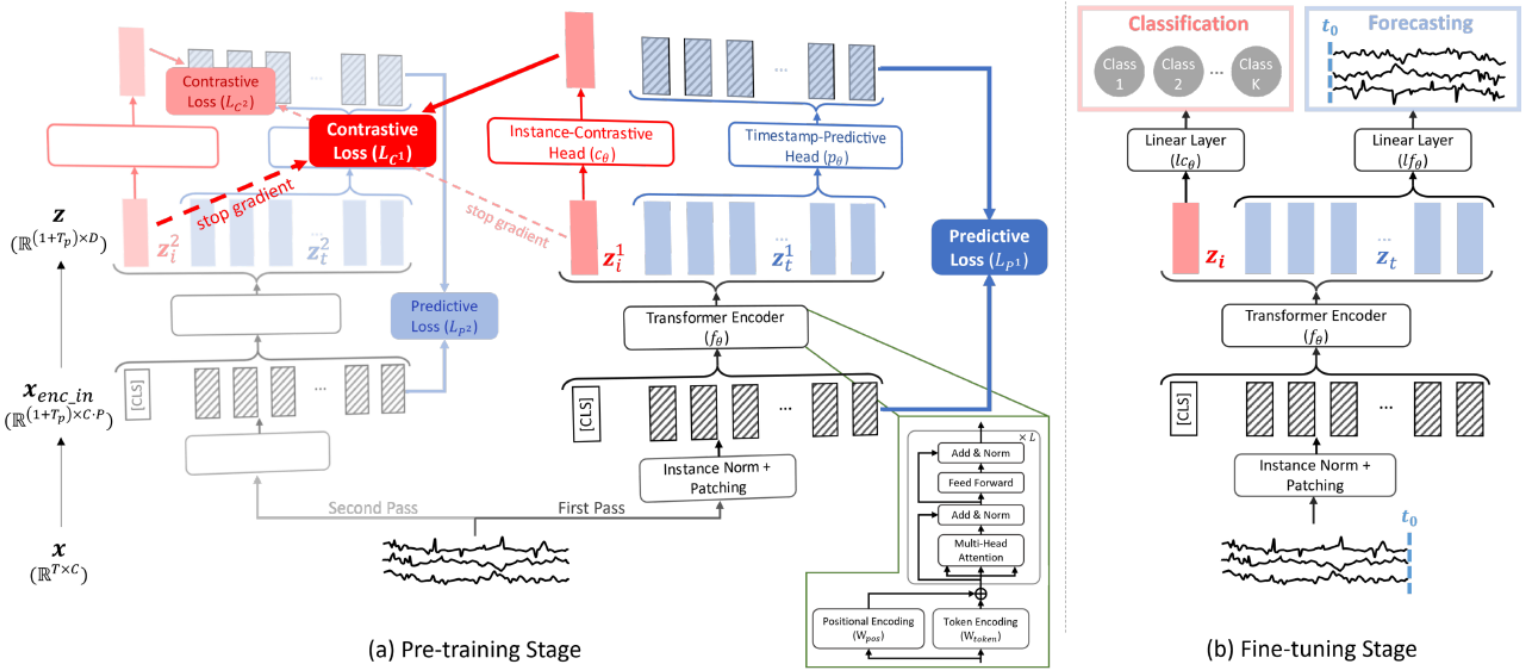
Classification



Segmentation

Self-supervised Learning with Time-Series

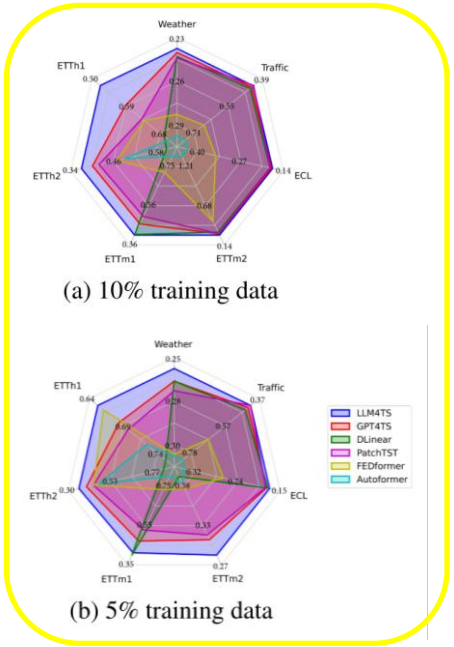
- Disentangled Representation Learning

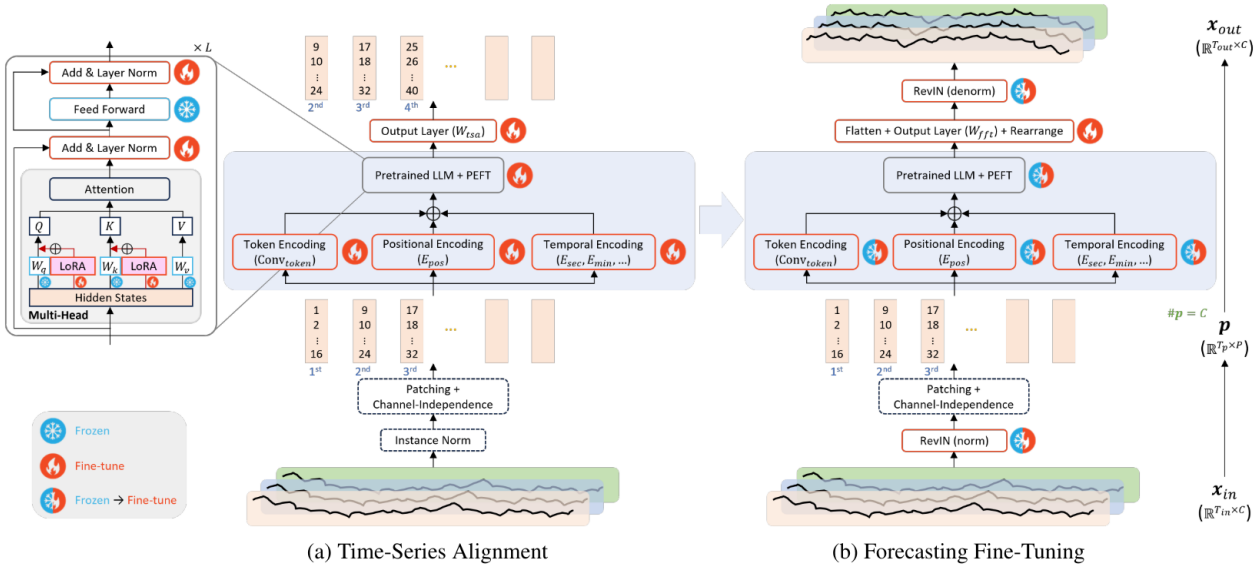


Ching Chang, Chiao-Tung Chan, Wei-Yao Wang, Wen-Chih Peng, Tien-Fu Chen, "TimeDRL: Disentangled Representation Learning for Multivariate Time-Series", ICDE 2024

Leverage LLMs in Time-Series Forecasting

- Strong few-shot learning capability





Ching Chang, Wei-Yao Wang, Wen-Chih Peng, Tien-Fu Chen, "LLM4TS: Aligning Pre-Trained LLMs as Data-Efficient Time-Series Forecasters", under review

Explain AI in simple terms

AI, or Artificial Intelligence, refers to the ability of computers or machines to perform tasks that typically require human intelligence, such as understanding language, recognizing images, making decisions, and solving problems. This is achieved by training algorithms on large amounts of data and adjusting their parameters so that they can learn to perform these tasks with a high level of accuracy. There are many different types of AI, including supervised learning, unsupervised learning, and reinforcement learning, each with its own strengths and weaknesses. Overall, the goal of AI is to create systems that can perform tasks with human-like or even superhuman intelligence.