AI Research and Applications in the Era of Big Data and "Large" Models

Dejing Dou (窦德景）

Distinguished Professor, Fudan University

Chief Scientist, BEDI Cloud

Nov 5, 2024 @SMU, Singapore

Outline

- A Short Self Introduction and Summary
- AI, Big Data, and “Large” (Foundation) Models
- Research and Case Studies

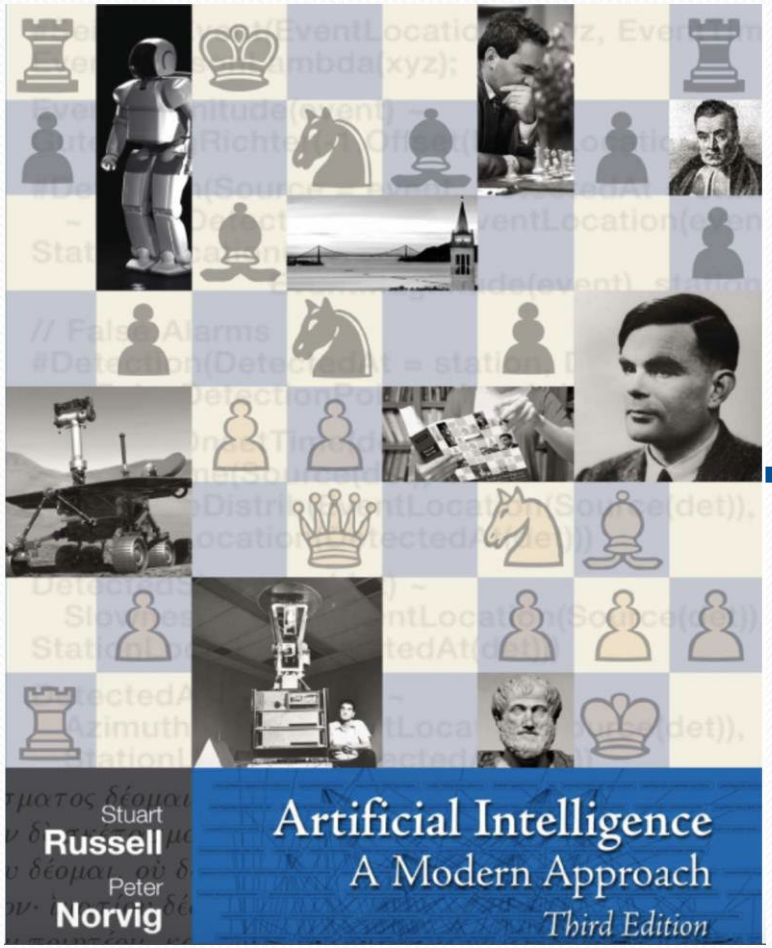
Previous Background

- B.E. degree in Electronic Engineering, Tsinghua University, 1996
- M.S. degree in Electrical Engineering, Yale University, 2000
- Ph.D. degree in Artificial Intelligence, Yale University, 2004
- Assistant, Associate（tenured）, Full（tenured）Professors, Computer and Information Science, University of Oregon, 2004-2022
- Visiting Associate Professor, Biomedical Informatics Research Center, Stanford University, 2012-2013
- Director of the NSF IUCRC Center for Big Learning (CBL) , University of Oregon, 2018-2020
- Head of Big Data Lab and Business Intelligence Lab, Baidu Research, 2019-2022
- Chief Data Scientist, Partner and Vice President, BCG GC, 2022-2024
- Adjunct Professor, Electronic Engineering Department, Tsinghua University, 2023-2026

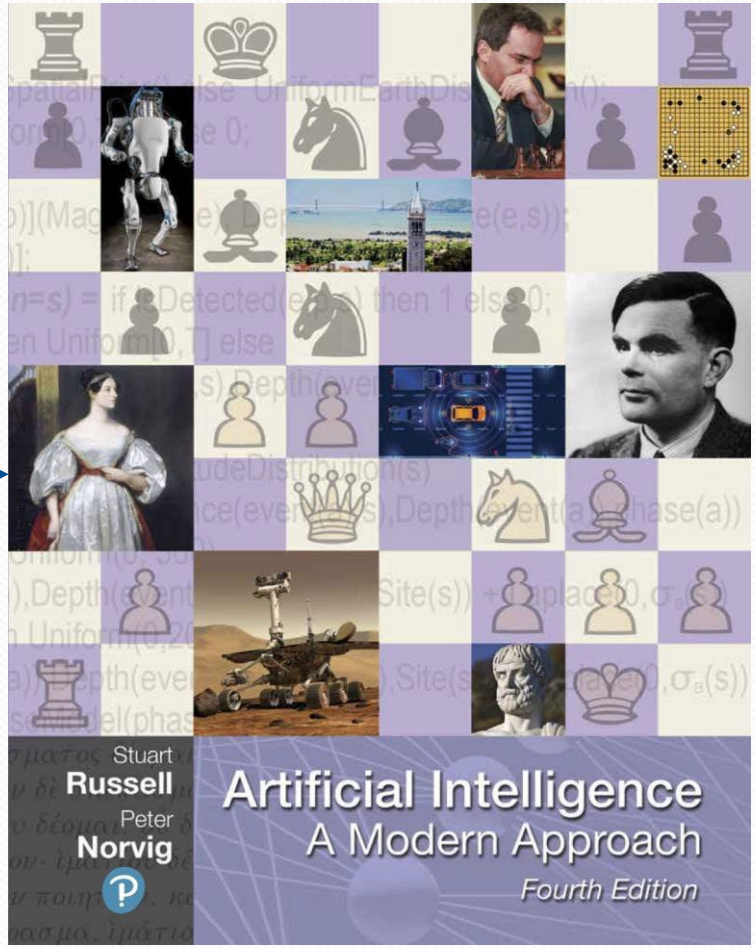
Academic Achievements

- **Research**
  - Published more than 250 research papers, including AAAI, IJCAI, ICML, NeurIPS, ICLR, KDD, ICDM, ICDE, ACL, EMNLP, CVPR, ICCV, CIKM, ISWC, AIJ, JMLR, TPAMI, TKDE, TKDD, KAIS,
    JIIS, and Nature Sustainability, with more than 10000 Google Scholar citations.
  - Received over $5 million PI research grants from the NSF and the NIH (two R01s)
- **Teaching and Student Supervising**
  - Taught AI, Machine Learning, Data Mining, Data Science, Database classes at UO.
  - Graduated 7 Ph.D.s and 2 Posdocs, 3 of them became tenured professors in US, 2 of their students became tenure track assistant professors in US in 2023
  - Supervised other 10 Ph.D. students and 9 MS students, they went to top IT companies including Google, Facebook, Microsoft, Amazon, Apple, Intel, Cisco…
- **Services**
  - PC co-chairs for IEEE BigData 2023 (Industry Track), KDD Cup 2022 (leading Organizer), IEEE ICMLA 2020, FFSE 2018, ODBASE 2013, IEEE IPCCC 2011, IEEE ISDPE 2010
  - Senior member of AAAI, ACM, and IEEE

AI Textbooks



2009

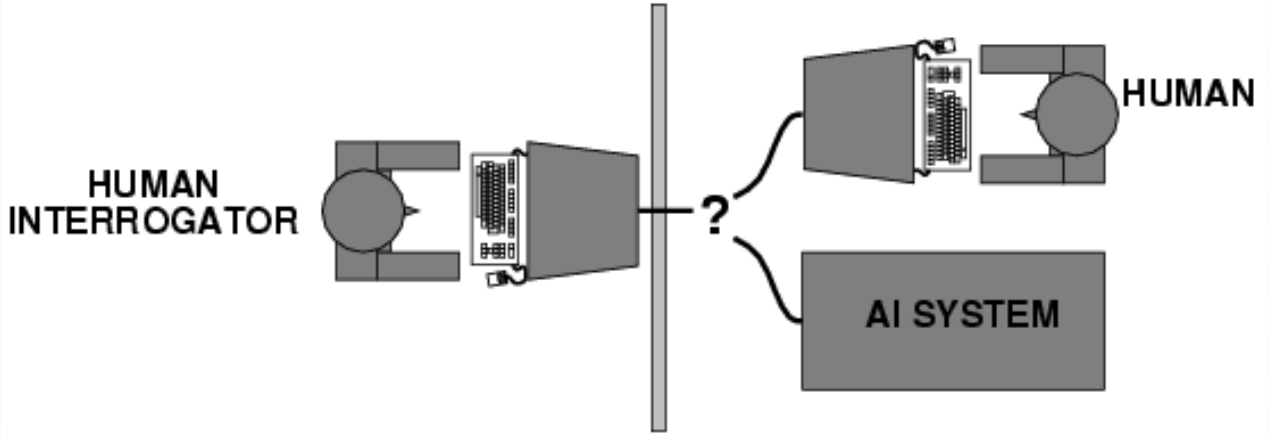


2020

21 Deep Learning. (40 pages)
24 Deep Learning for Natural Language Processing (26 pages)

Acting humanly: Turing Test

- Turing (1950) "Computing machinery and intelligence":
- "Can machines think?" ➔ "Can machines behave intelligently?"
- Operational test for intelligent behavior: the Imitation Game



- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in following 50 years
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

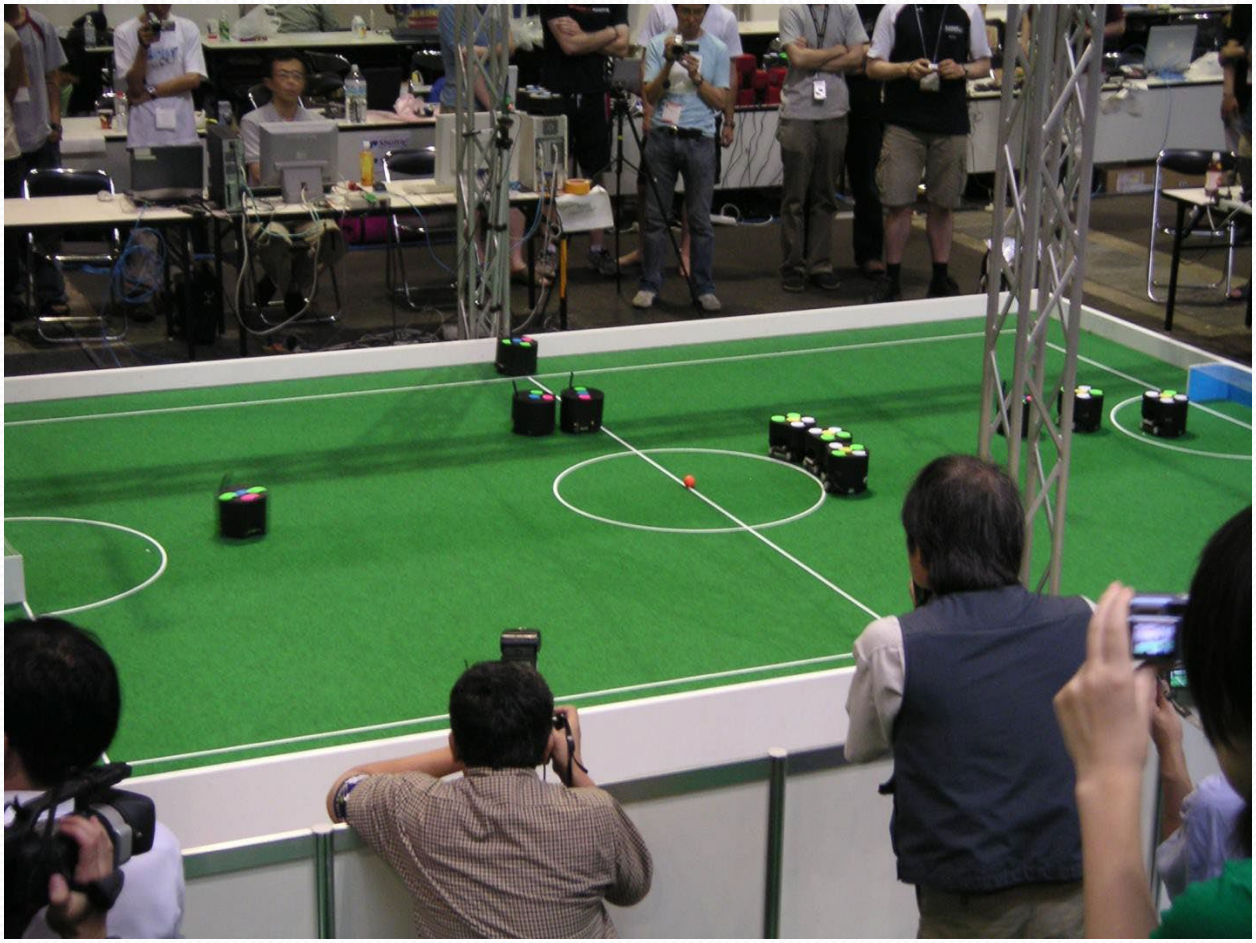
Football Version of Turing Test

- *By the year 2050, develop a team of fully autonomous humanoid robots that can win against the human world cup champion team.*




Robocup: Small Size Group

- *I joined AI in 2000 because of I knew how to do Radio Communication ☺*

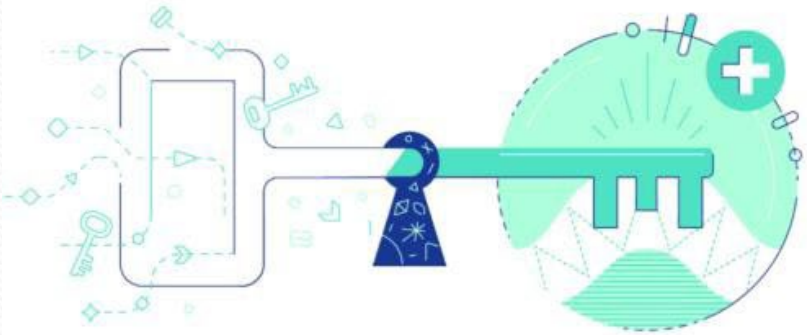


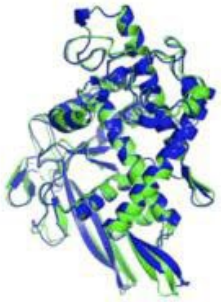
AlphaGo (GO, 2016)

AlphaGo Zero (2017)

AlphaFold2 (2020) and AlphaFold3 (2024)

DeepMind

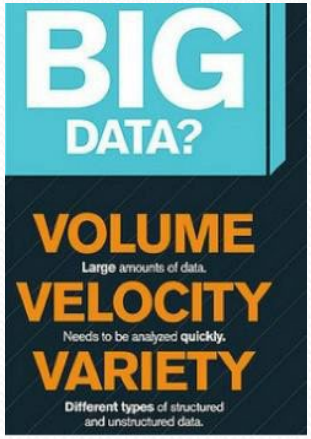


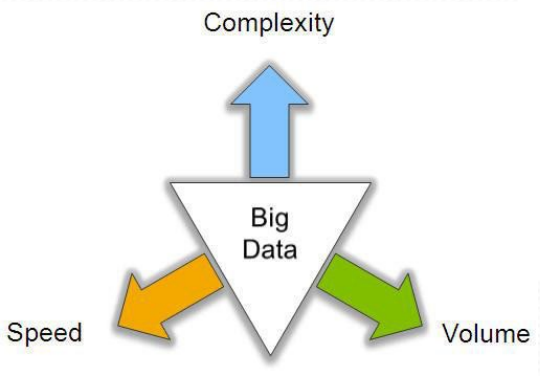


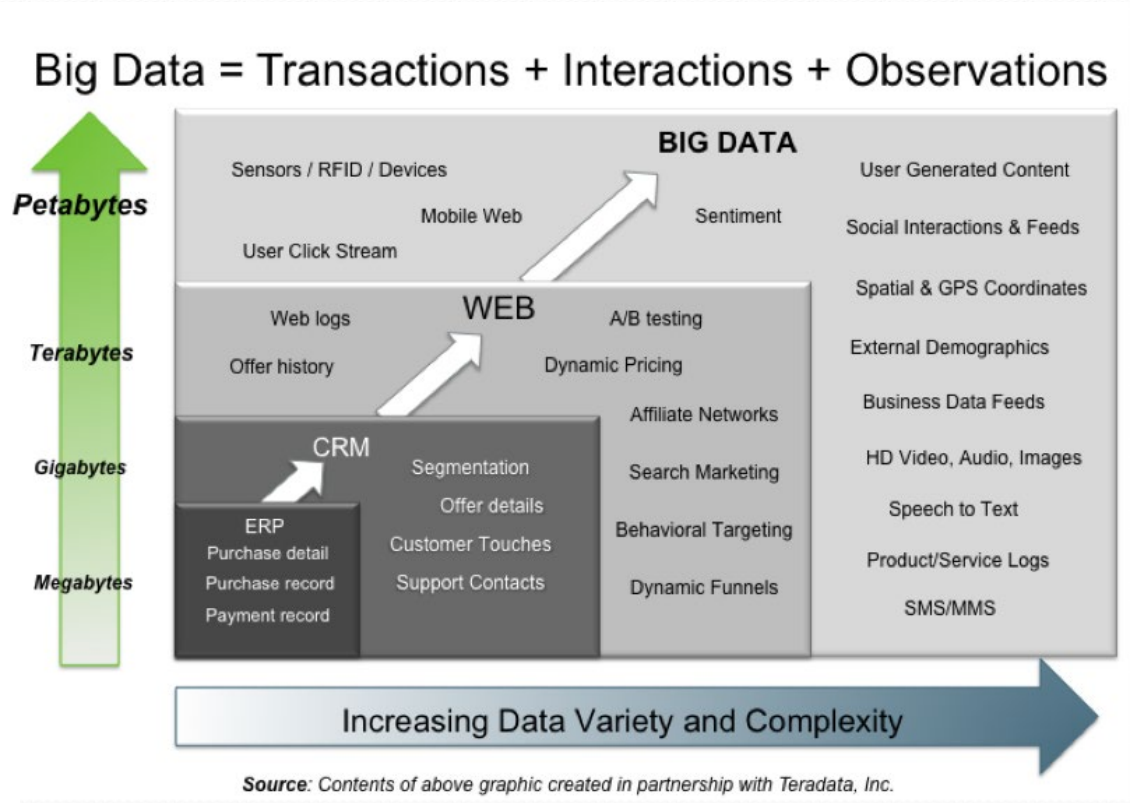
T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)

T1049 / 6y4f
93.3 GDT
(adhesin tip)

Big Data: 3V's

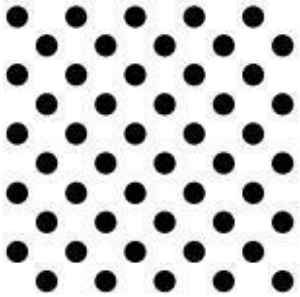
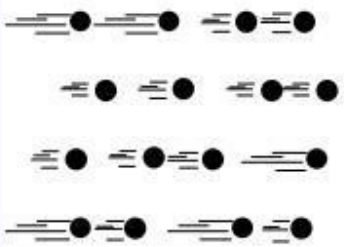
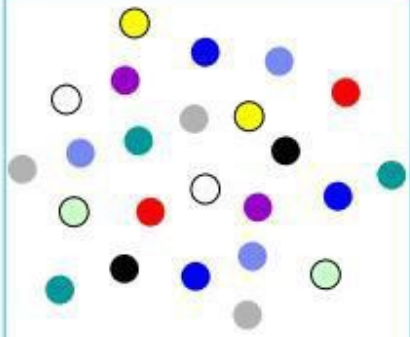
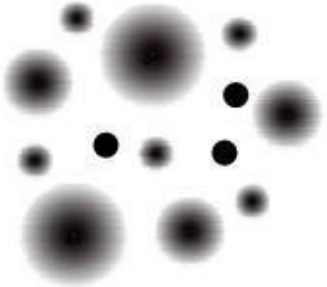




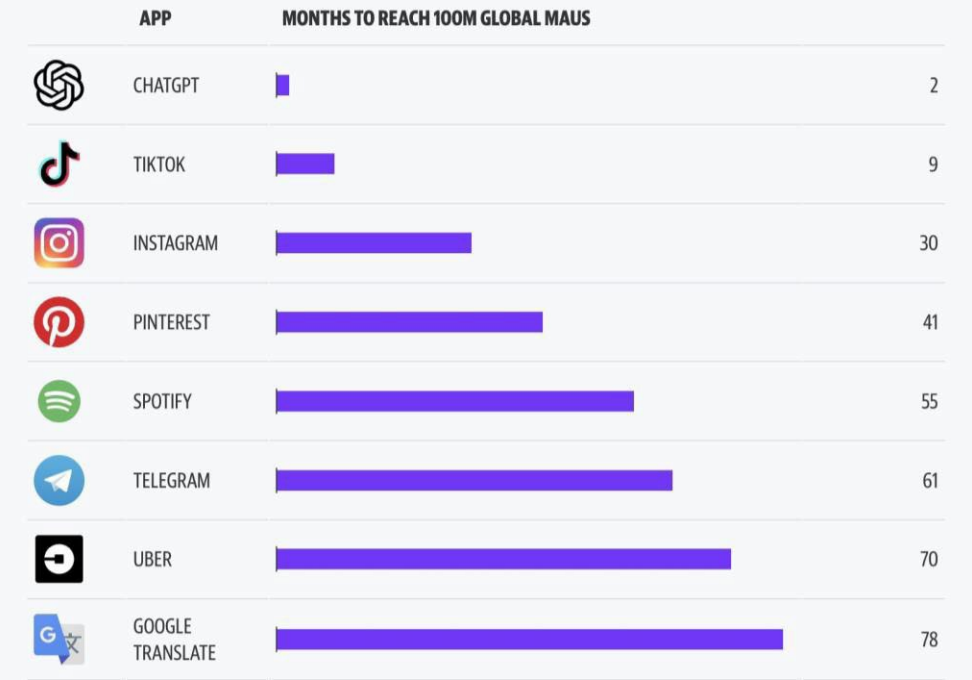


***Source**: Contents of above graphic created in partnership with Teradata, Inc.*

Some Make it 4V's

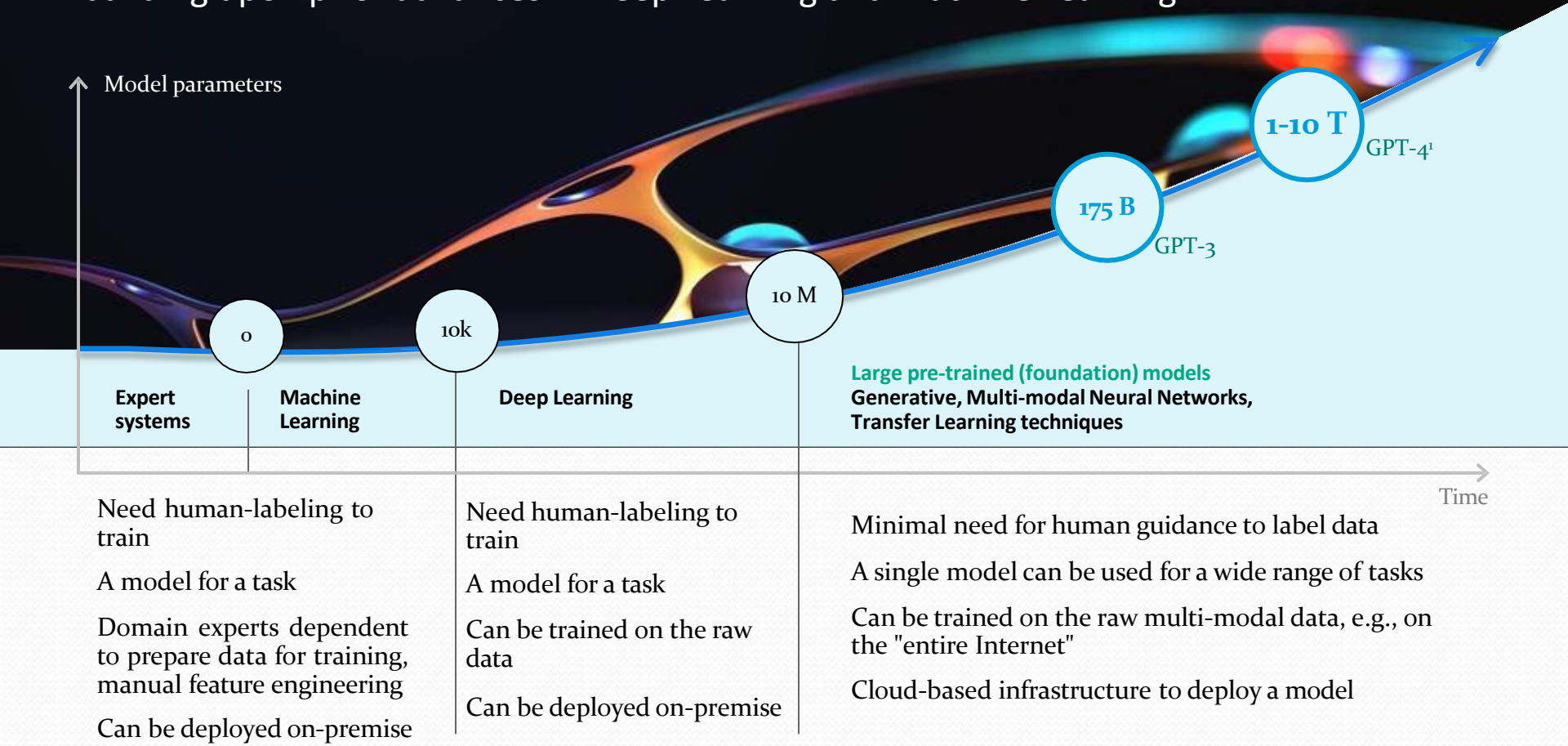
Volume	Velocity	Variety	Veracity*
			
Data at Rest	**Data in Motion**	**Data in Many Forms**	**Data in Doubt**
Terabytes to exabytes of existing data to process	Streaming data, milliseconds to seconds to respond	Structured, unstructured, text, multimedia	Uncertainty due to data inconsistency & incompleteness, ambiguities, latency, deception, model approximations

ChatGPT reached 100M users in less than 2 months, a fraction of the time it took previous viral hits



Sources: UBS, yahoo finance

Foundation ("Large") Models are an emerging paradigm for Artificial Intelligence, building upon prior advances in Deep Learning and Machine Learning



Need human-labeling to train

A model for a task

Domain experts dependent to prepare data for training, manual feature engineering

Can be deployed on-premise

Need human-labeling to train

A model for a task

Can be trained on the raw data

Can be deployed on-premise

Minimal need for human guidance to label data

A single model can be used for a wide range of tasks

Can be trained on the raw multi-modal data, e.g., on the "entire Internet"

Cloud-based infrastructure to deploy a model

1. # parameters used in GPT-4 not publicly available, reliable estimates range from 1-20 T

Topic 1: Spatiotemporal Data Mining in Smart City, Climate, and Clean Energy

Spatio-temporal Big Data of Baidu

6 billion queries per day

150 million POI,
10 million kilometer roads

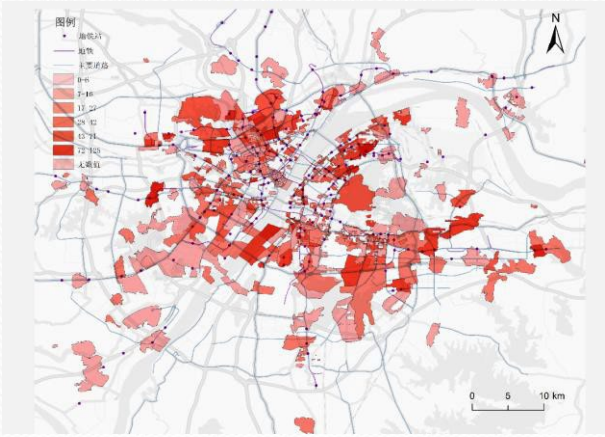
130 billion positioning request per day

Urban Analysis with Data Mining and Machine Learning

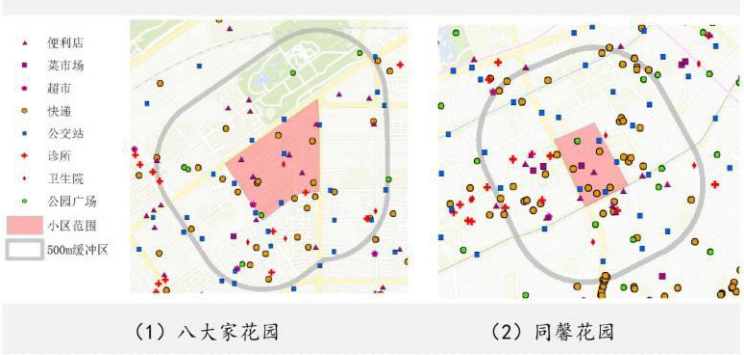
- Build urban cognitive ability based on Baidu bigdata including search query, user profile data and satellite image.
- Continue to optimize Happy City Index with Xinhua News agency
- Develop automatic city report generation system, and support the report of Xiongan and Wenzhou.

Urban Analysis – Transfer Learning

- It is very useful to detect the high-risk neighborhood for potential epidemic
- Our objective：
  - Design a city transfer learning framework
  - Learn the characteristics of high-risk neighborhoods in epicenter cities
  - Transfer that knowledge to the target city to make prediction



high-risk neighborhood in Wuhan



The distribution of living facility in a neighborhood

Congxi Xiao, Jingbo Zhou, Jizhou Huang, An Zhuo, Ji Liu, Haoyi Xiong, and Dejing Dou. “C-Watcher: A Framework for Early Detection of High- Risk Neighborhoods Ahead of COVID-19 Outbreak." In Proceedings of AAAI 2021. pp. 4892-4900, 2021.

Urban Analysis – Climate Change

- Work on research report about Health and Climate Change in China
- Using the Baidu query data to analyze the individual engagement in health and climate change.
- The report is published in *The Lancet Public Health*

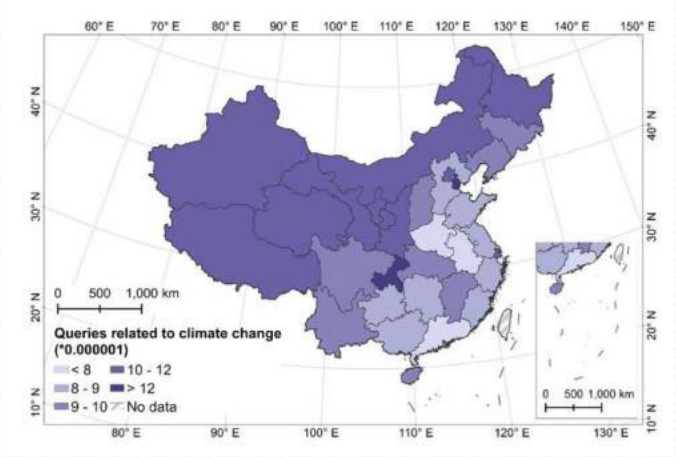


Figure 44: The distribution of the proportion of the queries related to climate change in different provinces in China in 2019.

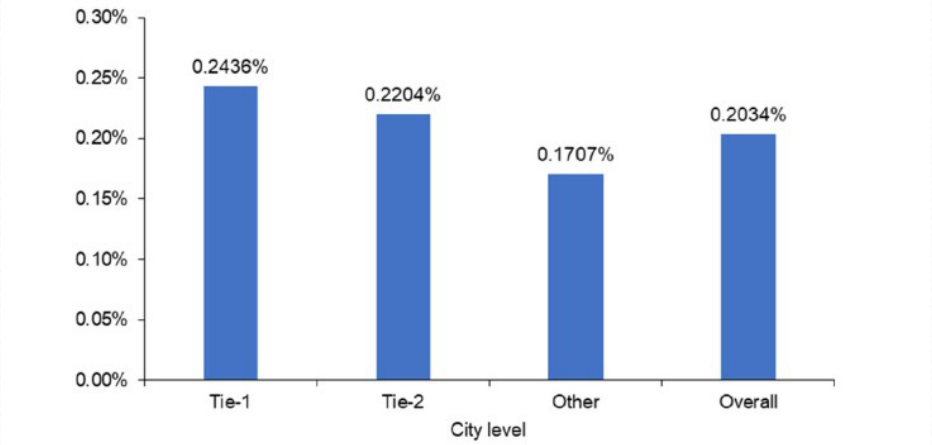


Figure 43: Share of health&climate change co-queries out of total climate change queries in tie-1, tie-2 and other cities

Urban Analysis – Air Pollution

- Air pollution is likely to exacerbate the risk factors for resident health
- There is still a lack of a large-scale nationwide quantification of
- the mental health risks posed by air pollution
- Use internet search data across 252 cities in China to estimate the impact of short-term and long-term exposure to air pollution on urbanites'mental health at the city level.
- The paper has been published in Nature Sustainability

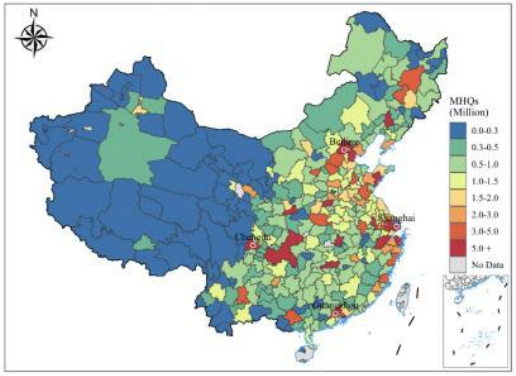
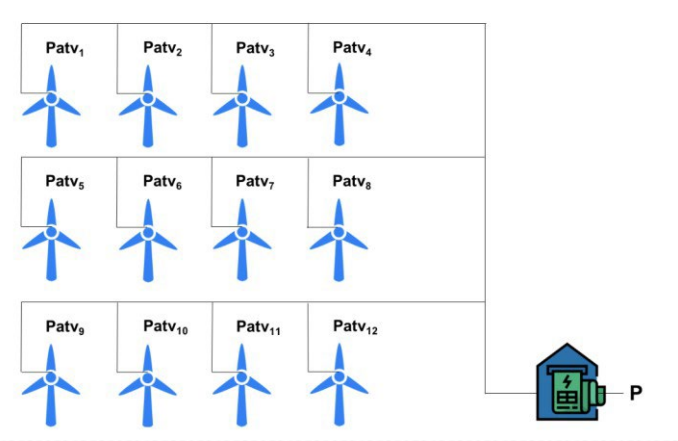


Fig. 1 The density distribution of MHQs in China. The spatial distribution of geo-tagged search queries about mental health on Baidu. Four major cities are marked—Beijing, Shanghai, Guangzhou, and Chengdu.

Zhi Cao, Jingbo Zhou, Meng Li, Jizhou Huang, and Dejing Dou. "Urbanites' mental health undermined by air pollution." Nature Sustainability. 6, 470-478, 2023.

Wind Power Forecasting (with Longyuan)



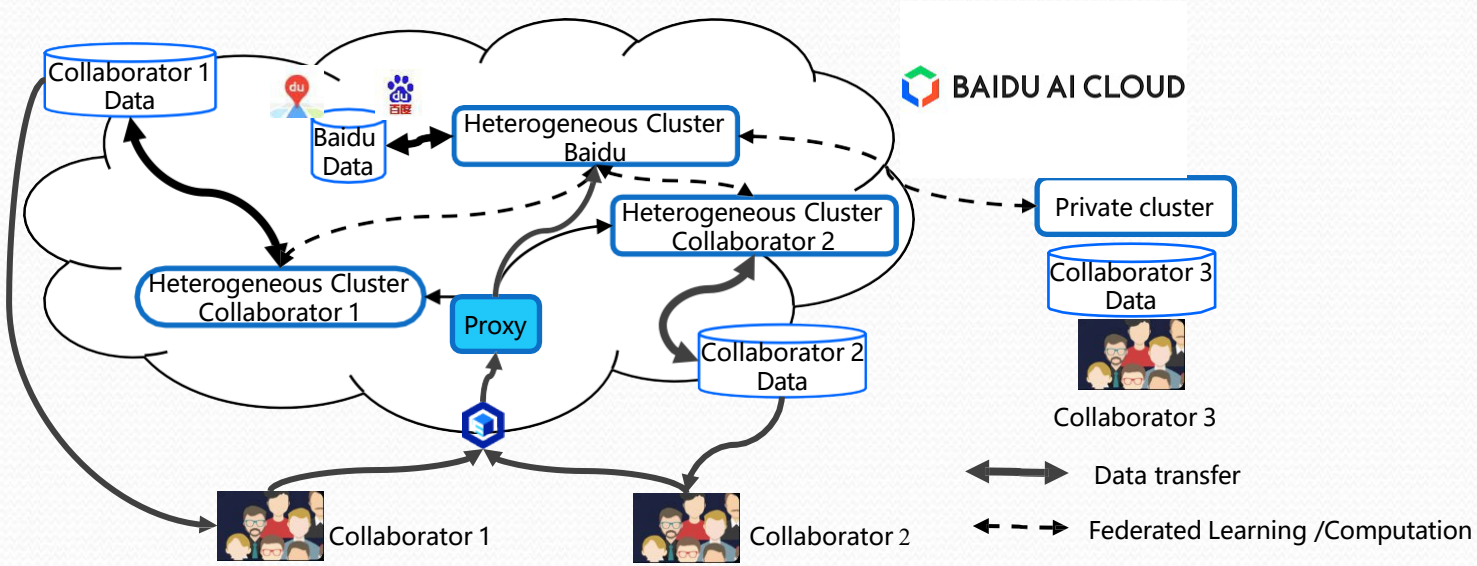
- Wind power plays a leading role in electricity production in the renewable energy sector;
- However, the uncertainties and fluctuations of wind power lay significant obstacles to its use in practice;
- Wind Power Forecasting (WPF) thus becomes one of the most critical issues in wind power integration & operation.

- Holds KDD Cup 2022 (“Spatial Dynamic Wind Power Forecasting Challenge")
- Yan Li, Xinjiang Lu, Yaqing Wang, Dejing Dou. “Generative Time Series Forecasting with Diffusion, Denoise, and Disentanglement." In NeurIPS 2022, 2022.

Topic 2: Federated Learning in data privacy protection and Trustworthy AI

Data federation

- A data federation system for the collaboration among research institutes (e.g., universities and Baidu) while ensuring data privacy.
  - Data privacy & model security & efficient federated learning & distributed machine learning & Interpretability



- Ji Liu, Jizhou Huang, Yang Zhou, Xuhong Li, Shilei Ji, Haoyi Xiong, and Dejing Dou. “From distributed machine learning to federated learning: a survey." Knowl. Inf. Syst. 64(4): 885-917, 2022
- Ji Liu, Lei Mo, Sijia Yang, Jingbo Zhou, Shilei Ji, Haoyi Xiong, Dejing Dou. “Data Placement for Multi-Tenant Data Federation on the Cloud." IEEE Trans. Cloud Comput. 11(2): 1414-1429, 2023.

FedCube: Safety-aware Data Collaboration for the Fight against COVID-19

Our Research Contributions in Federated Learning (1)

- Multi-Job Federated Learning: A cost model based on the training time and the data fairness. A reinforcement learning and optimization- based method to schedule devices for multiple jobs. (AAAI'2022, TPDS'2023)

- Federated Learning on Server Data: A dynamic server update algorithm to determine the optimal steps of the server update. A layer-adaptive model pruning method to achieve a balance between efficiency and effectiveness. (IJCAI'2022)

- An approach to enable data processing on the cloud with the data from different organizations. An algorithm in order to partition and store data on the cloud so as to achieve multiple objectives while satisfying the constraints based on a multi-objective cost model. (TCC'2023)

- Cross-Device Federated Learning." IEEE Trans. Parallel Distributed Syst. 34(2): 535-551, 2023.constraints based on a multi-objective cost model. (TCC'2023)
- Chendi Zhou, Ji Liu, Juncheng Jia, Jingbo Zhou, Yang Zhou, Huaiyu Dai, and Dejing Dou. "Efficient device scheduling with Multi-Job Federated Learning." In Proceedings of AAAI 2022. pp. 9971-9979, 2022.
- Hong Zhang, Ji Liu, Juncheng Jia, Yang Zhou, Huaiyu Dai, and Dejing Dou. "FedDUAP: Federated Learning with Dynamic Update and Adaptive Pruning Using Shared Data on the Server." In Proceedings of IJCAI 2022. pp. 2776-2782, 2022.
- Ji Liu, Lei Mo, Sijia Yang, Jingbo Zhou, Shilei Ji, Haoyi Xiong, Dejing Dou. "Data Placement for Multi-Tenant Data Federation on the Cloud." IEEE Trans. Cloud Comput. 11(2): 1414-1429, 2023.

Our Research Contributions in Federated Learning (2)

- Multi-Job Federated Learning: A cost model based on the training time and the data fairness. A reinforcement learning and optimization- based method to schedule devices for multiple jobs. (AAAI'2022, TPDS'2023)

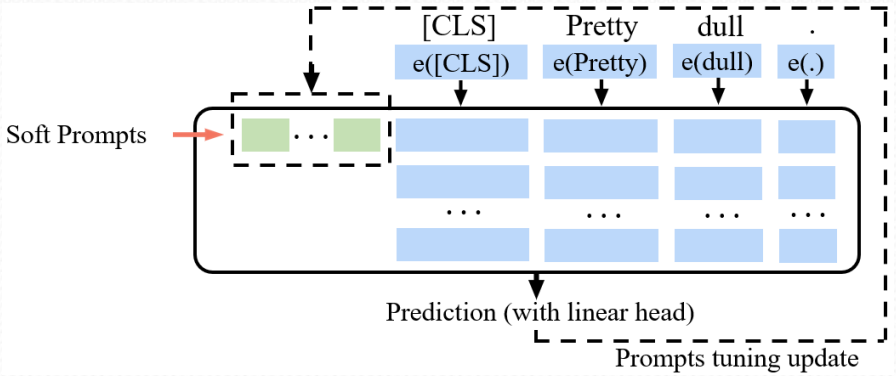
- Federated Learning on Server Data: A dynamic server update algorithm to determine the optimal steps of the server update. A layer-adaptive model pruning method to achieve a balance between efficiency and effectiveness. (IJCAI'2022)

- An approach to enable data processing on the cloud with the data from different organizations. An algorithm in order to partition and store data on the cloud so as to achieve multiple objectives while satisfying the constraints based on a multi-objective cost model. (TCC'2023)

- Jiayin Jin, Jiaxiang Ren, Yang Zhou, Lingjuan Lyu, Ji Liu, and Dejing Dou. "Accelerated Federated Learning with Decoupled Adaptive Optimization." In ICML 2022. pp. 10298-10322, 2022.
- Guanghao Li, Yue Hu, Miao Zhang, Ji Liu, Quanjun Yin, Yong Peng, and Dejing Dou."FedHiSyn: A Hierarchical Synchronous Federated Learning Framework for Resource and Data Heterogeneity." In ICPP 2022. pp. 8:1-8:11, 2022.
- Tianshi Che, Yang Zhou, Zijie Zhang, Lingjuan Lyu, Ji Liu, Da Yan, Dejing Dou, Jun Huan. "Fast Federated Machine Unlearning with Nonlinear Functional Theory." In ICML 2023. pp. 4241-4268, 2023.

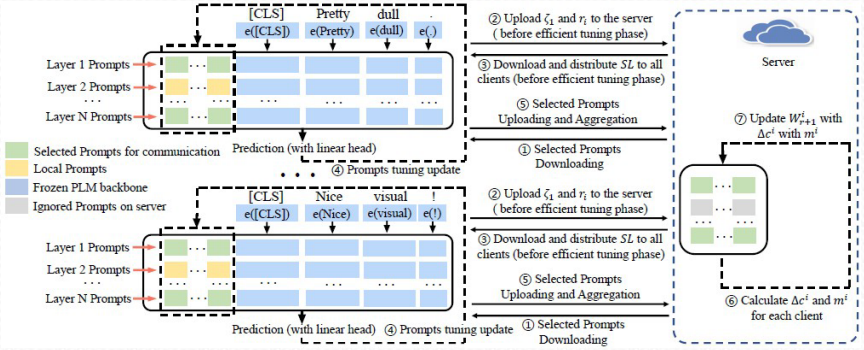
Most Recent: Federated Learning for LLMs

- The Large Language Models (LLMs) correspond to huge communication and computation costs in fine-tuning
- The pre-training or the fine-tuning process is almost inapplicable in Federated learning (FL) scenarios
- The prompt design can lead to excellent performance while freezing the original LLMs
  - Hard prompt methods: search proper prompts within a discrete space of words
  - Soft prompt methods: optimize continuous prompts in tuning
- Adaptive optimization methods
  - Adaptive moment estimation (Adam)
  - SGD with momentum (SGDM)
  - Can be deployed on server side or on device side



Our Contributions (in EMNLP 2023 and AAAI 2024)

- **Parameter-efficient prompt tuning**
  - Selects a proper set of layers to communicate the prompt parameters
  - We propose a scoring method to measure the importance of each layer
  - We propose a lossless layer selection method to determine the optimal number of selected layers
- **Communication-efficient adaptive optimization**
  - Applies adaptive optimization on:
    - Server based on SGDM
    - Device based on Adam
  - We reset the momentum buffer zero during local updates
    - To avoid extra communication
  - We maintain a state for each de on the server
    - To mitigate client drift problems

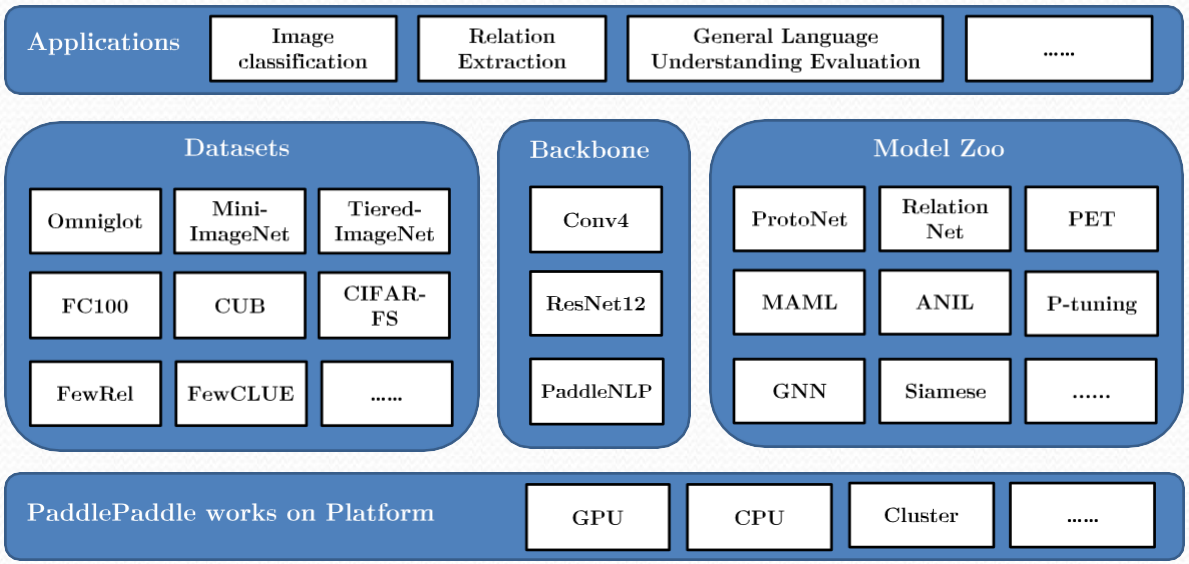


- Tianshi Che, Ji Liu, Yang Zhou, Jiaxiang Ren, Jiwen Zhou, Victor Sheng, Huaiyu Dai, Dejing Dou. “Federated Learning of Large Language Models with Parameter-Efficient Prompt Tuning and Adaptive Optimization” In EMNLP 2023， pp. 7871-7888
- Ji Liu, Juncheng Jia, Tianshi Che, Chao Huo, Jiaxiang Ren, Yang Zhou, Huaiyu Dai, Dejing Dou. "FedASMU: Efficient Asynchronous Federated Learning with Dynamic Staleness-aware Model Update" (to appear) AAAI 2024.

Topic 3: Few Shot Learning and Bio-computing

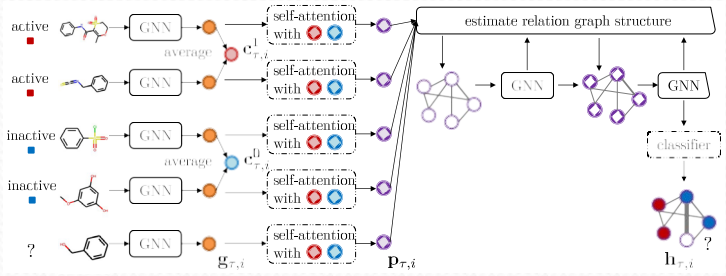
PaddleFSL: Few Shot Learning Toolkit

- **What is it**: A python library for few-shot learning which builds upon PaddlePaddle
- **Target audience**: people who look for fast prototyping FSL solutions
- Our solution built upon PaddleFSL won the fourth place in FewCLUE challenge organized by NLPCC
- Appear at ICML 2021 Baidu Expo

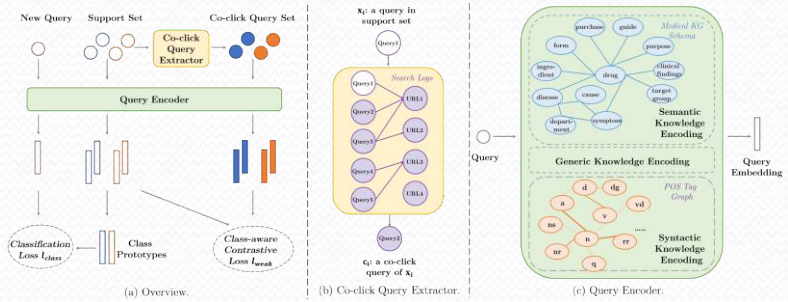


https://github.com/tata1661/FSL-Mate/tree/master/PaddleFSL

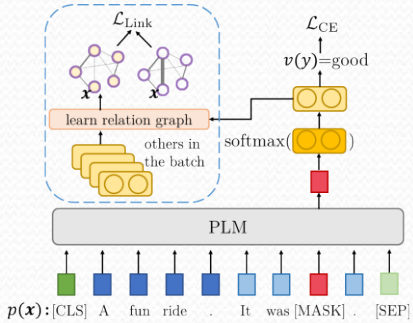
Published Work



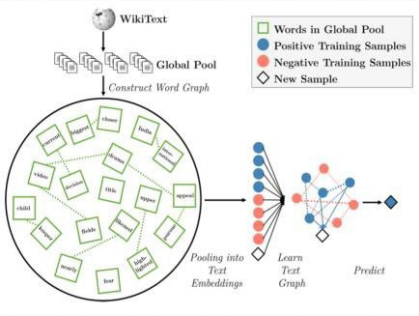
Few-shot Molecular Property Prediction (NeurIPS 2021, TPAMI 2024)



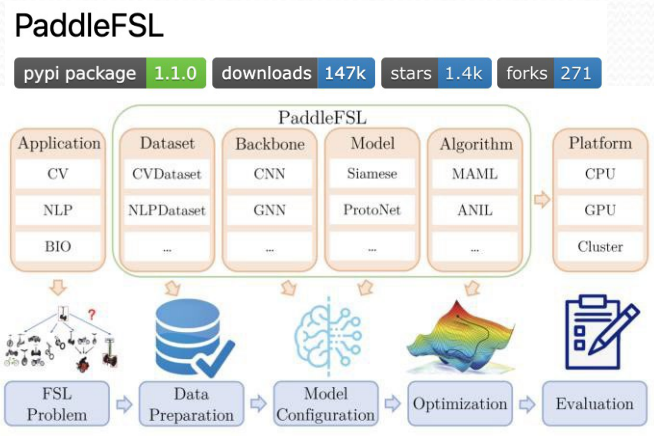
Few-shot Intent Recognition (SIGIR 2021)



Prompt-tuning for Pretrained Language Models (NAACL Findings 2022)



Few-shot Short Text Classification (EMNLP 2021、2022)



FSL Toolkit in PaddlePaddle

Property-aware Adaptive Relation Networks for Molecular Property Prediction

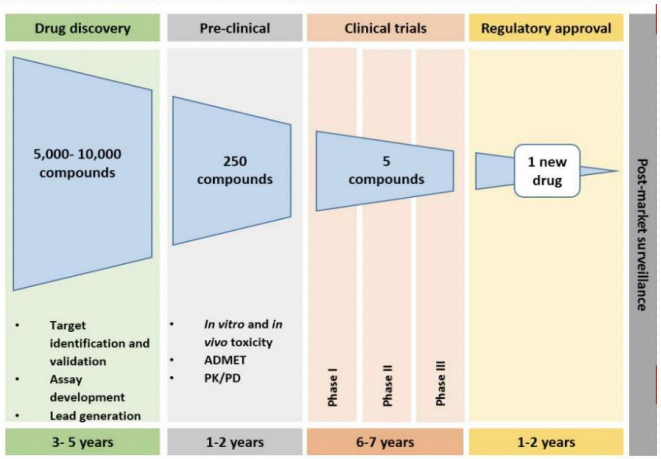
Motivation

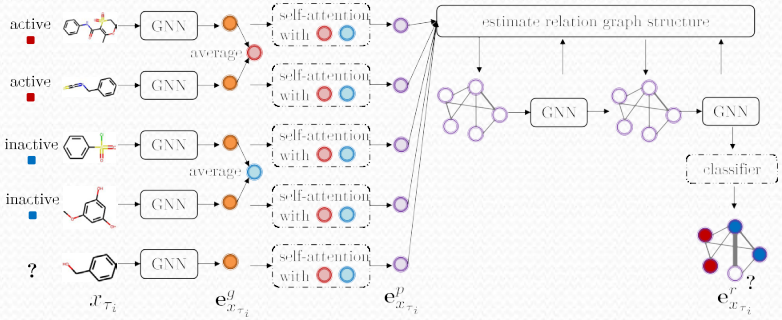
- Molecular property prediction is essentially a few-shot problem
- Existing works ignore property specific structure and relationship

Our Solutions

- Models substructures and molecule relationships w.r.t the target property
- Adopts a selective-update training strategy to separately capture generic and property- aware knowledge
- Consistently outperforms the others under both standard few-shot learning settings and transfer learning across different datasets setting

IJCAI WSRL 2021 Best Paper Runner-up.

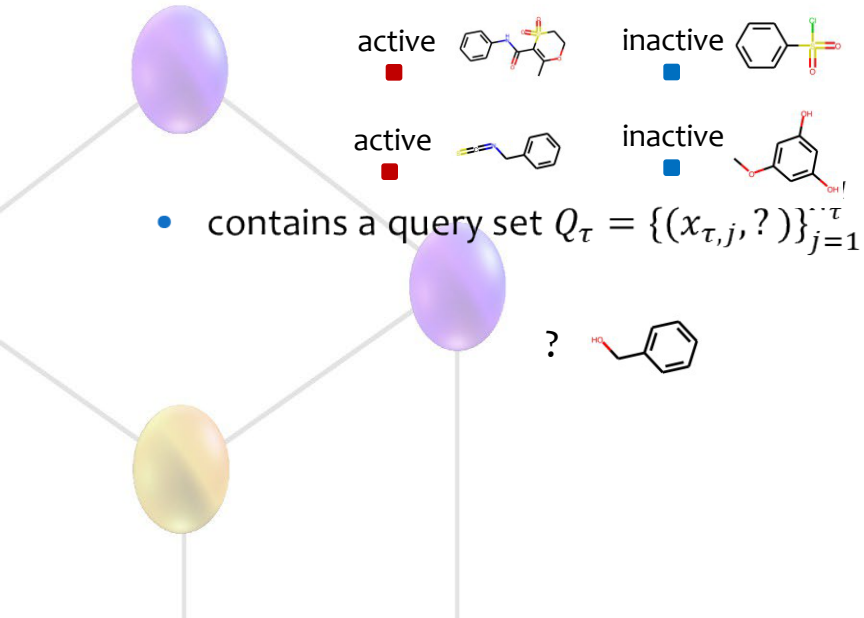




Yaqing Wang, Abulikemu Abuduweili, Quanming Yao, and Dejing Dou. "Property-Aware Relation Networks for Few-Shot Molecular Property Prediction." In Proceedings of NeurIPS 2021. pp. 17441-17454.

Problem Formulation

- Target: learning a predictor from a set of property prediction tasks and generalize to predict new properties with a few labeled molecules
- Each task T_τ is a 2-way K-shot classification task
  - corresponds to an experimental assay testing on whether each molecule $x_{\tau,i}$ is active ($y_{\tau,i} = 1$) or inactive ($y_{\tau,i} = 0$) on a target property
  - contains a support set $S_\tau = \{(x_{\tau,i}, y_{\tau,i})\}_{i=1}^{2K}$, where K is small

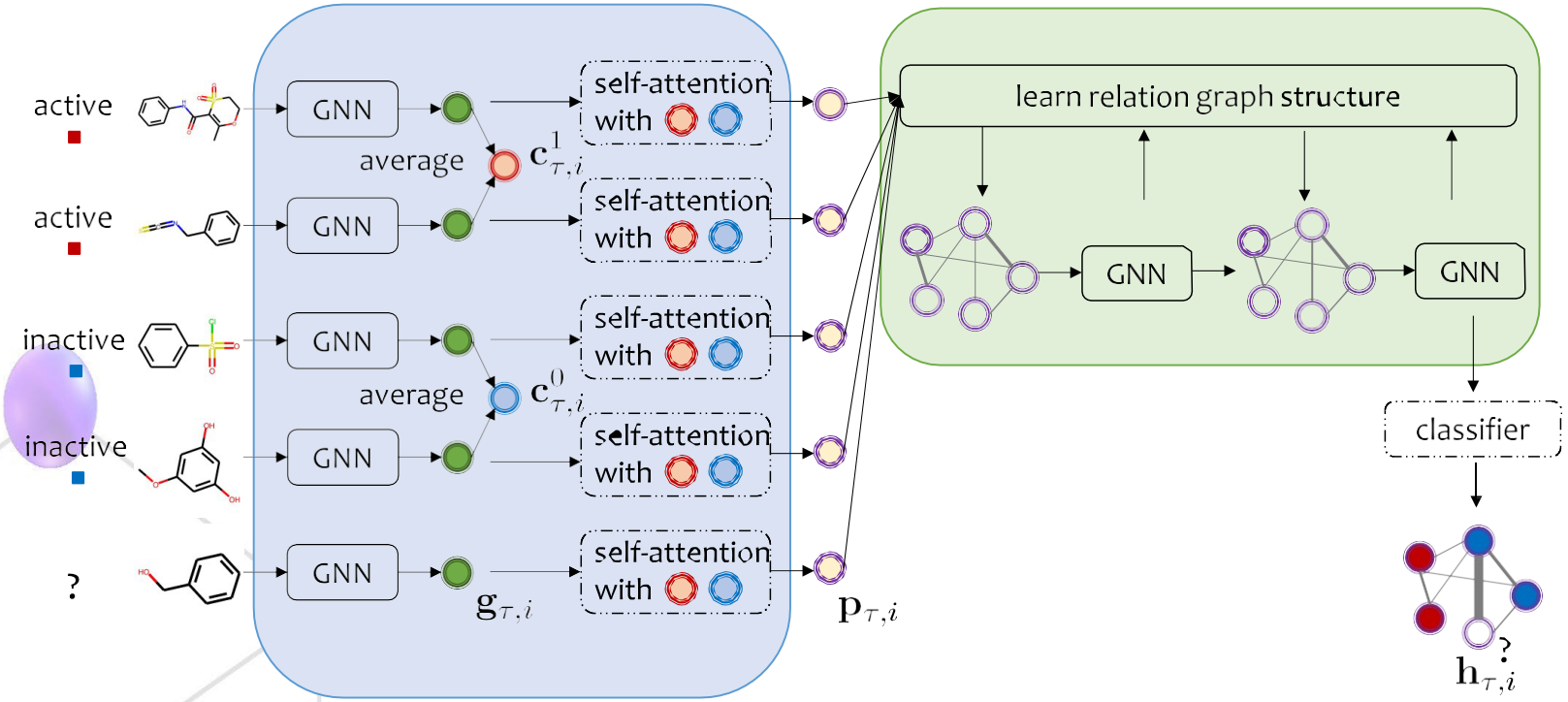


- contains a query set $Q_\tau = \{(x_{\tau,j}, ?)\}_{j=1}^{\cdots\tau}$

An example
2-way: active and inactive
2-shot: each class has two labeled samples

PAR Framework

We propose Property-Aware Relation networks (PAR).



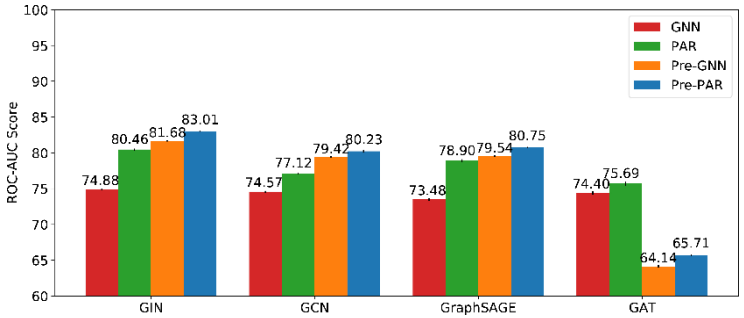
Varying Molecular Encoders

We compare PAR with fine-tuning the encoder (denote as GNN)

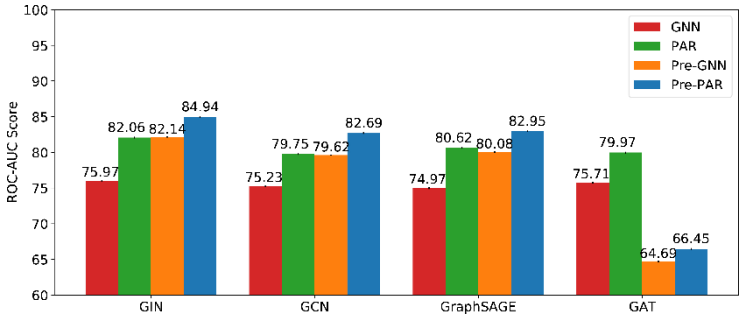
- GIN [Xu et al., 2018] (used)
- GCN [Duvenaud et al., 2015]
- GraphSAGE [Hamilton et al., 2017]
- GAT [Veli˘ckovi´c et al., 2017]

GIN is the consistently better than the others

PAR consistently outperforms GNN



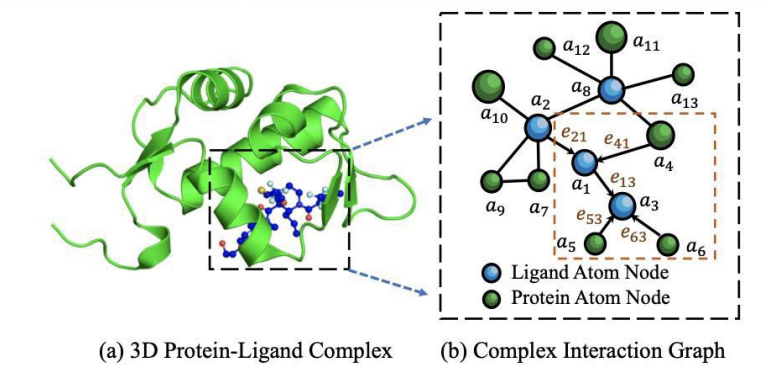
1-shot



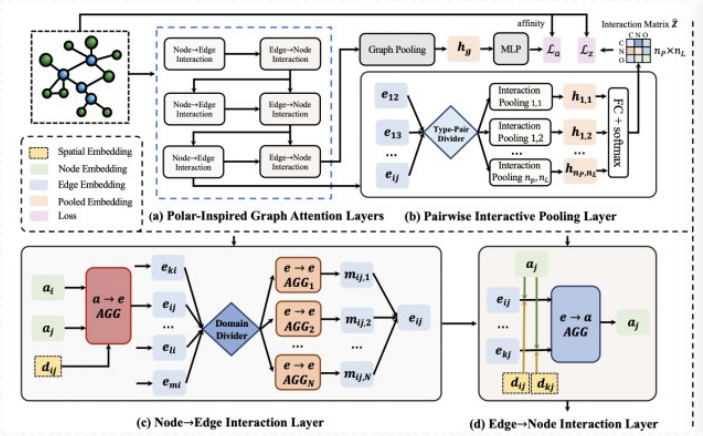
10-shot

Drug-Target Interaction (KDD'21, TKDE'23)

- Predicting Protein – Ligand binding affinity strongly depend on 3D-Structures
- Challenges
  - Rotation Invariant
  - Long-Range Interactions
- SIGN
  - Polar-Inspired Graph Attention Layers
  - Pairwise Interactive Pooling



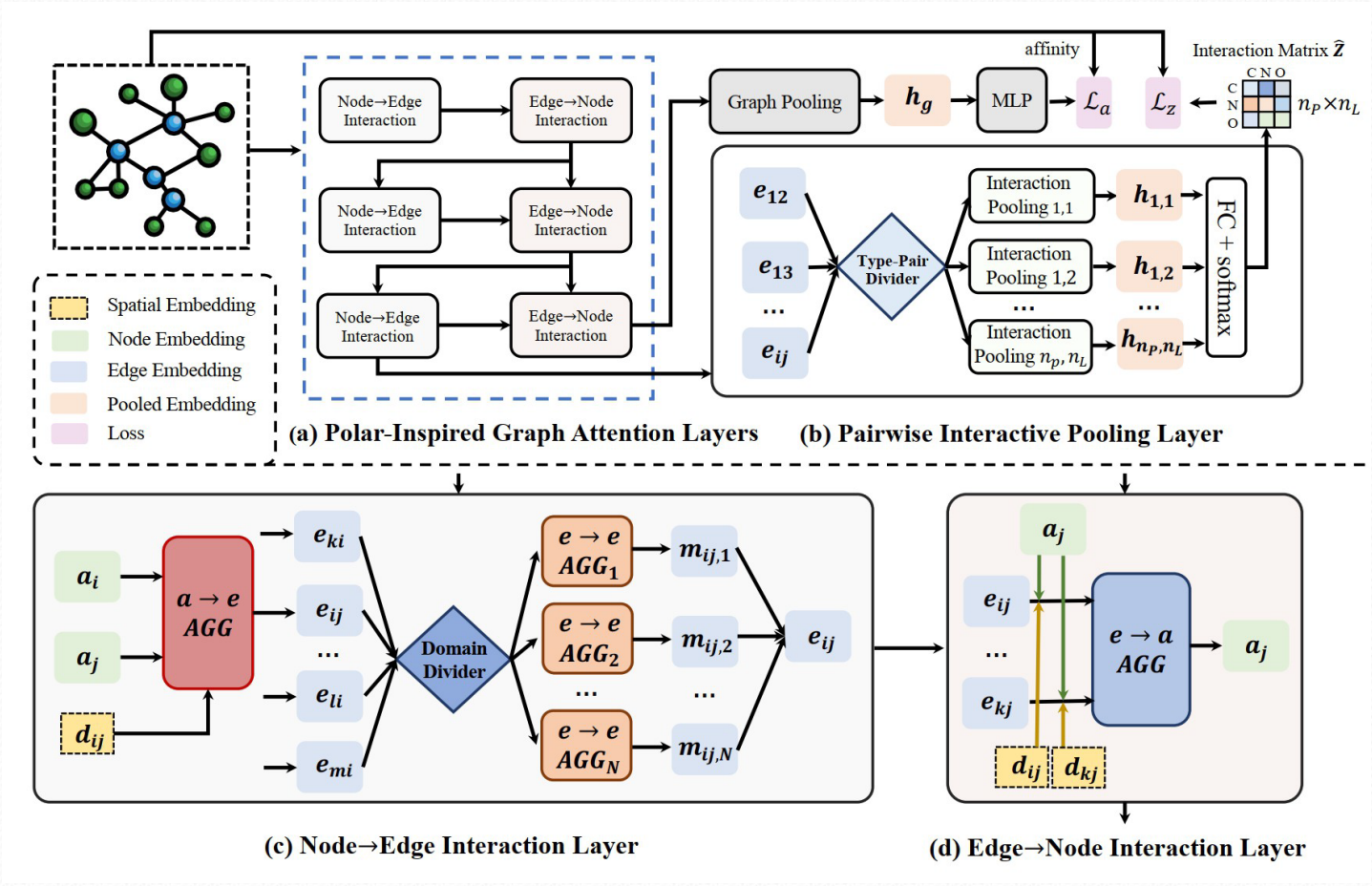
Method		PDBbind core set				CSAR-HiQ set			
		RMSE ↓	MAE ↓	SD ↓	R ↑	RMSE ↓	MAE ↓	SD ↓	R ↑
ML-based Methods	LR	1.675 (0.000)	1.358 (0.000)	1.612 (0.000)	0.671 (0.000)	2.071 (0.000)	1.622 (0.000)	1.973 (0.000)	0.652 (0.000)
	SVR	1.555 (0.000)	1.264 (0.000)	1.493 (0.000)	0.727 (0.000)	1.995 (0.000)	1.553 (0.000)	1.911 (0.000)	0.679 (0.000)
	RF-Score	1.446 (0.008)	1.161 (0.007)	1.335 (0.010)	0.789(0.003)	1.947 (0.012)	1.466 (0.009)	1.796 (0.020)	0.723 (0.007)
CNN-based Methods	Pafnucy	1.585 (0.013)	1.284 (0.021)	1.563 (0.022)	0.695 (0.011)	1.939 (0.103)	1.562 (0.094)	1.885 (0.071)	0.686 (0.027)
	OnionNet	1.407 (0.034)	1.078 (0.028)	1.391 (0.038)	0.768 (0.014)	1.927 (0.071)	1.471 (0.031)	1.877 (0.097)	0.690 (0.040)
GraphDTA Methods	GCN	1.735 (0.034)	1.343 (0.037)	1.719 (0.027)	0.613 (0.016)	2.324 (0.079)	1.732 (0.065)	2.302 (0.061)	0.464 (0.047)
	GAT	1.765 (0.026)	1.354 (0.033)	1.740 (0.027)	0.601 (0.016)	2.213 (0.053)	1.651 (0.061)	2.215 (0.050)	0.524 (0.032)
	GIN	1.640 (0.044)	1.261 (0.044)	1.621 (0.036)	0.667 (0.018)	2.158 (0.074)	1.624 (0.058)	2.156 (0.088)	0.558 (0.047)
	GAT-GCN	1.562 (0.022)	1.191 (0.016)	1.558 (0.018)	0.697 (0.008)	1.980 (0.055)	1.493 (0.046)	1.969 (0.057)	0.653 (0.026)
GNN-based Methods	SGCN	1.583 (0.033)	1.250 (0.036)	1.582 (0.320)	0.686 (0.015)	1.902 (0.063)	1.472 (0.067)	1.891 (0.077)	0.686 (0.030)
	GNN-DTI	1.492 (0.025)	1.192 (0.032)	1.471 (0.051)	0.736 (0.021)	1.972 (0.061)	1.547 (0.058)	1.834 (0.090)	0.709 (0.035)
	DMPNN	1.493 (0.016)	1.188 (0.009)	1.489 (0.014)	0.729 (0.006)	1.886 (0.026)	1.488 (0.054)	1.865 (0.035)	0.697 (0.013)
	MAT	1.457 (0.037)	1.154 (0.037)	1.445 (0.033)	0.747 (0.013)	1.879 (0.065)	1.435 (0.058)	1.816 (0.083)	0.715 (0.030)
	DimeNet	1.453 (0.027)	1.138 (0.026)	1.434 (0.023)	0.752 (0.010)	1.805 (0.036)	1.338 (0.026)	1.798 (0.027)	0.723 (0.010)
	CMPNN	1.408 (0.028)	1.117 (0.031)	1.399 (0.025)	0.765 (0.009)	1.839 (0.096)	1.411 (0.064)	1.767 (0.103)	0.730 (0.052)
Ours	SIGN	**1.316 (0.031)**	**1.027 (0.025)**	**1.312 (0.035)**	**0.797 (0.012)**	**1.735 (0.031)**	**1.327 (0.040)**	**1.709 (0.044)**	**0.754 (0.014)**



Shuangli Li, Jingbo Zhou*, Tong Xu, Liang Huang, Fan Wang, Haoyi Xiong, Weili Huang, Dejing Dou*, Hui Xiong* 2021. "Structure-aware Interactive Graph Neural Networks for the Prediction of Protein-Ligand Binding Affinity." In Proceedings of KDD 2021. pp. 975-985, 2021.

The Proposed Model

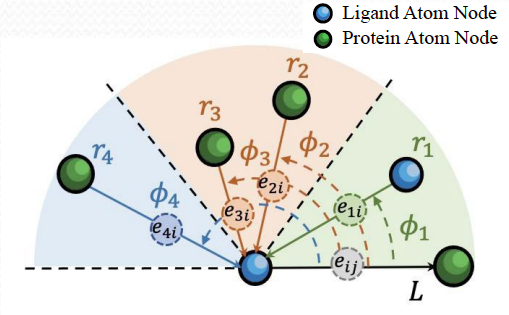
- Structure-aware Interactive Graph Neural Network (SIGN)

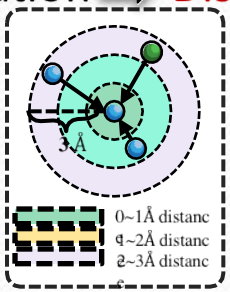


The Proposed Model

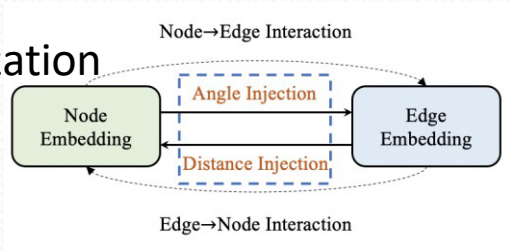
- Polar Coordinate-Inspired Graph Attention

  - Establish a local polar coordinate system in GNN
    - Preserve angle information ⇨ Angle domain division
    - Preserve distance information ⇨ Distance discretization





Angle domain divisionDistance discretization



  - Apply a node↔edge interaction schema

Experimental Results

- Comparison with baselines

Table 2: Performance comparison on *PDBbind core set* and *CSAR-HiQ set*.

Method		PDBbind core set				CSAR-HiQ set			
		RMSE ↓	MAE ↓	SD ↓	R ↑	RMSE ↓	MAE ↓	SD ↓	R ↑
ML-based Methods	LR	1.675 (0.000)	1.358 (0.000)	1.612 (0.000)	0.671 (0.000)	2.071 (0.000)	1.622 (0.000)	1.973 (0.000)	0.652 (0.000)
	SVR	1.555 (0.000)	1.264 (0.000)	1.493 (0.000)	0.727 (0.000)	1.995 (0.000)	1.553 (0.000)	1.911 (0.000)	0.679 (0.000)
	RF-Score	1.446 (0.008)	1.161 (0.007)	1.335 (0.010)	0.789(0.003)	1.947 (0.012)	1.466 (0.009)	1.796 (0.020)	0.723 (0.007)
CNN-based Methods	Pafnucy	1.585 (0.013)	1.284 (0.021)	1.563 (0.022)	0.695 (0.011)	1.939 (0.103)	1.562 (0.094)	1.885 (0.071)	0.686 (0.027)
	OnionNet	1.407 (0.034)	1.078 (0.028)	1.391 (0.038)	0.768 (0.014)	1.927 (0.071)	1.471 (0.031)	1.877 (0.097)	0.690 (0.040)
GraphDTA Methods	GCN	1.735 (0.034)	1.343 (0.037)	1.719 (0.027)	0.613 (0.016)	2.324 (0.079)	1.732 (0.065)	2.302 (0.061)	0.464 (0.047)
	GAT	1.765 (0.026)	1.354 (0.033)	1.740 (0.027)	0.601 (0.016)	2.213 (0.053)	1.651 (0.061)	2.215 (0.050)	0.524 (0.032)
	GIN	1.640 (0.044)	1.261 (0.044)	1.621 (0.036)	0.667 (0.018)	2.158 (0.074)	1.624 (0.058)	2.156 (0.088)	0.558 (0.047)
	GAT-GCN	1.562 (0.022)	1.191 (0.016)	1.558 (0.018)	0.697 (0.008)	1.980 (0.055)	1.493 (0.046)	1.969 (0.057)	0.653 (0.026)
GNN-based Methods	SGCN	1.583 (0.033)	1.250 (0.036)	1.582 (0.320)	0.686 (0.015)	1.902 (0.063)	1.472 (0.067)	1.891 (0.077)	0.686 (0.030)
	GNN-DTI	1.492 (0.025)	1.192 (0.032)	1.471 (0.051)	0.736 (0.021)	1.972 (0.061)	1.547 (0.058)	1.834 (0.090)	0.709 (0.035)
	DMPNN	1.493 (0.016)	1.188 (0.009)	1.489 (0.014)	0.729 (0.006)	1.886 (0.026)	1.488 (0.054)	1.865 (0.035)	0.697 (0.013)
	MAT	1.457 (0.037)	1.154 (0.037)	1.445 (0.033)	0.747 (0.013)	1.879 (0.065)	1.435 (0.058)	1.816 (0.083)	0.715 (0.030)
	DimeNet	1.453 (0.027)	1.138 (0.026)	1.434 (0.023)	0.752 (0.010)	1.805 (0.036)	1.338 (0.026)	1.798 (0.027)	0.723 (0.010)
	CMPNN	1.408 (0.028)	1.117 (0.031)	1.399 (0.025)	0.765 (0.009)	1.839 (0.096)	1.411 (0.064)	1.767 (0.103)	0.730 (0.052)
Ours	SIGN	**1.316 (0.031)**	**1.027 (0.025)**	**1.312 (0.035)**	**0.797 (0.012)**	**1.735 (0.031)**	**1.327 (0.040)**	**1.709 (0.044)**	**0.754 (0.014)**

Our proposed model SIGN achieves the best performance on two benchmarks

Thank you!

dejingdou@gmail.com