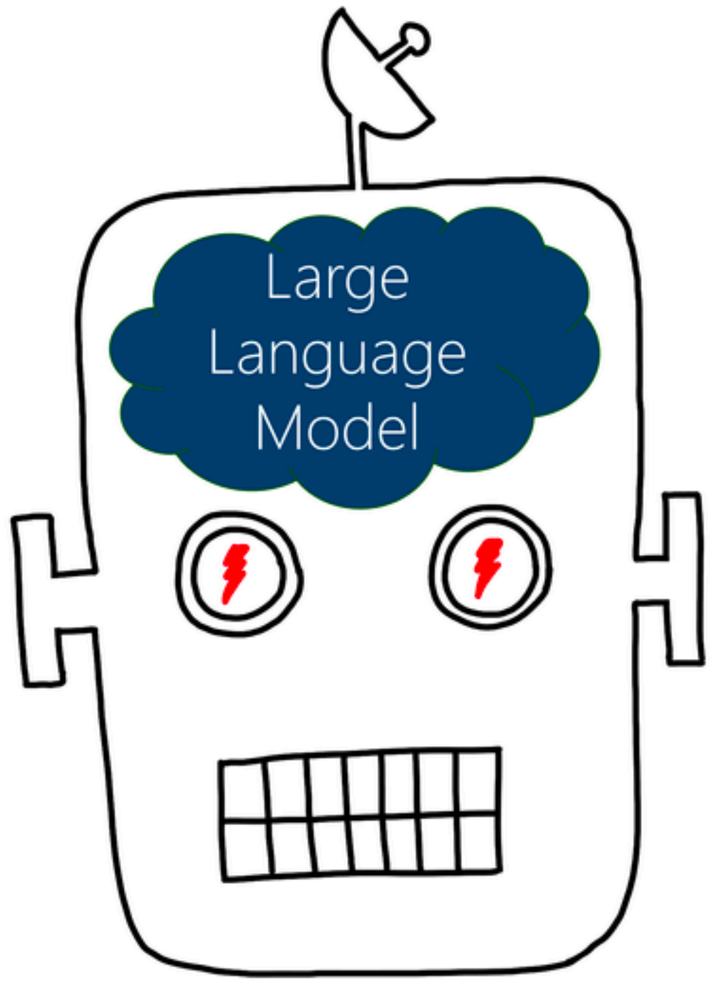
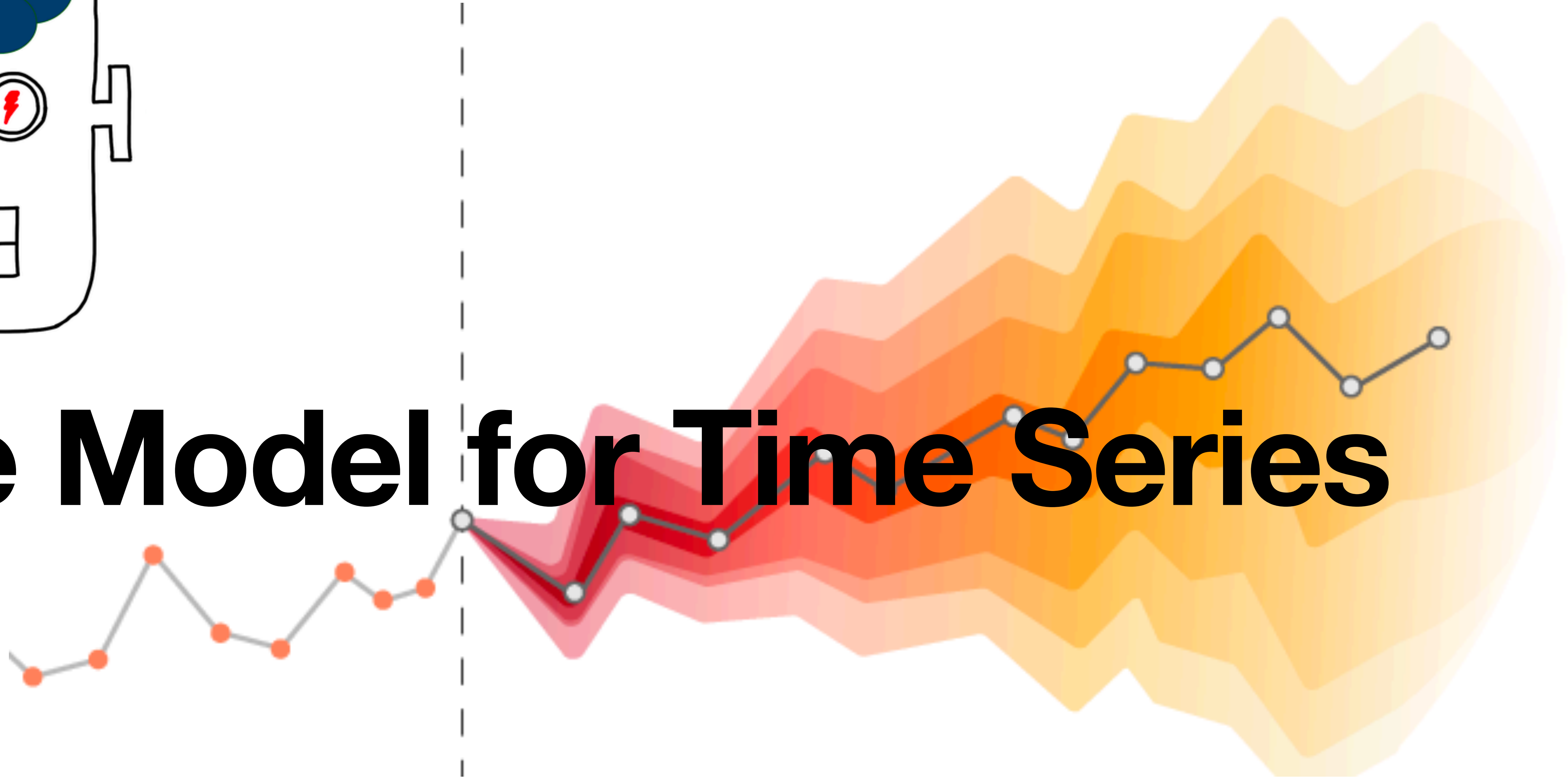




Large Language Model for Time Series

Survey and Outlook

Xiyuan Zhang, Ranak Roy Chowdhury, Rajesh K. Gupta, Jingbo Shang

University of California, San Diego

Outline

- Introduction
- Taxonomy
- Datasets
- Future Directions
- Conclusion

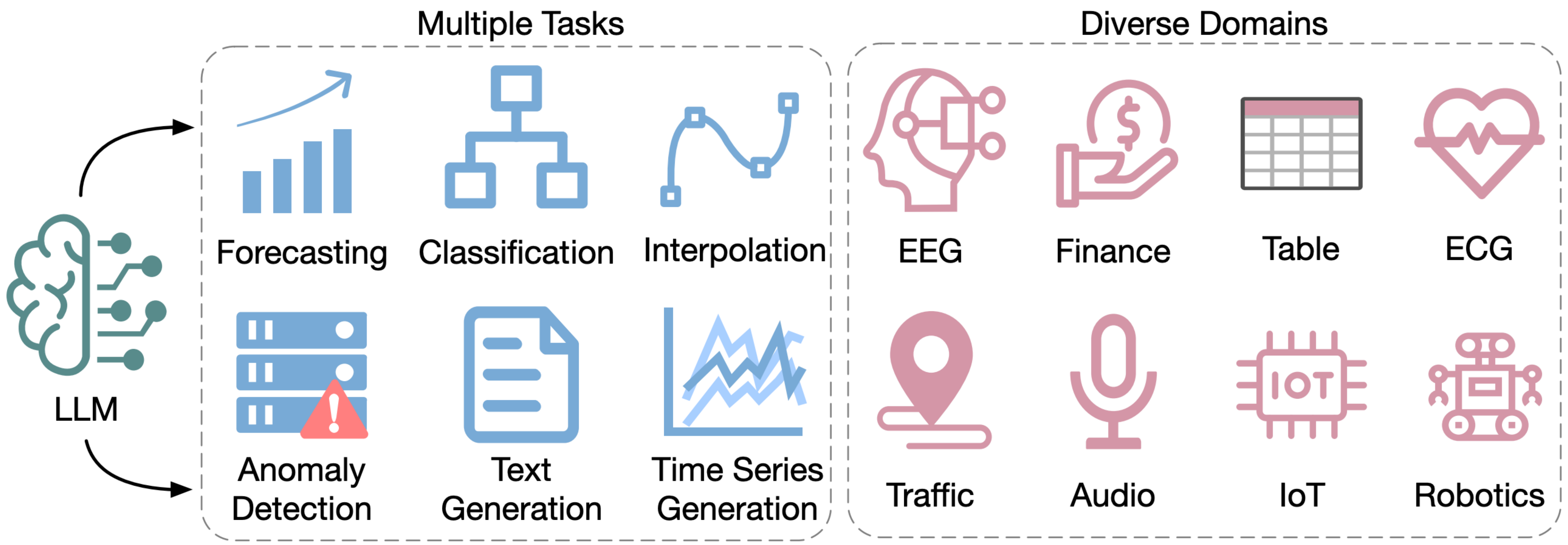
Outline

- Introduction
- Taxonomy
- Datasets
- Future Directions
- Conclusion

Introduction

From Text to Multimodal Analysis

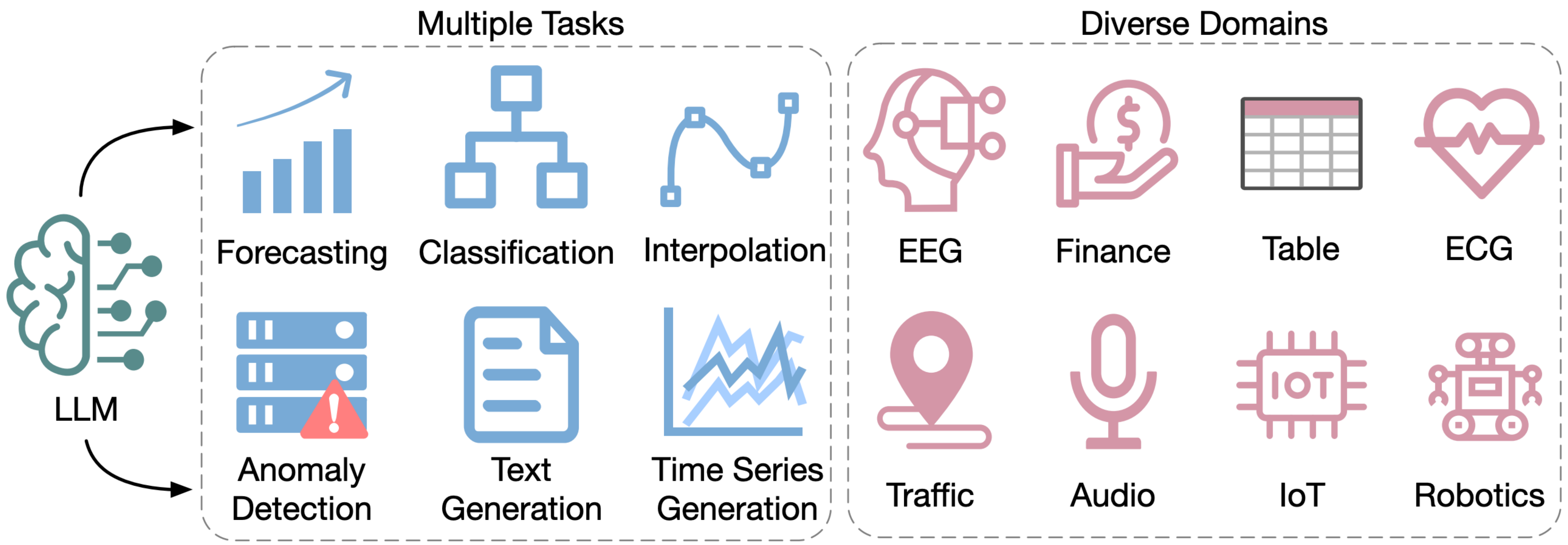
- LLMs have significantly revolutionized NLP and CV domains
- How can LLMs benefit time series analysis?

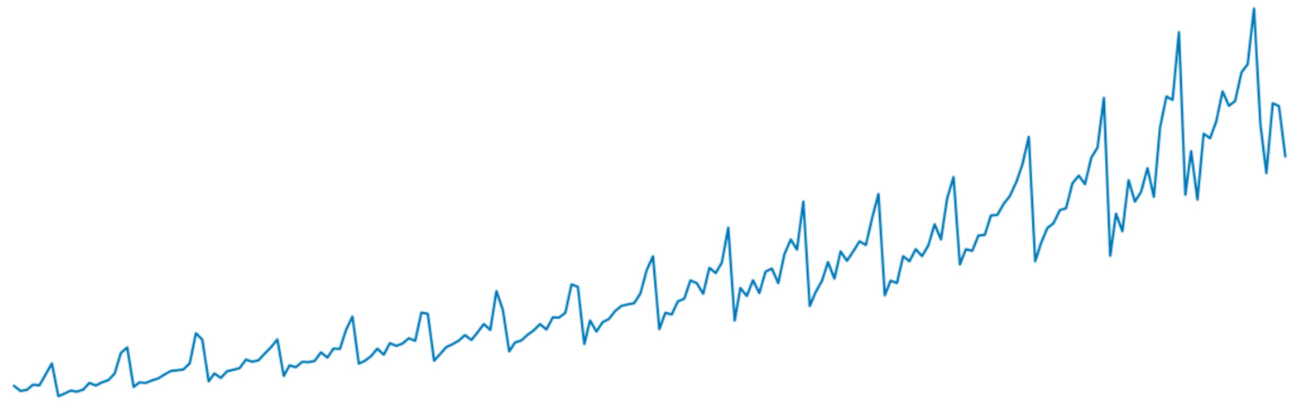


Introduction

From Text to Multimodal Analysis

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





Time series are **continuous** numerical data

How to bridge the modality gap?

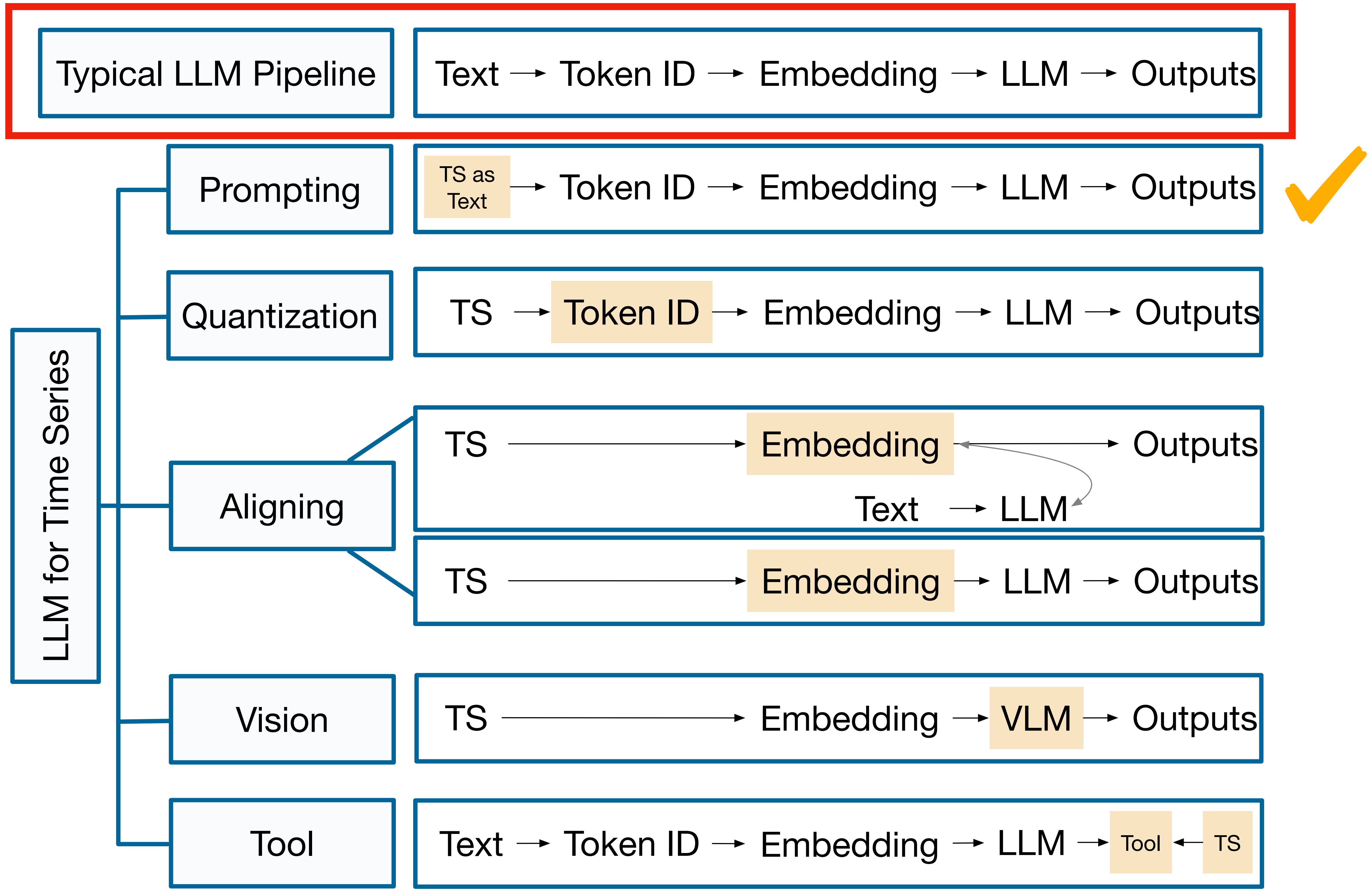
Today Is Wed

Large language models are originally trained on **discrete** text data

Introduction

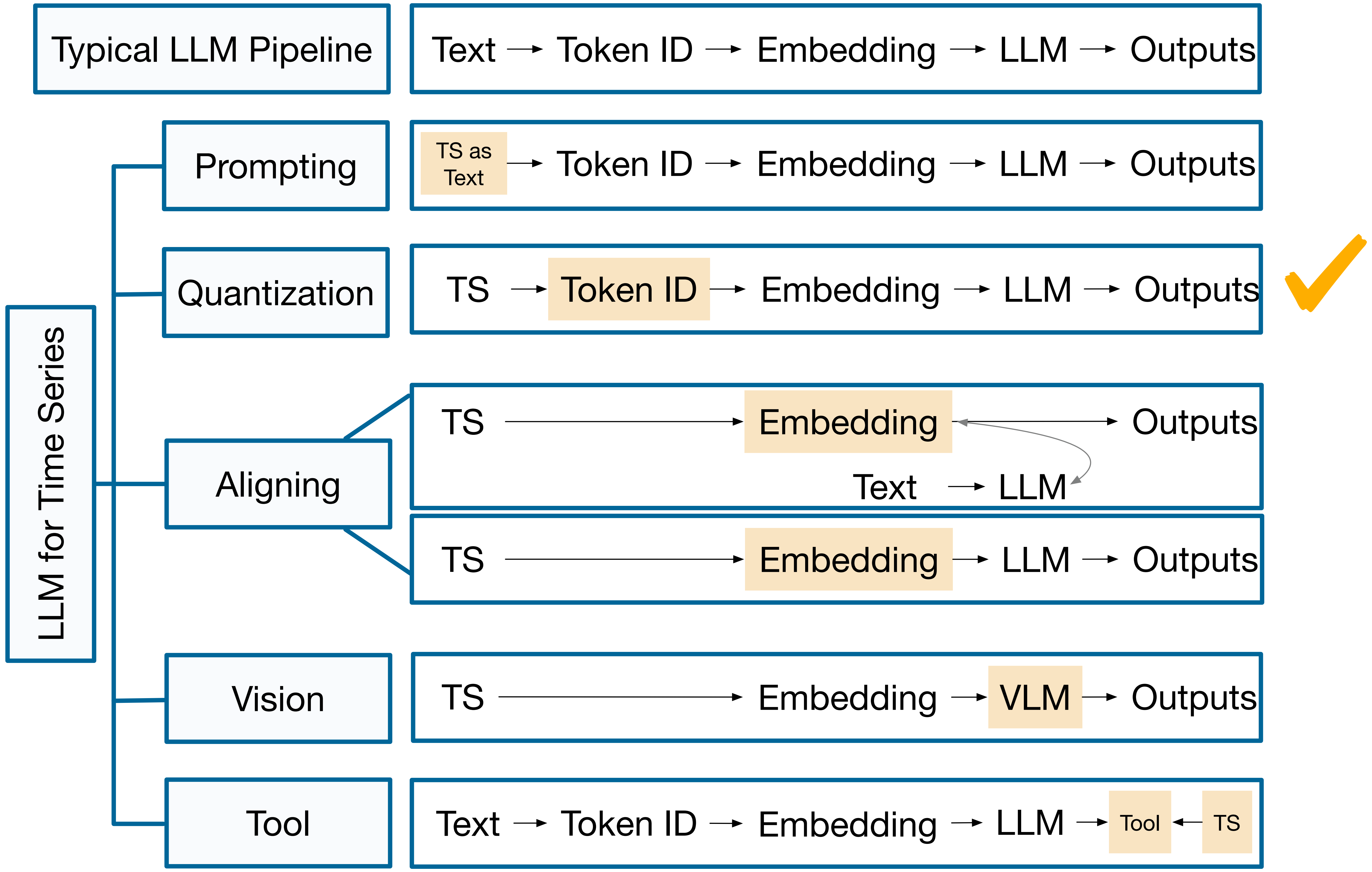
Taxonomy

- Each category targets one stage in typical LLM-driven NLP pipelines
  - Prompting: input stage
  - Quantization: tokenization stage
  - Aligning: embedding stage
  - Vision as bridge: LLM stage
  - Tool: output stage



Introduction
Taxonomy

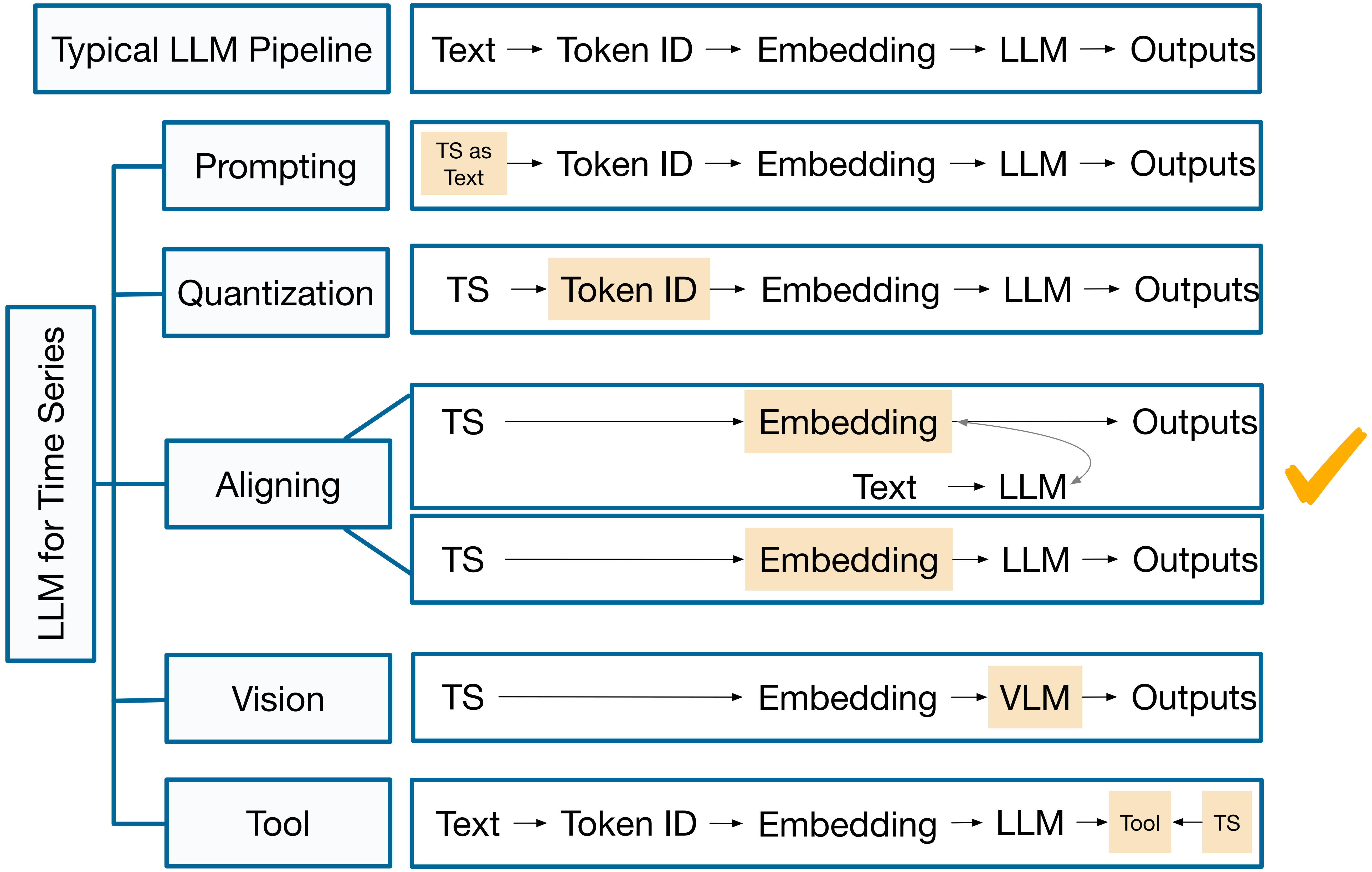
- Each category targets one stage in typical LLM-driven NLP pipelines
  - Prompting: input stage
  - Quantization: tokenization stage
  - Aligning: embedding stage
  - Vision as bridge: LLM stage
  - Tool: output stage



Introduction

Taxonomy

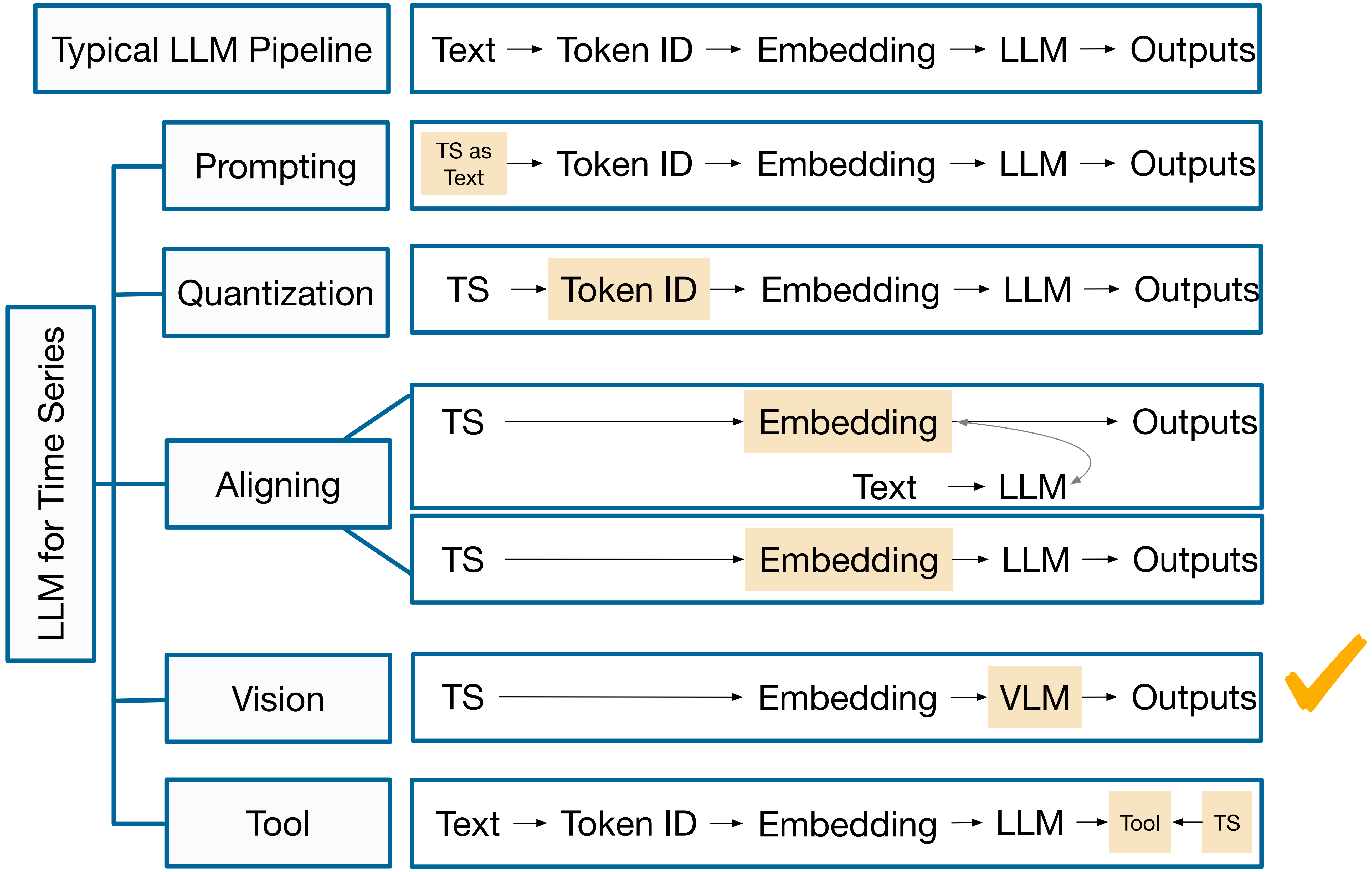
- Each category targets one stage in typical LLM-driven NLP pipelines
  - Prompting: input stage
  - Quantization: tokenization stage
  - ✓ Aligning: embedding stage
  - Vision as bridge: LLM stage
  - Tool: output stage



Introduction

Taxonomy

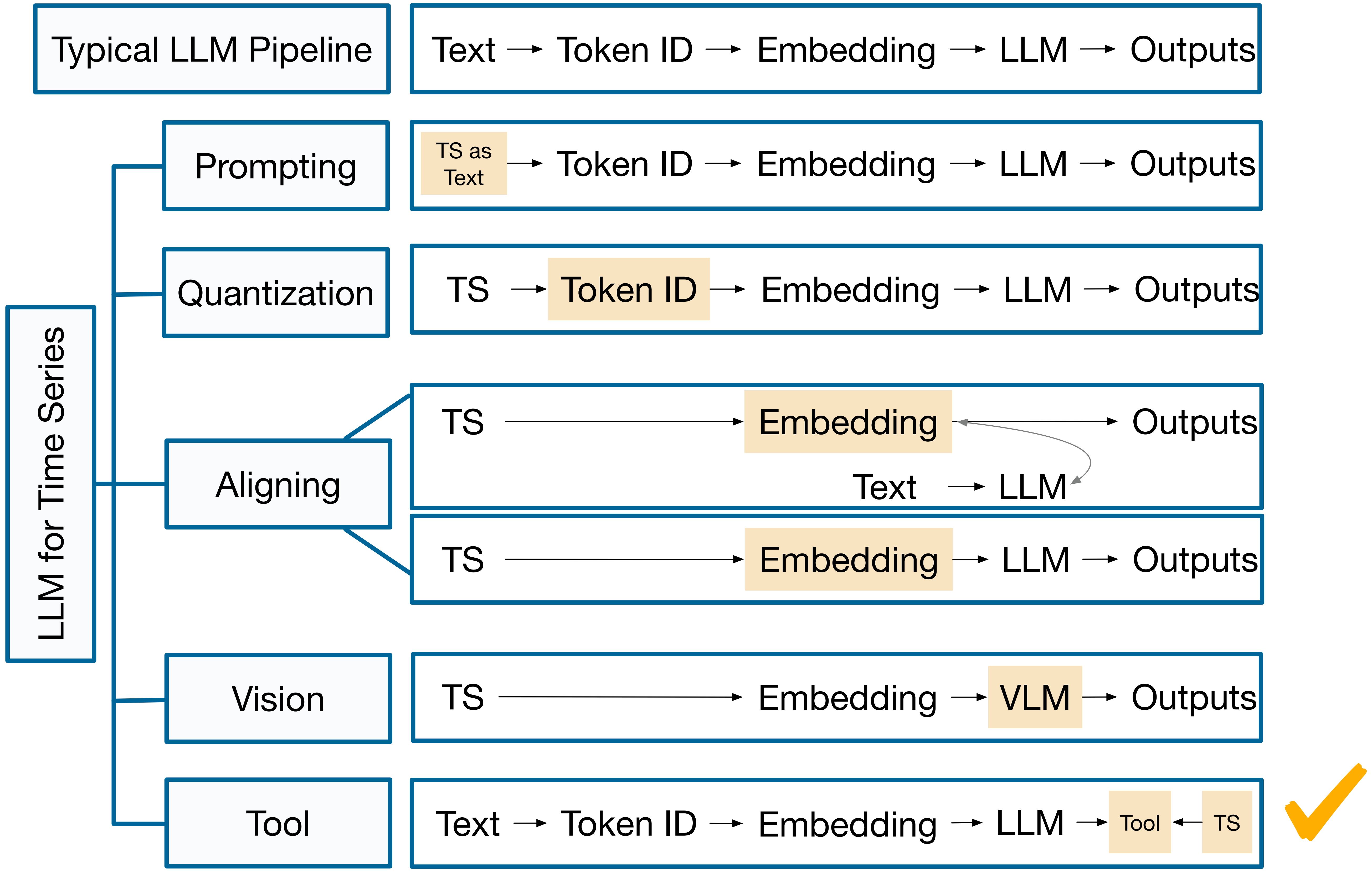
- Each category targets one stage in typical LLM-driven NLP pipelines
  - Prompting: input stage
  - Quantization: tokenization stage
  - Aligning: embedding stage
  - Vision as bridge: LLM stage
  - Tool: output stage



Introduction

Taxonomy

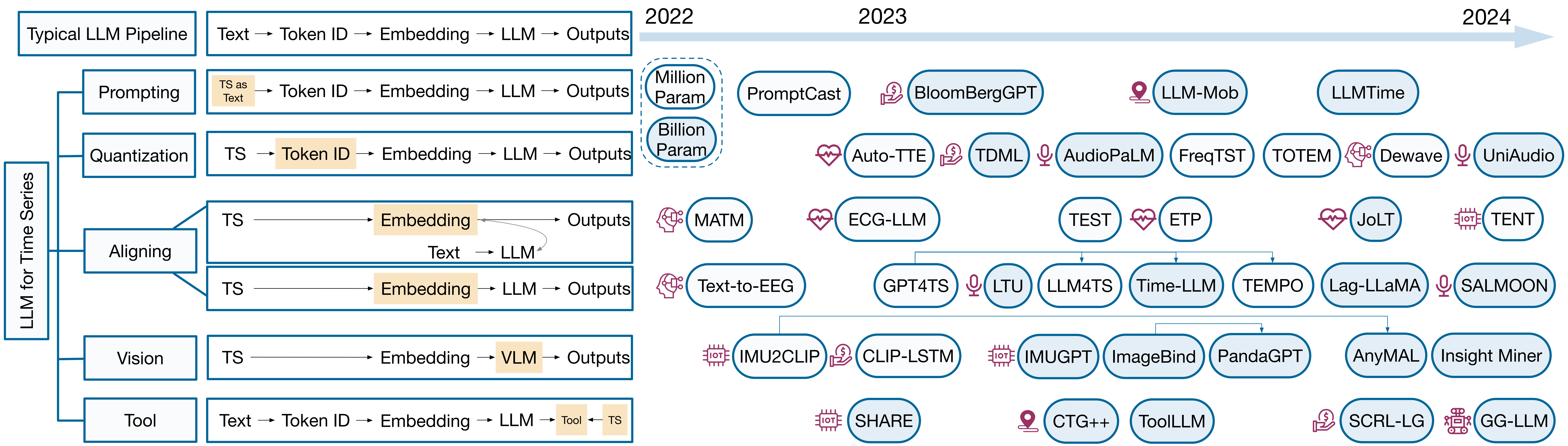
- Each category targets one stage in typical LLM-driven NLP pipelines
  - Prompting: input stage
  - Quantization: tokenization stage
  - Aligning: embedding stage
  - Vision as bridge: LLM stage
  - Tool: output stage



Introduction

Resources

- Up-to-date Github repo summarizing LLM4TS papers + datasets
  - https://github.com/xiyuanzh/awesome-llm-time-series



Outline

- Introduction
- Taxonomy
- Datasets
- Future Directions
- Conclusion

Taxonomy

Prompting

- Number-agnostic tokenization
- Number-specific tokenization

Number-agnostic

Number-specific

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" $\rightarrow$ "1 2 , 1 2 3 , 1 2 3 0 , 1 2 3 0 0"

Examples of representative direct prompting methods.

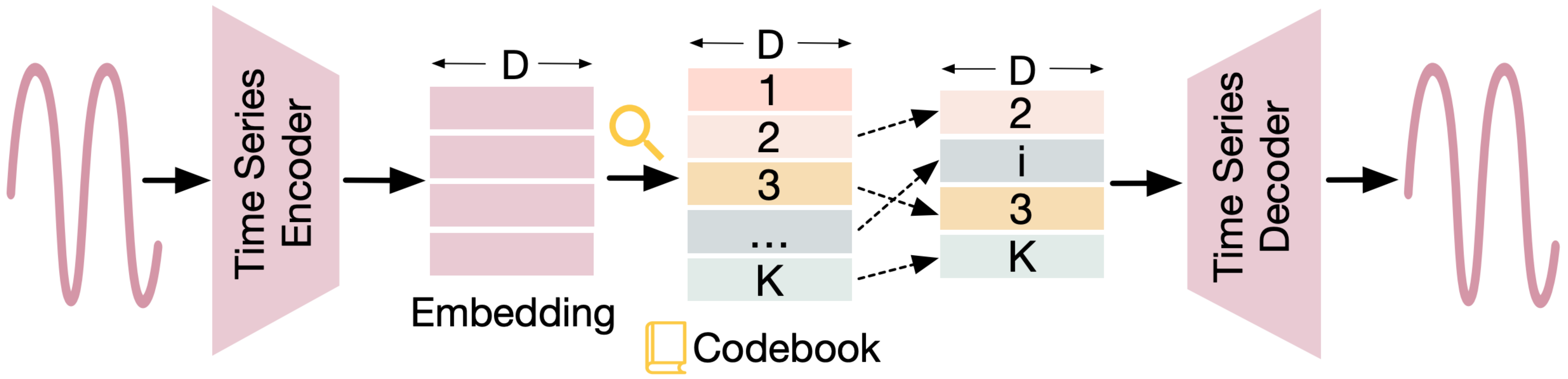
Taxonomy

Quantization

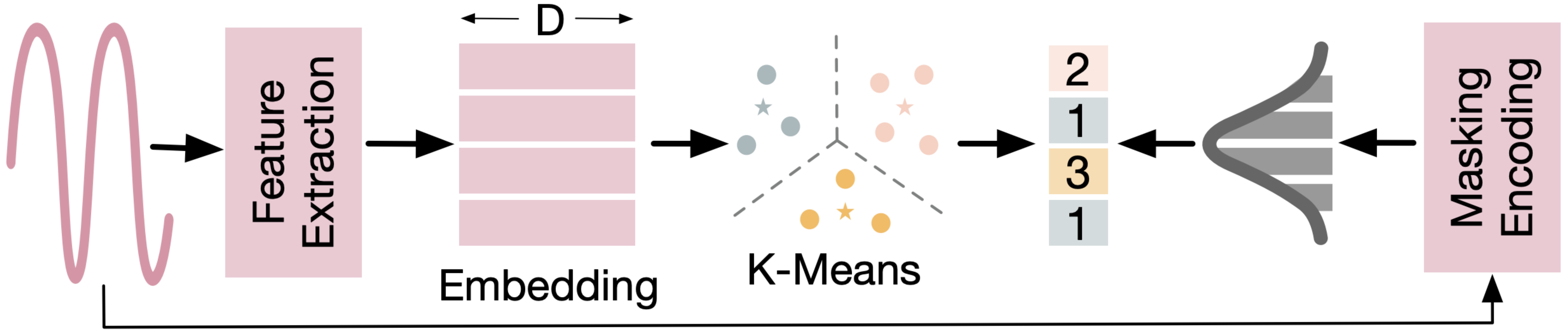
- Discrete indices
  - From VQ-VAE

$$\mathbf{q}_i = \mathbf{c}_{k_i}, k_i = \arg\min_j \|g_\phi(\mathbf{x}_s)_i - \mathbf{c}_j\|_2, \mathbf{k} = [k_i]_{i=1}^{\frac{T}{S}}$$

  - From K-Means



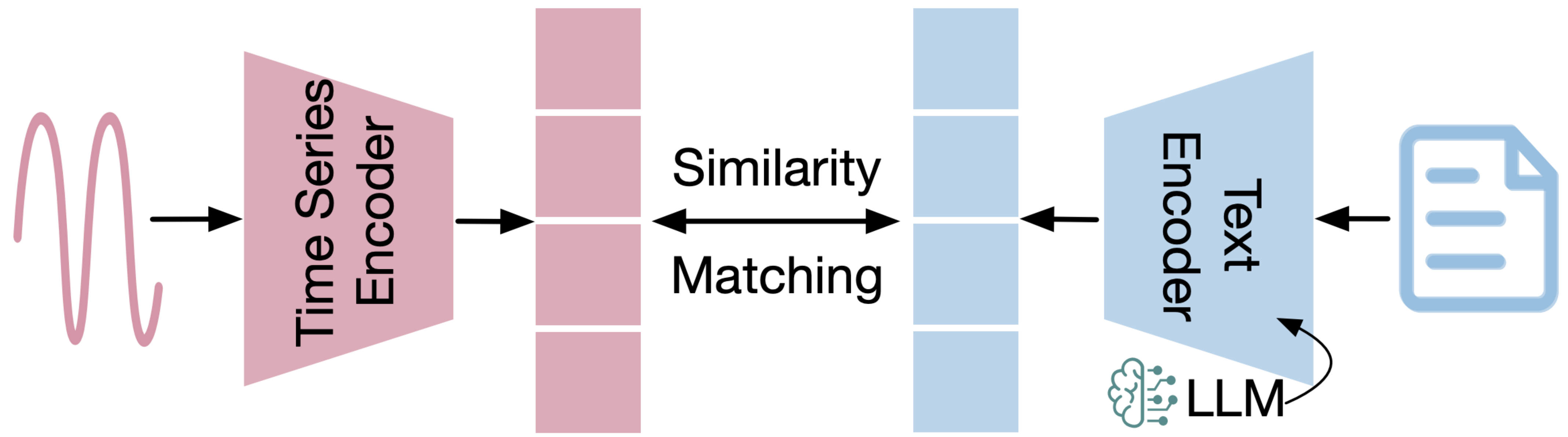
VQ-VAE based quantization method



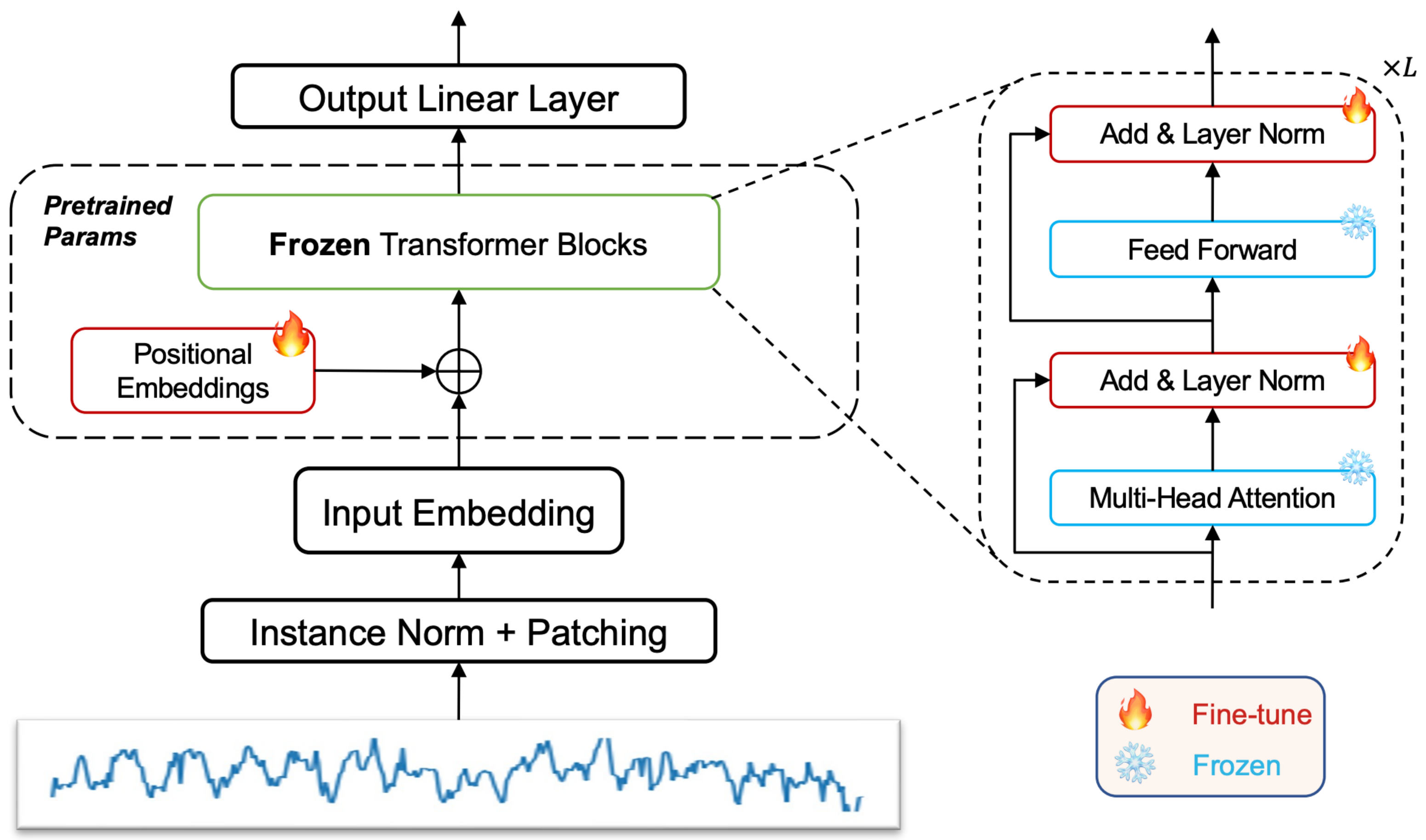
K-Means based quantization method

Taxonomy
Aligning

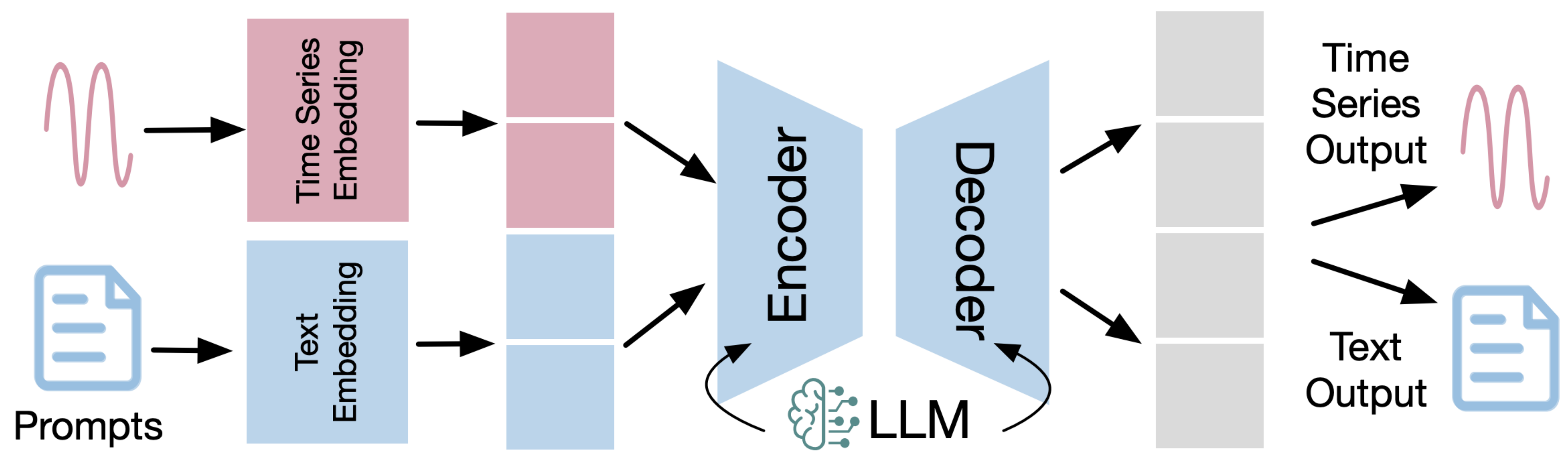
- Similarity matching
  - Contrastive loss
- LLMs as backbones



Aligning by similarity matching



One Fits All: Power General Time Series Analysis by Pretrained LM, NeurIPS 2023

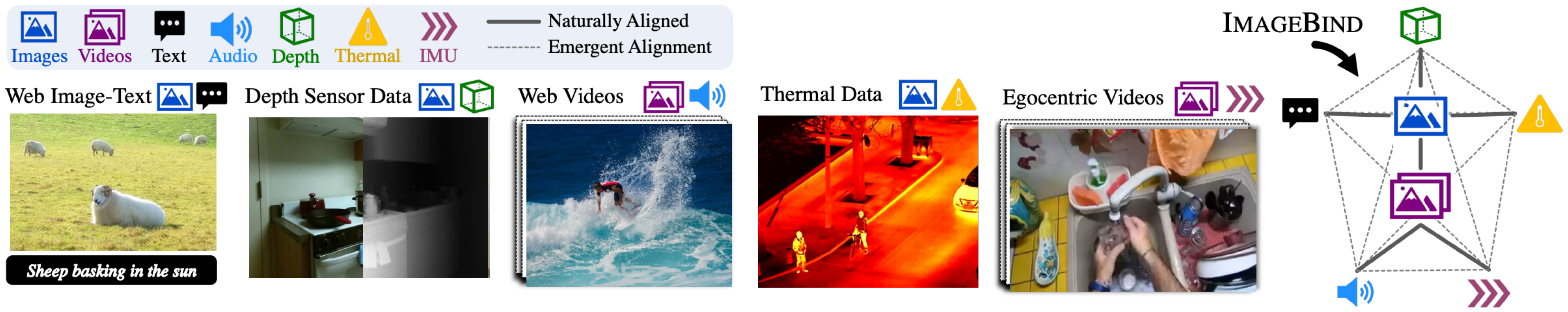


Aligning with LLMs as backbones

Taxonomy

Vision as Bridge

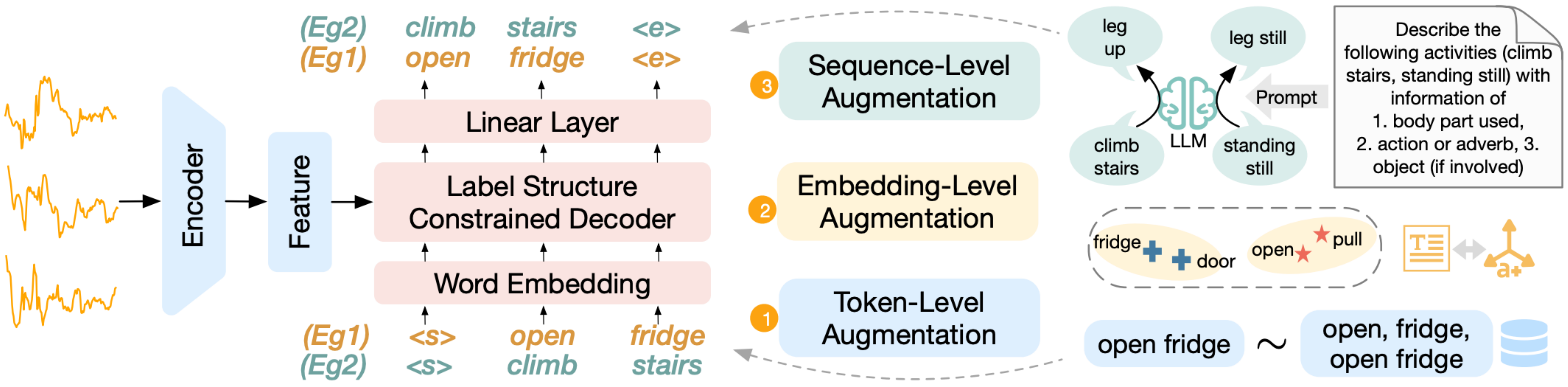
- Paired data
- Physics relationships
- Time series plots as images



ImageBind: One Embedding Space to Bind Them All, CVPR 2023

Taxonomy
Tool

- Code
- API call
- Text domain knowledge



Unleashing the Power of Shared Label Structures for Human Activity Recognition, CIKM 2023

Taxonomy

Comparison

- Data: zero (prompt), visual (vision)
- Model: billion (prompt/tool), million (aligning, quantization)
- Efficiency: quantization and aligning more efficient than prompting
- Optimization: two stage, indirect tools

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; zero-shot capability	lose semantics; not efficient
	Number-Specific	LLMTime [Gruver *et al.*, 2023]			
Quantization	VQ-VAE	DeWave [Duan *et al.*, 2023]	$k_i = \arg\min_j \|g_\phi(\mathbf{x}_s)_i - \mathbf{c}_j\|_2$	flexibility of index and time series conversion	may require two-stage training
	K-Means	AudioLM [Borsos *et al.*, 2023]	$\mathbf{k} = [k_i]_{i=1}^{\frac{T}{S}}, \mathbf{y} = f_\theta(\mathbf{k}, \mathbf{x}_t)$		
	Text Categories	TDML [Yu *et al.*, 2023]	$\mathbf{y} = f_\theta(q(\mathbf{x}_s), \mathbf{x}_t)$		
Aligning	Similarity Match	ETP [Liu *et al.*, 2023a] MATM [Han *et al.*, 2022]	$\mathbf{y} = g_\phi(\mathbf{x}_s)$ $\mathcal{L} = \mathrm{sim}(g_\phi(\mathbf{x}_s), f_\theta(\mathbf{x}_t))$	align semantics of different modalities; end-to-end training	complicated design and fine-tuning
	LLM Backbone	GPT4TS [Zhou *et al.*, 2023a]	$\mathbf{y} = f_\theta(g_\phi(\mathbf{x}_s), \mathbf{x}_t)$		
Vision as Bridge	Paired Data	ImageBind [Girdhar *et al.*, 2023]	$\mathcal{L} = \mathrm{sim}(g_\phi(\mathbf{x}_s), h_\psi(\mathbf{x}_v))$	additional visual knowledge	not hold for all data
	TS Plots as Images	Wimmer and Rekabsaz [2023]	$\mathbf{y} = h_\psi(\mathbf{x}_s)$		
Tool	Code	CTG++ [Zhong *et al.*, 2023]	$z = f_\theta(\mathbf{x}_t)$ $\mathbf{y} = z(\mathbf{x}_s)$	empower LLM with more abilities	optimization not end-to-end
	API	ToolLLM [Qin *et al.*, 2023]			

Outline

- Introduction
- Taxonomy
- Datasets
- Future Directions
- Conclusion

Datasets

Summary

- Internet of Things (IoT): IMU
- Finance: stock
- Healthcare: EEG, ECG
- Audio/Music/Speech

Domain	Dataset	Size	Major Modalities	Task
Internet of Things	Ego4D[2] [Grauman *et al.*, 2022]	$3,670$h 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] [Anonymous, 2023a]	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

Outline

- Introduction
- Taxonomy
- Datasets
- Future Directions
- Conclusion

Future Directions
Summary

- Theoretical understanding
- Multimodal and multitask analysis
- Efficient algorithms
- Combining domain knowledge
- Customization and privacy

Outline

- Introduction
- Taxonomy
- Datasets
- Future Directions
- Conclusion

Conclusion

Summary

- The first survey that builds a taxonomy for how to transfer knowledge from LLMs for time series analysis
- Multimodal text and time series datasets
- Paper link: https://arxiv.org/abs/2402.01801
- GitHub repo: https://github.com/xiyuanzh/awesome-llm-time-series