Towards Graph Foundation Models

WWW 2024 Tutorial

Philip S. Yu, Chuan Shi, Cheng Yang, Yuan Fang, Lichao Sun

Welcome to Big AI era!

- **Driving Forces:**
  - Technology advances
  - Availability of big data for training
  - Availability of powerful GPU

- **Performance improves with size.**
  - "The race to scale" begins…

- **The new thing (2021--)**
  - HUGE neural networks
  - VAST amounts of training data
  - MASSIVE compute power for training

On the Opportunities and Risks of Foundation Models

Rishi Bommasani* Drew A. Hudson Ehsan Adeli Russ Altman Simran Arora
Sydney von Arx Michael S. Bernstein Jeannette Bohg Antoine Bosselut Emma Brunskill
Erik Brynjolfsson Shyamal Buch Dallas Card Rodrigo Castellon Niladri Chatterji
Annie Chen Kathleen Creel Jared Quincy Davis Dorottya Demszky Chris Donahue
Moussa Doumbouya Esin Durmus Stefano Ermon John Etchemendy Kawin Ethayarajh
Li Fei-Fei Chelsea Finn Trevor Gale Lauren Gillespie Karan Goel Noah Goodman
Shelby Grossman Neel Guha Tatsunori Hashimoto Peter Henderson John Hewitt
Daniel E. Ho Jenny Hong Kyle Hsu Jing Huang Thomas Icard Saahil Jain
Dan Jurafsky Pratyusha Kalluri Siddharth Karamcheti Geoff Keeling Fereshte Khani
Omar Khattab Pang Wei Koh Mark Krass Ranjay Krishna Rohith Kuditipudi
Ananya Kumar Faisal Ladhak Mina Lee Tony Lee Jure Leskovec Isabelle Levent
Xiang Lisa Li Xuechen Li Tengyu Ma Ali Malik Christopher D. Manning
Suvir Mirchandani Eric Mitchell Zanele Munyikwa Suraj Nair Avanika Narayan
Deepak Narayanan Ben Newman Allen Nie Juan Carlos Niebles Hamed Nilforoshan
Julian Nyarko Giray Ogut Laurel Orr Isabel Papadimitriou Joon Sung Park Chris Piech
Eva Portelance Christopher Potts Aditi Raghunathan Rob Reich Hongyu Ren
Frieda Rong Yusuf Roohani Camilo Ruiz Jack Ryan Christopher Ré Dorsa Sadigh
Shiori Sagawa Keshav Santhanam Andy Shih Krishnan Srinivasan Alex Tamkin
Rohan Taori Armin W. Thomas Florian Tramèr Rose E. Wang William Wang Bohan Wu
Jiajun Wu Yuhuai Wu Sang Michael Xie Michihiro Yasunaga Jiaxuan You Matei Zaharia
Michael Zhang Tianyi Zhang Xikun Zhang Yuhui Zhang Lucia Zheng Kaitlyn Zhou
Percy Liang*[1]

Center for Research on Foundation Models (CRFM)
Stanford Institute for Human-Centered Artificial Intelligence (HAI)
Stanford University

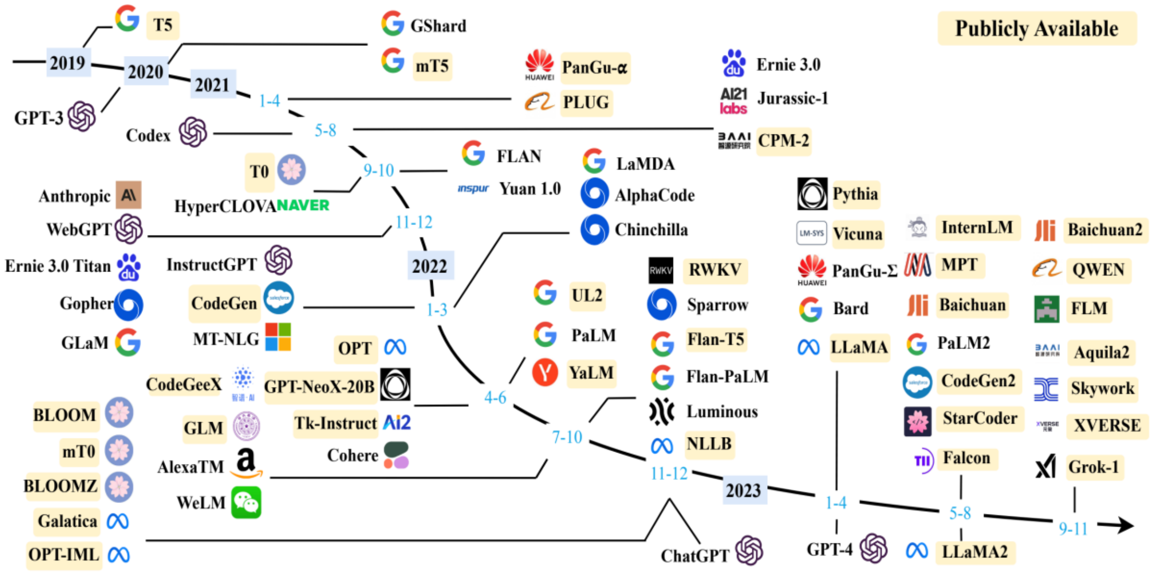
AI is undergoing a paradigm shift with the rise of models (e.g., BERT, DALL-E, GPT-3) trained on broad data (generally using self-supervision at scale) that can be adapted to a wide range of downstream tasks. We call these models foundation models to underscore their critically central yet incomplete character. This report provides a thorough account of the opportunities and risks of foundation models, ranging from their capabilities (e.g., language, vision, robotic manipulation, reasoning, human interaction) and

Foundation Models

A foundation model is a model that is trained on broad data and can be adapted to a wide range of downstream tasks.

➢ Big Idea

- Pretrain model, then fine-tune
- Revolutionize many research domains
  - Language
  - Vedio…



➢ Representative Examples

- Large Language Models (LLMs)
  - E.g., ELMo with millions of parameters to GPT-4 with trillions of parameters.
- Vedio Models: SORA

Graph Foundation Models

A graph foundation model (GFM) is a model pre-trained on extensive graph data, adapted for diverse downstream graph tasks.

➢ Motivation

- Existing LLMs struggle to model graph data
  - Euclidean data v.s. non-Euclidean data
- Existing LLMs struggle to handle graph tasks
  - node/edge/graph-level tasks

➢ Scope of this tutorial

- Concept of graph foundation model
- Recent progress
  - GNN-based methods
  - LLM-based methods
  - GNN+LLM-based methods
- Future directions

Towards Graph Foundation Models: A Survey and Beyond

JIAWEI LIU, CHENG YANG*, Beijing University of Posts and Telecommunications, China
ZHIYUAN LU, JUNZE CHEN, YIBO LI, Beijing University of Posts and Telecommunications, China
MENGMEI ZHANG, TING BAI, Beijing University of Posts and Telecommunications, China
YUAN FANG, Singapore Management University, Singapore
LICHAO SUN, Lehigh University, USA
PHILIP S. YU, University of Illinois Chicago, USA
CHUAN SHI†, Beijing University of Posts and Telecommunications, China

Foundation models have emerged as critical components in a variety of artificial intelligence applications, and showcase significant success in natural language processing and several other domains. Meanwhile, the field of graph machine learning is witnessing a paradigm transition from shallow methods to more sophisticated deep learning approaches. The capabilities of foundation models to generalize and adapt motivate graph machine learning researchers to discuss the potential of developing a new graph learning paradigm. This paradigm envisions models that are pre-trained on extensive graph data and can be adapted for various graph tasks. Despite this burgeoning interest, there is a noticeable lack of clear definitions and systematic analyses pertaining to this new domain. To this end, this article introduces the concept of Graph Foundation Models (GFMs), and offers an exhaustive explanation of their key characteristics and underlying technologies. We proceed to classify the existing work related to GFMs into three distinct categories, based on their dependence on graph neural networks and large language models. In addition to providing a thorough review of the current state of GFMs, this article also outlooks potential avenues for future research in this rapidly evolving domain.

Outline

Philip S. Yu University of Illinois Chicago
09:00-09:05 Introduction (5mins)

Chuan Shi Beijing University of Posts and Telecommunications
09:05-09:40 Overview (35mins)

Cheng Yang Beijing University of Posts and Telecommunications
09:40-10:30 GNN-based Methods (50mins)

10:30-11:00 Break (30mins)

Yuan Fang Singapore Management University
11:00-12:00 LLM/GNN+LLM-based Methods (50mins)

Host: Chuan Shi Beijing University of Posts and Telecommunications
12:00-12:30 Panel (30mins)

Towards Graph Foundation Models
Part I: Overview

Prof. Chuan Shi

shichuan@bupt.edu.cn

BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

Outline

√ Graph Foundation Models

- Progress in Related Work

- Challenges and Future Direction

Foundation Models

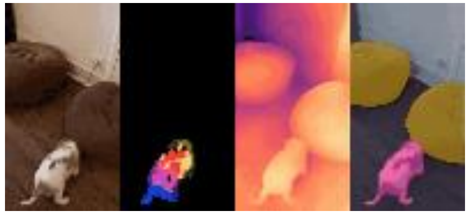
A foundation model is any model that is trained on broad data and can be adapted to a wide range of downstream tasks.[1]

Language

OpenAI × GPT4

Language foundation models show initial signs of universal AI capabilities.

Vision



Meta × DINOv2

Vision foundation models exhibit strong image understanding capabilities.

Speech

Google × USM

Speech foundation models show the capability to recognize hundreds of languages.

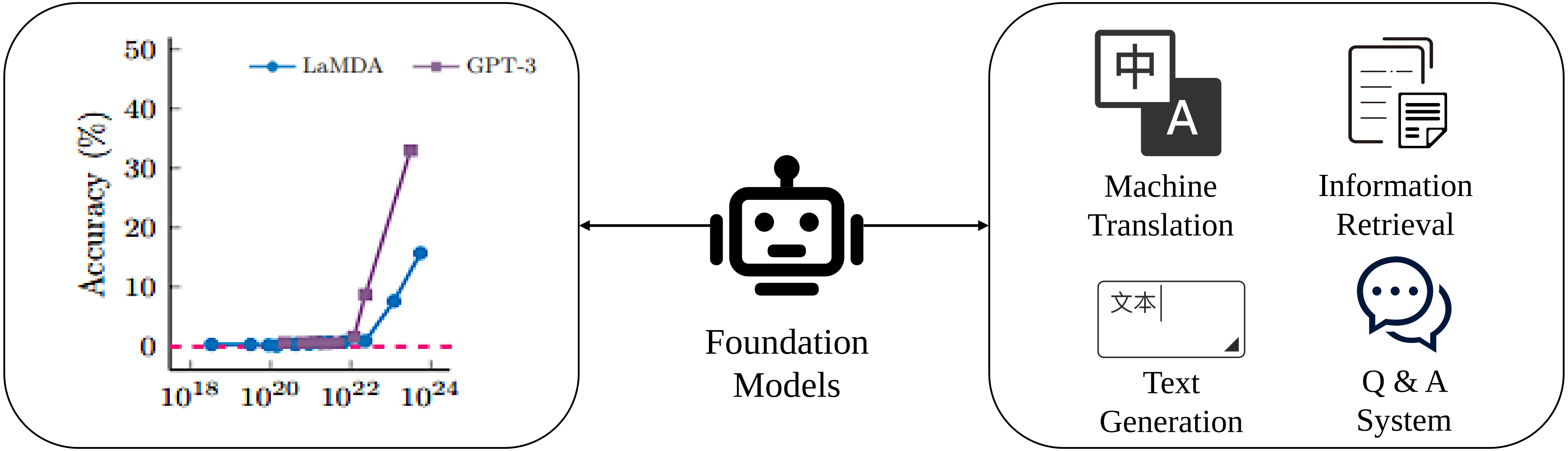
Foundation models have become a reality in domains like language, vision, and speech.

[1] R. Bommasani, D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brun-skill, et al., "On the opportunities and risks of foundation models," arXiv preprint arXiv:2108.07258, 2021.

Characteristics of Foundation Models

Two Characteristics of Foundation Models:

- Emergence: As a foundation model scales up, it spontaneously manifests novel capabilities.
- Homogenization: The model's versatility enables its deployment across diverse applications.



Emergence

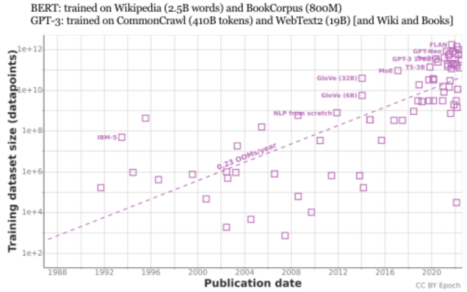
Homogenization

Wei J, Tay Y, Bommasani R, et al. Emergent abilities of large language models[J]. arXiv preprint arXiv:2206.07682, 2022.

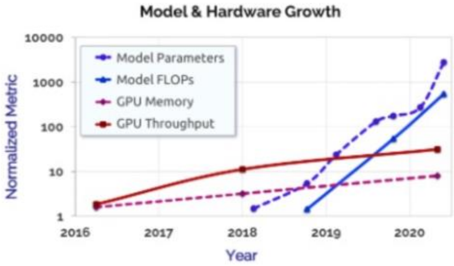
Factors Driving Foundation Model Success

Data

- The increasing number of data-collecting devices results in a massive growth in data volume.



Data Growth



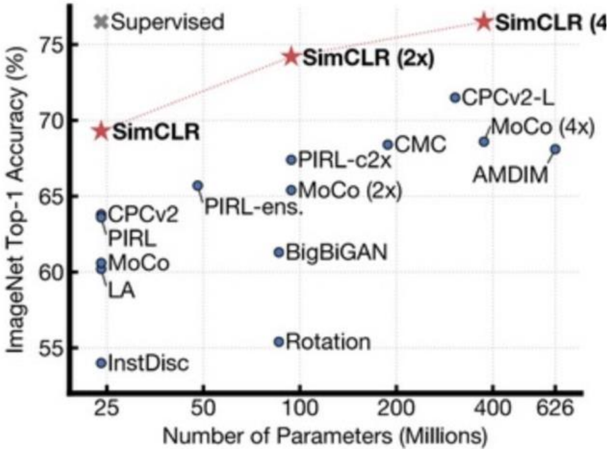
GPU Development

Hardware

- the rapid advancement of GPU hardware

Self-supervised Learning (SSL)

- exploiting raw unlabeled data



SSL

Transformer Architectures

- attention mechanism

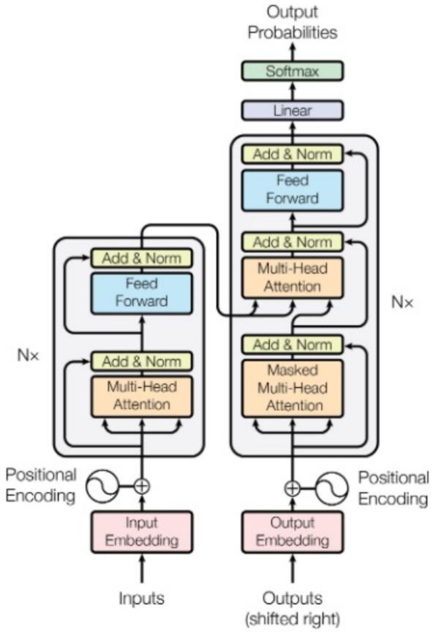


Figure 1: The Transformer - model architecture.

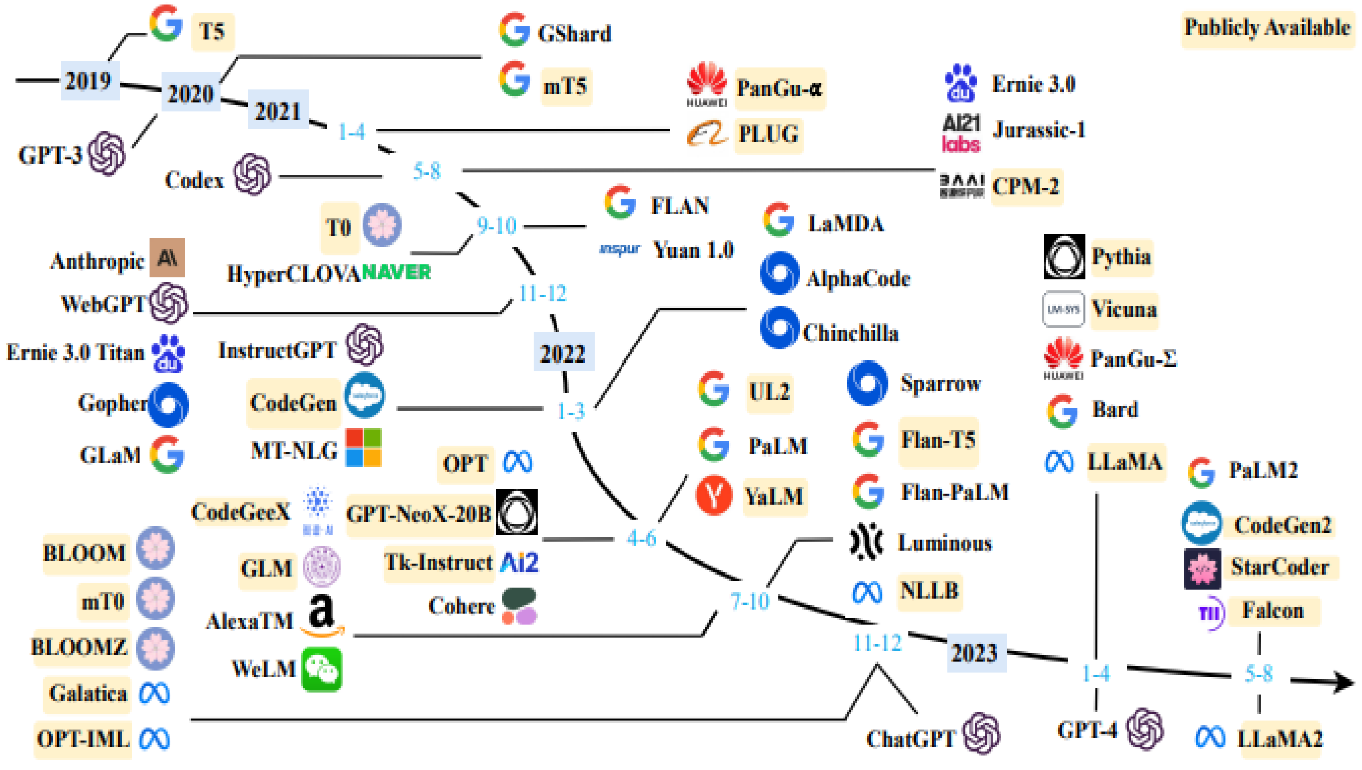
Transformer

Language Foundation Models

Large Language Models (LLMs) refer to pre-trained language models with massive parameters and are typical representatives of foundation models.

- LLMs have progressed from models like ELMo with millions of parameters to GPT-4 with trillions of parameters.

- LLMs showcase key AI abilities like comprehension, generation, logic, and memory, hinting at the path towards artificial general intelligence (AGI).

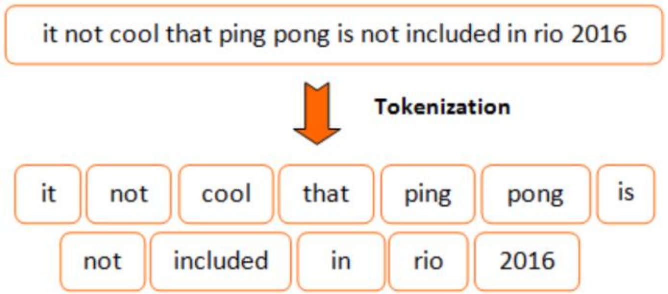


Zhao W X, Zhou K, Li J, et al. A survey of large language models[J]. arXiv preprint arXiv:2303.18223, 2023.

Large Language Models

Data

- ➢ Language data: text or spoken content in a human language
  - sequential data
  - Euclidean data



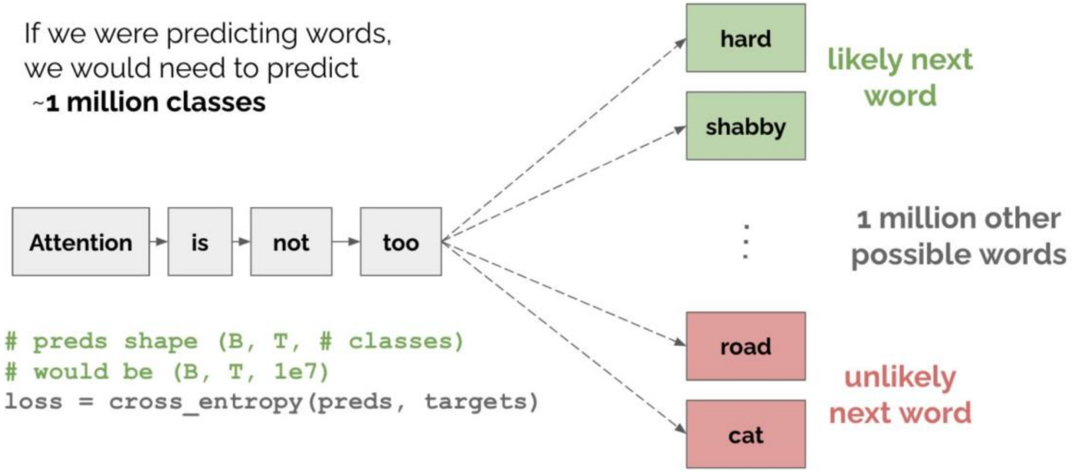
Language Data

Backbone Architectures

- ➢ Mostly based on Transformer
  - e.g., BERT[1], GPT-3[2]
- ➢ Pre-trained with pretext tasks:
  - next word prediction (NWP)
  - masked language modeling (MLM)…

Downstream Tasks

- ➢ Hundreds of downstream tasks
  - e.g., machine translation, sentiment analysis…

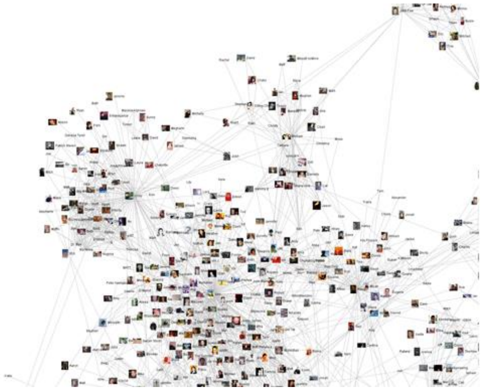


Next Word Prediction (NWP)

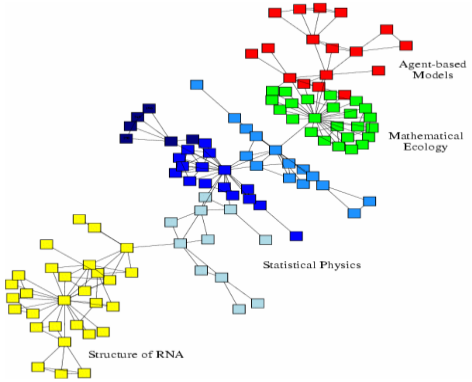
[1] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.
[2] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners[C]. NeurIPS 2020, 33: 1877-1901.

Graphs

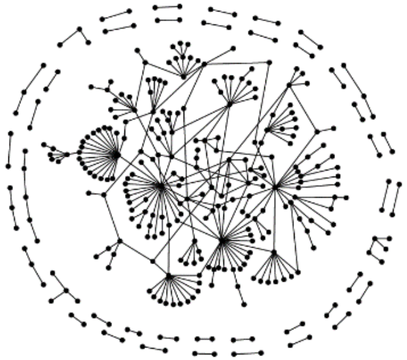
Graphs are a general language for describing and modeling complex systems.



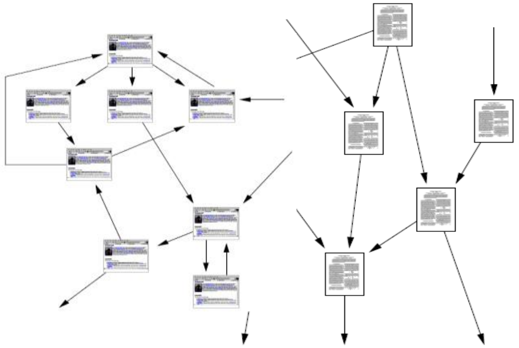
Social networks



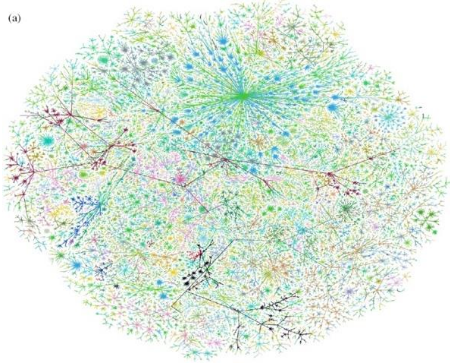
Economic networks



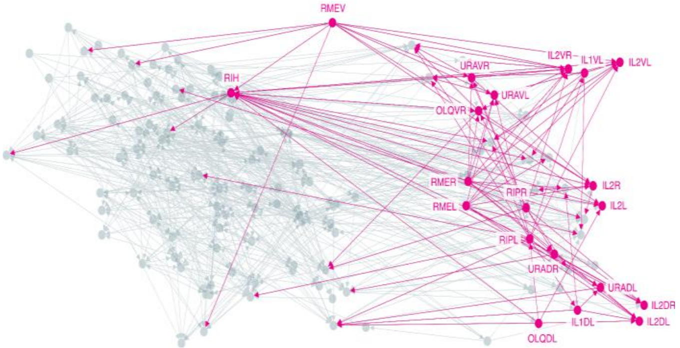
Biomedical networks



Information networks



Internet

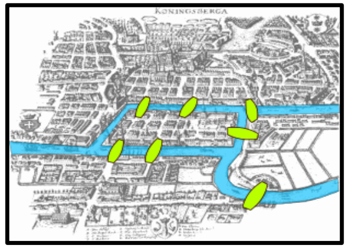


Networks of neurons

Graph Machine Learning

- Graph G is an ordered pair (V, E), where V is the node set and E is the edge set.
- Graph machine learning refers to the application of machine learning to graph data, commonly known as graph learning or graph models.

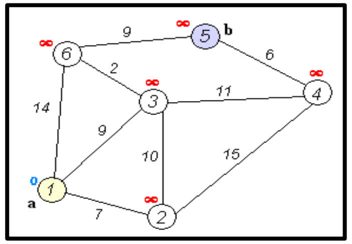
Seven Bridges of Königsberg



Graph theory
- Euler

1736

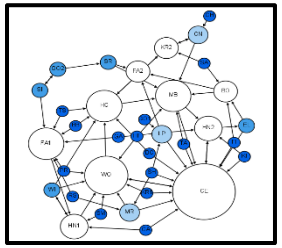
Shortest Path Problem



Graph algorithms
- Dijkstra

1956

Long Tail Distribution

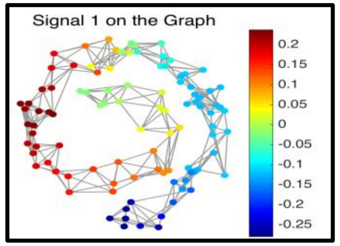


Network Science
- Barabasi

2002

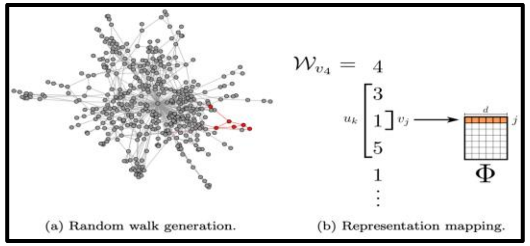
2013

Graph signal processing
- Shuman



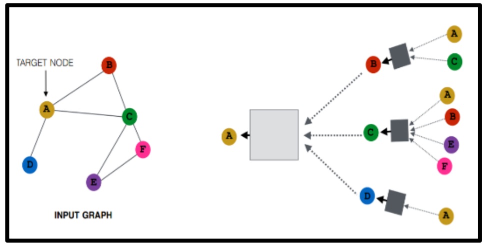
2014

Graph embedding
- DeepWalk



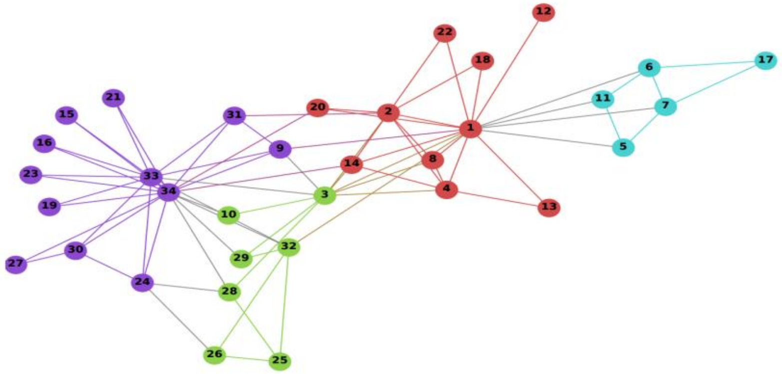
2017

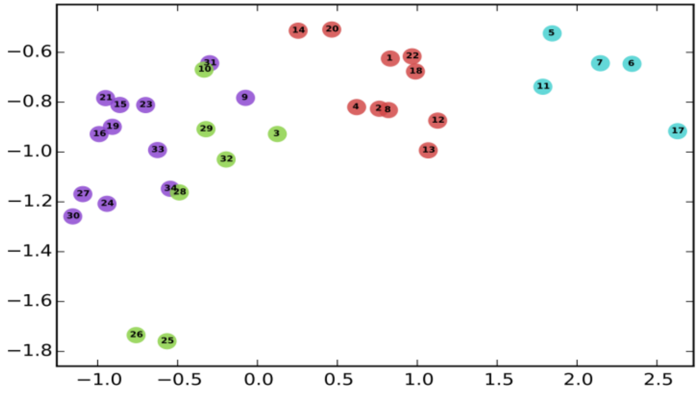
Graph neural networks
- GCN



Graph Representation Learning

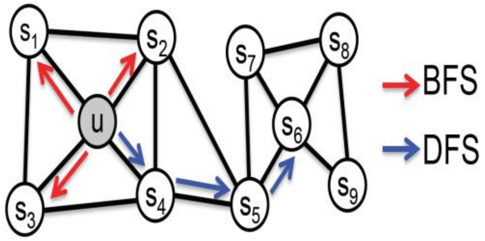
Graph Representation Learning (GRL): embed each node of a graph into a low-dimensional vector space





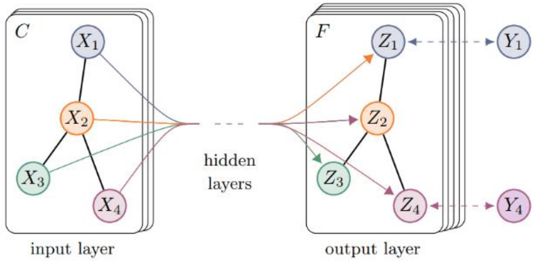
Shallow model

- ➢ Random walk based
  - e.g., DeepWalk, node2vec



Deep model

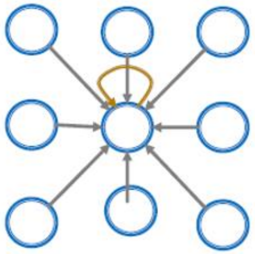
- ➢ GNN based
  - e.g., GCN, GraphSage, GAT



Data in GNN

Data

- **Graph data**
  - non-Euclidean data
- **Various domains**
  - social networks
  - molecules
  - E-commerce…
- **Various types**
  - homogenous graph
  - heterogenous graph
  - hypergraph…



Graph

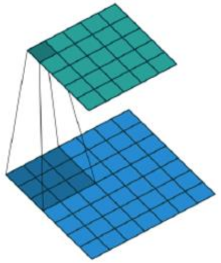
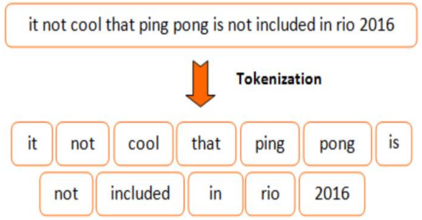
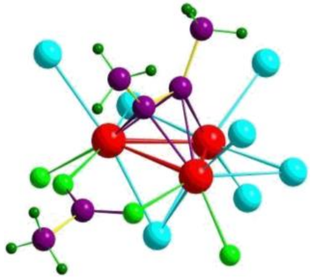


Image (Grid)

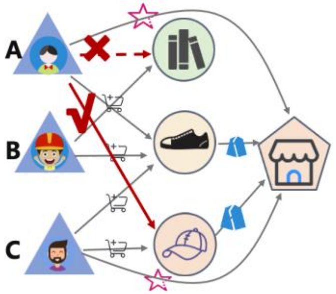


Language (Seq.)

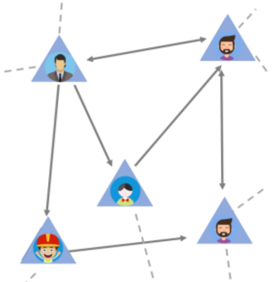
Social Networks



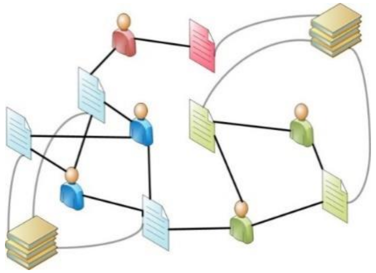
Molecules



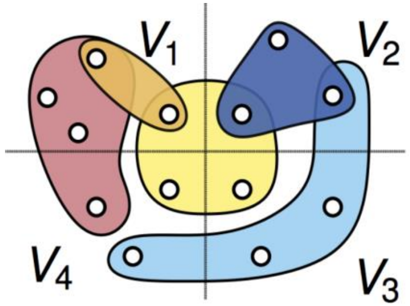
E-commerce



Homogeneous



Heterogeneous

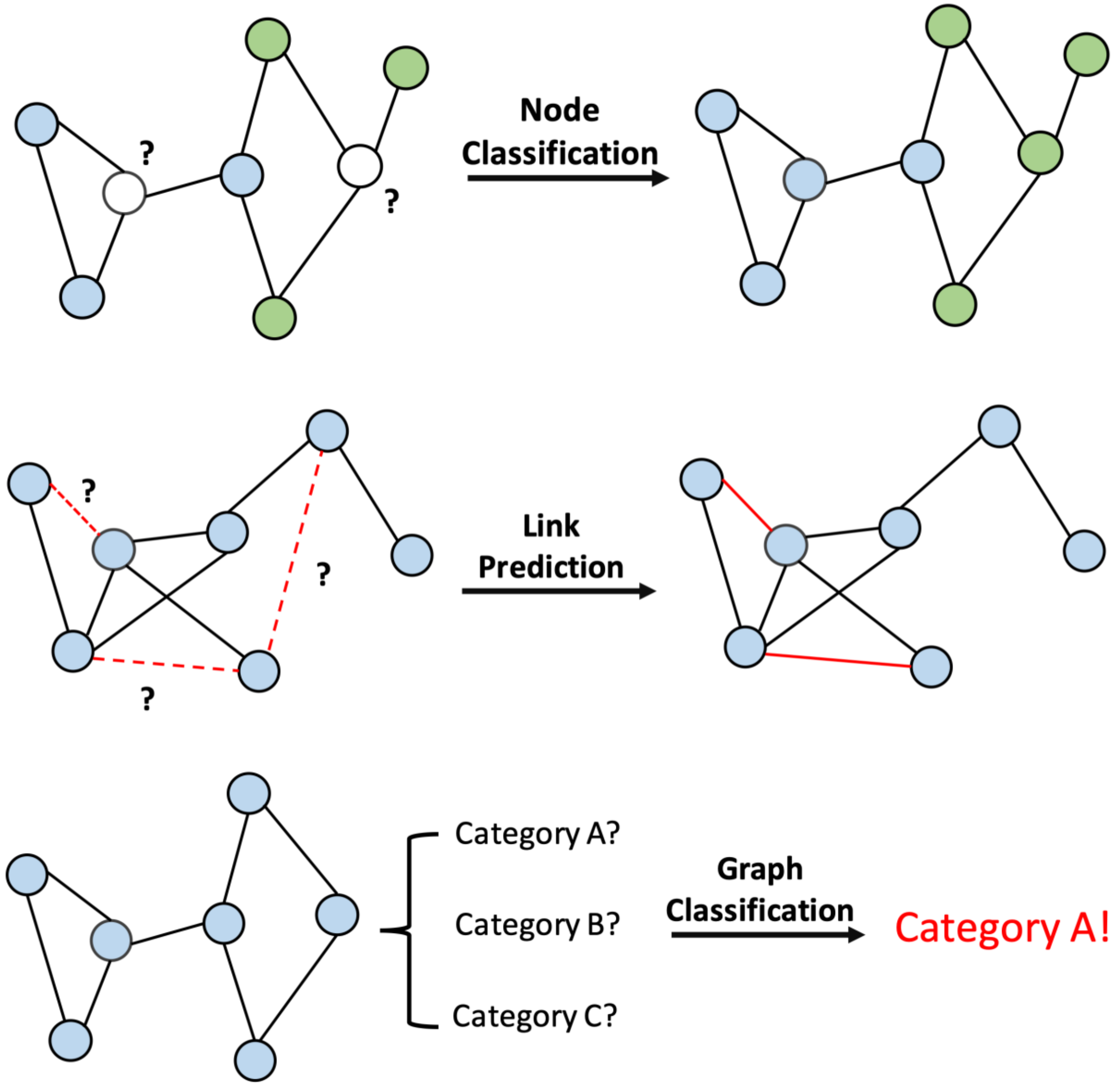


Hypergraph

Tasks in GNN

Downstream Tasks

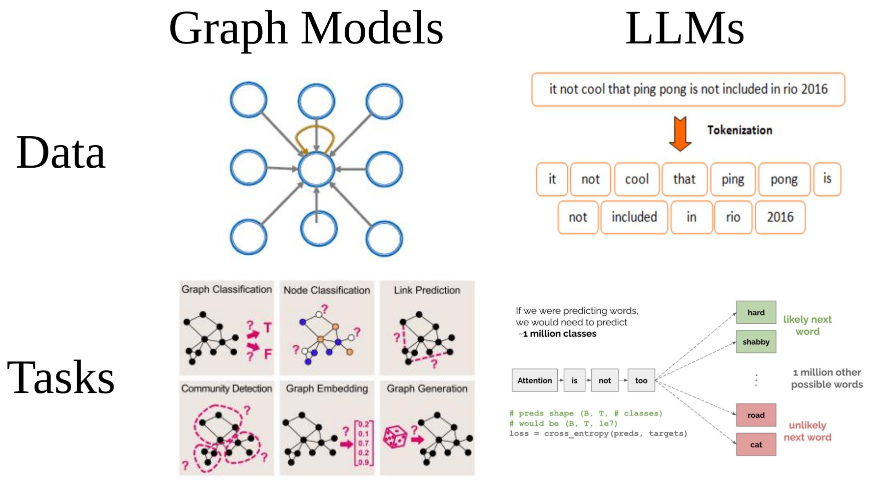
- **Node-level tasks**
  - node classification
  - node regression
  - node clustering…
- **Edge-level tasks**
  - link prediction
  - shortest path prediction
  - maximum flow prediction…
- **Graph-level tasks**
  - graph classification
  - graph generation
  - graph condensation…



Graph Models Meet Large Language Models

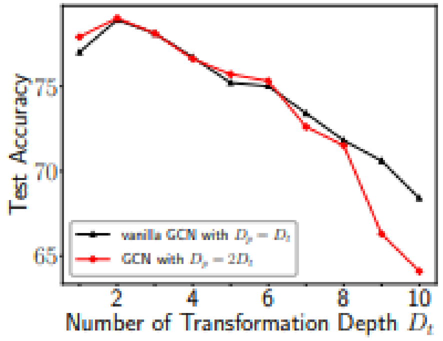
LLMs cannot solve graph-related problems.

- LLMs struggle to model graph structure semantics.
- LLMs struggle to handle diverse graph tasks.

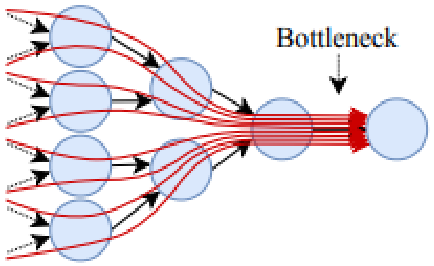


Graph models do not possess the capabilities of LLMs.

- Limited expressive power
- Deep GNNs: over-smoothing/over-squassion issues
- Lack emergence capability
- Cannot support multiple tasks



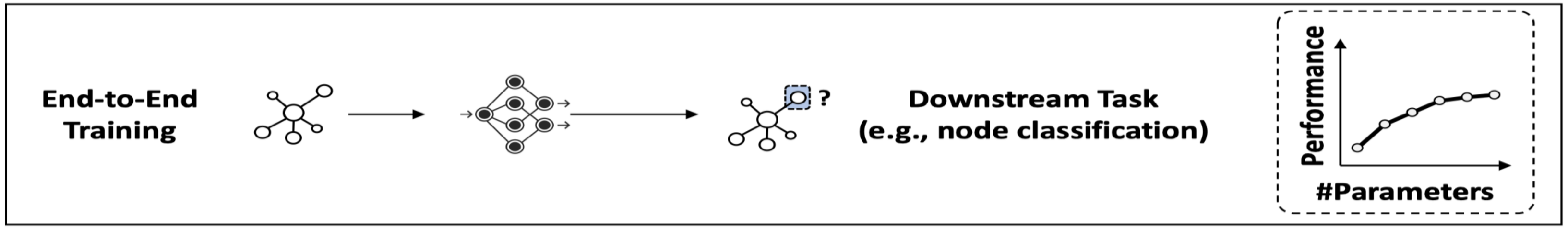
Performance Decline of Deep GNNs



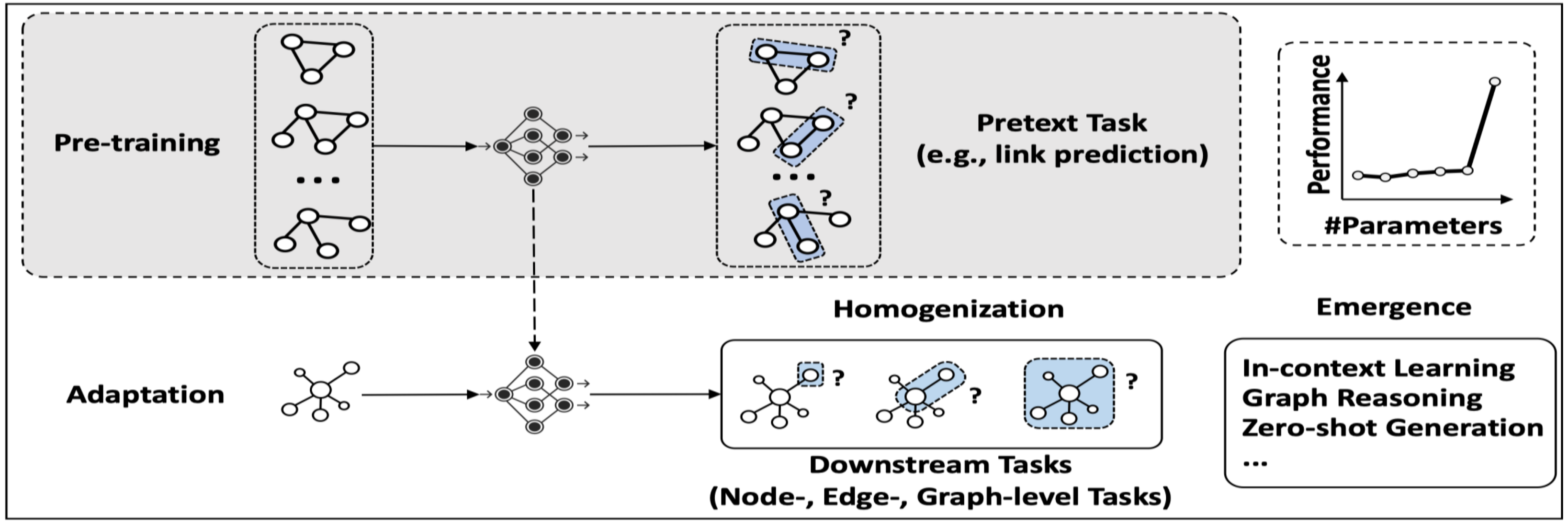
Information Bottleneck in GNNs

Graph Foundation Models

A graph foundation model (GFM) is a model pre-trained on extensive graph data, adapted for diverse downstream graph tasks.



(a) Deep Graph Learning.



(b) Graph Foundation Models.

Jiawei Liu, Cheng Yang, Zhiyuan Lu, Junze Chen, Yibo Li, Mengmei Zhang, Ting Bai, Yuan Fang, Lichao Sun, Philip S. Yu, Chuan Shi. Towards Graph Foundation Models: A Survey and Beyond. arXiv 2023.

Characteristics of Graph Foundation Models

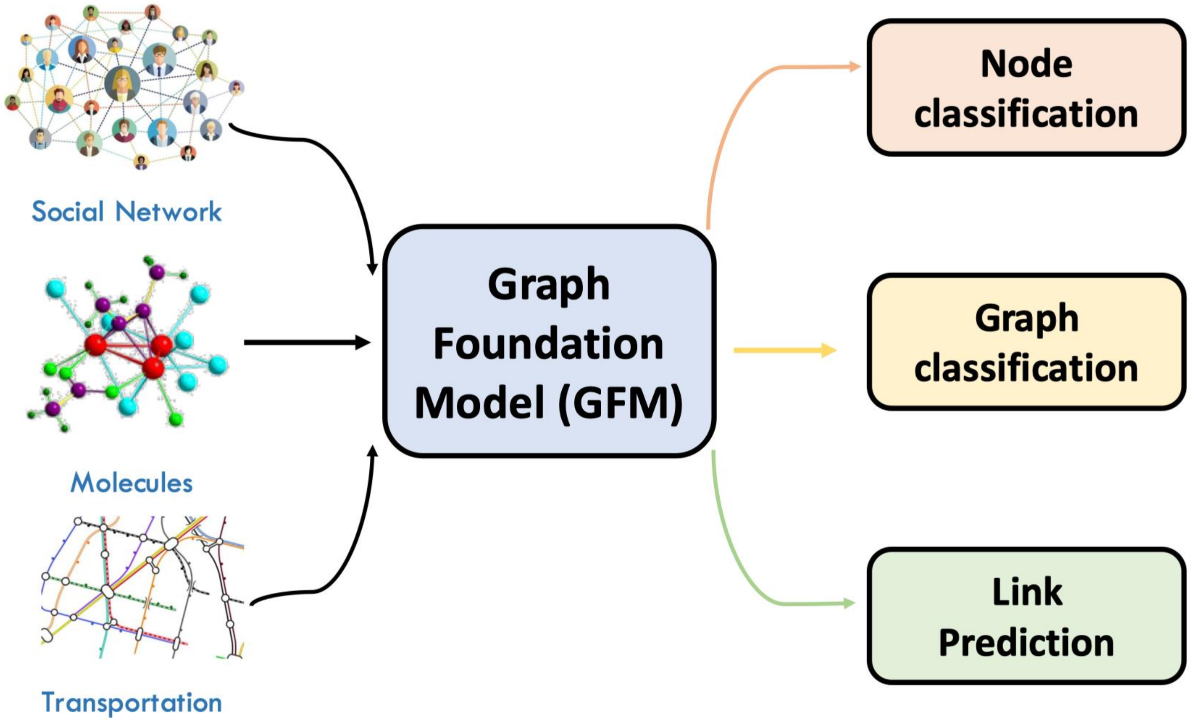
Two Characteristics

Emergence

- ➢ Novel capbility when larger model or more graph data
  - graph reasoning
  - graph generation…

Homogenization

- ➢ Apply to different formats of tasks
  - node/edge/graph tasks

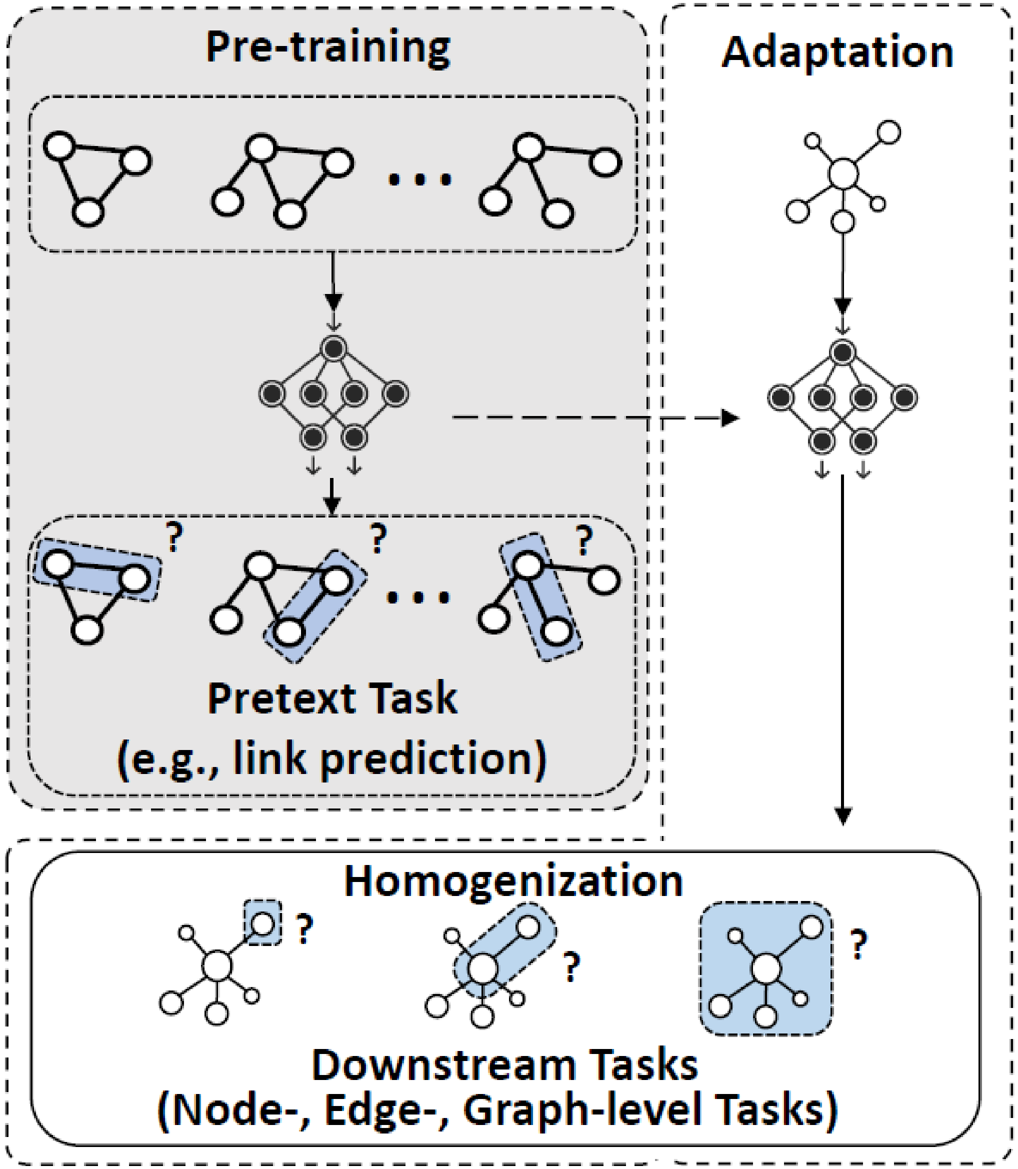


Key Techniques of Graph Foundation Models

Key Techniques of GFMs

- **Pre-training**: neural networks are trained on a large graph dataset in a self-supervised manner
  - contrastive pre-training: contrastive positive samples against negative samples
  - generative pre-training: reconstruct or predict original feature

- **Adaptation**: adapt pre-trained models to specific downstream tasks or domains to enhance their performance
  - fine-tuning
  - prompt-tuning



GFMs v.s. LLMs

Similarities: common goal and similar learning paradigm

Differences: (1) different data and tasks; (2) technological differences

		Language Foundation Model	Graph Foundation Model
Similarities	Goal	Enhancing the model's expressive power and its generalization across various tasks	
	Paradigm	Pre-training and Adaptation	
Intrinsic differences	Data	Euclidean data (text)	Non-Euclidean data (graphs) or a mixture of Euclidean (e.g., graph attributes) and non-Euclidean data
	Task	Many tasks, similar formats	Limited number of tasks, diverse formats
Extrinsic differences	Backbone Architectures	Mostly based on Transformer	No unified architecture
	Homogenization	Easy to homogenize	Difficult to homogenize
	Domain Generalization	Strong generalization capability	Weak generalization across datasets
	Emergence	Has demonstrated emergent abilities	No/unclear emergent abilities as of the time of writing

Outline

- Graph Foundation Models

√ Progress in Related Work

- Challenges and Future Direction

Taxonomy of Related Work

No GFMs until now, but a lot of explorations is on the way.

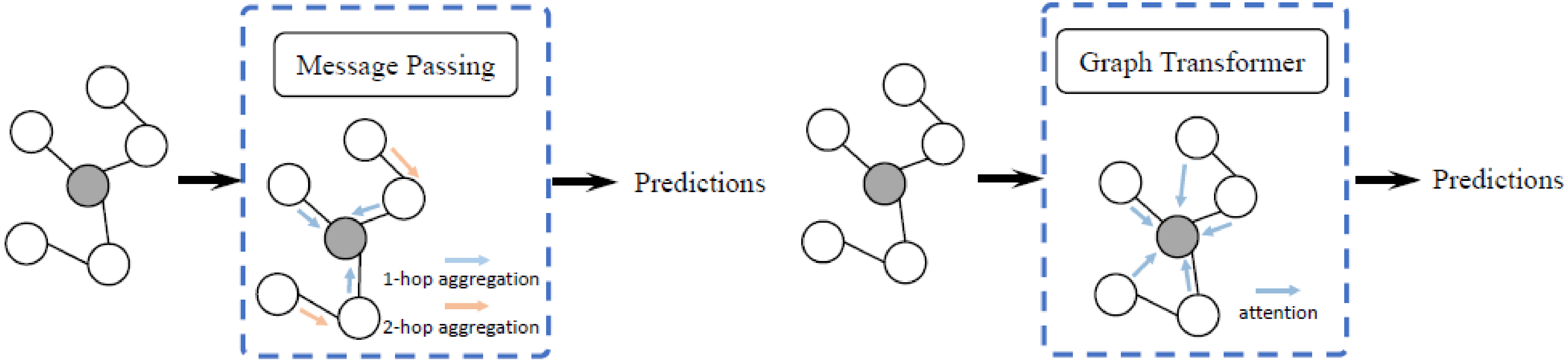
Categorize existing explorations into three distinct groups according to the dependence on GNNs and LLMs

- **Towards Graph Foundation Models**
  - GNN-based Models
    - Backbone Architectures
      - Message Passing-based 4.1.1
      - Graph Transformer-based 4.1.2
    - Pre-training
      - Contrastive Methods 4.2.1
      - Generative Methods 4.2.2
    - Adaptation
      - Fine-Tuning 4.3.1
      - Prompt-Tuning 4.3.2
  - LLM-based Models
    - Backbone Architectures
      - Graph-to-Token 5.1.1
      - Graph-to-Text 5.1.2
    - Pre-training
      - Language Modelling 5.2.1
      - Masked Language Modelling 5.2.2
    - Adaptation
      - Manual Prompting 5.3.1
      - Automatic Prompting 5.3.2
  - GNN+LLM-based Models
    - Backbone Architectures
      - GNN-centric 6.1.1
      - Symmetric 6.1.2
      - LLM-centric 6.1.3
    - Pre-training
      - GNN or LLM-based 6.2.1
      - Alignment-based 6.2.2
    - Adaptation
      - Fine-Tuning 6.3.1
      - Prompt-Tuning 6.3.2

GNN-based Models

Seeking to enhance current graph learning through innovative approaches in GNN model architectures, pre-training, and adaptation.

- Architectures: Graph Transformer, e.g., Specformer (ICLR23), CoBFormer (ICML24)
- Pre-training: Graph Pretraining, e.g., PT-HGNN (KDD21), GraphPAR (WWW24)
- Adaptation: Graph Prompt, e.g., All In One (KDD23), MultiGPrompt (WWW24)



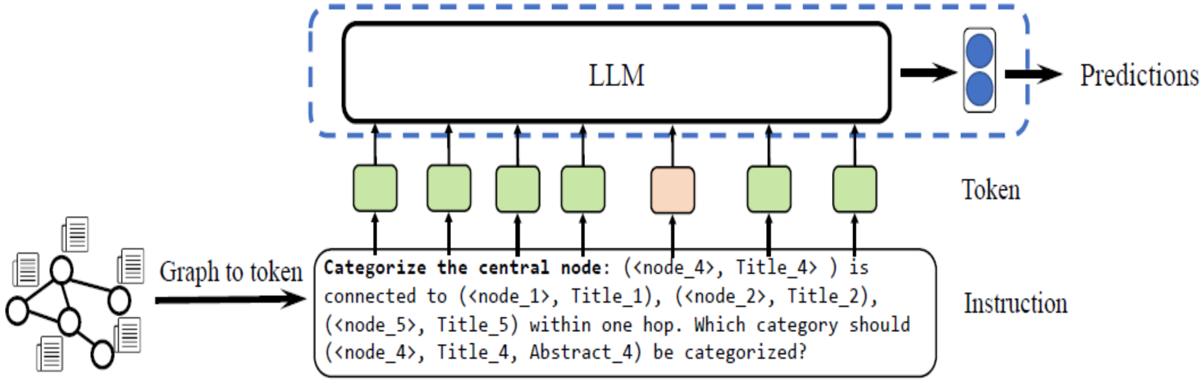
(a) Message Passing.

(b) Graph Transformer.

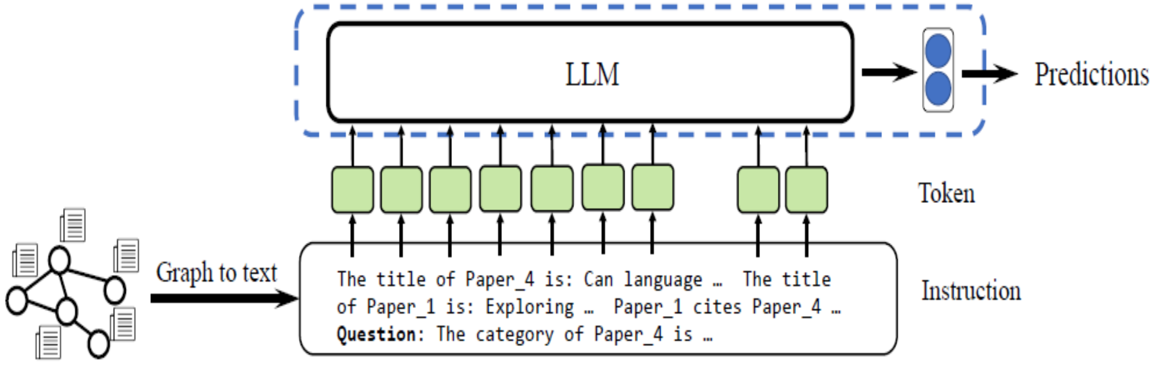
LLM-based Models

Exploring the feasibility of transforming graphs into text or tokens to leverage LLMs as foundation models.

- Graph-to-Token: transform graphs into tokens and then input them into LLMs
  - e.g., InstructGLM
- Graph-to-Text: transform graphs into texts and then input them into LLMs
  - e.g., NLGraph (NIPS24), LLM4Mol



(a) Graph-to-token.

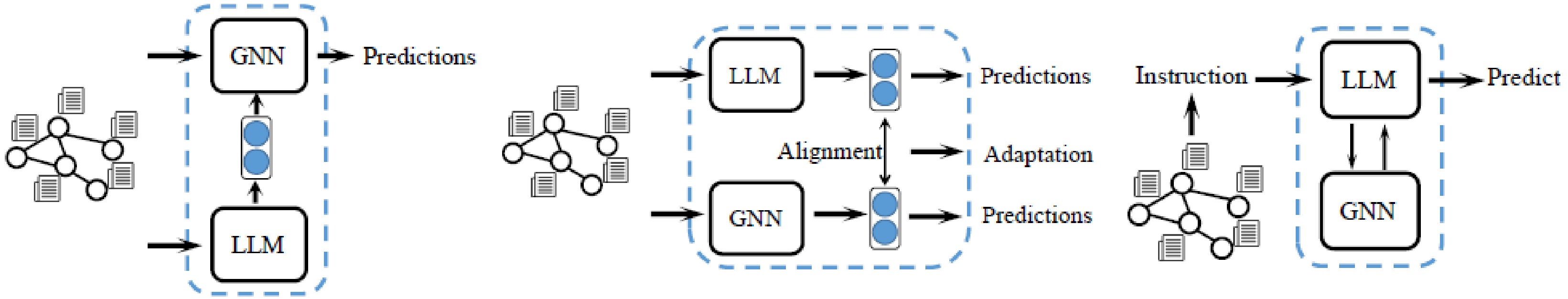


(b) Graph-to-text.

GNN+LLM-based Models

Exploring synergies between GNNs and LLMs to enhance graph learning.

- GNN-centric Models: utilize LLM to extract node feature and make predictions using GNN
  - e.g., SimTeG, TAPE
- Symmetric Models: align the embeddings of GNN and LLM
  - e.g., GraphTranslator (WWW24), G2P2 (SIGIR23), ConGrat
- LLM-centric Models: utilize GNNs to enhance the performance of LLM
  - e.g., Graph-Toolformer



Outline

- Graph Foundation Models
- Progress in Related Work

√ Challenges and Future Direction

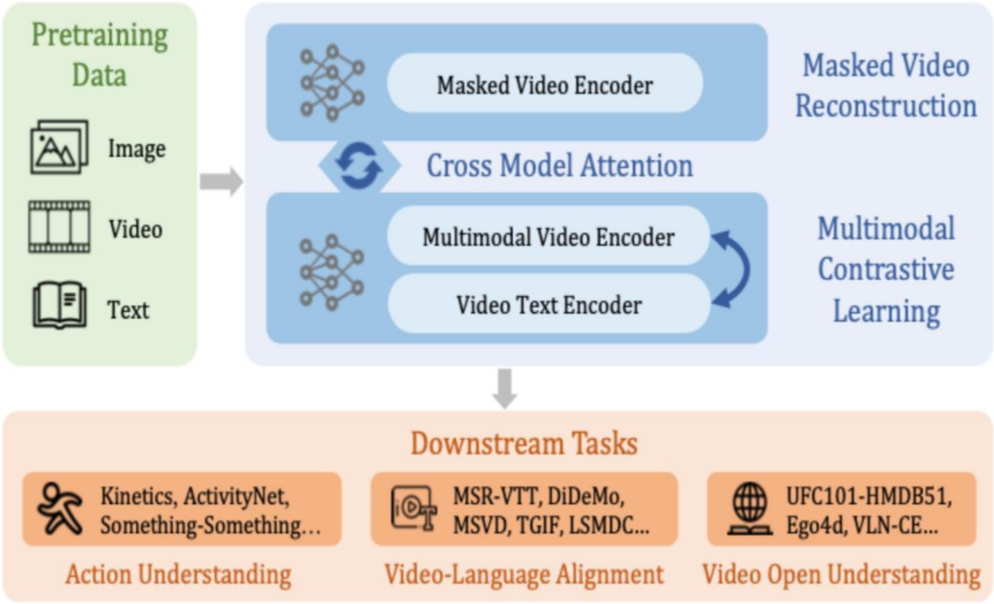
Challenges in Model

Model Architectures

- ➢ It remains unknown whether current architectures are optimal choices.
- ➢ Multimodal foundation models
  - Using graph to extend the multiple modalities…

Model Training

- ➢ Is there uniform pretext tasks for graph
- ➢ Some ideas from other directions
  - knowledge distillation
  - reinforcement learning from human feedback
  - model editing…

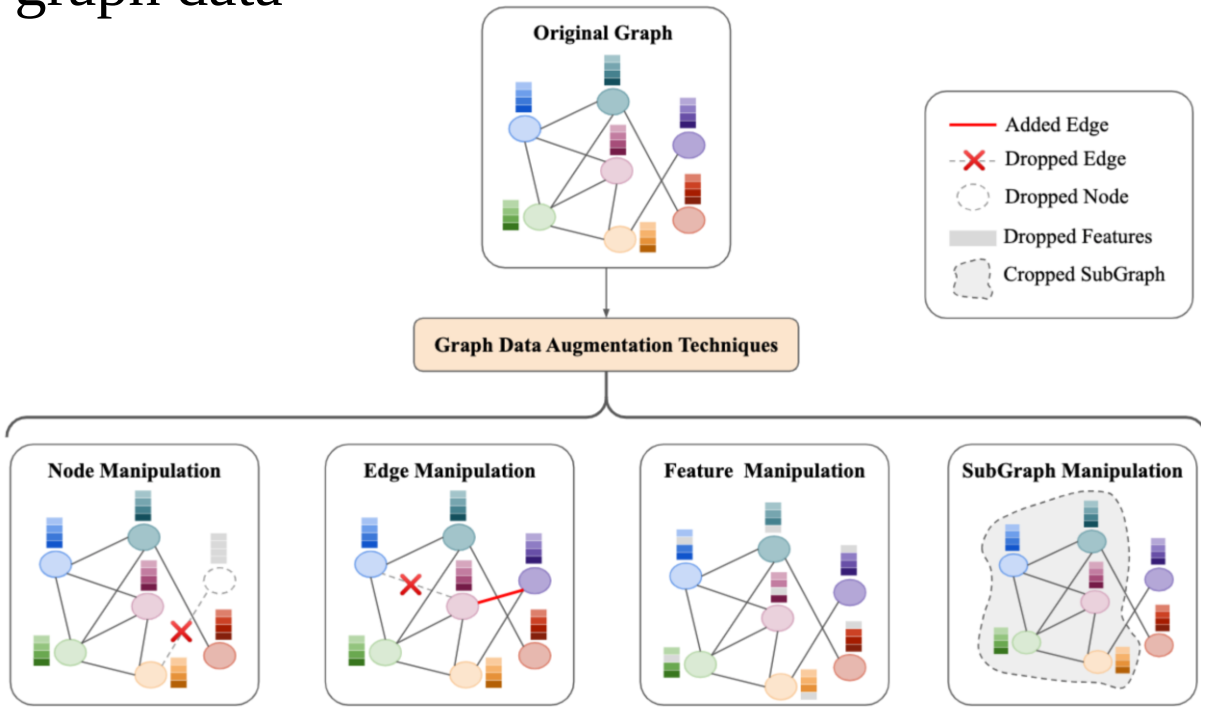


Multimodal Foundation Models

Challenges in Data and Evluation

Data Quantity and Quality

- Limited amount of open-source large-scale graph data
  - concentrated in a single domain
- Using augmentation strategies
  - graph structure learning
  - feature completion
  - label mixing…



Graph Augmentation

Evaluation

- Lacking labels in open-ended tasks
  - human evaluation
  - meta-evaluation
- Evaluating robustness, trustworthiness, holistic performance…

Challenges in Applications

Killer Applications

- It is not yet clear that graph foundation models can similarly catalyze groundbreaking applications in graph tasks.
- Promising fields
  - urban computing
  - drug development…

Safety

- Black-box nature introduces safety concerns.
  - hallucination
  - privacy leaks
- Promising technologies
  - counterfactual reasoning…

图数据挖掘和机器学习

扫码关注我们

www.shichuan.org

Thanks

Q&A