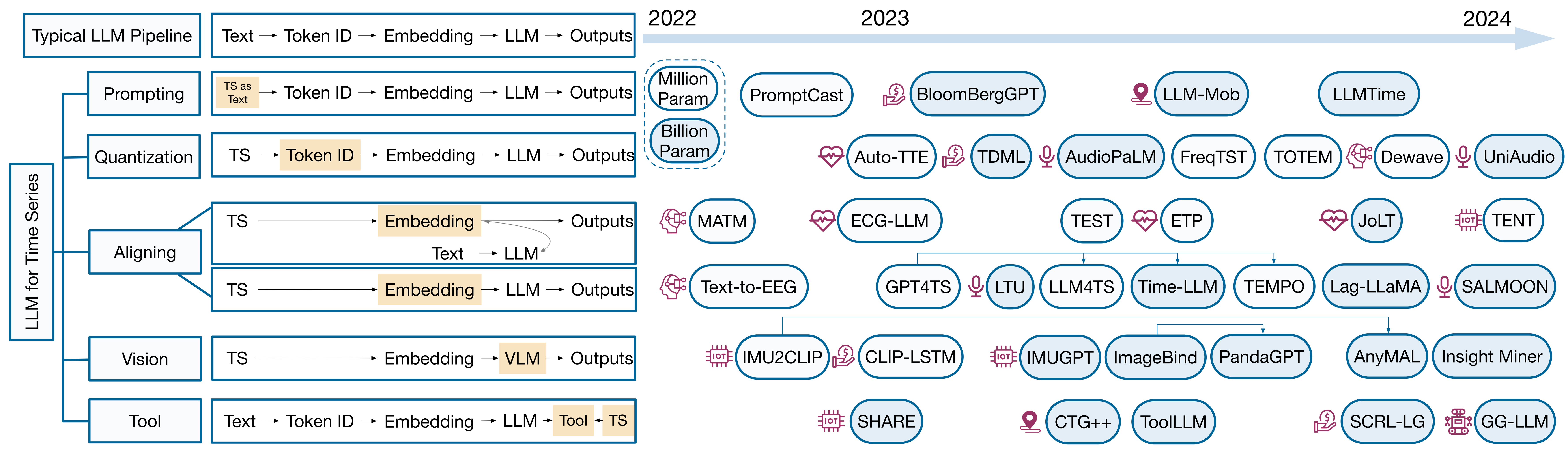
Large Language Models for Time Series: A Survey

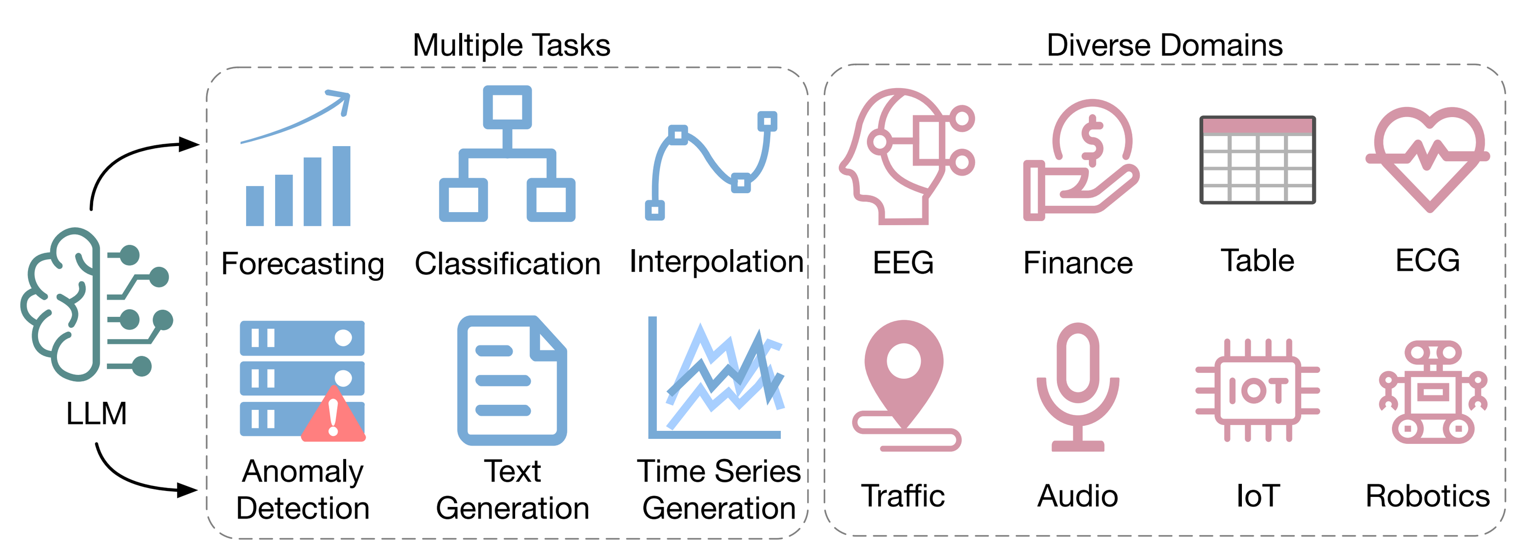
Xiyuan Zhang, Ranak Roy Chowdhury, Rajesh K. Gupta, Jingbo Shang
University of California, San Diego

IJCAI JEJU 2024



Introduction

- LLMs have significantly revolutionized NLP and CV domains.
- How can LLMs benefit time series analysis?



- Key challenge
  - How to bridge the modality gap between LLMs trained on ***discrete textual data*** and ***continuous numerical time series***?

Taxonomy

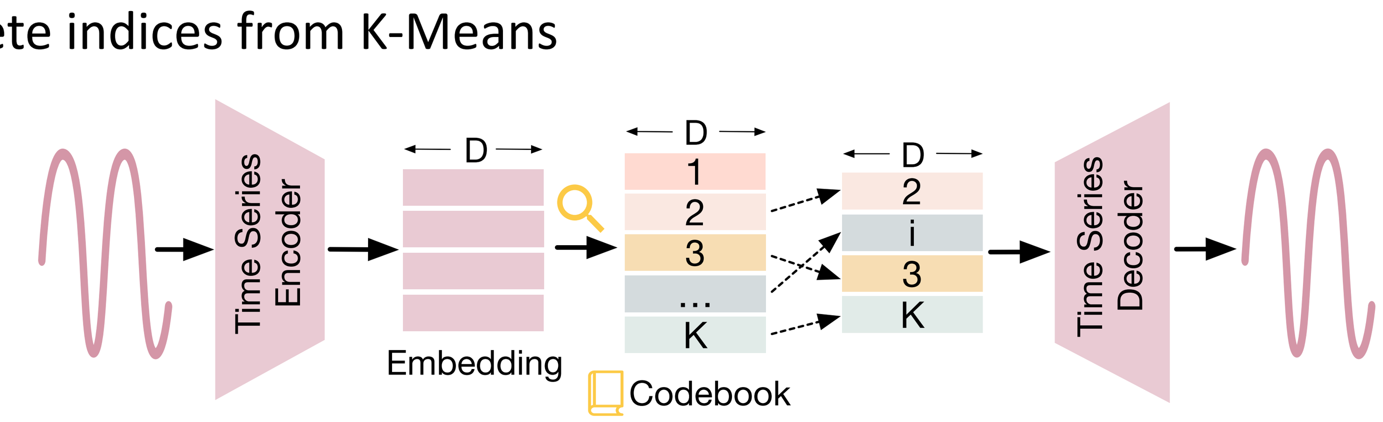
- If we outline typical LLM-driven NLP pipelines in five stages
  - Input text, tokenization, embedding, LLM, output
- Then each category of our taxonomy targets one specific stage in this pipeline

- **Prompting (Input Stage)**
  - Number-agnostic tokenization
  - Number-specific tokenization

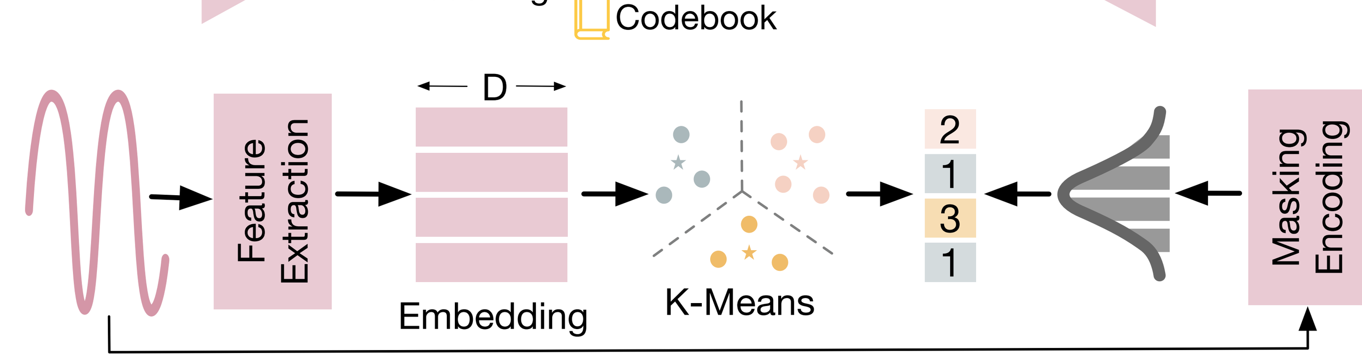
Method	Example
PromptCast [Xue and Salim, 2022]	"From $\{t_1\}$ to $\{t_{\text{obs}}\}$, the average temperature of region $\{U_m\}$ was $\{x_t^m\}$ degree on each day. What is the temperature going to be on $\{t_{\text{obs}}\}$?"
Liu *et al.* [2023d]	"Classify the following accelerometer data in meters per second squared as either walking or running: 0.052,0.052,0.052,0.051,0.052,0.055,0.051,0.056,0.06,0.064"
TabLLM [Hegselmann *et al.*, 2023]	"The person is 42 years old and has a Master's degree. She gained $594. Does this person earn more than 50000 dollars? Yes or no? Answer:"
LLMTime [Gruver *et al.*, 2023]	"0.123, 1.23, 12.3, 123.0" → "1 2 , 1 2 3 , 1 2 3 0 , 1 2 3 0 0"

- **Quantization (Tokenization Stage)**
  - Discrete indices from VQ-VAE
  - Discrete indices from K-Means

VQ-VAE

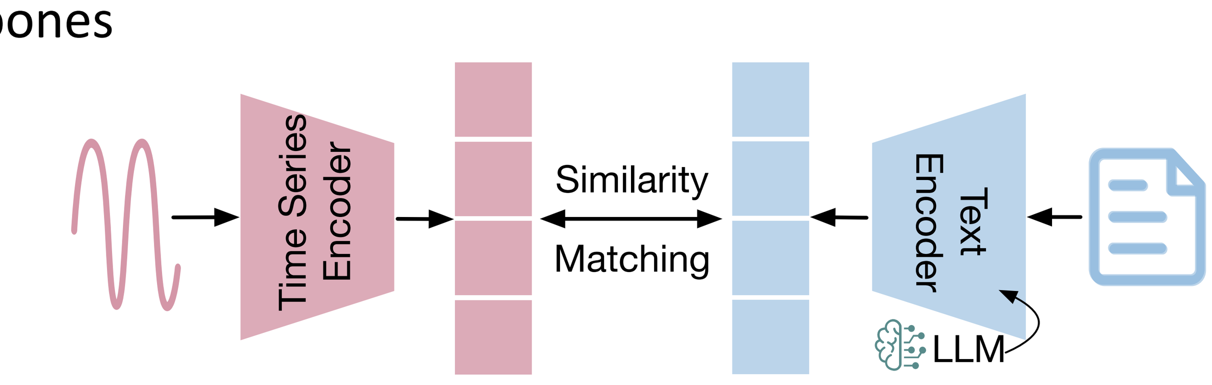


K-Means

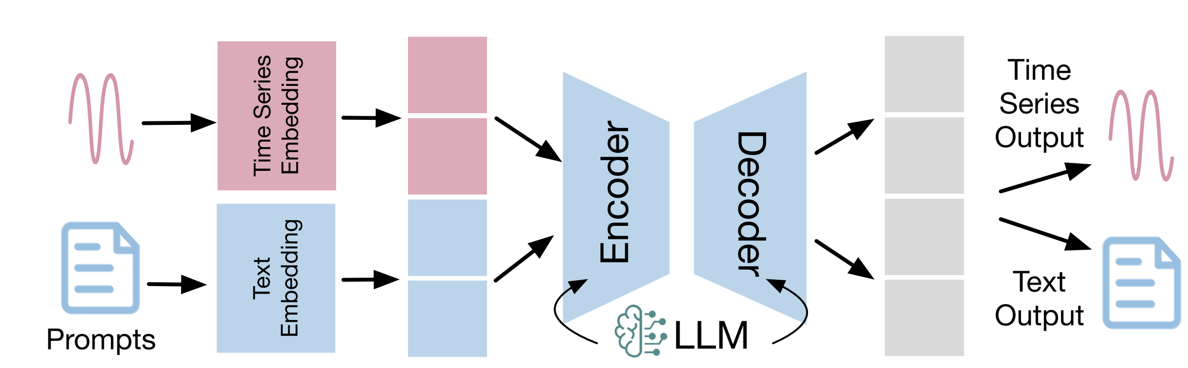


- **Aligning (Embedding Stage)**
  - Similarity matching
  - LLMs as backbones

Similarity matching



LLMs as backbones



- **Vision as Bridge (LLM Stage)**
  - Paired data, physics relationships, time series plots as images
- **LLMs as Tools (Output Stage)**
  - Code, API call, text domain knowledge

Comparison within the Taxonomy

Method	Subcategory	Representative Works	Equations	Advantages	Limitations
Prompting	Number-Agnostic	PromptCast [Xue and Salim, 2022]	$\mathbf{y} = f_\theta(\mathbf{x}_s, \mathbf{x}_t)$	easy to implement;	lose semantics;
	Number-Specific	LLMTime [Gruver *et al.*, 2023]		zero-shot capability	not efficient
Quantization	VQ-VAE	DeWave [Duan *et al.*, 2023]	$k_i = \arg\min_j \lVert g_\phi(\mathbf{x}_s)_i - \mathbf{c}_j \rVert_2$	flexibility of	may require
	K-Means	AudioLM [Borsos *et al.*, 2023]	$\mathbf{k} = [k_i]_{i=1}^{\frac{T}{S}}, \mathbf{y} = f_\theta(\mathbf{k}, \mathbf{x}_t)$	index and time	two-stage
	Text Categories	TDML [Yu *et al.*, 2023]	$\mathbf{y} = f_\theta(q(\mathbf{x}_s), \mathbf{x}_t)$	series conversion	training
Aligning	Similarity Match	ETP [Liu *et al.*, 2023a]	$\mathbf{y} = g_\phi(\mathbf{x}_s)$	align semantics of	complicated
		MATM [Han *et al.*, 2022]	$\mathcal{L} = \text{sim}(g_\phi(\mathbf{x}_s), f_\theta(\mathbf{x}_t))$	different modalities;	design and
	LLM Backbone	GPT4TS [Zhou *et al.*, 2023a]	$\mathbf{y} = f_\theta(g_\phi(\mathbf{x}_s), \mathbf{x}_t)$	end-to-end training	fine-tuning
Vision as Bridge	Paired Data	ImageBind [Girdhar *et al.*, 2023]	$\mathcal{L} = \text{sim}(g_\phi(\mathbf{x}_s), h_\psi(\mathbf{x}_v))$	additional visual	not hold
	TS Plots as Images	Wimmer and Rekabsaz [2023]	$\mathbf{y} = h_\psi(\mathbf{x}_s)$	knowledge	for all data
Tool	Code	CTG++ [Zhong *et al.*, 2023]	$z = f_\theta(\mathbf{x}_t)$	empower LLM	optimization
	API	ToolLLM [Qin *et al.*, 2023]	$\mathbf{y} = z(\mathbf{x}_s)$	with more abilities	not end-to-end

Datasets

- Internet of Things (IMU), healthcare (EEG, ECG), finance (stock), audio/music/speech

Domain	Dataset	Size	Major Modalities	Task
Internet of Things	Ego4D[2] [Grauman *et al.*, 2022]	3, 670h data, 3.85M narrations	text, IMU, video, audio, 3D	classification, forecasting
	DeepSQA[3] [Xing *et al.*, 2021]	25h data, 91K questions	text, imu	classification, question answering
Finance	PIXIU[4] [Xie *et al.*, 2023b]	136K instruction data	text, tables	5 NLP tasks, forecasting
	MoAT[5] [Lee *et al.*, 2023]	6 datasets, 2K timesteps in total	text, time series	forecasting
Healthcare	Zuco 2.0[6] [Hollenstein *et al.*, 2019]	739 sentences	text, eye-tracking, EEG	classification, text generation
	PTB-XL[7] [Wagner *et al.*, 2020]	60h data, 71 unique statements	text, ECG	classification
	ECG-QA[8] [Oh *et al.*, 2023]	70 question templates	text, ECG	classification, question answering
Audio	OpenAQA-5M[9] [Gong *et al.*, 2023]	5.6M (audio, question, answer) tuples	text, audio	tagging, classification
Music	MusicCaps[10] [Agostinelli *et al.*, 2023]	5.5K music clips	text, music	captioning, generation
Speech	CommonVoice[11] [Ardila *et al.*, 2019]	7, 335 speech hours in 60 languages	text, speech	ASR, translation

Resources

- Check out our paper and Github repo (**awesome-llm-time-series**) for more details!

Paper

Github